The Long-Run Phillips Curve

Luca Benati
University of Bern*

Abstract

I use Bayesian structural VARs identified based on a combination of long-run and sign restrictions to investigate the long-run trade-off between inflation and the unemployment rate in the United States, the Euro area, the U.K., and Canada over the post-WWII period.

Evidence suggests that either of the four structural shocks which are here allowed to introduce a unit root component in the inflation rate has generated a vertical long-run Phillips curve. The only exception is the United Kingdom, for which monetary policy shocks have induced a statistically significant, albeit small, negative long-run Phillips trade-off.

For the Euro area, the U.K., and Canada Johansen's cointegration tests point towards the presence of cointegration between either inflation and unemployment, or inflation, unemployment, and a short-term interest rate, with the long-run Phillips trade-off implied by the estimated cointegrating vectors being negative and sizeable. I argue however that this evidence should be discounted, as, conditional on the estimated structural VARs—which, by construction, do not feature cointegration between any variable—Johansen's procedure tends to spuriously detect cointegration a non-negligible fraction of the times.

Keywords: Inflation; unemployment; Phillips curve; unit roots; cointegration; Bayesian VARs; structural VARs; long-run restrictions; sign restrictions.

*Department of Economics, University of Bern, Schanzeneckstrasse 1, CH-3001 Bern, Switzerland. Email: luca.benati@vwi.unibe.ch

1 Introduction

In spite of the central role played by the unemployment-inflation trade-off in shaping the evolution of both macroeconomic thinking[1] and policymaking over the last several decades, surprisingly little econometric work has been devoted to investigating the nature of the long-run trade-off. Further, the only existing investigation of the long-run Phillips curve based on structural VAR methods—King and Watson (1994)—has produced evidence of a *negative* long-run trade-off conditional on aggregate demand-side shocks for the post-WWII United States, thus suggesting that, in line with the pre-Phelps-Friedman-Lucas-Sargent consensus, it is indeed possible to permanently decrease the unemployment rate by accepting a permanently higher inflation rate.

Nearly two decades after King and Watson (1994), there are however several reasons to reconsider this issue.

First, King and Watson's (1994) finding of a negatively sloped long-run Phillips curve, if correct, would have radical implications for the conduct of monetary policy, as it would imply that the current consensus, within both academia and central banks, that there is no long-run trade-off between inflation and economic activity—with its corollary that the central bank should focus on delivering low and stable inflation—is misplaced. Current monetary frameworks have been built around the notion that there is no long-run trade-off which can be exploited by monetary policy: in spite of its strong conceptual appeal, it is important to know whether such a notion is in fact supported by empirical evidence.

Second, in the years since 1994 structural VAR econometrics has seen important developments in terms of identification. When King and Watson wrote, short-run restrictions were still either of the 'inertial' type—that is, based on imposing zeros in the impact matrix of the structural shocks at t=0—or they were based on the notion of 'calibrating' some of these impacts based on information extraneous to the VAR.[2] In recent years, several contributions have highlighted the dangers associated with the former approach,[3] whereas the reliability of the latter crucially hinges, as a matter of logic, on just how credible the numbers the researcher is imposing in the VAR's structural impact matrix truly are. Since imposing a specific number entails making a very strong assumption—implying an extent of knowledge we typically

[1] See in particular Lucas (1972a), Lucas (1972b), and Lucas (1973).

[2] Indeed, this is how King and Watson (1994) achieved identification in their preferred specification. As stressed by Evans (1994) in his comment on King and Watson (1994), *'[i]dentification of the supply and demand shocks is achieved by imposing a value for* λ [the parameter which determines the impact of demand shocks on the unemployment rate at $t = 0$] a priori *in the empirical analysis.'*. (See Evans, 1994, p. 222.)

[3] The work of Fabio Canova and his co-authors, in particular (see, first and foremost, Canova and Pina (2005)) has demonstrated that, since inertial restrictions are, in general, incompatible with the structure of general equilibrium models—in the specific sense that, within DSGE models, the impact matrix of the structural shocks at t=0 is, in general 'full', i.e., it has no zero entries—imposing such zeros can lead to dramatically distorted inference, for example 'identifying' price puzzles which are not in the data generation process.

do not have—an alternative style of identification based on weaker informational requirements might be regarded as preferrable. Several researchers[4] have therefore proposed sign restrictions—that is, restrictions on the signs of the impacts of the structural shocks at t=0, and possibly on their impulse-response functions at longer horizons—as the best (or least bad ...) way of achieving identification based on short-run restrictions. As shown by Canova and Paustian (2011), indeed, DSGE models often imply a robust pattern of signs for the impacts of the structural shocks at t=0 (where 'robust' means that such pattern holds true for alternative sub-classes of DSGE models, and for a wide range of plausible parameters' configurations), which is often sufficient to disentangle the structural shocks from one another. In fact, when seen from the perspective of DSGE models, a specific pattern of signs for the impacts of the structural shocks at t=0 is typically the *only* kind of information we can be reasonably confident about, whereas the specific values taken by such impacts are, in general, much more uncertain, thus raising doubts on the reliability of an approach to identification based on the notion of calibrating such impacts.[5]

Third, King and Watson's analysis was entirely based on a bivariate VAR for the first differences of inflation and the unemployment rate, but, as shown by Evans (1994) in his comment,[6] even accepting their identification strategy, evidence based on trivariate VARs was sometimes significantly different, pointing in some cases towards a vertical long-run Phillips curve.[7] Further, for the reason discussed, e.g., by Sargent (1987)—the first-difference filter largely wipes out variance at the business-cycle frequency—since the *level* of the U.S. unemployment rate is highly informative about the state of the business cycle, its *first difference*, as a matter of logic, is not (this argument holds to an even greater extent for inflation, which is less informative about the state of the business cycle than the unemployment rate). This is potentially problematic since correctly identifying permanent shocks to macroeconomic variables ultimately hinges on the VAR containing sufficient information about transitory dynamics.

For all of these reasons, it is of interest to reconsider the issue of the long-run Phillips trade-off taking into account of the developments in structural VAR econometrics since 1994.

---

[4] See in particular Faust (1998), Canova and de Nicolo (2002), and Uhlig (2005).

[5] An approach to identification based on sign restrictions is not without problems of its own. As extensively discussed by Fry and Pagan (2007), in particular, sign restrictions suffers from the shortcoming that they are intrinsically 'weak information', and therefore they should not be expected to produce strong inference.

[6] See Evans (1994, Section 3.2, and in particular the results reported in Figure 2).

[7] More generally, it is unlikely that the first differences of inflation and the unemployment rate provide a 'minimal statistical summary' of the economy, thus capturing the economy's fundamental driving forces. In particular, King and Watson's VAR, by eschewing either the Federal Funds rate, or an *ex post* real rate, does not contain strong information about the monetary policy stance, which is potentially problematic given that the existence (or not) of a long-run negative trade-off which can be exploited by monetary policy is a main object of interest.

1.1 This paper: methodology and main results

In this paper I use structural VARs identified *via* a combination of long-run and sign restrictions in order to investigate the long-run trade-off between inflation and the unemployment rate induced by both demand- and supply-side shocks in the United States, the Euro area, United Kingdom, and Canada over the post-WWII period. Specifically, since—as originally pointed out by Sargent (1971) and Lucas (1972a)—a *necessary* condition in order to be able to investigate the slope of the long-run Phillips trade-off is that inflation does contain permanent shocks,

(i) I separate the VAR's structural shocks into two *sets*, depending on the fact that they do, or they do not have a permanent impact on inflation, and
(ii) I further disentangle the four shocks which are here allowed to exert a permanent impact on inflation into demand- and supply-side ones, by imposing Canova and Paustian's (2011) DSGE-based 'robust sign restrictions' on their impact on the endogenous variables at t=0.

The main findings from the benchmark set of results, which are based on the GDP deflator, and on imposing sign restrictions only on impact—can be summarized as follows.

For the United States, over the full sample period (1953Q2-2011Q4) augmented Dickey-Fuller tests strongly reject the null of a unit root in the unemployment rate, thus implying that the U.S. long-run Phillips curve is vertical.[8] Further, evidence of a unit root in the inflation rate is weak, as it cannot be rejected based on the GDP deflator, but it can instead be strongly rejected based on the CPI, thus implying that, even if the unemployment rate did in fact contain a unit root component, analyses of the long-run Phillips trade-off based on this sample should be viewed with suspicion. On the other hand, results for the pre-Volcker period, for which it is not possible to reject the null of a unit root in either GDP deflator or CPI inflation, and in the unemployment rate, suggest that neither of the four structural shocks which are here allowed to exert a permanent impact on inflation has generated a non-vertical long-run Phillips curve. The extent of uncertainty associated with estimated long-run trade-offs is however substantial. This is due to the fact that we are here estimating a feature of the data pertaining to the infinite long-run, and, as it is well known,[9] this inevitably produces imprecise estimates, unless the researcher is willing to impose upon the data very strong informational assumptions (which is not advisable). Within the present context, this problem is compounded by the use of sign restrictions, which, as stressed by Fry and Pagan (2007), are intrinsically 'weak information', and should therefore not be expected to produce strong inference.

[8] King and Watson (1994) could not reject the null of a unit root in either inflation or the unemployment rate for the period up to 1993. Indeed, for their sample period I get exactly their same results.

[9] See e.g. Faust and Leeper (1998).

Turning to the Euro area and the two inflation-targeting countries—for all of which evidence of a unit root in either inflation or the unemployment rate is strong—results for the Euro area and Canada are even weaker than for the United States, with the fractions of draws from the posterior distribution for which the estimated long-run trade-offs are negative ranging between 40.6 and 77.9 per cent for the Euro area, and between 41.9 and 78.0 per cent for Canada. For the United Kingdom, on the other hand, there is a significant difference between monetary shocks—for which the fraction of draws for which the associated long-run trade-off is estimated to be negative is equal to 90.4 per cent, and is therefore marginally significant at the 10 per cent level—and other shocks, for which the corresponding fractions of draws range between 33.5 and 59.4 per cent. One possible way to interpret this result is as it being largely due to the Thatcher disinflation of the early 1980s, which, in the presence of hysteresis in the labor market, slashed inflation by causing, at the same time, a permanent increase in the unemployment rate.

For the Euro area, the U.K., and Canada Johansen's cointegration tests point towards the presence of cointegration between either inflation and unemployment, or inflation, unemployment, and a short-term interest rate, with the long-run Phillips trade-off implied by the estimated cointegrating vectors being negative and sizeable. I argue however that this evidence should be discounted, as, conditional on the estimated structural VARs—which, by construction, do not feature cointegration between *any* variable—Johansen's procedure tends to spuriously detect cointegration a non-negligible fraction of the times. For example, for the Euro area and the United Kingdom, the fractions of bootstrapped p-values for Johansen's trace statistic for testing the null of no cointegration between inflation and the unemployment rate which are smaller than 10 per cent are equal to 22.9 and 24.2 per cent respectively. This means that, if the estimated structural VARs were the true data-generation process, Johansen's trace test would incorrectly reject the null of no cointegration between inflation and unemployment at the 10 per cent level between one-fifth and one-fourth of the times. As for cointegration between inflation, unemployment, and the short rate, results are even worse. For Canada, for example, conditional on the estimated structural VAR being the true data-generation process, Johansen's trace test statistic would incorrectly reject the null of no cointegration between the three variables at the 10 per cent level a remarkable 31.5 per cent of the times. As I show *via* Monte Carlo, this is *not* the product of the comparatively short samples I am working with, as the fraction of simulations for which the bootstrapped trace statistic incorrectly rejects the null of no cointegration between two independent random walks at a given significance level ranges between 11.3 and 11.9 per cent at the 10 per cent level; between 5.5 and 6.0 per cent at the 5 per cent level; and between 1.1 and 1.3 per cent at the 1 per cent level. So it truly appears that, if the estimated structural VARs were the authentic data-generation processes, cointegration tests would be biased towards spuriously detecting cointegration between inflatiion and unemployment (and possibly, the short rate).

Summing up, the benchmark set of results based on the GDP deflator, and on imposing sign restrictions only on impact, produces evidence of a non-zero long-run Phillips trade-off only for the United Kingdom conditional on monetary shocks. Results based on the CPI are slightly stronger. For the Euro area for example, the fraction of draws from the posterior for which the long-run trade-off induced by monetary shocks is estimated to have been negative is equal to 82.3 per cent, compared to the 77.9 per cent based on the GDP deflator. However, since, in general, the GDP deflator should be regarded as a better measure of inflationary pressures than the CPI, which by construction is affected by food and energy price shocks to a much greater extent, I argue that this set of results should be somehow discounted.

The paper is organized as follows. The next two sections discuss the choice of the sample periods and present results from unit root tests for inflation, the unemployment rate, and the short rate, whereas Section 4 discusses the Bayesian methodology for reduced-form VAR estimation, and the identification strategy I use, based on a combination of long-run and sign restrictions imposed on impact. Section 5 discusses the evidence, whereas Section 6 discusses the issue of the possible (in principle) presence of cointegration between either inflation and the unemployment rate, or inflation, the unemployment rate, and the short rate. Section 7 concludes.

2 Choosing the Sample Periods

As extensively discussed by King and Watson (1994), building upon Sargent (1971) and Lucas (1972a), a necessary condition in order to be able to identify the long-run trade-off between inflation and the unemployment rate is that both series contain permanent shocks. Under this respect, for the Euro area and the two inflation-targeting countries considered herein the choice of the sample period is crucial. As I have documented elsewhere, indeed—see Benati (2008)—either in the Euro area under European Monetary Union (henceforth, EMU), or in the U.K., Canada, Sweden, and Australia under inflation targeting, inflation has exhibited essentially no persistence, and it has often been statistically indistinguishable from white noise. On the other hand, for either country inflation had exhibited very high persistence, to the point that the null of a unit root could not be rejected, for the periods before the introduction of the current monetary regimes. In what follows I therefore consider, for the Euro area, the period 1970Q1-1998Q4;[10] for the United Kingdom, the period 1972Q2-1992Q3;[11] and for Canada, the period 1961Q2-1990Q4.[12]

[10] EMU started in January 1999, whereas Euro area data are only available starting from 1970Q1.

[11] June 23, 1972 marks the floating of the pound *vis-à-vis* the U.S. dollar, whereas inflation-targeting was introduced on October 8, 1992. As shown by Benati (2008), before the June 1972 floating of the pound U.K. inflation exhibited quite significantly lower persistence.

[12] Canada introduced inflation targeting in February 1991.

3 Results from Unit Root Tests

Table 1 reports bootstrapped p-values for augmented Dickey-Fuller (henceforth, ADF) tests for inflation, the unemployment rate, a short-term interest rate, and several other series of interest, for the United States, the Euro area, the United Kingdom, Japan, Canada, and Australia. For either series, p-values have been computed by bootstrapping 10,000 times estimated ARIMA(p,1,0) processes. In all cases, the bootstrapped processes are of length equal to the series under investigation.[13] 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, once again for reasons of robustness I consider three alternative lag orders, two, four, and eight.

3.1 Inflation and the unemployment rate

Starting from inflation, the null of a unit root cannot be rejected based on either the GDP deflator or the CPI for the Euro area, the United Kingdom, and Canada. For Australia, it cannot be rejected based on the CPI, but it can be strongly rejected based on the GDP deflator. By the same token, evidence for Japan is not strong, with two p-values just above 5 per cent based on the GDP deflator. Because of such a comparatively weak evidence of a unit root in inflation for either Japan or Canada, in what follows I exclude the two countries from the analysis, and I uniquely focus on the remaining four. Finally, for the United States evidence based on the full sample period is weak, with the p-values not being statistically significant at conventional levels based on the GDP deflator, but being consistently below 5 per cent based on the CPI. This suggests that results for the U.S. based on the full sample period should be viewed with suspicion,[14] and in what follows I will therefore focus on the pre-Volcker sample period, for which evidence of a unit root in inflation is instead strong based on either the CPI or the GDP deflator.

Evidence of a unit root in the unemployment rate is strong for either the Euro area, the United Kingdom, or Canada. For the United States it is strong for the pre-Volcker sample period, but it is weak for the full sample period, with the p-values being consistently below 10 per cent. Rejection of a unit root in the unemployment rate automatically implies that the U.S. long-run Phillips curve is vertical, so that even if there have indeed been permanent shocks to the inflation rate, they have not exerted any permanent impact on unemployment. If, based on the results reported in Table 1, one is willing to accept the notion that the U.S. unemployment rate does not contain any permanent component, then, for the United States, there is no point to

[13]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.

[14]For the sample period 1948Q1-2011Q4 evidence against the notion of a unit root in inflation is even stronger, with bootstrapped p-values based on two, four, and eight lags being equal to 0.002, 0.004, and 0.013 based on the CPI, and to 0.005, 0.008, and 0.103 based on the GDP deflator.

proceed further. In what follows, however, I want to give the U.S. long-run Phillips curve its best chance of revealing itself, and I therefore proceed by focusing on the pre-Volcker sample period. As we will see, even doing this I am not able to identify a non-vertical long-run Phillips curve in the United States.

3.2 Other series

One possible limitation of King and Watson's (1994) analysis was its bivariate nature, and their eschewing of potentially important information on the monetary policy stance and the state of the business cycle.[15] Beyond the first differences of inflation and the unemployment rate, in what follows I therefore also include in the VAR real GDP growth, either the level or the first difference of a short-term interest rate, and a variable capturing the state of the business cycle.

As for real GDP growth, rejection of a unit root is uniformly strong for either country and lag length, with the bootstrapped p-values ranging between zero and 6.1 per cent.

As for the short rate, evidence of a unit root is strong for all countries, and based on either lag order, with the single exception of the United Kingdom for which p-values range between 1.8 and 2.7 per cent. In what follows, the short rate will therefore enter the VAR in levels for the United Kingdom, and in first differences for all other countries.

Turning to variables which contain information on the state of the business cycle, the consumption-GDP ratio exhibits obvious trends in both the United Kingdom and Canada, and indeed bootstrapped p-values are uniformly high, and do not allow to reject the null of a unit root. For the United States, rejection of a unit root is very strong based on the full sample period, whereas for the pre-Volcker period p-values range between 4.8 and 15.4 per cent, and therefore do not allow, strictly speaking, to reject a unit root with a high confidence. An important point to stress, however, is that, as extensively discussed by Cochrane (1994), failure to reject a unit root in the consumption-GDP ratio over comparatively short sample periods should be dismissed, as (i) economic theory *does* imply that consumption and GDP are cointegrated, so that their ratio *should* be stationary as a matter of logic, and (ii) precisely because the consumption-GDP ratio is such a good proxy for the transitory component of GDP, it is very highly persistent, and in short samples unit root tests may therefore spuriously point towards the presence of a unit root. Because of this, in what follows I use, for the pre-Volcker United States, the consumption-GDP ratio as the stationary variable capturing the state of the business cycle even if, strictly speaking, results

[15] As pointed out by Evans (1994) in his comment '*[...] it would be interesting to know how well these econometric methods actually recover the fundamental driving processes of an economy. For example, in general equilibrium models with technology shocks, monetary policy shocks, and fiscal shocks, what do these bivariate identifications actually recover from simulated data?.*' (See Evans, 1994, p. 229.)

from ADF tests do not allow to reject a unit root over this specific sample. By the same token, for the Euro area p-values range between 2.7 and 11.1 per cent, but for the longer period 1970Q1-2011Q4 they allow us to confidently reject a unit root, being equal to 0.005, 0.008, and 0.062 respectively. In this case, too, I discount lack of a strong rejection for the shorter period based on Cochrane's (1994) argument, and I include the consumption-GDP ratio in the VAR for the Euro area. Things are more problematic for the United Kingdom and Canada: as I pointed out, for these countries the consumption-GDP ratio exhibits obvious trends, and results from unit root tests cannot therefore be confidently dismissed based on Cochrane's (1994) argument. I therefore consider three other variables containing information about the state of the business cycle: the investment-GDP ratio, a long-short sprerd, and the ratio between the nominal change in inventories and nominal sales. For Canada, a unit root can be rejected for the investment-GDP ratio,[16] whereas for the United Kingdom only the ratio between the change in inventories and sales is clearly stationary, with p-values ranging between 1.4 and 2.6 per cent.

4 Methodology

4.1 VAR estimation

Since identification will be based on a mixture of long-run restrictions, which can be effectively imposed within either a Classical or Bayesian framework, and sign restrictions, which are conceptually easier to impose within the latter setup,[17] I implement VAR estimation within a Bayesian framework as in Uhlig (1998) and Uhlig (2005). Specifically, I *exactly* follow Uhlig (1998, 2005) in terms of both distributional assumptions—the distributions for the VAR's coefficients and its covariance matrix are postulated to belong to the Normal-Wishart family—and of priors. For estimation details the reader is therefore referred to either the Appendix of Uhlig (1998), or to Appendix B of Uhlig (2005). Following Uhlig (1998, 2005), who sets $p = 12$ with monthly data, I correspondingly set $p = 4$ with quarterly data. Finally, for each estimated VAR I consider 10,000 draws from the posterior distribution of the VAR's coefficients and covariance matrix of innovations (the draws are computed exactly as in Uhlig (1998, 2005)).

[16] The p-value is equal to 10.3 per cent with a lag order equal to 2, but it decreases to 4.8 and 1.0 per cent for four and eight lags, respectively. Overall, since, as a very general rule, results based on longer lag lengths should be regarded as more reliable than those based on shorter lag lengths (the reason being that the feature we are here investigating—the presence or absence of a unit root—pertains to the infinite long run, which longer lag lengths should reasonably be expected to capture with greater reliability) I regard evidence as pointing towards rejection of the null of a unit root.

[17] On this, see the discussion in Uhlig (1998), or Uhlig (2005).

4.2 Series entering the VAR

Summing up, for all countries the VAR features a constant, the first differences of inflation and the unemployment rate, real GDP growth, and either the level of short rate (for the United Kingdom) or its first difference (for all other countries). As for the stationary series capturing the state of the business cycle, for both the United States and the Euro area they are the consumption-GDP ratio and the spread between the long- and the short-term interest rate; for Canada it is the investment-GDP ratio; and for the United Kingdom it is the ratio between the change in inventories and sales.

4.3 Identification

My identification strategy is based on a combination of long-run and sign restrictions. Since a necessary condition in order to be able to meaningfully investigate the slope of the long-run Phillips trade-off is that inflation does contain permanent shocks, I start by separating the VAR's structural shocks into two *sets*, depending on the fact that they do, or they do not have a permanent impact on inflation.

Let the structural VAR(p) model be given by

$$Y_t = B_0 + B_1 Y_{t-1} + ... + B_p Y_{t-p} + A_0 \epsilon_t \tag{1}$$

where $Y_t \equiv [\Delta y_t,\ \Delta\pi_t,\ \Delta U_t,\ R_t,\ S_t]'$ for the United Kingdom, and $Y_t \equiv [\Delta y_t,\ \Delta\pi_t,\ \Delta U_t,\ \Delta R_t,\ S_t]'$ for all other countries, with y_t, π_t, U_t and R_t being log real GDP, inflation, the unemployment rate, and the short rate, respectively, and S_t being (the vector of) the additional stationary variable(s); A_0 being the impact matrix of the structural shocks at $t = 0$; and $\epsilon_t \equiv [\epsilon_t^{TE},\ \epsilon_t^{MO},\ \epsilon_t^{TA},\ \epsilon_t^{MK},\ \epsilon_t^{TR\prime}]'$ being the structural shocks, which, as standard practice, are assumed to be unit-variance and orthogonal to one another, with ϵ_t^{TE}, ϵ_t^{MO}, ϵ_t^{TA}, ϵ_t^{MK} being Canova and Paustian's (2011) 'technology', 'monetary policy', 'taste', and 'markup' shocks (to be discussed shortly), which are here allowed to exert a permanent impact on inflation, and $\epsilon_t^{TR\prime}$ being instead a (vector of) shock(s) which, by construction, has (have) a transitory impact on inflation. The second row of the matrix of long-run impacts of the structural shocks, $[I_N - B(1)]^{-1} A_0$—i.e., the row corresponding to inflation—is therefore postulated to have the following structure,

$$\text{Long-run impacts of the structural shocks on } \pi_t\text{:}\ \begin{array}{c} \begin{array}{ccccc} \epsilon_t^{TE} & \epsilon_t^{MO} & \epsilon_t^{TA} & \epsilon_t^{MK} & \epsilon_t^{TR\prime} \end{array} \\ \left[\begin{array}{ccccc} x & x & x & x & 0_{1\times K} \end{array}\right] \end{array} \tag{2}$$

—where a '0' means that the corresponding long-run impact has been restricted to zero, whereas an 'x' means that it has been left unrestricted, and $K = 2$ for the U.S. and the Euro area, and $K = 1$ for the U.K. and Canada—thus implying that ϵ_t^{TE}, ϵ_t^{MO}, ϵ_t^{TA}, and ϵ_t^{MK} have a permanent impact on inflation, whereas $\epsilon_t^{TR\prime}$ does not. The restriction that $\epsilon_t^{TR\prime}$ is the only (vector of) shock(s) which does not have a permanent impact on inflation is sufficient to disentangle it from the other four shocks. As for separating ϵ_t^{TE}, ϵ_t^{MO}, ϵ_t^{TA}, and ϵ_t^{MK} from one another, I achieve that by imposing the following set of sign restrictions on impact:

	Shock:			
Variable:	ϵ_t^{TE}	ϵ_t^{MO}	ϵ_t^{TA}	ϵ_t^{MK}
Real GDP growth	+	-	+	-
Inflation	-	-	+	+
Unemployment rate	+	+	-	+
Short rate	-	+	+	+

where '+' means 'greater than, or equal to zero', and '-' means 'smaller than, or equal to zero'. In words, a technology shock causes real GDP growth and the unemployment rate not to decrease, and inflation and the short rate not to increase; a monetary shock causes real GDP growth and inflation not to increase, and unemployment and the short rate not to decrease; a taste shock causes unemployment not to increase, and all other variables not to decrease; and a markup shock causes real GDP growth not to increase, and all other variables not to decrease. These restrictions are the same as the 'robust sign restrictions' reported by Canova and Paustian (2011) in their Table 2 for their benchmark DSGE model featuring sticky prices, sticky wages, and several standard frictions (see the column labelled as 'M'), with the only obvious difference that, since their model features employment, instead of the unemployment rate, the signs we are here imposing on the unemployment rate are the opposite of those reported by Canova and Paustian for employment.

In what follows I impose these sign restrictions only on impact. The reason for doing this is that, as stressed by Canova and Paustian (2011), whereas sign restrictions on impact are, in general, robust—in the specific sense that they hold for the vast majority of sub-classes within a specific class of DSGE models, and for the vast majority of plausible parameters' configurations—restrictions at longer horizons are instead, as they put it, 'whimsical', meaning that they are hard to pin down, and in general, they are not robust across sub-classes of models, and for alternative plausible parameters' configurations.[18]

[18] One obvious limitation of imposing the sign restrictions only on impact is that we are here using a comparatively limited amount of information in order to achieve identification. As a consequence, our results necessarily end up being less sharp than they could have been had we been reasonably confident about imposing a specific pattern of sign restrictions at horizons greater than zero. This compounds a well-known limitation of sign restrictions which has been extensively discussed by Fry and Pagan (2007): as these authors stress, sign restrictions are intrinsically 'weak information', since they are based on the notion of uniquely imposing a specific pattern of signs on the IRFs. The rationale behind our decision of imposing sign restrictions only on impact is that it is better to impose a limited amount of information about which we can be reasonably confident than a greater amount of information about which we have limited confidence.

4.4 Computing the structural impact matrix A_0

For each draw from the posterior distribution, I compute the structural impact matrix, A_0 by combining the procedure proposed by Rubio-Ramirez, Waggoner, and Zha (2005) for imposing sign restrictions[19] with the imposition of the previously discussed zero restrictions on the matrix of long-run impacts of the structural shocks, $[I_N - B(1)]^{-1}A_0$, by means of a deterministic rotation matrix. Specifically, for draw j from the posterior distribution of the VAR's estimates, for j = 1, 2, 3, ..., 10,000, let $P_j D_j P_j'$ be the eigenvalue-eigenvector decomposition of the VAR's covariance matrix, Ω_j, and let $\tilde{A}_{0,j} \equiv P_j D_j^{\frac{1}{2}}$. I draw an $N \times N$ matrix, K, from the $N(0, 1)$ distribution, I take the QR decomposition of K—that is, I compute matrices Q and R such that $K = Q \times R$—and I compute the 'starting estimate' of the structural impact matrix as $\bar{A}_{0,j} = \tilde{A}_{0,j} \cdot Q'$. I then impose the zero restrictions in the second row of the matrix of the long-run impacts of the structural shocks *via* an appropriate Householder matrix[20] H. If the resulting structural impact matrix $A_{0,j} = \bar{A}_{0,j}H$ satisfies the sign restrictions I keep it, otherwise I discard it and I repeat the procedure until I obtain an impact matrix which satisfies both the sign restrictions and the long-run restriction at the same time.

4.5 Definition of the long-run Phillips curve

In line with King and Watson (1994), I define the slope of the long-run Phillips curve as the ratio between the long-run impact of the relevant shock on U_t and its long-run impact on π_t. For each draw from the posterior distribution I compute this ratio, thus obtaining the posterior distribution of the slope of the long-run Phillips curve.

5 Evidence

In this section I discuss the empirical evidence. I start by discussing the benchmark set of results based on the GDP deflator, and on imposing sign restrictions only on impact, and I then briefly mention results based on the CPI, and on imposing sign restrictions not only at zero, but also at longer horizons.

5.1 Baseline results

Figure 2 shows, for either country, the posterior distributions of the slopes of the long-run Phillips trade-offs induced by individual shocks, whereas Figures 3 and 4 show the posterior distributions of the fractions of the long-run variance of inflation and the unemployment rate, respectively, explained by each shock. Table 2 reports

[19] See at http://home.earthlink.net/~tzha02/ProgramCode/SRestrictRWZalg.m.

[20] I compute the Householder matrix *via* Algorithm 5.5.1 of Golub and VanLoan (1996).

the modes, the medians, and the 90%-coverage percentiles of the posterior distributions of the slopes of the long-run Phillips trade-offs induced by individual shocks, together with the fractions of draws for which the slopes are estimated to be negative. Finally, Tables 3 and 4 report, for inflation and the unemployment rate, respectively, the medians and the 90%-coverage percentiles of the posterior distributions of the fractions of long-run variance of either variable explained by each individual shock, together with the fractions of the mass of the posterior distribution which are below three selected 'cut-off points', 0.1, 0.05, and 0.01 (this is in order to provide a numerical measure of how strongly clustered towards zero such distributions are).

The main findings can be summarised as follows.

First, the extent of uncertainty associated with estimated long-run trade-offs is uniformly quite substantial. This originates from the fact that the feature of the data we are here estimating pertains to the infinite long-run, and, as it is well known (see, first and foremost, Faust and Leeper (1998)), this inevitably produces imprecise estimates, unless the researcher is willing to impose upon the data very strong informational assumptions (which, in general, is not advisable to do). Within the present context, this problem is compounded by our use of sign restrictions, which, as stressed by Fry and Pagan (2007), are intrinsically 'weak information', and should therefore not be expected to produce strong inference. On the other hand, it has to be stressed that, for the reasons discussed in the Introduction, the approach to identification adopted herein is, most likely, the most credible (or, to be more precise, the least incredible) one given the current state-of-the-art in structural VAR econometrics. To put it differently, given the lack of reliability of inertial restrictions extensively documented, e.g., by Canova and Pina (2005), and the unattractiveness of 'calibrating' elements of the VAR's structural impact matrix, it is not clear *what* kind of short-run restrictions can be combined with long-run ones in order to disentangle the various shocks exerting a permanent impact on inflation, other than sign restrictions. In a sense, this is the best we can do given the current state of knowledge, and if the price to be paid for using the least incredible identification strategy is a significant extent of uncertainty, we ought to live with it ...

Second, largely as a result of such a significant extent of uncertainty, with a *single* exception estimated trade-offs are never significantly different from zero at conventional levels for either country and either shock. For the United States, for example, the fractions of draws for which the slope is estimated to be negative range between 24.6 per cent for the markup shock to 81.7 per cent for the monetary shock, whereas for Canada they range between 41.9 per cent for the markup shock to 78.0 per cent for the monetary shock. The only exception is the trade-off induced by monetary shocks in the United Kingdom, for which 90.4 per cent of the mass of the posterior distribution of the estimated long-run trade-off is below zero, and is therefore marginally significant at the 10 per cent level. For other shocks, on the other hand, the corresponding fractions of draws range between 33.5 and 59.4 per cent, and are therefore far from being significant at any conventional level. One possible way to interpret this result is as it being largely due to the Thatcher disinflation of the early 1980s, which, in the presence of hysteresis in the labor market, slashed inflation by causing, at the same time, a permanent increase in the unemployment rate. For the United States, the fact that 81.7 per cent of the mass of the posterior distribution of the long-run trade-off induced by monetary shocks lies below zero—although not statistically significant at conventional levels—might be interpreted as weak evidence that, conceptually in line with King and Watson (1994), monetary policy shocks have indeed induced a small negative long–run trade-off between inflation and the unemployment rate (the modal and median estimates of the trade-off are equal to -0.284 and -0.39, respectively). Evidence for the other aggregate-demand side, non-monetary shock, on the other hand, is insignificant, with the fraction of draws for which the associated long-run trade-off is estimated to have ben negative being equal to 59.8 per cent.

Third, either modal or median estimates of the long-run trade-offs induced by individual shocks are uniformly quite small, thus suggesting that, even disregarding the previously discussed, significant extent of uncertainty, evidence consistently points towards weak-to-non-existent trade-offs. Focusing on modal estimates, for example, they range between -0.284 and 0.317 for the United States; between -0.367 and 0.100 for the Euro area; between -0.170 and 0.010 for the United Kingdom; and between -0.253 and -0.036 for Canada.

Fourth, the evidence reported in Figure 3 suggests that for either the United States or Canada, the shock which has been most likely to introduce a unit root component in the inflation rate is the markup shock. Evidence is especially clear for Canada, for which the posterior distribution of the fraction of the long-run variance of inflation explained by the markup shock is is quite significantly spread out, with a mode around 37 per cent, and a non-negligible fraction of the posterior associated with values beyond 50 per cent. Since, within the context of the DSGE model employed by Canova and Paustian (2011) to derive their robust sign restrictions, 'markup shocks' represent, among other things, inflationary impulses due to food and commodity price shocks, this suggests that the unit root component in U.S. and Canadian inflation rates was due, to a comparatively greater extent, to a set of disturbances of which food and commodity price shocks were part. Evidence for the Euro area is less clear-cut, as no single shock clearly stands out in terms of the fraction of the long-run variance of inflation it explains. Finally, for the United Kingdom, compatible with the previously discussed statistically significant evidence of a negative long-run trade-off conditional on monetary shocks, these shocks explain indeed the comparatively greater fraction of the long-run variance of inflation. In particular, the posterior distribution of the fraction of the long-run variance of inflation explained by monetary shocks is quite significantly spread out, and its second, smaller mode (the bigger one is at zero) is around 55 per cent.

Fifth, turning to unemployment, the most notable feature of the evidence reported in Figure 4 is the similarity, for the United Kingdom, of the shape of the posterior distribution of the fraction of the long-run variance of the unemployment rate explained by monetary shocks with the just-discussed correponding one for the inflation rate. This suggests that monetary shocks, beyond generating a negative long-run trade-off, also explain, most likely, a comparatively greater fraction of the frequency-zero variance of the unemployment rate than othet shocks. For other countries results are much less clear-cut, with no shock clearly standing out. This is especially clear for Canada, for which the posterior distributions are, in three cases out of four, nearly indistinguishable from one another.

Summing up, evidence for the United States suggests that, even if one really wants to give the long-run Phillips curve its best chance of manifesting itself, and therefore decides to narrowly focus on the pre-Volcker period, for which evidence of a unit root in inflation is stronger, still, it is not possible to detect statistically significant evidence of a non-vertical long-run trade-off conditional on any shock. The same holds true for either of the other three countries, with the single exception of the United Kingdom, for which monetary shocks appear to have generated a negative long-run trade-off between inflation and unemployment. Finally, it is important to stress that my results provide *no* evidence whatsoever in support of the notion, articulated for example by Friedrich Von Hayek,[21] that permanent increases in inflation are associated with corresponding permanent increases in the unemployment rate, which would imply that the long-run Phillips curve is positively sloped.

5.2 Results based on the CPI

Results based on the CPI are sometimes slightly stronger. For reasons of space, I do not report the entire set of results, and in what follows I only discuss some of them, but they are all available from upon request.[22] For the United States, for example, the fraction of draws from the posterior for which the long-run trade-off induced by monetary shocks is estimated to have been negative is equal to 83.4 per cent, compared to the 81.5 per cent based on the GDP deflator. By the same token, for the Euro area the corresponding fraction is equal to 82.3 per cent, compared to the 77.9 per cent based on the GDP deflator. Two things, however, ought to be stressed. First, the difference between the the two sets of results based on the CPI and the GDP deflator is uniformly modest. Second, since, in general, the GDP deflator should be regarded as a better measure of inflationary pressures than the CPI—which by construction is affected by food and energy price shocks to a much greater extent—this alternative set of results should be somehow discounted.

[21] In condemning the inflationary policies which led to the Great Inflation of the 1970s, von Hayek (1974) pointed out that *'the chief harm that inflation causes [is] that it gives the whole structure of the economy a distorted, lopsided character, which sooner or later makes a more extensive unemployment inevitable than that which that policy was intended to prevent.'*.

[22] In fact, I think it's likely I will include most of them in a future version of the paper.

5.3 Results based on imposing restrictions at longer horizons

One obvious way of obtaining sharper, and statistically stronger results is to impose more information on the data. Within the present context, one possibility is to impose sign restrictions not only on impact, but also at longer horizons.[23] If one is willing to do that, results get indeed slightly stronger. For example, imposing the sign restrictions not only on zero, but also up to two quarters after the impact, the fraction of draws for which the long-run trade-off for the Euro area conditional on monetary shocks is estimated to have been negative increases, based on the GDP deflator, from 77.9 to 79.4 per cent, whereas based on the CPI it reaches 87.9 per cent. For the United Kingdom, on the other hand, results are essentially the same, with, e.g., the fraction of draws based on the CPI rising to 92.0 per cent, compared to the 91.9 per cent obtained when restrictions are only imposed on impact. Once again, however, these results should be discounted, since, as pointed out by by Canova and Paustian (2011), whereas sign restrictions on impact are, in general, robust—in the specific sense that they hold for the vast majority of sub-classes within a specific class of DSGE models, and for the vast majority of plausible parameters' configurations—restrictions at longer horizons are instead, as they put it, 'whimsical', meaning that they are hard to pin down, and in general, they are not robust across sub-classes of models, and for alternative plausible parameters' configurations.

6 What About Cointegration?

Up until now I have completely ignored the issue of cointegration. On logical/conceptual grounds, the notion of cointegration between inflation and the unemployment rate—which implies that *all* shocks exerting a permanent impact on inflation have an equi-proportionate permanent impact on the unemployment rate[24]—might appear (at least, to me ...) as bizarre. Why should the ratio between the permanent impacts on unemployment and inflation induced by (say) a technology shock be *identical* to the ratio between the permanent impacts on the two variables induced by a monetary policy shock? From a statistical point of view, however, it is important to ascertain what the data actually tell us, and in this section I therefore proceed to perform either bivariate or trivariate cointegration tests between inflation, the unemployment rate or the short rate based on the procedure proposed by Johansen.

Table 5 reports bootstrapped p-values for Johansen's trace test of the null of no cointegration between either inflation and unemployment, inflation and the short rate

[23] Once again, for reasons of space I do not report the entire set of results, and in what follows I only discuss some of them, but they are all available from upon request.

[24] To be clear, this implies that, if (e.g.) the permanent impact on the unemployment rate of a one per cent permanent increase in the inflation rate induced by monetary shocks is equal to -2, it ought to be the case that the same holds true for *all* of the other shocks: for each of them, the ratio between the two permanent impacts on the unemployment rate and the inflation rate ought to be, likewise, equal to -2.

(with the single exception of the United Kingdom, for which results from the ADF tests reported in Table 1 strongly rejected the null of a unit root in the short rate), or inflation, the unemployment rate, and the short rate. p-values have been computed by bootstrapping the VAR estimated for the first difference of the relevant vector of series. To be clear, this means that, given (e.g.) the vector $Y_t = [\pi_t, U_t]'$, we start by selecting the lag order for cointegration tests as the maximum between the lag orders selected based on the Schwartz and the Hannan-Quinn criteria;[25] we perform Johansen's trace test of the null of no cointegration. Then, we estimate the VAR[26] for ΔY_t, we bootstrap it 10,000 times, thus generating bootstrapped artificial series $\Delta\tilde{Y}_t^j$, based on each of them we compute corresponding bootstrapped artificial series $\tilde{Y}_t^j$—that is, those for the levels of the series—and based on each of them we perform the same trace test we 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, we compute critical values (not reported here) and p-values. For the United States we detect no evidence of cointegration whatsoever, which implies that, under this respect, the previously discussed SVAR-based results are not subject to any *caveat*. More generally, bivariate cointegration tests do not detect any evidence of cointegration between inflation and the short rate. for either the United States, the Euro area, or Canada. Although at first sight puzzling—taken at face value, these results imply a rejection of the Fisher hypothesis that permanent shifts in inflation should map one-to-one into correspoinding shifts in interest rates—it is important to stress that the empirical violation of the Fisher hypothesis is a well-established stylized fact, so that these results should not be seen as surprising at all. On the other hand, for both the Euro area and Canada we detect strong evidence of cointegration between inflation, unemployment, and the short rate, whereas for the United Kingdom evidence points, at the 10 per cent level, towards cointegration between inflation and the unemployment rate.

For the Euro area and Canada, for which we detected strong evidence of cointegration between the three series, we also report bootstrapped p-values for testing the null of one single cointegrating vector, versus the alternative of two cointegrating vectors. Details of the bootstrapping procedure are the same as before, with the only difference that, instead of bootstrapping the estimated VAR for ΔY_t under the null of no cointegration, we bootstrap the VECM estimated conditional in there being one single cointegrating vector. Whereas for Canada the p-value, at 66.8 per cent, is very far from being signifiant at any conventional significant level, for the Euro area, at 1.2 per cent, it suggests the presence of an additional cointegrating vector. Figure 5 reports, for the Euro area and the United Kingdon—for which we detected evidence of cointegration in the bivariate representation for $Y_t = [\pi_t, U_t]'$, the bootstrapped distribution of the slope of the long-run Phillips curve implied by the estimated coin-

[25] [Here put reference on lag selection for cointegrated VARs based on the book by Lutkepohl]

[26] To be clear, the VAR we are estimating here is not a cointegrated VAR, that is, it is equual to the VECM representation without the error-correction term.

tegrating vector between inflation and unemployment, and of the elements of the loading vector of the cointegrating residual in the VECM representation. Results are completely different from those we previously discussed based on the structural VAR. The slope of the long-run Phillips curve is here not only highly statistically significantly different from zero—as implied by the results from the trace test, and as testified by the fact that mass of the bootstrapped distribution is pretty much away from zero—but it is also sizeable, with modal estimates around -1.1 for the Euro area, and -1.95 for the United Kingdom.

How should we interpret these results? One possibility is that they are just a fluke, possibly due to small-sample problems. As Table 6 shows, this is actually not the case *at all*. The table reports results from four sets of 10,000 Monte Carlo simulations of lengths equal to the actual sample lengths I am working with for the four countries. For each simulation, I randomly generate two independent random walks, and I apply *exactly* the same procedure I previously applied to the actual data, computing the p-values by bootstrapping the estimated VAR for the first differences of two random walks. Ideally, out of the 10,000 simulations, the fraction of bootstrapped p-values below x per cent should be equal to x per cent. As the results reported in the table show, the bootstrapped Johansen procedure I am using herein gets quite remarkably close to this ideal: the fraction of simulations for which the bootstrapped trace statistic incorrectly rejects the null of no cointegration between the two independent random walks at a given significance level ranges between 11.3 and 11.9 per cent at the 10 per cent level; between 5.5 and 6.0 per cent at the 5 per cent level; and between 1.1 and 1.3 per cent at the 1 per cent level. Quite remarkably, the performance for the United Kingdom, for which we have just 81 quarterly observations, is not dramatically different from that for Canada, for which we have instead 118 observations. This testifies to the power of bootstrapping, which can effectively take into account of the specific characteristics of the data the researcher is working with.

Although the bootstrapped Johansen procedure used herein performs remarkably well conditional on a data-generation process in which the series of interest are independent random walks, it is an open question how well such a procedure performs conditional on data-generation processes such as the previously estimated structural VARs. To put it differently, suppose that the structural VARs we previously estimated—in which, by construction, there is *no* cointegration whatsoever between any series—are, for either country, the true model of the economy: how often would the bootstrapped Johansen procedure incorrectly reject the null of no cointegration? Table 7 reports evidence on this, by showing the fraction of simulations for which Johansen's bootstrapped trace statistic incorrectly rejects the null of no cointegration at a given significance level, based on takig the estimated structural VARs as the data-generation processes. As the table shows, Johansen's procedure tends to spuriously detect cointegration a non-negligible fraction of the times. For example, for the Euro area and the United Kingdom, the fractions of bootstrapped p-values for Johansen's trace statistic for testing the null of no cointegration between inflation and the unemployment rate which are smaller than 10 per cent are equal to 22.9 and 24.2 per cent respectively. This means that the trace test would incorrectly reject the null of no cointegration between inflation and unemployment at the 10 per cent level between one-fifth and one-fourth of the times. As for cointegration between inflation, unemployment, and the short rate, results are even worse. For Canada, for example, conditional on the estimated structural VAR being the true data-generation process, Johansen's trace test statistic would incorrectly reject the null of no cointegration between the three variables at the 10 per cent level a remarkable 31.5 per cent of the times. This implies that evidence such as that reported in Table 5 should be discounted, as those results might as well be due to the limitations of cointegration tests conditional on this specific data generation process.

7 Conclusions

In this paper I have used Bayesian structural VARs identified based on a combination of long-run and sign restrictions to investigate the long-run trade-off between inflation and the unemployment rate in the United States, the Euro area, the U.K., and Canada over the post-WWII period. Evidence suggests that either of the four structural shocks which are here allowed to introduce a unit root component in the inflation rate has generated a vertical long-run Phillips curve. The only exception is the United Kingdom, for which monetary policy shocks have induced a statistically significant, albeit small, negative long-run Phillips trade-off. For the Euro area, the U.K., and Canada Johansen's cointegration tests point towards the presence of cointegration between either inflation and unemployment, or inflation, unemployment, and a short-term interest rate, with the long-run Phillips trade-off implied by the estimated cointegrating vectors being negative and sizeable. I argue however that this evidence should be discounted, as, conditional on the estimated structural VARs—which, by construction, do not feature cointegration between any variable—Johansen's procedure tends to spuriously detect cointegration a non-negligible fraction of the times.

References

Benati, L. (2008): "Investigating Inflation Persistence Across Monetary Regimes," *Quarterly Journal of Economics*, 123(3), 1005–1060.

Canova, F., and G. de Nicolo (2002): "Monetary disturbances matter for business fluctuations in the G-7," *Journal of Monetary Economics*, 49(6), 1131Ű–1159.

Canova, F., and M. Paustian (2011): "Measurement with Some Theory: Using Sign Restrictions to Evaluate Business Cycle Models," *Journal of Monetary Economics*, 58, 345–361.

Canova, F., and J. Pina (2005): "What VAR Tell Us About DSGE Models," in Diebolt, C. and Kyrtsou, C. New Trends In Macroeconomic, Springer Verlag.

Cochrane, J. H. (1994): "Permanent and Transitory Components of GNP and Stock Prices," *Quarterly Journal of Economics*, 109(1), 241–265.

Evans, C. (1994): "Comment on: The Post-War U.S. Phillips Curve, A Revisionist Econometric History," *Carnegie-Rochester Conference Series on Public Policy.*

Faust, J. (1998): "The Robustness of Identified VAR Conclusions About Money," *Carnegie-Rochester Conference Series on Public Policy*, 49, 207–244.

Faust, J., and E. Leeper (1998): "When Do Long-Run Identifying Restrictions Give Reliable Results?," *Journal of Business and Economic Statistics*, 15(3), 345–353.

Fry, R., and A. Pagan (2007): "Some Issues in Using Sign Restrictions for Identifying Structural VARs," NCER Working Paper n. 14, April 2007.

Golub, G., and C. VanLoan (1996): *Matrix Computations, III Edition.* Johns Hopkins University Press, Baltimore.

King, R., and M. Watson (1994): "The Post-War U.S. Phillips Curve: A Revisionist Econometric History," *Carnegie-Rochester Conference Series on Public Policy.*

Lucas, R. E. (1972a): "Econometric Testing of the Natural Rate Hypothesis," *in Eckstein, O., ed. The Econometrics of Price Determination, Washington, D.C., Board of Governors of the Federal Reserve System*, pp. 50–59.

——— (1972b): "Expectations and the Neutrality of Money," *Journal of Economic Theory*, 4(April 1972), 103–24.

——— (1973): "Some International Evidence on Output-Inflation Trade-Offs," *American Economic Review*, 63(3), 326–334.

Rubio-Ramirez, J., D. Waggoner, and T. Zha (2005): "Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference," *Review of Economic Studies*, 77(2), 665–696.

Sargent, T. J. (1971): "A Note on the Accelerationist Controversy," *Journal of Money, Credit and Banking*, 3(3), 721–725.

Sargent, T. J. (1987): *Macroeconomic Theory, II Edition*. Orlando, Academic Press.

Uhlig, H. (1998): "Comment On: The Robustness of Identified VAR Conclusions About Money," *Carnegie-Rochester Conference Series on Public Policy*, 49, 245–263.

——— (2005): "What are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure," *Journal of Monetary Economics*, 52(2), 381–419.

von Hayek, F. (1974): "Inflation's Path to Unemployment," *The Daily Telegraph, 15-16 October 1974, reprinted in Friedrich von Hayek (1978), New Studies in Philosophy, Politics, Economics, and the History of Ideas, The University of Chicago Press.*

Table 1 Bootstrapped p-values[a] for augmented Dickey-Fuller tests without trend

	Inflation						Unemployment rate		
	(based on GDP deflator)			*(based on CPI)*					
	p=2	p=4	p=8	p=2	p=4	p=8	p=2	p=4	p=8
United States									
pre-Volcker (1954Q3-1979Q3)	0.464	0.640	0.753	0.658	0.624	0.789	0.093	0.219	0.373
full sample (1954Q3-2011Q4)	0.118	0.192	0.239	0.019	0.041	0.043	0.010	0.044	0.083
Euro area (1970Q1-1998Q4)	0.515	0.744	0.845	0.693	0.663	0.763	0.238	0.258	0.289
Japan (1966Q4-2011Q4)	0.052	0.054	0.310	0.031	0.110	0.329	0.575	0.572	0.485
United Kingdom (1972Q1-1992Q3)	0.222	0.199	0.471	0.136	0.272	0.292	0.169	0.329	0.500
Canada (1961Q1-1990Q4)	0.136	0.193	0.161	0.150	0.229	0.225	0.527	0.633	0.558
Australia (1969Q3-1994Q2)	0.000	0.001	0.001	0.123	0.304	0.479	0.286	0.307	0.324
	Short rate			Real GDP growth			Consumption-GDP ratio		
United States									
pre-Volcker (1954Q3-1979Q3)	0.364	0.261	0.545	0.000	0.000	0.008	0.048	0.093	0.154
full sample (1954Q3-2011Q4)	0.200	0.122	0.281	0.000	0.000	0.003	0.002	0.004	0.027
Euro area (1970Q1-1998Q4)	0.328	0.362	0.581	0.000	0.003	0.016	0.027	0.063	0.111
United Kingdom (1972Q1-1992Q3)	0.018	0.024	0.027	0.002	0.003	0.061	0.874	0.868	0.726
Canada (1961Q1-1990Q4)	0.333	0.350	0.438	0.000	0.004	0.052	0.127	0.112	0.247
	Investment-GDP ratio			Long-short spread			Inventories-sales ratio		
United States, pre-Volcker (1954Q3-1979Q3)				0.045	0.007	0.008			
Euro area (1970Q1-1998Q4)	0.107	0.079	0.061	0.004	0.003	0.018			
United Kingdom (1972Q1-1992Q3)	0.564	0.413	0.391	0.106	0.148	0.439	0.014	0.026	0.026
Canada (1961Q1-1990Q4)	0.103	0.048	0.010	0.046	0.170	0.206			

[a] Based on 10,000 bootstrap replications of estimated ARIMA processes.

Table 2 Posterior distribution of the slope of the long-run Phillips curve induced by individual shocks

	Mode, median, and 90%-coverage percentiles			Fraction of draws for which slope is negative
United States, *pre-Volcker (1954Q3-1979Q3)*				
technology shock	0.084	0.016	[-2.696; 3.028]	0.488
monetary shock	-0.284	-0.394	[-3.076; 1.624]	0.815
taste shock	-0.084	-0.110	[-2.198; 1.747]	0.597
mark-up shock	0.317	0.290	[-1.147; 1.319]	0.245
Euro area (1970Q1-1998Q4)				
technology shock	-0.033	-0.044	[-3.520; 3.327]	0.527
monetary shock	-0.367	-0.367	[-3.048; 2.365]	0.779
taste shock	-0.033	0.038	[-1.288; 1.335]	0.454
mark-up shock	0.100	0.154	[-1.445; 1.618]	0.403
United Kingdom (1972Q1-1992Q3)				
technology shock	-0.070	-0.060	[-1.261; 1.161]	0.594
monetary shock	-0.170	-0.204	[-0.947; 0.301]	0.904
taste shock	-0.010	-0.028	[-1.291; 1.506]	0.546
mark-up shock	0.010	0.067	[-0.590; 0.637]	0.335
Canada (1961Q1-1990Q4)				
technology shock	-0.253	-0.215	[-2.386; 2.025]	0.680
monetary shock	-0.229	-0.222	[-1.257; 0.749]	0.780
taste shock	-0.301	-0.237	[-1.959; 1.433]	0.723
mark-up shock	-0.036	0.042	[-0.399; 0.602]	0.419

Table 3 Fractions of the long-run variance of GDP deflator inflation explained by individual shocks

	Median, and 90%-coverage percentiles		Fraction of the mass of the posterior distribution below: 0.1	0.05	0.01
United States, pre-Volcker (1954Q3-1979Q3)					
technology shock	0.070	[0.001; 0.483]	0.588	0.429	0.200
monetary shock	0.211	[0.003; 0.682]	0.312	0.211	0.089
taste shock	0.194	[0.003; 0.670]	0.345	0.233	0.096
mark-up shock	0.325	[0.049; 0.787]	0.126	0.051	0.005
Euro area (1970Q1-1998Q4)					
technology shock	0.053	[0.000; 0.368]	0.645	0.488	0.228
monetary shock	0.201	[0.003; 0.716]	0.342	0.233	0.101
taste shock	0.305	[0.007; 0.745]	0.221	0.148	0.059
mark-up shock	0.275	[0.048; 0.695]	0.142	0.054	0.005
United Kingdom (1972Q1-1992Q3)					
technology shock	0.068	[0.001; 0.453]	0.593	0.435	0.200
monetary shock	0.384	[0.009; 0.845]	0.184	0.122	0.052
taste shock	0.080	[0.001; 0.546]	0.551	0.407	0.187
mark-up shock	0.280	[0.026; 0.762]	0.198	0.097	0.021
Canada (1961Q1-1990Q4)					
technology shock	0.081	[0.001; 0.531]	0.5467	0.400	0.190
monetary shock	0.193	[0.003; 0.688]	0.343	0.233	0.099
taste shock	0.097	[0.001; 0.591]	0.507	0.368	0.170
mark-up shock	0.419	[0.059; 0.842]	0.092	0.040	0.007

Table 4 Fractions of the long-run variance of the unemployment rate explained by individual shocks

	Median, and 90%-coverage percentiles		Fraction of the mass of the posterior distribution below: 0.1	0.05	0.01
United States, pre-Volcker (1954Q3-1979Q3)					
technology shock	0.048	[0.000; 0.333]	0.677	0.507	0.238
monetary shock	0.176	[0.002; 0.649]	0.359	0.243	0.104
taste shock	0.079	[0.001; 0.496]	0.554	0.401	0.188
mark-up shock	0.224	[0.004; 0.714]	0.300	0.194	0.084
Euro area (1970Q1-1998Q4)					
technology shock	0.046	[0.000; 0.335]	0.685	0.519	0.252
monetary shock	0.169	[0.002; 0.671]	0.367	0.249	0.102
taste shock	0.059	[0.000; 0.417]	0.620	0.465	0.228
mark-up shock	0.229	[0.003; 0.722]	0.302	0.210	0.093
United Kingdom (1972Q1-1992Q3)					
technology shock	0.055	[0.001; 0.396]	0.640	0.478	0.224
monetary shock	0.390	[0.015; 0.846]	0.161	0.102	0.040
taste shock	0.073	[0.001; 0.498]	0.573	0.425	0.196
mark-up shock	0.116	[0.001; 0.655]	0.470	0.339	0.157
Canada (1961Q1-1990Q4)					
technology shock	0.123	[0.001; 0.637]	0.449	0.321	0.145
monetary shock	0.135	[0.001; 0.660]	0.431	0.307	0.137
taste shock	0.118	[0.001; 0.670]	0.463	0.332	0.148
mark-up shock	0.081	[0.001; 0.553]	0.550	0.399	0.184

fs

Table 5 Bootstrapped p-values[a] for Johansen's cointegration tests

	Trace test of the null of no cointegration between:		
	π_t and U_t	π_t and R_t	π_t, U_t and R_t
United States, pre-Volcker (1954Q3-1979Q3)	0.181	0.149	0.168
Euro area (1970Q1-1998Q4)	0.023	0.568	0.005
United Kingdom (1972Q1-1992Q3)	0.094	NA[b]	NA[b]
Canada (1961Q1-1990Q4)	0.509	0.493	0.010
	Test of the null of one cointegrating vector, versus the alternative of two, for π_t, U_t and R_t:		
Euro area (1970Q1-1998Q4)	0.012		
Canada (1961Q1-1990Q4)	0.668		

[a] Based on 10,000 bootstrap replications. [b] For the United Kingdom, results from the ADF tests reported in Table 1 strongly reject the null of a unit root in the short rate.

Table 6 Monte Carlo evidence on the performance of the Johansen procedure: fraction of simulations for which the bootstrapped trace statistic incorrectly rejects the null of no cointegration between two independent random walks at a given significance level[a]

	Fractions of bootstrapped p-values for Johansen's trace statistic below:		
Sample size (in quarters)	0.1	0.05	0.01
$T = 100$ *(U.S., pre-Volcker: 1954Q3-1979Q3)*	0.119	0.059	0.013
$T = 115$ *(Euro area: 1970Q1-1998Q4)*	0.118	0.060	0.012
$T = 81$ *(United Kingdom: 1972Q1-1992Q3)*	0.118	0.059	0.011
$T = 118$ *(Canada: 1961Q1-1990Q4)*	0.113	0.055	0.011

[a] Based on 10,000 simulations.

Table 7 Fraction of simulations for which Johansen's trace test statistic incorrectly rejects the null of no cointegration at a given significance level, based on estimated structural VARs

Based on the estimated structural VAR for:	Fractions of bootstrapped p-values for Johansen's trace statistic below: 0.1	0.05	0.01
	π_t and U_t		
Euro area (1970Q1-1998Q4)	0.229	0.140	0.059
United Kingdom (1972Q1-1992Q3)	0.242	0.167	0.074
	π_t, U_t and R_t		
Euro area (1970Q1-1998Q4)	0.230	0.152	0.062
Canada (1961Q1-1990Q4)	0.315	0.237	0.100

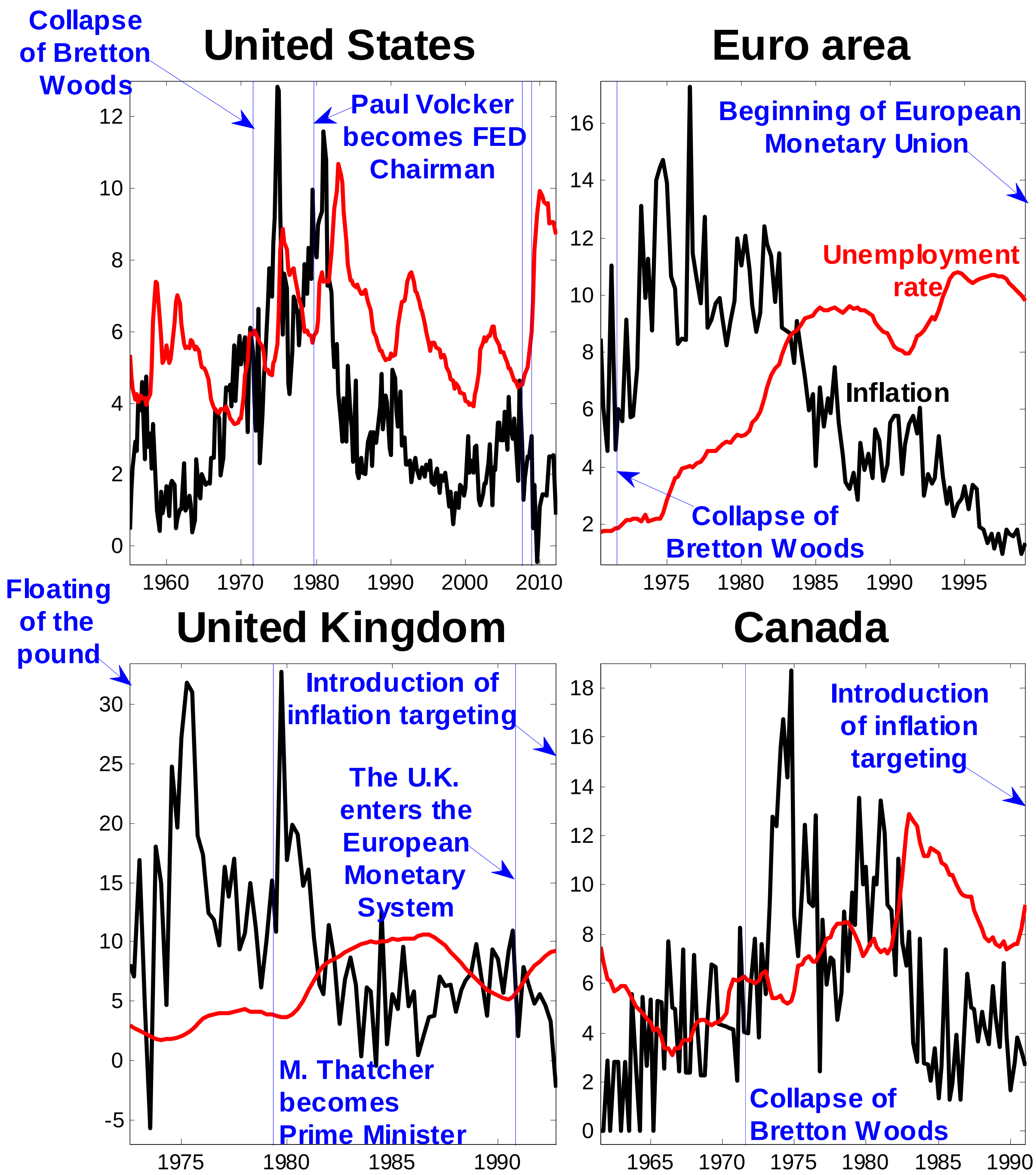

Figure 1 Inflation and the unemployment rate

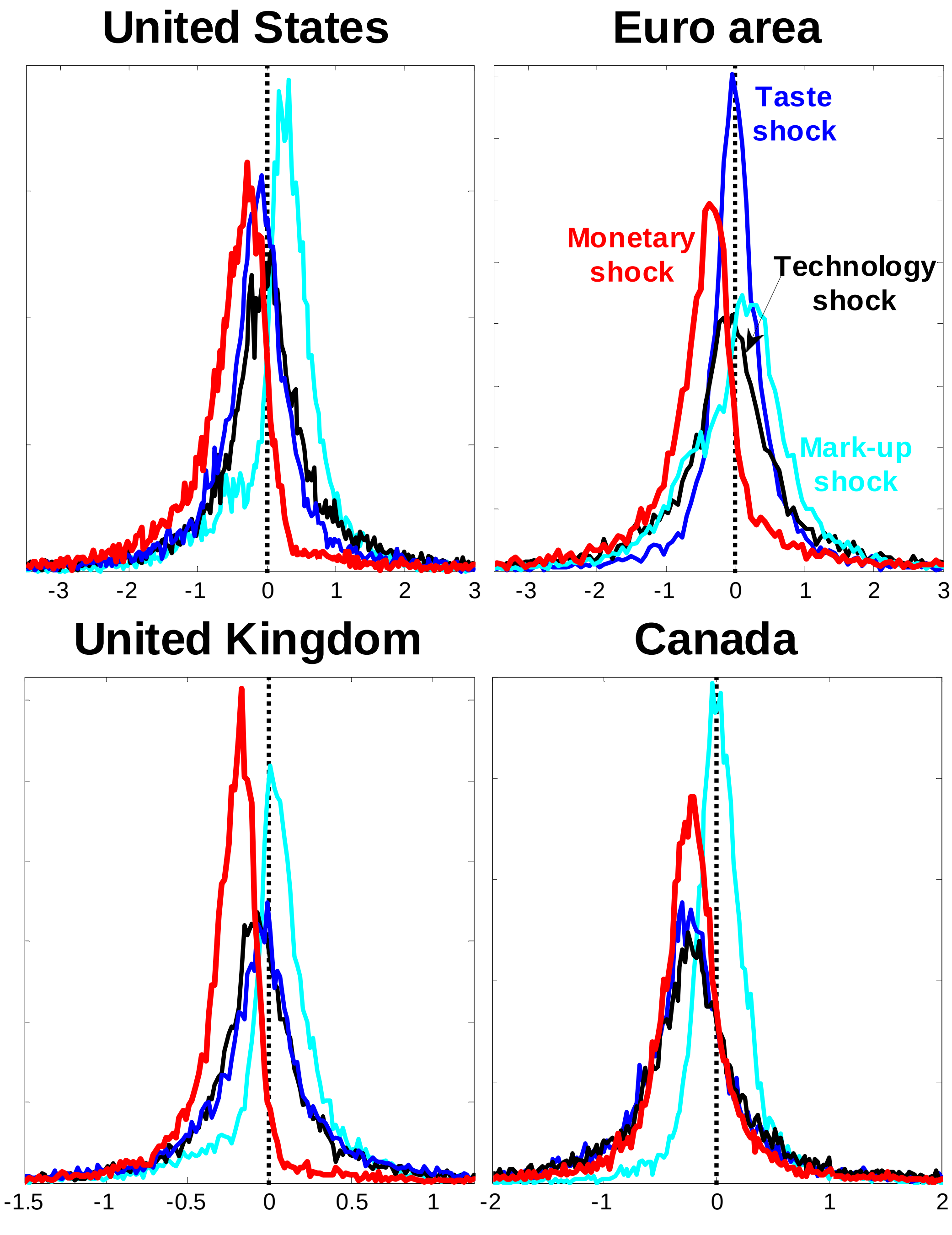

Figure 2 Posterior distributions of the slopes of the long-run Phillips curve induced by individual shocks

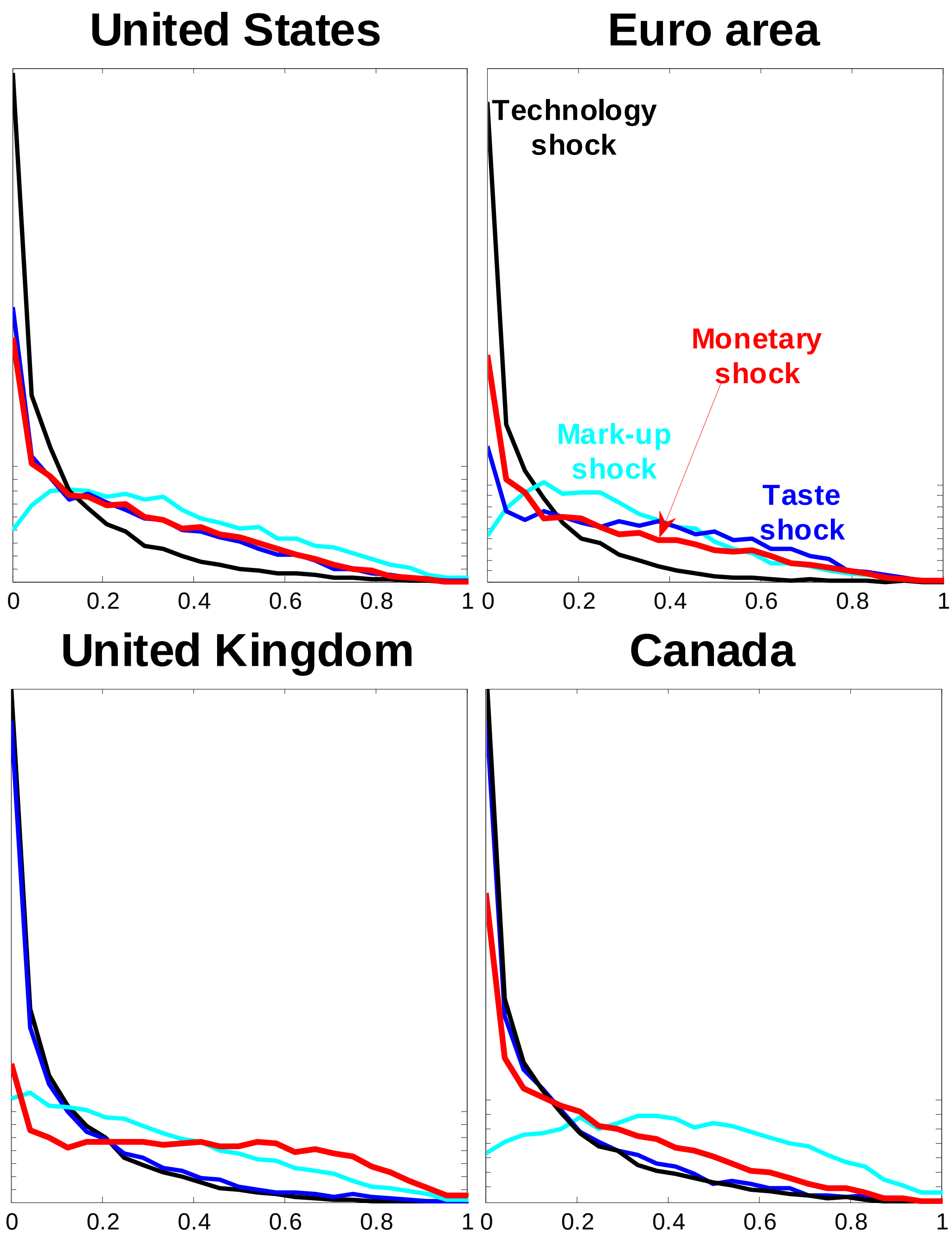

Figure 3 Posterior distributions of the fractions of the long-run variance of inflation explained by individual shocks

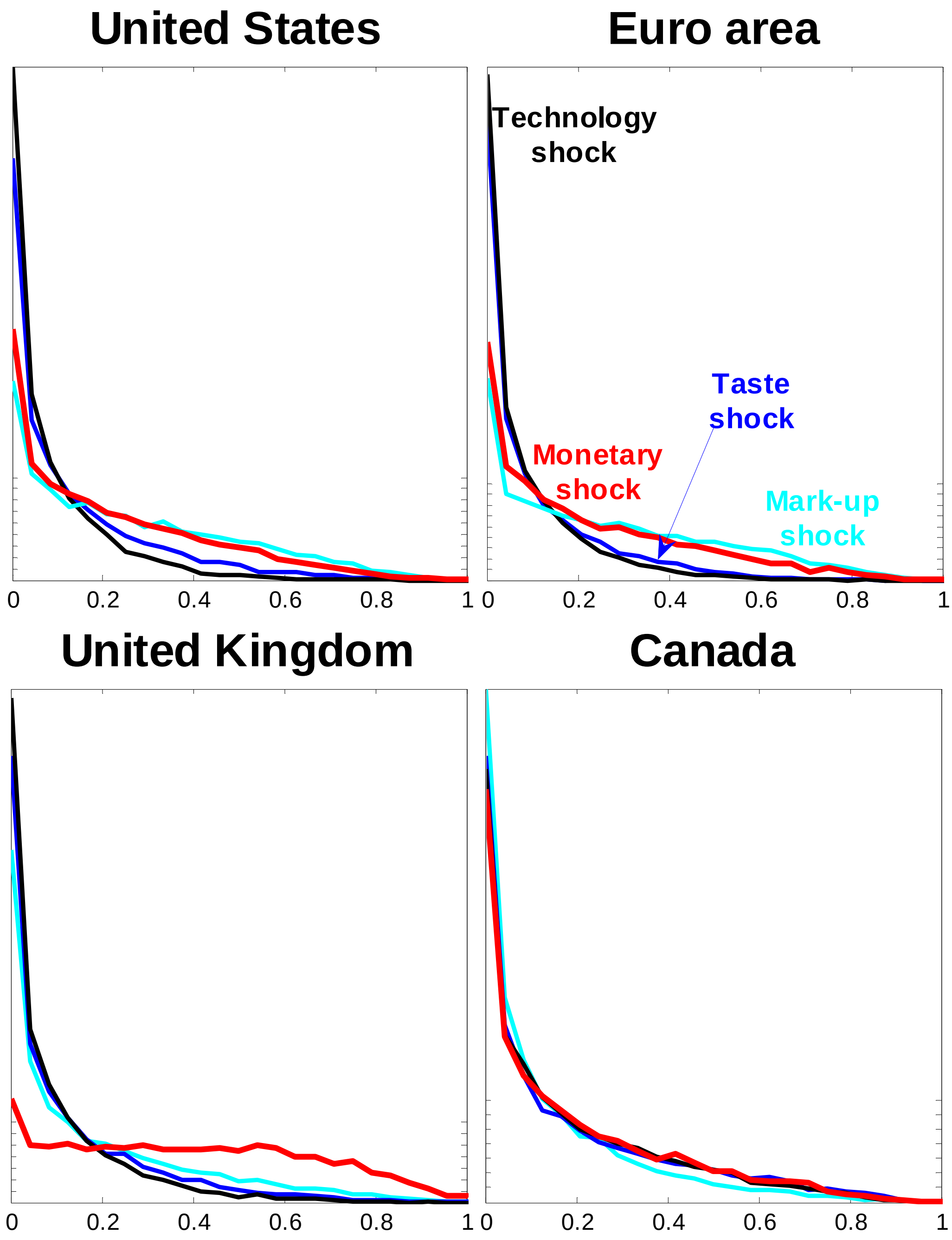

Figure 4 Posterior distributions of the fractions of the long-run variance of the unemployment rate explained by the various shocks

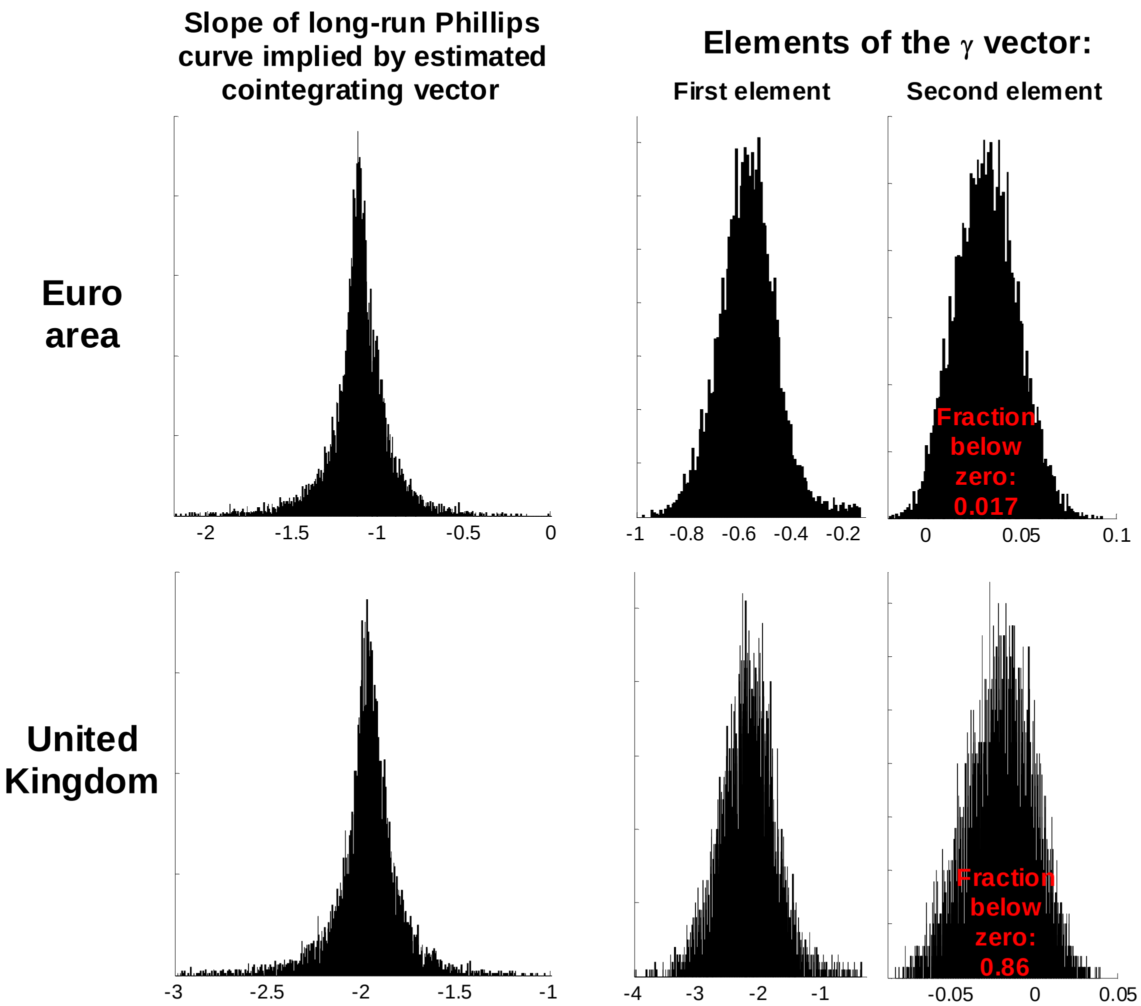

Figure 5 Bootstrapped distributions of the slope of the long-run Phillips curve implied by the estimated cointegrating vector between inflation and unemployment, and of the elements of the γ vector