

Testing the New Keynesian Model on U.S. and Euro Area Data

Mikael Juselius*

*Swedish School of Economics, and RUESG,
University of Helsinki*

Abstract

I apply the Johansen and Swensen (1999, 2004) method of testing exact rational expectations within the cointegrated VAR (Vector Auto-Regressive) model, to testing the New Keynesian (NK) model. This method permits the testing of rational expectation systems, while allowing for non-stationary data. The NK-model is tested on quarterly U.S. and Euro area time series data. I find that the restrictions implied by the core equations of the NK-model are rejected regardless of sample periods or measures of real marginal costs. I also provide a tentative explanation of the results favored by previous researches.

Key words: New Keynesian Phillips curve, cointegration, vector autoregressive model.

JEL Classification: C32, C52, E31, E52.

*Address: University of Helsinki, Department of Economics, P.O. Box 17, FIN - 00014 Helsinki. E-mail: mikael.juselius@helsinki.fi.

1 Introduction

The popularity of the New Keynesian model in recent years has led to numerous empirical attempts to evaluate the performance of the model. A recent overview of this literature is provided by Henry and Pagan (2004). The typical study has focused on one of the two structural equations of the NK-model, usually the New Keynesian Phillips (NKP) curve since it captures the relevant features of price stickiness. For instance Gali and Gertler (1999), and Gali et al. (2001) find strong evidence in favor of the Phillips curve, using single equation GMM (General Method of Moments). Sbordone (2002) also reports favorable results by a slightly different approach,¹ while Fuhrer (1997) obtains less favorable results using ML (Maximum Likelihood). More recent contributions include Matheron and Maury (2004), McAdam and Willman (2004), Roberts (2005), and Nelson and Lee (2007). The second equation, the expectational “IS” curve, has been investigated by Fuhrer (2000) and more recently in Kara and Nelson (2004) and Fuhrer and Rudebusch (2004).

Since the early contributions, a number of empirical issues have been raised. The use of single equation estimation procedures has been criticized on the grounds that empirical identification requires a system approach. Furthermore, due to problems such as weak instruments, GMM estimates are likely to be very imprecise. Thorough discussions on these issues can be found in Ma (2002), Mavroeidis (2004), Rudd and Whelan (2005a,b).² These difficulties has led authors such as Linde (2005) and Giordani (2004) to consider a full system approach. Another, largely neglected, issue is the apparent non-stationary behavior of the data. This problem is noted and discussed by Bardsen et al. (2004), among others. Empirical versions of the NK-model are typically obtained by assuming stationarity and log-linearizing around the steady states. Non-stationary data invalidates such linearizations as is discussed by Dees et al. (2008).³ Moreover, studies that overlook non-stationarities in the data run the risk of reporting spurious results.

¹Sbordone (2002) uses the method of testing present value models, proposed by Campbell and Shiller (1987). She assumes that the data is stationary, although the method also allows for special cases when the data is non-stationary. The Campbell and Shiller (1987) technique with non-stationary data has recently been employed to the NKP-curve by Demery and Duck (2003) and Tillmann (2005).

²However, see also Gali et al. (2005) and Sbordone (2005) for answers to some of this criticism.

³Section 2 provides a brief discussion of some conditions under which the NK-model can be linearized even when data is non-stationary.

The aim of this paper is to demonstrate how one can test the validity of the restrictions implied by the NK-model when the key variables are difference stationary. The procedure is illustrated by testing the core equations of the NK-model within a cointegrated VAR (Vector Auto-Regressive) model on quarterly U.S. and Euro area time series data. The sample period is 1960:1-2005:2 for the U.S. and 1970:1-2003:4 for the Euro area. The restrictions implied by the core equations of the NK-model are tested by the exact linear rational expectations (RE) method proposed by Johansen and Swensen (1999, 2004). The advantage of this method is that it permits formal testing of the restrictions implied by RE systems, at the same time allowing the data to be non-stationary.

The tests proposed here are related to the single equation tests of the NKP-curve by Fanelli (2008) and Barkbu and Batini (2005). Fanelli (2008) uses the three-step method of Fanelli (2002), to formally test the NKP-curve within a cointegrated VAR model on Euro area data.⁴ and rejects the NKP-curve specification. Barkbu and Batini (2005) apply the Johansen and Swensen method to the NKP-curve on Euro area data. They obtain favorable results for the NKP-curve using a minimal information set. However, it is doubtful whether the information used by Barkbu and Batini (2005) captures the main features of the inflation process (see discussion by Bardsen et al. (2004)). This paper differs from Fanelli (2008) and Barkbu and Batini (2005) in basically two ways. First, an extended information set is used which combined with the Johansen and Swensen method, allows for testing both core equations of the NK-model instead of only the NKP-curve equation. Second, the model here is also tested on U.S. data.

The results suggest that the evidence in favor of the core equations of the NK-model, the IS curve and the new Keynesian Phillips curve, is weak. The restrictions implied by the equations are rejected on both U.S. and Euro area data. Sensitivity analysis of different sample periods and of different measures of marginal costs do not change the results. Furthermore, by only considering the cointegration restrictions implied by the NK-model (*cointegration implications*, henceforth) a less demanding test of the NK-model is provided. These restrictions form a subset of the complete set of restrictions implied by the NK-model and, hence, constitute a necessary condition for

⁴There are some drawbacks to this method. First, it is only possible to test one RE equation and, second, a condition similar to strong exogeneity of the forcing variables is needed in the estimations.

the NK-model. The latter are rejected in most cases. Interestingly, the cointegration implications of the NKP-curve are not rejected on Euro area data when labor's share is used as a measure of marginal costs. It is precisely for this case that favorable results on the NKP-curve have been reported by, for instance, Gali and Gertler (1999). But, since the overall restrictions of the equation are rejected, the support for the NKP-curve is nevertheless not overwhelming. Methods that rely on less formal evaluations of the NKP-curve, such as the size and significance of the marginal costs and forward terms, run the risk of claiming success when only the cointegration implications are met.

The next section introduces a baseline New Keynesian model. The data and information sets are discussed in section 3, while section 4 introduces the Johansen and Swensen method. The estimation results are presented in section 5 followed by a discussion in section 6. Section 7 concludes.

2 The New Keynesian model

This section introduces a standard version of the closed economy New Keynesian model.⁵ The NK-model belongs to a class of “miniature” dynamic stochastic general equilibrium (DSGE) models that are based on optimizing households and firms, rational expectations, and nominal price rigidities. The model consists of three non-linear equations, a forward-looking “IS curve” that relates output to the real rate of interest, a New Keynesian Phillips curve that relates inflation to real marginal costs, and a central bank policy rule for the nominal interest rate. In a stationary setting, empirical variants of these equations are obtained by log-linearizing around the steady states of the key variables and adding lags.⁶ The two structural equations, the IS curve and

⁵The standard closed economy model has been extended in several ways, for instance by incorporating labor market imperfections (Erceg et al., 2000) or by accounting for investments in capacity (Razin, 2005). Open economy issues have been investigated by several authors, for example Clarida et al. (2002), Galí and Monacelli (2005), Svensson (2000), Batini et al., 2005, and Matheson (2008). The core equations of the NK-model often have the same form in such extensions of the model.

⁶The lagged terms can be motivated for example by rule of thumb pricing and habit persistence as in Gali and Gertler (1999) and Fuhrer (2000). Detailed derivations and discussions of the equations are provided by McCallum and Nelson (1999), Clarida et al. (1999), Yun (1996), Walsh (2003), and Woodford (2003), among others.

the NKP-curve commonly take the linearized representations

$$\tilde{y}_t = \varphi_{11}E_t\tilde{y}_{t+1} - \varphi_{12}(i_t - E_t\Delta p_{t+1}) + \varphi_{13}\tilde{y}_{t-1} + v_t \quad (1)$$

$$\Delta p_t = \varphi_{21}E_t\Delta p_{t+1} + \varphi_{22}x_t + \varphi_{23}\Delta p_{t-1} \quad (2)$$

where constants representing equilibrium values are suppressed, $\tilde{y}_t = y_t - y_t^f$ is the flexible price output gap, y_t is real output and y_t^f is the level of output that would prevail under flexible prices, i_t is the nominal short-term interest rate, p_t is the price level, $v_t = \varphi_{11}E_t y_{t+1}^f + \varphi_{13}y_{t-1}^f - y_t^f$, x_t is real marginal costs, E_t is the expectations operator conditional on the agent's information set at time t , and the coefficients, $\varphi_{ij} \geq 0$ for all i and j are functions of the structural parameters from the underlying theory.⁷ The purely forward looking versions of the two structural equations are obtained by setting $\varphi_{11} = 1$ and $\varphi_{13} = \varphi_{23} = 0$.

In addition to equations (1) and (2), a policy rule for the nominal interest rate is usually derived by specifying a policy objective and solving under discretion or commitment. For example, a frequently used specification is the interest rate smoothing Taylor rule

$$i_t = \phi_1 i_{t-1} + (1 - \phi_1)(\phi_2(\Delta p_t - \Delta p^*) + \phi_3 \tilde{y}_t + \phi_4 x_t)$$

where Δp^* is a constant inflation target (*e.g.* Clarida et al., 1999), and ϕ_i are parameters. However, there are several problems associated with such policy rules. First, the empirical support for rules with constant targets have been mixed and these rules are generally not robust to small alterations in the specifications (see for example Kozicki, 1999). In contrast, (1) and (2) are more robust to alternative ways of specifying the underlying theoretical structure and have received more support in the literature. Hence, imposing the restrictions from some (empirically incorrect) policy rule may cause rejection of the complete NK-model system even when the structural equations (1) and (2) are data consistent. Second, the optimal inflation target may have varied over time as argued in Cogley and Sargent (2005), Ireland (2007), and Sbordone (2007). Allowing for a time varying policy rule introduces a latent variable into the system, causing a violation of the exactness of the

⁷In addition, stationary error terms could be added to (1) and (2) to account for measurement errors, deviations from linearity *et cetera*, but are left out since they do not add to the discussion.

policy equation and thereby rendering the Johansen and Swensen framework inapplicable.⁸

For these reasons, I do not impose any policy rule restrictions in the empirical analysis. Ignoring these restrictions does not invalidate the tests of (1) and (2), and has the advantage of simplifying the analysis considerably. But, of course, empirical identification of the parameters of the NK-model requires that the processes of all forcing variables are specified. Therefore, if the tests of (1) and (2) are not rejected, the next step should be to estimate a reasonable policy rule for the interest rate prior to parameter identification. Exploiting such policy rule restrictions would also increase the degrees of freedom of the test and, hence, its efficiency.

The empirical counterpart of the variables in (1) and (2) typically display a high degree of persistence (see for example Bardsen et al. 2004 and Dees et al. 2008) suggesting non-stationarity rather than stationarity. In this case, two important issues that have to be addressed:

First, the log-linearization of (1) and (2) is typically achieved by assuming that the key variables are stationary around their (deterministic) steady states. If the variables are non-stationary these derivations are no longer valid.⁹ By linearizing around cointegration relations instead of variables, as in Altug (1989) and Ireland (2004) it is, however, possible to derive (1) and (2) from a DSGE model when the variables are difference stationary.¹⁰ This

⁸Equations (1) and (2) are still exact even when the optimal inflation target is time varying as noted by Juselius (2008). However, the parameters may not be identifiable in this case.

⁹Even if stationarity is assumed, as is common in the literature, the variance of the key variables tend to be very large which makes linearizing around their respective steady states a very poor approximation. Moreover, inference is highly unreliable in small samples with near unit roots as demonstrated by Johansen (2006).

¹⁰In fact, this easily achieved for (1). A typical Euler equation for consumption is given by

$$E_t (C_t/C_{t+1})^{-\lambda_1} = \lambda_2(1 + i_t)E_t (P_t/P_{t+1})$$

where C_t is consumption, P_t is the price level, and λ_i are parameters. If $\ln(C_t) \sim I(1)$, $\ln(P_t) \sim I(2)$, and $\ln(1 + i_t) \approx i_t \sim I(1)$, and the log of the real interest rate $1 + r_t = (1 + i_t)E_t (P_t/P_{t+1})$ is $I(0)$, the Euler equation can be linearized around the log-normally distributed stationary variables C_t/C_{t+1} and $1 + r_t$. Using the equilibrium condition $C_t = Y_t$ and subtracting y_t^f from both sides of the linearized equation yields the purely forward looking version of (1). Note that subtracting y_t^f is strictly not needed since we do not linearize around it. A similar derivation for the NKP-curve is much more involved but should in principle yield an equation similar to (2) for some models of price stickiness (*e.g.* Roberts, 1995).

possibility is not fully explored here. Instead, I will take the view in Fanelli (2008) and Dees et al. (2008) that (1) and (2) are likely to capture the essential features of inflation and output dynamics regardless of the properties of the individual series.

Second, the sources of stochastic trends in (1), (2), and the policy rule need to be specified. The only possible sources of stochastic trends are the exogenous variables (see for instance Framroze Møller, 2008), i.e. x_t and y_t^f . In addition, the nominal interest rate can be an additional source in (1) and (2) if we allow for a time varying inflation target, $(\Delta p)_t^*$. Once stochastic trends in the exogenous variables are permitted, restrictions on the parameters of (1) and (2) must be considered on order to ensure consistency.¹¹

Consider first the restriction $\varphi_{21} + \varphi_{23} = 1$, which is frequently imposed on the NKP-curve. When x_t contains a stochastic trend, this restriction implies the implausible result that $\Delta p_t \sim I(2)$ unless $\varphi_{22} = 0$. Thus, to avoid this problem, $\varphi_{21} + \varphi_{23} < 1$ must be imposed instead.

Similarly, the restriction $\varphi_{11} + \varphi_{13} = 1$ is often imposed in the literature. Under this restriction, (1) can be expressed in terms of $\Delta \tilde{y}_t$ and v_t is stationary. In this case, the real interest rate, $r_t = i_t - E_t \Delta p_{t+1}$, must be stationary if the NK-model is true, implying that both i_t and Δp_t must share the stochastic trend in x_t . On the other hand, the case $\varphi_{11} + \varphi_{13} < 1$ is not possible unless y_t is stationary, since otherwise the non-stationary latent variable v_t does not cancel in (1). Hence, the restriction $\varphi_{11} + \varphi_{13} = 1$ must be applied when (1) is linearized around a non-stationary flexible price level of output.

An alternative version of (1), in terms of y_t rather than \tilde{y}_t , can be obtained by linearizing around the difference of y_t or even around a stationary non-linear combination of y_t and the real interest rate. In this case $\varphi_{11} + \varphi_{13} < 1$ can be permitted allowing for more interesting dynamics. For example, if the inflation target is time varying, the real interest rate can be non-stationary and still be consistent with the NK-model. Another advantage is that no measure of y_t^f is needed in the analysis. For these reasons, the specifications considered in the empirical analysis are

$$y_t = E_t y_{t+1} - \varphi_{32} (i_t - E_t \Delta p_{t+1}) \quad (3)$$

$$\Delta p_t = \varphi_{41} E_t \Delta p_{t+1} + \varphi_{42} x_t. \quad (4)$$

¹¹Here, the discussion is restricted to such restrictions that have direct implications for the stochastic trends. For a more general discussion of restrictions on the parameters of the NK-model, see for instance Bardsen et al. (2004).

corresponding to the purely theoretical NK-model, and

$$y_t = \varphi_{51}E_t y_{t+1} - \varphi_{52}(i_t - E_t \Delta p_{t+1}) + \varphi_{53}y_{t-1} \quad (5)$$

$$\Delta p_t = \varphi_{61}E_t \Delta p_{t+1} + \varphi_{62}x_t + \varphi_{63}\Delta p_{t-1} \quad (6)$$

corresponding to the hybrid version of the model.

Two popular measures of x_t have been used in empirical work; labor's share of income, i.e. $x_t = w_t n_t / y_t p_t$, where w_t is wages and n_t is the number of employed, and the output gap, i.e. $x_t = y_t - y_t^n$, where y_t^n is some measure of potential output. The implicit assumption here is that the flexible price output gap is approximately equal to some measure of the gap between output and its potential y_t^n . The output gap should, in principle, be stationary by construction. But this implies that both inflation and nominal interest rates should be stationary as well, making it hard to reconcile (3) and (4) with the high degree of persistence typically found in empirical studies of these variables. This may also help to explain why φ_{42} or φ_{62} have been found to be insignificant in previous empirical applications with the output gap as a measure of marginal costs (see for example Galí and Gertler, 1999). However, the empirical results of section 5 show that unit-roots cannot be rejected in output gap measures for samples of up to 30 years of quarterly data. Thus, (3) and (4) may still be reasonable descriptions of short-run to medium-run price and output dynamics, even when output gaps are used as real marginal costs.

Finally, it should be noted that money is not absent from the New Keynesian model. However, the NKP-curve implies that money has no explanatory power over inflation. Moreover, if the central bank targets the interest rate, the money stock is completely determined by the relationship

$$m_t - p_t = \varphi_{51}y_t - \varphi_{52}i_t$$

implying that money has no interesting role to play in the model. This 'unimportance of money' consequence of interest rate targeting will also be tested empirically.

3 Data and information

This section introduces the data and discusses potential information sets that can be used to evaluate the NK-model. The data consists of quarterly U.S.

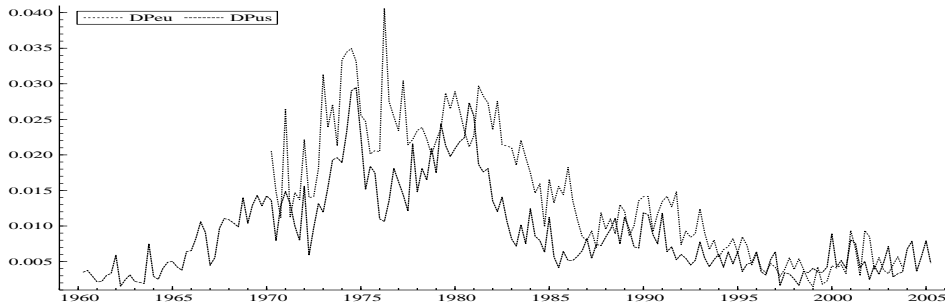


Figure 1: The Euro area and U.S. inflation rates.

and Euro area time series on the following variables (in logs): a price index, p_t , a nominal short-run interest rate, i_t , a real money aggregate, m_t , real output, y_t , potential output, y_t^n , and real aggregate wages, w_t . The Euro area data spans the years 1970:01-2003:04 and the U.S. data 1960:01-2005:02 (apart from a production function based measure of potential output which spans 1973:02-2003:04 and 1964:2-2005:02 respectively). Figure 1 plots Euro area and U.S. inflation rates. Detailed descriptions of the data are provided in appendix A.

3.1 Information sets

The minimal theory consistent information set that can be used to test the baseline NK-model is clearly $\mathcal{I}_0 = \{\Delta p, i, y, x\}$. This information set can be extended to include money, $\mathcal{I}_1 = \{\Delta p, i, m, y, x\}$, but if interest targeting in the NK-model provides a good description of central bank behavior the money stock should be completely determined by the other variables. As this is an interesting and testable hypothesis, \mathcal{I}_1 will be the basic information set in this paper.

Most empirical studies on the NKP-curve work with a smaller information set $\mathcal{I}_{npc} = \{\Delta p, x\}$ since the focus is on the Phillips curve.¹² Bardsen et al. (2004) have criticized the use of this kind of information set on the grounds that it is too small to account for the variation in the data, potentially leading to misspecified models. They show that this type of misspecification may explain the favorable results on the NKP-curve in the literature. When they

¹²Although, when GMM is used, the set of instruments usually contain other variables as well.

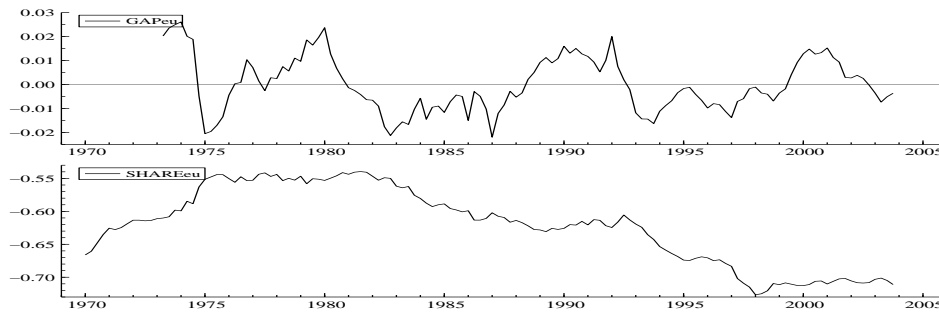


Figure 2: The production function based measure of the output gap and labor's share for the Euro area data.

extend the information set to include more variables, they find no support for the NKP-curve and that the results are consistent with a non-stationary inflation rate. However, the information set of Bardsen et al. is not exclusively motivated by the NK-model, whereas \mathcal{I}_1 is and, as shown below, sufficient to ensure a well-specified model.

The discussion in section 2 showed that different measures of real marginal costs have different implications for the core equations of the model. Which measure to use has been debated to some extent see Gali and Gertler (1999), Sbordone (2002) and Rudd and Whelan (2005a)). This paper takes a pragmatic approach and uses several output gap measures and labor's share as proxies for real marginal costs. To facilitate a comparison between the models that use different measures of x_t , only the extra information needed to replicate the relevant measure is added to the model's information set. For example, if the preferred measure of marginal costs is the output gap, the information set, $\mathcal{I}_{11} = \{\Delta p, i, m, y, y^n\}$, is modeled. In this case the output gap is defined as the *restriction*, $x_t = y_t - y_t^n$, on the statistical model. If labor's share is used, $\mathcal{I}_{12} = \{\Delta p, i, m, y, w\}$, is modeled and labor's share is defined as the *restriction*, $x_t = w_t - y_t$. Note that this information should be in the agents' information sets since they can always deduce y_t^n or w_t from y_t and x_t . Hence, there is no particular reason to restrict the information from the outset. Figure 2 plots the EU output gap and labor's share and figure 3 plots the corresponding U.S. measures.

It is clear from the figures that the two measures describe very different dynamics. In particular, the output gap measure appears to be in line with common views of the business cycle, while labor's share does not appear to

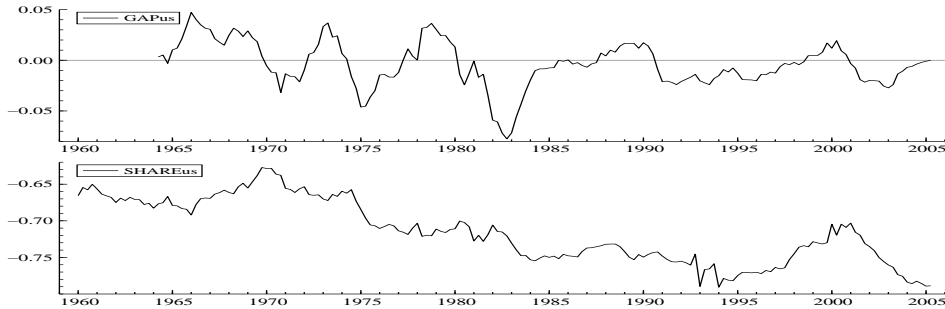


Figure 3: The production function based measure of the output gap and labor's share for the U.S. data.

capture cyclical variation to any noticeable degree.¹³

As a final note, given the difficulties to obtain a reasonable measure for potential output, an alternative measure based on the Hodrick and Prescott (1997) filter was also used. There were no significant differences in the results.¹⁴

4 Testing exact rational expectations within a cointegrated VAR model

This section describes the main results from Johansen and Swensen (2004) on testing rational expectations in a cointegrated VAR model when a linear trend is restricted to the cointegration space. The simpler case, with no deterministic trend in the model is similar and described in Johansen and Swensen (1999). The exact restrictions implied by the NK-model are presented at the end of the section.

The baseline statistical model is the p -dimensional VAR model with k lags in error correction form

$$\Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \mu_0 + \mu_1 t + \Phi D_t + \varepsilon_t \quad (7)$$

where the vector process X_t is assumed to be at most $I(1)$, $\varepsilon_t \sim N_p(0, \Sigma)$, Φ is a $p \times m$ matrix, and D_t consist of the other deterministic components.

¹³See Rudd and Whelan, 2005a for a discussion of this point.

¹⁴Giorno et al. (1995) discusses the relative merits of different potential output measures.

Cointegration can be investigated as the hypothesis that the matrix Π is of reduced rank, r . If $0 < r < p$ then at least some of the variables cointegrate and

$$\Pi = \alpha\beta'$$

where α and β are two $p \times r$ matrices of full column rank. Let the subscript \perp denote the orthogonal complement of a matrix. The deterministic trend is restricted to the cointegration space, i.e. $\alpha'_{\perp}\mu_1 = 0$, to avoid quadratic trends in the data. Thus, we can write $\mu_1 = \alpha\kappa_1$ where κ_1 is an r -dimensional vector. These assumptions imply that (7) can be written as

$$\Delta X_t = \alpha\beta^* X_{t-1}^* + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \mu_0 + \Phi D_t + \varepsilon_t \quad (8)$$

where $\beta^* = (\beta', \kappa_1)'$ is a $(p+1) \times r$ matrix and $X_{t-1}^* = (X'_{t-1}, t)'$.

Johansen and Swensen consider expectations of the form

$$E[c'_1 X_{t+1} | \Theta_t] + c'_0 X_t + c'_{-1} X_{t-1} + \dots + c'_{-k+1} X_{t-k+1} + c_c + c_{\tau}(t+1) + c_{\phi} D_{t+1} = 0 \quad (9)$$

where the $p \times q$ ($0 < q < r$) matrices c_i ($i = -k+1, \dots, 1$) are known as are the matrices c_{τ} and c_{ϕ} . The q -dimensional vector c_c can contain unknown parameters. The expectational equation (9) can be reformulated so that it corresponds to (8) by

$$E[c'_1 \Delta X_{t+1} | \Theta_t] - d'_1 X_t + d'_{-1} \Delta X_{t-1} + \dots + d'_{-k+1} \Delta X_{t-k+2} + c_c + c_{\tau}(t+1) + c_{\phi} D_{t+1} = 0 \quad (10)$$

where $d_{-i+1} = -\sum_{j=i-1}^{k-1} c_{-j}$, $i = 0, \dots, k$. Defining $d_1^* = (d'_1, -c_{\tau})'$, the restrictions on the statistical model (8) implied by (10) are

$$\begin{aligned} \beta^* \alpha' c_1 &= d_1^* \\ \Gamma'_i c_1 &= -d_{-i} \\ \mu'_0 c_1 &= -c'_c \\ \Phi' c_1 &= -c'_{\phi}. \end{aligned} \quad (11)$$

The maximum likelihood under the restrictions is

$$L_{H,max}^{-2/T} = |\tilde{\Sigma}_{22}^*| |S_{11}^*| \prod_{i=1}^{r-q} (1 - \tilde{\lambda}_i^*) / |c'_1 c_1| |c'_{1\perp} c_{1\perp}| \quad (12)$$

where $\tilde{\Sigma}_{22}^*$ is the likelihood of the marginal model, $c_1' \Delta X_t$, and the remaining terms are the likelihood of the conditional model, $c_{1\perp}' \Delta X_t$. The product in (12) is taken to be 1 if $q = r$. The maximum likelihood of the unconstrained model (8) is

$$L_{max}^{-2/T} = |S_{00}^*| \prod_{i=1}^r (1 - \hat{\lambda}_i^*).$$

The LR test statistic, given as -2 times the log of the ratio between the restricted and the unrestricted likelihoods, is

$$\begin{aligned} -2\ln Q &= T \left(\ln |\tilde{\Sigma}_{22}^*| + \ln |S_{11}^*| + \sum_{i=1}^{r-q} \ln(1 - \tilde{\lambda}_i^*) \right) \\ &\quad - T \left(\ln |S_{00}^*| + \sum_{i=1}^r \ln(1 - \hat{\lambda}_i^*) + \ln(|c_1' c_1| |c_{1\perp}' c_{1\perp}|) \right). \end{aligned}$$

The test statistic is asymptotically χ^2 -distributed with $kpq + q(m+1)$ degrees of freedom. Estimates of unknown parameters, φ , in the c_i matrices can be obtained by numerical optimization, provided the cointegrating relations can be expressed as smooth functions, $\beta(\varphi)$, of the parameters. In that case, the degrees of freedom turn out to be $kpq + q(m+1) - w$, where w is the number of additional unknown parameters. The core equations of the NK-model satisfy this condition as is evident from the representations of d_1^* below.

4.1 Restrictions implied by the NK-model

Let $X_t = (\Delta p_t, i_t, m_t, y_t, y_t^n)'$ and $k = 1$. In terms of (9) the pure NKP-curve, equation (4), takes the form

$$(-\varphi_{41}, 0, 0, 0, 0)E_t \begin{pmatrix} \Delta p_{t+1} \\ i_{t+1} \\ m_{t+1} \\ y_{t+1} \\ y_{t+1}^n \end{pmatrix} + (1, 0, 0, -\varphi_{42}, \varphi_{42}) \begin{pmatrix} \Delta p_t \\ i_t \\ m_t \\ y_t \\ y_t^n \end{pmatrix} = 0.$$

Hence, $c_1 = (-\varphi_{41}, 0, 0, 0, 0)'$ and $c_0 = (1, 0, 0, -\varphi_{42}, \varphi_{42})'$ which implies $d_1 = (\varphi_{41} - 1, 0, 0, \varphi_{42}, -\varphi_{42})'$. It is now straightforward to derive the restrictions on the parameters of (8) by using (11). Similarly, the pure IS curve in equation (3) can be expressed in terms of (9) by, $c_1 = (-\varphi_{32}, 0, 0, -1, 0)'$ and $c_0 = (0, \varphi_{32}, 0, 1, 0)'$. The extensions to equations (5) and (6) are obvious,

provided $k \geq 2$. It is also straightforward to use $X_t = (\Delta p_t, i_t, m_t, y_t, w_t)'$ as the IS curve restrictions are the same. However, the signs on φ_{42} are interchanged in c_0 and d_1 for the NKP-curve.

The simultaneous test of (3) and (4) can be performed by

$$c_1 = \begin{pmatrix} -\varphi_{32} & -\varphi_{41} \\ 0 & 0 \\ 0 & 0 \\ -1 & 0 \\ 0 & 0 \end{pmatrix}, c_0 = \begin{pmatrix} 0 & 1 \\ \varphi_{32} & 0 \\ 0 & 0 \\ 1 & -\varphi_{42} \\ 0 & \varphi_{42} \end{pmatrix}$$

provided that $r \geq 2$. Similar extensions as above are again obvious.

5 Testing the NK-model

In this section, the restrictions implied by the NK-model are tested on Euro area and U.S. data. Initial modeling of the data is performed prior to testing the restrictions, since information about cointegration rank is a prerequisite in the Johansen and Swensen method. We begin by analyzing Euro area data and then proceed with U.S. data.

5.1 Euro area data

This section reports the results of fitting the cointegrated VAR model (8) to Euro area data with $X_t^* = (\Delta p_t, i_t, m_t, y_t, y_t^n, t)'$ for the information set \mathcal{I}_{11}^{EU} and $X_t^* = (\Delta p_t, i_t, m_t, y_t, w_t, t)'$ for \mathcal{I}_{12}^{EU} . In the following, the former is referred to as the ‘‘gap model’’ and the latter as the ‘‘share model’’. Initial model analysis suggested that the lag length $k = 2$ is adequate in both models and that linear trends should be included in the cointegration spaces.

The reduced rank test statistic reported in Table 1 suggest that the rank is three in both models, though $r = 2$ was borderline accepted in the share model.¹⁵ Nevertheless, $r = 3$ seems to be the best choice based on further information in the model, such as the magnitude of the characteristic roots of the model, the graphs of the CI relations under the two choices, and the significance of the adjustment coefficients. This implies that there are

¹⁵A sensitivity analysis was conducted with respect to this choice but it did not change any of the results significantly. The results are available upon request from the author.

r	\mathcal{I}_{11}^{EU} , 1973:3-2003:4				\mathcal{I}_{12}^{EU} , 1970:1-2003:4			
	λ_i	trace	trace95	p-value	λ_i	trace	trace95	p-value
0	0.41	158.42**	88.55	0.00	0.34	152.30**	88.55	0.00
1	0.27	95.83**	63.66	0.00	0.32	96.15**	63.66	0.00
2	0.25	57.76**	42.77	0.00	0.14	45.34*	42.77	0.03
3	0.14	23.43	25.73	0.10	0.12	25.50	25.73	0.05
4	0.04	5.18	12.48	0.58	0.07	9.08	12.48	0.18

Table 1: The rank test statistic (trace test) for the full sample Euro area data. In the table, λ_i are the eigenvalues from the reduced rank regression (see Johansen, 1995). Trace95 are the 95%-quantiles of the trace distribution and (**) denotes rejection at the 1% significance level and (*) denotes rejection at the 5% significance level.

two common stochastic trends in the system. The only possible sources of these trends, given that the NK-model is true, are marginal costs and a time varying inflation rate.

Standard misspecification tests indicated some deviations from normality in both models,¹⁶ as well as minor problems with autocorrelation and ARCH in the share model. None of the variables were found to be stationary, nor long-run excludable in the two models. The results of these tests are reported in appendix C. Finally, recursive tests for constant parameters were also performed on both models.¹⁷ The results from these tests indicated two possible structural breaks, one in the early 1980's and one at around the middle of 1993. For this reason, separate analyzes for the full sample and for the subsample 1982:1-2003:4 were conducted.¹⁸ Similar breaks have been found, for example, by Batini (2006) and Barkbu and Batini (2005).

Table 2 reports the rank test statistic for the subsample 1982:1-2003:4. Again, the appropriate choice of rank seems to be three in the gap model, though $r = 2$ was almost accepted. For the share model the rank test sug-

¹⁶This is mainly due to some large outliers in the turbulent seventies. These outliers can be accounted for by dummy variables, but doing so does not change the results.

¹⁷The full description of these recursive test can be found in Hansen and Johansen (1999) and include two tests for the constancy of the β -vectors, a test for the constancy of the log-likelihood, a fluctuation test of the eigenvalues, among others. These results are available upon request.

¹⁸The subsamples 1970:1-1981:4 and 1993:2-2003:4 are too small for reliable estimation and hence not considered in the main text. However, the latter subsample is discussed in appendix C.

r	\mathcal{I}_{11}^{EU} , 1982:1-2003:4				\mathcal{I}_{12}^{EU} , 1982:1-2003:4			
	λ_i	trace	trace95	p-value	λ_i	trace	trace95	p-value
0	0.41	133.54**	88.55	0.00	0.45	112.46**	88.55	0.00
1	0.27	88.67**	63.66	0.00	0.24	60.50	63.66	0.09
2	0.25	46.32*	42.77	0.02	0.19	37.04	42.77	0.17
3	0.14	21.03	25.73	0.18	0.11	18.58	25.73	0.31
4	0.04	7.17	12.48	0.34	0.09	8.40	12.48	0.23

Table 2: The rank test statistic (trace test) for the 1982:1-2003:2 Euro area data. In the table, λ_i are the eigenvalues from the reduced rank regression (see Johansen, 1995). Trace95 is the 95%-quantiles of the trace distribution and (**) denotes rejection at the 1% significance level and (*) denotes rejection at the 5% significance level.

gests $r = 1$. Based on additional information in the model we find that $r = 2$ is the appropriate choice.¹⁹ Sensitivity analysis was conducted with respect to these choices but it did not change any of the results significantly. Note that the finding of $r = 2$ in the share model implies three common stochastic trends, which is inconsistent with the NK-model in section 2. No serious misspecification was detected in either model, except for some small deviations from normality. As before, stationarity was rejected for all variables in both models, but now long-run exclusion of w_t could not be rejected with a p-value of 0.51. Finally, recursive tests of parameter stability were performed and did not signal parameter instability over the period.

An interesting additional result is that the money stock is needed in the information sets. Long-run exclusion was rejected for this variable and removing it from the information sets considerably worsened the fit of each model. Moreover, the hypothesis that money has a unit vector in α was rejected (formally tested in appendix C), implying that money has some explanatory power over the other variables in the system. This result may that European central banks did not solely target the interest rate in the period before the EMU.

The results of testing the restrictions implied by the core equations of the NK-model are reported in table 3. The details of the estimations are provided in appendix B. The single equation restrictions are first considered

¹⁹The trace test point unambiguously toward $r = 2$ if w_t is excluded from the model. Since additional information should not in principle reduce the CI rank, this provides an additional reason for maintaining $r = 2$ despite the evidence for $r = 1$ in table 2. It seems that the inclusion of w_t in the information set “muddles the water”.

\mathcal{I}	T	Equ i	φ_{i1}	φ_{i2}	φ_{i3}	$-2\ln Q$	df	p-value
\mathcal{I}_{11}^{EU}	73:3-	3	–	-0.18	–	49.68	13	0.00
	“	4	1.11	-0.06	–	48.98	12	0.00
	“	5	0.79	-0.09	0.21	41.10	12	0.00
	“	6	1.79	-0.11	-0.66	30.44	11	0.01
	82:1-	3	–	-0.18	–	58.45	13	0.00
	“	4	1.16	-0.05	–	41.27	12	0.00
	“	5	0.83	-0.14	0.17	54.52	12	0.00
	“	6	1.96	-0.10	-0.87	20.86	11	0.04
\mathcal{I}_{12}^{EU}	70:1-	3	–	-0.33	–	55.83	13	0.00
	“	4	1.49	-0.05	–	55.72	12	0.00
	“	5	0.81	-0.20	0.19	48.73	12	0.00
	“	6	2.12	-0.04	-0.72	40.69	11	0.00
	82:1-	3	–	-0.22	–	44.10	13	0.00
	“	4	2.13	-0.10	–	14.28	12	0.28
	“	5	0.83	-0.17	0.17	40.24	12	0.00
	“	6	2.98	-0.11	-0.72	7.22	11	0.78

Table 3: Tests of the restrictions implied by the core equations of the NK-model (3)-(6) on Euro area data. The column “Equ i ” indicates that the restrictions implied by equation (i) is being tested and φ_{ij} are the corresponding estimates. In equation (5) we have the additional restriction $\varphi_{51} + \varphi_{53} = 1$ (hence, we have 12 degrees of freedom).

separately.

As can be seen from table 3, almost all restrictions are strongly rejected. Furthermore, the coefficient estimates are clearly not sensible within the NK-model. For instance, in the cases of the NKP-curve, equations (4) and (6), the coefficients on the forward terms, φ_{41} and φ_{61} are above one and the coefficient on the forcing variable is small and negative *regardless of the measure used for marginal costs*. Also, for the IS curve, the coefficient on the real interest rate has the wrong sign. Only the coefficients of the forward and backward terms can be considered sensible. In the few cases where the restrictions are not rejected, the coefficients are not plausible. If the coefficients are restricted to the unit interval in these cases, the restrictions are strongly rejected.

Finally, the restrictions from both (3)-(4) and (5)-(6) were tested simultaneously on all periods and all information sets. These restrictions were

r	$\mathcal{I}_{11}^{US}, 1964:2-2005:2$				$\mathcal{I}_{12}^{US}, 1960:1-2005:2$			
	λ_i	trace	trace95	p-value	λ_i	trace	trace95	p-value
0	0.24	114.22**	88.55	0.00	0.22	108.36**	88.55	0.000
1	0.19	70.06*	63.66	0.01	0.11	64.04*	63.66	0.046
2	0.13	36.86	42.77	0.18	0.09	42.66	42.77	0.051
3	0.06	14.55	25.73	0.62	0.07	24.40	25.73	0.074
4	0.03	5.15	12.48	0.58	0.06	10.18	12.48	0.121

Table 4: The rank test statistic (trace test) for the full sample U.S. data. In the table, λ_i are the eigenvalues from the reduced rank regression (see Johansen, 1995). Trace95 is the 95%-quantiles of the trace distribution and (**) denotes rejection at the 1% significance level and (*) denotes rejection at the 5% significance level.

strongly rejected in all cases, as should be expected, given the rejection of the single equation restrictions above.

These results imply that the evidence for the IS curve and the New Keynesian Phillips curve on Euro area data must be considered weak. The results of testing the NKP-curve are similar to those of Fanelli (2008) in this respect. In section 6 we discuss some reasons for this failure of the model.

5.2 U.S. data

Initial modeling of the two information sets for U.S. data suggested $k = 3$ that a linear trend should be restricted to the cointegration space in both models. The rank test statistic reported in Table 4 suggest that the rank is two in the first model. However, $r = 1$ is close to acceptance in the share model and there is also some uncertainty between the choice of $r = 2$ and $r = 3$. It appears that the inclusion of w_t in the information set again “muddles the water”. The choice $r = 3$ can be disregarded if one takes into account other information in the model as done before. Thus, only the results for $r = 2$ are reported below (see footnote 15). Again, this result implies three common stochastic trends which is inconsistent with the NK-model.

Standard misspecification tests indicated deviations from normality, due to some very large outliers, and problems with ARCH, stemming from the short-term interest rate series, in both models. None of the variables were found to be stationary nor long-run excludable in the two models. Finally, recursive tests for constant parameters were also performed. Both models showed evidence of a structural break at around 1979, marking the beginning

r	$\mathcal{I}_{11}^{US}, 1982:1-2005:2$				$\mathcal{I}_{12}^{US}, 1982:1-2005:2$			
	λ_i	trace	trace95	p-value	λ_i	trace	trace95	p-value
0	0.53	147.41**	88.55	0.00	0.51	137.34**	88.55	0.00
1	0.35	78.80**	63.66	0.00	0.36	73.83**	63.66	0.00
2	0.21	39.61	42.77	0.10	0.17	34.01	42.77	0.29
3	0.10	18.71	25.73	0.31	0.15	17.44	25.73	0.39
4	0.09	8.92	12.48	0.19	0.03	2.80	12.48	0.88

Table 5: The rank test statistic (trace test) for the subsample, 1982:1-2005:2, U.S. data. In the table, λ_i are the eigenvalues from the reduced rank regression (see Johansen, 1995). Trace95 is the 95%-quantiles of the trace distribution and (**) denotes rejection at the 1% significance level and (*) denotes rejection at the 5% significance level.

of the Volcker-Greenspan era. Similar structural breaks have previously been found in the empirical literature, for example in Roberts (2005) and Romer and Romer (2004). Because the ARCH problems in the short-run interest rate disappeared after 1982, the sample was split at that point. Roberts (2005) considers a similar split, but leaves out the years 79-83 corresponding to the Volcker disinflation era. A sensitivity analysis of this choice showed that the main results did not change. Thus, separate analyzes of the full sample and of the subsample 1982:1-2005:2 were conducted. In the share model, there was some evidence of a structural break around 1993.

Table 5 reports the rank test statistic for the sample 1982:1-2005:2. The appropriate choice of rank is two in both models. There was no serious misspecification in the gap model, apart from some small deviations from normality, while there were evidence of small problems with ARCH, autocorrelation, and deviations from normality in the second model. Stationarity was rejected for all variables in both models and again the long-run exclusion of w_t could not be rejected (p-value 0.40) in the second model. Finally, recursive tests for parameter stability were re-performed for the models. The tests did not show any serious parameter instability over the period in the gap model, while there were still some evidence of a break around 1993 in the share model.

The results of testing the restrictions implied by the core equations of the NK-model on the U.S. data are reported in table 6. Almost all restrictions are rejected, as can be seen from table 6. Furthermore, the coefficient estimates are very similar to those of the Euro area data. Finally, the restrictions from

\mathcal{I}	T	Equ i	φ_{i1}	φ_{i2}	φ_{i3}	$-2\ln Q$	df	p-value
\mathcal{I}_{11}^{US}	64:2-	3	–	-0.40	–	86.41	18	0.00
	“	4	1.13	-0.03	–	49.13	17	0.00
	“	5	0.80	-0.28	0.20	75.01	17	0.00
	“	6	1.52	-0.04	-0.54	24.43	16	0.08
	82:1-	3	–	0.06	–	72.63	18	0.00
	“	4	1.60	-0.02	–	50.71	17	0.00
	“	5	0.80	-0.27	0.20	75.53	17	0.00
	“	6	1.98	-0.07	-0.75	39.02	16	0.00
\mathcal{I}_{12}^{US}	60:1-	3	–	-0.43	–	73.71	18	0.00
	“	4	1.24	-0.05	–	45.02	17	0.00
	“	5	0.80	-0.29	0.20	61.34	17	0.00
	“	6	1.66	-0.04	-0.42	32.94	16	0.01
	82:1-	3	–	0.02	–	57.65	18	0.00
	“	4	1.66	-0.02	–	44.17	17	0.00
	“	5	0.70	0.00	0.30	38.54	17	0.00
	“	6	2.27	-0.01	-0.55	38.04	16	0.00

Table 6: Tests of the restrictions implied by the core equations of the NK-model (3)-(6) on the U.S. data. The column “Equ i ” indicates that the restrictions implied by equation (i) are being tested and φ_{ij} are the corresponding estimates. In equation (5) we have the additional restriction $\varphi_{52} + \varphi_{53} = 1$ (hence, we have 17 degrees of freedom).

both (3)-(4) and (5)-(6) were tested simultaneously on both periods and both models. These restrictions were strongly rejected in all cases. Hence, the evidence in favor of the NK-model on U.S data must also be considered weak.

6 Explaining the results

The reasons for the empirical failure of the NK-model are investigated in this section. We begin by discussing a particular condition on cointegration for the NK-model. This condition is then tested and interpreted in light of previous findings in the literature. The estimated coefficients in tables 3 and 6 are also given an interpretation.

6.1 Cointegration implications of the NK-model under $I(1)$ data

It is easy to provide a necessary condition on cointegration implied by the NK-model, provided that the data is non-stationary and well described by model (8). In this case inflation must be cointegrated with the measure of marginal costs and output must be cointegrated with the real rate of interest.²⁰ This can be seen directly, by observing that if the first restriction

$$\beta^* \alpha' c_1 = d_1^* \quad (13)$$

in (11) holds, then $d_1^* \in sp(\beta^*)$. That $d_1^* \in sp(\beta^*)$ is only a necessary condition is clear, since (11) includes several other restrictions. The advantage of the condition $d_1^* \in sp(\beta^*)$ is that it is very easy to verify on data.

For each of equations (3)-(6), d_1^* will take an explicit form (the d_1^* corresponding to equation i is denoted by d_{i1}^*). Thus, if $x_t = y_t - y_t^n$ we get

$$\begin{aligned} d_{31}^* &= (-\varphi_{32}, \varphi_{32}, 0, 0, 0, 0)' \\ d_{41}^* &= (1 - \varphi_{41}, 0, 0, -\varphi_{42}, \varphi_{42}, 0)' \\ d_{51}^* &= (-\varphi_{52}, \varphi_{52}, 0, 1 - \varphi_{51} - \varphi_{53}, 0, 0)' \\ d_{61}^* &= (1 - \varphi_{61} - \varphi_{63}, 0, 0, -\varphi_{62}, \varphi_{62}, 0)' \end{aligned}$$

and if $x_t = w_t - y_t$, the signs on the coefficients φ_{42} and φ_{62} are changed. Assuming $\varphi_{ij} > 0$ for all i and j , it can be seen that d_{31}^* is nested in d_{51}^* by the restriction $\varphi_{51} + \varphi_{53} = 1$, and that d_{41}^* and d_{61}^* are similar. Note also the theoretically interesting cases $\varphi_{41} \neq 1$ and $\varphi_{61} + \varphi_{63} = 1$. Table 7 reports the results of testing whether these relations are in the estimated cointegration space. Attention here is restricted to the subsample starting in 1982:1.

It can be seen from the table, that the cointegration implications of the ‘‘IS’’ curve, the rows of d_{31}^* and d_{51}^* , are rejected in all cases. Hence, it is not surprising that the corresponding hypotheses in tables 3 and 6 were rejected. Table 7 also reveals some interesting facts about the NKP-curve (d_{41}^* and d_{61}^*). The cointegration implications of the NKP-curve on Euro area data are rejected if we use the output gap as a measure of marginal cost. Furthermore, it can easily be seen that the signs of the estimates are wrong in this case. However, if labor’s share is used as a measure instead, the

²⁰A similar necessary condition on cointegration for present value models is discussed by Campbell and Shiller (1987).

\mathcal{I}^{EU}		$\hat{\beta}_{\Delta p}$	$\hat{\beta}_i$	$\hat{\beta}_m$	$\hat{\beta}_y$	$\hat{\beta}_{y^n/w}$	$\hat{\beta}_t$	p-value
\mathcal{I}_{11}^{EU}	d_{31}^*	-1	1	0	0	0	0	0.00
	d_{41}^*	1	0	0	7.78	-7.78	0	0.01
	d_{51}^*	-5.17	5.17	0	1	0	0	0.00
	d_{61}^*	0	0	0	1	-1	0	0.02
\mathcal{I}_{12}^{EU}	d_{31}^*	-1	1	0	0	0	0	0.00
	d_{41}^*	1	0	0	0.08	-0.08	0	0.37
	d_{51}^*	-26.42	26.42	0	1	0	0	0.03
	d_{61}^*	0	0	0	-1	1	0	0.00
\mathcal{I}_{11}^{US}	d_{31}^*	-1	1	0	0	0	0	0.00
	d_{41}^*	1	0	0	-0.18	0.18	0	0.33
	d_{51}^*	-41.10	41.10	0	1	0	0	0.00
	d_{61}^*	0	0	0	1	-1	0	0.10
\mathcal{I}_{12}^{US}	d_{31}^*	-1	1	0	0	0	0	0.00
	d_{41}^*	1	0	0	0.09	-0.09	0	0.00
	d_{51}^*	-43.54	43.54	0	1	0	0	0.02
	d_{61}^*	0	0	0	-1	1	0	0.00

Table 7: Tests of the cointegration implications of the NK-model. Estimated β coefficients are denoted by $\hat{\beta}_x$, where x indicates the variable. The hypotheses in d_{61}^* are derived under the additional restriction $\varphi_{61} + \varphi_{63} = 1$.

condition holds and the signs on the coefficients are correct and of reasonable magnitude. This, then, provides a possible explanation for the success of the labor's share measure in the previous literature. In essence, since the data has been assumed to be stationary, what has been estimated is the necessary condition on cointegration implied by the NKP-curve. However, this finding is not sufficient to conclude that the NKP-curve is a good description of inflation. Methods that rely on less formal tests of the NKP-curve, and more generally the NK-model, run the risk of claiming success when only the cointegration implications are met.

Generally, cointegration between the key variables of any expectational equation of the form (9), is a necessary condition when the data is $I(1)$. Hence, investigating cointegration between the variables in the system provides valuable information for future theoretical developments. This, potentially interesting avenue, is not explored further in the present paper.

The opposite finding holds on U.S. data. The cointegration implications

are not rejected when the output gap measure is used, but rejected when the labor's share measure is used. This finding is contrary to the argument by Gali and Gertler (1999), among others, that labor's share is the correct measure of marginal costs. The result may also account for the poor performance of the labor's share based NKP-curve on U.S. data that has been previously observed.

Finally, note that d_{61}^* in table 7 essentially tests if the output gap or labor's share is stationary. The stationarity of the measures is rejected in most cases, with the exception of the U.S. output gap. However, output gap measures should be structurally stationary by construction and it is puzzling that stationarity is rejected for the European output gap. It appears that the persistence of the European business cycle creates a near unit-root problem in the output gap series.²¹

6.2 Coefficient estimates and solution of the RE system

The fact that neither Δp_t and x_t nor y_t and $r_t = i_t - \Delta p_t$ are cointegrated in most cases, may account for the implausible and strange coefficients in tables 3 and 6. For instance, note that the coefficients on x_t in the NKP-curve are consistently small compared to the coefficients on Δp_t . If the coefficients on x_t are not statistically different from zero, it seems reasonable that the coefficients capture the unit root behavior of inflation, rather than being meaningful in terms of the NKP-curve. Such results were found by Bardsen et al. (2004) on Euro area data.

The stability of the RE system can be investigated by the method proposed by Blanchard and Kahn (1980). To this end, (3)-(4) and (5)-(6), with $x_t = y_t - y_t^n$, are written in the form

$$\begin{pmatrix} X_{t+1} \\ E_t P_{t+1} \end{pmatrix} = A \begin{pmatrix} X_t \\ P_t \end{pmatrix} + \gamma Z_t \quad (14)$$

where γZ_t collects the exogenous variables y_t^n and i_t .²² In terms of (14), (3) and (4) are represented by $X_t = \emptyset$, $P_t = (y_t, \Delta p_t)'$, and

$$A = \begin{pmatrix} 1 - \frac{\varphi_{32}\varphi_{42}}{\varphi_{41}} & -\frac{\varphi_{32}}{\varphi_{41}} \\ -\frac{\varphi_{42}}{\varphi_{41}} & \frac{1}{\varphi_{41}} \end{pmatrix}. \quad (15)$$

²¹The unit-root in the output gap variable is rejected if the full sample, 1973:2-2003:4, is investigated.

²²Alternatively, both y_t^n and i_t could be treated as predetermined with roots less or equal to one in absolute value. Doing so does not add anything to the analysis.

Likewise, for equations (5) and (6) we have $X_t = (y_{t-1}, \Delta p_{t-1})'$, $P_t = (y_t, \Delta p_t)'$, and

$$A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \frac{-\varphi_{53}}{\varphi_{51}} & \frac{\varphi_{52}\varphi_{63}}{\varphi_{51}\varphi_{61}} & \frac{1}{\varphi_{51}} + \frac{\varphi_{52}\varphi_{62}}{\varphi_{51}\varphi_{61}} & \frac{-\varphi_{52}}{\varphi_{51}\varphi_{61}} \\ 0 & \frac{-\varphi_{63}}{\varphi_{61}} & \frac{-\varphi_{62}}{\varphi_{61}} & \frac{1}{\varphi_{61}} \end{pmatrix}. \quad (16)$$

The corresponding A matrices when $x_t = w_t - y_t$ are similar apart from some changes in the signs. Using the values from tables 3 and 6, the roots of (15) and (16) can be calculated. For all cases in the tables, one root is very close to unity, while the remaining roots are within the unit circle. In the different gap models the roots cluster around 1.02 and for the share models around 0.99. This suggest that there is no unique stable forward solution to the system and, since it is unlikely that we could reject the unit root statistically, that the solution is non-stationary. Hence, the proper interpretation of the estimates in tables 3 and 6 is that they confirm that the data is non-stationary and that the cointegration implications do not hold.

7 Conclusions

This paper applies the Johansen and Swensen (1999, 2004) method of testing linear rational expectations models, to testing the New Keynesian Model on U.S. and Euro area data. The tests were conducted on both the individual equations separately and on system as a whole. The NK-model was rejected on both U.S. and Euro area data. Several sensitivity analyzes with respect to the choice of measures, sample periods, etc. were also preformed but they did not change the results. Hence, the evidence for the NK-model must be considered weak.

Some potential reasons for the empirical shortcomings of the model were also discussed. Much of the previous literature has assumed stationarity on behalf of the key variables in the NK-model. However, this is empirically implausible, as shown in this paper among others. When non-stationarity is allowed, the equations of the NK-model do not satisfy, in most cases, a particular necessary condition, namely that the key variables must be cointegrated. Interestingly, the cointegration implications are satisfied when labor's share is used as a measure of marginal costs on Euro area data. This might explain the success of the NKP-curve previously reported for this measure and data.

In essence, what has previously been estimated is a cointegration relationship. This has then been interpreted as evidence in favor of the NKP-curve, although a formal test of this hypothesis is rejected. The cointegration implications are not satisfied on U.S. data, which accounts for the previously reported poorer performance of the model on the U.S. data.

The results also suggest a potential way forward. Cointegration between the key variables, is necessary condition of any linear rational expectation hypothesis, when the data is non-stationary. Thus, any exploratory investigation on cointegration between the variables within a economically meaningful information set clearly provides valuable information on potential extensions of the theoretical models.

A Data definitions and sources

This appendix provides the precise definitions and sources of the data that was used in the analysis. All data is available from the sources below (membership required for the AWM), or upon request from the author.

A.1 Euro area data

The main data source for the European data is the Area Wide Model (AWM) dataset, available from the Euro Area Business Cycle Network (EABCN, www.eabcn.org, see Fagan et al., 2001). Additional data was obtained from OECD databases. The data spans the years 1970:1-2003:4, with the notable exception of the production function based potential output series which spans 1973:2-2003:4.

p_t = (log of) GDP deflator, base year 1995 (AWM series YED). Sensitivity analyzes were conducted by using the CPI index (OECD, economic outlook) and a deflator at basic prices (AWM series YFD) but they did not change the results significantly.

r_t = Short-run interest rate (AWM series STN).

m_t = (log of) Real EMU monetary aggregate M3 in millions of EUR. The nominal series was obtained from OECD, main economic indicators, and deflated by the index used for p_t .

y_t = (log of) Real GDP (AWM series YER).

y_t^n = (log of) Potential real output. The main measure used was the production function based measure from AWM series YET. This measure was available from 1973:2-2003:4. Sensitivity analysis was conducted with Hodric-Prescott filtered real GDP (using scale parameters 400, 1600).

w_t = (log of) Total real compensation to employees (AWM series WIN deflated by p_t).

It should be pointed out that the transformation, $x_t = w_t - y_t$, is identical to the labor's share measure used in Clarida et al. (1999), apart from scaling.

A.2 U.S. data

The main source for the U.S. data is the OECD database (www.oecd.org). The data spans the years 1960:1-2005:2, with the notable exception of the production function based potential output series which spans 1964:2-2005:2.

p_t = (log of) GDP deflator, base year 2000. A sensitivity analysis was conducted by using the CPI index but it did not change the results significantly. Both series can be found in the OECD economic outlook database.

r_t = 3 month LIBOR, obtained from the OECD economic outlook database.

m_t = (log of) Real money stock M2 in millions of US dollars (OECD, economic outlook). Deflated by p_t .

y_t = (log of) Real GDP (OECD, economic outlook).

y_t^n = (log of) Potential real output. The main measure used in the analysis was the production function based measure (available from OECD, economic outlook). This measure was available from 1964:2-2005:2. Sensitivity analysis was conducted with Hodric-Prescott filtered GDP (using scale parameters 400, 1600).

w_t = (log of) Total real compensation to employees obtained from OECD, economic outlook (deflated by p_t). A sensitivity analysis was conducted by using total real wages and salaries. The transformation, $w_t - y_t$, corresponds very closely to the labor's share measure published by the Bureau of Labors Statistics (BLS, www.bls.gov). The results do not

change significantly if the BLS labor's share measure is used in the analysis.

B Optimization

This appendix describes the methods used to obtain the coefficient estimates of the unknown parameters in the c_i matrices of section 4. As noted by Johansen and Swensen (1999), as long as the functions of the parameters are smooth, numerical optimization techniques can be applied to maximize the likelihood function. To this end both grid search and the quasi Newton optimization algorithm by Broyden-Fletcher-Goldfarb-Shanno (BFGS) were used.

In some of the cases there were several local maxima, in which case a grid search over reasonable starting values was conducted. The reported parameters correspond to the maximum (in all cases, the other local maxima produced very low values of the likelihood and very extreme values of the parameters).

Restricting the parameters to the unit interval was conducted by setting $\varphi_{ij} = \frac{1}{1+|\mathcal{V}_{ij}|}$ and maximizing over \mathcal{V}_{ij} , and by grid search over the unit intervals. The hypotheses were strongly rejected in all cases.

C Miscellaneous results

Various results that are of interest, but strictly not needed in the main text, are reported in this appendix. Table 8 reports the tests for stationarity.²³

It was claimed in the text that a test for a unit vector in the α matrix for money was rejected. The results from testing this hypothesis on the subsample 1982:1- produced p-values 0.001 and 0.200 for the EU gap and share models respectively. Similarly, we get p-values 0.001 and 0.000 for the US gap and share models respectively. The results from the full sample tests were similar.

Finally, table 9 provides the results from testing the optimizing IS curve on Euro area data subsample 1993:3-2003:4.

²³The tests for trend stationarity were similar, apart from a few cases in the longer sample, where weak evidence for trend stationarity was found.

Test for stationarity								
\mathcal{I}	T	r	Δp_t	i_t	m_t	y_t	y_t^n	w_t
\mathcal{I}_{11}^{EU}	73:2-	3	0.00**	0.00**	0.00**	0.00**	0.00**	–
	82:1-	3	0.00**	0.00**	0.00**	0.00**	0.00**	–
\mathcal{I}_{12}^{EU}	70:1-	3	0.01**	0.00**	0.00**	0.01**	–	0.02*
	82:1-	2	0.01*	0.00**	0.00**	0.00**	–	0.01**
\mathcal{I}_{11}^{US}	64:2-	2	0.00**	0.00**	0.00**	0.00**	0.00**	–
	82:1-	2	0.00**	0.00**	0.00**	0.00**	0.00**	–
\mathcal{I}_{12}^{US}	60:1-	2	0.06	0.01*	0.00**	0.00**	–	0.00**
	82:1-	2	0.00**	0.00**	0.00**	0.00**	–	0.00**

Table 8: Test for stationarity. The table reports the p-values of the hypothesis. (*) and (**) indicates rejection at the 5% and 1% significance levels respectively.

\mathcal{I}	Equ i	φ_{i1}	φ_{i2}	φ_{i3}	$-2\ln Q$	df	p-value
\mathcal{I}_{11}^{EU}	3	–	0.29	–	24.19	13	0.03*
\mathcal{I}_{11}^{EU}	5	0.70	0.12	0.30	15.59	12	0.21
\mathcal{I}_{12}^{EU}	3	–	0.18	–	33.28	13	0.00**
\mathcal{I}_{12}^{EU}	5	0.69	0.06	0.31	24.26	12	0.02*

Table 9: Tests of the restrictions implied by the equations (3) and (5) on Euro area data subsample 1993:3-2003:4. The column “Equ i ” indicates that the restrictions implied by equation (i) is being tested and φ_{ij} are the corresponding estimates. In equation (5) we have the additional restriction $\varphi_{51} + \varphi_{53} = 1$ (hence, we have 12 degrees of freedom).

These results should be viewed with great caution since only 43 observations are used in the estimations. Nevertheless, the results point to the possibility of a structural break where-after output evolves according to (3) or (5), at least when the output gap is used. The coefficient φ_{32} , and to some extent φ_{52} , depending on the assumptions used to derive the hybrid version, is the inverse of a preference parameter σ , where σ stems from an utility function of the form $u(c, \cdot) = \frac{c^{1-\sigma}}{1-\sigma} + \dots$. The two first equations in table 9 imply the estimates $\sigma = 3.45$ and $\sigma = 8.33$, which are highly plausible. Furthermore, the weight to the forward variable is approximately 0.70 compared to 0.30 for the backward variable, in line with the beliefs of most researchers. However, the NKP-curve was rejected on this same sample, as were both the IS curve and the NKP-curve on the U.S. data.

References

- Altug, S., November 1989. Time-to-Build and Aggregate Fluctuations: Some New Evidence. *International Economic Review* 30 (4), 889–920.
- Bardsen, G., Jansen, E., Nymoen, R., 2004. Econometric Evaluation of the New Keynesian Phillips Curve. *Oxford Bulletin of Economics and Statistics* 66, 671–686, supplement.
- Barkbu, B., Batini, N., March 2005. The New-Keynesian Phillips Curve When Inflation is Non-Stationary. Ph.D. thesis, European University Institute, in PhD thesis "Essays on Prices" by Bergljot Barkbu (chapter 3).
- Batini, N., November 2006. Euro area inflation persistence. *Empirical Economics* 31 (4), 977–1002.
- Batini, N., Jackson, B., Nickell, S., 2005. An open-economy new Keynesian Phillips curve for the U.K. *Journal of Monetary Economics* 52, 1061–1071.
- Blanchard, J. O., Kahn, C. M., July 1980. The Solution of Linear Difference Models under Rational Expectations. *Econometrica* 48 (5), 1305–1312.
- Campbell, J. Y., Shiller, R. J., October 1987. Cointegration and Tests of Present Value Models. *The Journal of Political Economy* 95 (5), 1062–1088.
- Clarida, R., Gali, J., Gertler, M., December 1999. The Science of Monetary Policy: A New Keynesian Perspective. *Journal of Economic Literature* 37 (4), 1661–1707.
- Clarida, R., Gali, J., Gertler, M., 2002. A simple framework for international monetary policy analysis. *Journal of Monetary Economics* 49, 879–904.
- Cogley, T., Sargent, T., April 2005. Drift and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S. *Review of Economic Dynamics* 8 (2), 262–302.
- Dees, S., Pesaran, H., Smith, L., Smith, R., 2008. Identification of new keynesian phillips curves from a global perspective, institute for the Study of Labor (IZA) Discussion Papers No. 3298.

- Demery, D., Duck, N. W., 2003. Cointegration-Based Tests of the New Keynesian Model of Inflation, university of Bristol Discussion Paper No. 541.
- Erceg, C. J., Henderson, D. W., Levin, A. T., 2000. Optimal monetary policy with staggered wage and price contracts. *Journal of Monetary Economics* 46, 281–313.
- Fagan, G., Henry, J., Mestre, R., January 2001. An Area-Wide Model (AWM) for the Euro Area, ECB Working Paper No. 42.
- Fanelli, L., 2002. A new approach for estimating and testing the linear quadratic adjustment cost model under rational expectations and I(1) variables. *Journal of Economic Dynamics & Control* 26, 117–139.
- Fanelli, L., 2008. Testing the New Keynesian Phillips Curve Through Vector Autoregressive Models: Results from the Euro Area. *Oxford Bulletin of Economics and Statistics* 70 (2), 53–66.
- Framroze Møller, N., 2008. Bridging Economic Theory Models and the Cointegrated Vector Autoregressive Model, *Economics Discussion Paper*, 21.
- Fuhrer, J. C., August 1997. The (Un)Importance of Forward-Looking Behavior in Price Specifications. *Journal of Money, Credit and Banking* 29 (3), 338–350.
- Fuhrer, J. C., June 2000. Habit Formation in Consumption and Its Implications for Monetary-Policy Models. *The American Economic Review* 90 (3), 367–390.
- Fuhrer, J. C., Rudebusch, G. D., 2004. Estimating the Euler equation for output. *Journal of Monetary Economics* 51, 1133–1153.
- Galí, J., Monacelli, T., 2005. Monetary Policy and Exchange Rate Volatility in a Small Open Economy. *Review of Economic Studies* 72 (3), 707–734.
- Gali, J., Gertler, M., 1999. Inflation dynamics: A structural econometric analysis. *Journal of Monetary Economics* 44, 195–222.
- Gali, J., Gertler, M., Lopez-Salido, D. J., 2001. European inflation dynamics. *European Economic Review* 45, 1237–1270.

- Gali, J., Gertler, M., Lopez-Salido, D. J., 2005. Robustness of the estimates of the Hybrid New Keynesian Phillips Curve. *Journal of Monetary Economics* 52 (6), 1107–1118.
- Giordani, P., 2004. Evaluating New-Keynesian Models of Small Open Economy. *Oxford Bulletin of Economics and Statistics* 66, 713–733, supplement.
- Giorno, C., Richardson, P., Roseveare, D., van den Noord, P., 1995. Estimating Potential Output, Output Gaps and Structural Budget Balances, OECD Economics Department Working Papers, No. 152.
- Hansen, H., Johansen, S., 1999. Some tests for parameter constancy in cointegrated VAR-models. *Econometrics Journal* 2 (2), 306–333.
- Henry, B. S. G., Pagan, A. R., 2004. The Econometrics of the New Keynesian Policy Model: Introduction. *Oxford Bulletin of Economics and Statistics* 66, 581–607, supplement.
- Hodrick, R. J., Prescott, E. C., February 1997. Postwar U.S. Business Cycles: An Empirical Investigation. *Journal of Money, Credit and Banking* 29 (1), 1–16.
- Ireland, P., November 2004. Technology Shocks in the New Keynesian Model. *The Review of Economics and Statistics* 86 (4), 923–936.
- Ireland, P., December 2007. Changes in the federal reserve’s inflation target: Causes and consequences. *Journal of Money, Credit and Banking* 39 (8), 1851–1882.
- Johansen, S., 1995. *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press.
- Johansen, S., 2006. *Post Walrasian Macroeconomics*. Cambridge University Press, Cambridge, Ch. Confronting the economic model with the data, pp. 287–300.
- Johansen, S., Swensen, A. R., 1999. Testing exact rational expectations in cointegrated vector autoregressive models. *Journal of Econometrics* 93, 73–91.

- Johansen, S., Swensen, A. R., 2004. More on testing exact rational expectations in cointegrated vector autoregressive models: Restricted constant and linear term. *Econometrics Journal* 7, 389–397.
- Juselius, M., 2008. Cointegration Implications of Linear Rational Expectation Models, Bank of Finland Research Discussion Paper No. 6.
- Kara, A., Nelson, E., 2004. International Evidence on the Stability of the Optimizing IS Equation. *Oxford Bulletin of Economics and Statistics* 66, 687–712, supplement.
- Kozicki, S., 1999. How useful are taylor rules for monetary policy? *Economic Review*, Federal Reserve Bank of Kansas City QII, 5–33.
- Linde, J., September 2005. Estimating New-Keynesian Phillips curves: A full information maximum likelihood approach. *Journal of Monetary Economics* 52 (6), 1135–1149.
- Ma, A., 2002. GMM estimation of the new Phillips curve. *Economics Letters* 76, 411–417.
- Matheron, J., Maury, T.-P., 2004. Supply-side refinements and the New Keynesian Phillips curve. *Economic Letters* 82, 391–396.
- Matheson, T., February 2008. Phillips curve forecasting in a small open economy. *Economics Letters* 98 (2), 161–166.
- Mavroeidis, S., 2004. Weak Identification of Forward-looking Models in Monetary Economics. *Oxford Bulletin of Economics and Statistics* 66, 609–635, supplement.
- McAdam, P., Willman, A., 2004. Supply, Factor Shares and Inflation Persistence: Re-examining Euro-area New-Keynesian Phillips Curves. *Oxford Bulletin of Economics and Statistics* 66, 637–670, supplement.
- McCallum, B. T., Nelson, E., Aug. 1999. An Optimizing IS-LM Specification for Monetary Policy and Business Cycle Analysis. *Journal of Money Credit and Banking* 31 (3), 298–316.
- Nelson, C., Lee, J., 2007. Expectation horizon and the phillips curve: the solution to an empirical puzzle. *Journal of Applied Econometrics* 22 (1), 161–178.

- Razin, A., February 2005. Aggregate Supply, Investment in Capacity, and Potential Output. *Journal of Money, Credit and Banking* 37 (1), 179–190.
- Roberts, J. M., November 1995. New Keynesian Economics and the Phillips Curve. *Journal of Money, Credit and Banking* 27 (4), 975–984.
- Roberts, J. M., 2005. How Well Does the New Keynesian Sticky-Price Model Fit the Data? *Contributions to Macroeconomics* 5 (1), 1–37.
- Romer, C. D., Romer, D. H., 2004. Choosing the Federal Reserve Chair: Lessons from History. *Journal of Economic Perspectives* 18, 129–162.
- Rudd, J., Whelan, K., April 2005a. Does the Labor’s Share Drive Inflation? *Journal of Money, Credit and Banking* 37, 297–312.
- Rudd, J., Whelan, K., 2005b. New tests of the new-Keynesian Phillips curve. *Journal of Monetary Economics* 52 (6), 1167–1181.
- Sbordone, A., 2007. Inflation persistence: Alternative interpretations and policy implications. *Journal of Monetary Economics* 54, 1311–1339.
- Sbordone, A. M., 2002. Prices and unit labor costs: a new test of price stickiness. *Journal of Monetary Economics* 49, 265–292.
- Sbordone, A. M., 2005. Do expected future marginal costs drive inflation dynamics? *Journal of Monetary Economics* 52 (6), 1183–1197.
- Svensson, L. E. O., 2000. Open-economy inflation targeting. *Journal of International Economics* 50, 155–183.
- Tillmann, P., 2005. The New Keynesian Phillips Curve in Europe: does it fit or does it fail?, *Deutsche Bundesbank Discussion Paper Series* 1, No. 4.
- Walsh, C. E., 2003. *Monetary Theory and Policy*, 2nd Edition. The MIT Press.
- Woodford, M., 2003. *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.
- Yun, T., 1996. Nominal Price Rigidity, Money Supply Endogeneity, and Business Cycles. *Journal of Monetary Economics* 37, 345–370.