AARHUS UNIVERSITY
BUSINESS AND SOCIAL SCIENCES
DEPARTMENT OF ECONOMICS AND BUSINESS

CREATES
Center for Research in Econometric Analysis of Time Series

Global Hemispheric Temperature Trends and Co–Shifting: A Shifting Mean Vector Autoregressive Analysis

Matthew T. Holt and Timo Teräsvirta

CREATES Research Paper 2012-54

Department of Economics and Business
Aarhus University
Fuglesangs Allé 4
DK-8210 Aarhus V
Denmark

Email: oekonomi@au.dk
Tel: +45 8716 5515

Global Hemispheric Temperature Trends and Co–Shifting: A Shifting Mean Vector Autoregressive Analysis*

Matthew T. Holt†

Department of Economics, Finance & Legal Studies
University of Alabama

Timo Teräsvirta‡

CREATES, *Department of Economics and Business*
Aarhus University

November 21, 2012

Abstract

This paper examines trends in annual temperature data for the northern and southern hemisphere (1850–2010) by using variants of the shifting–mean autoregressive (SM–AR) model of González and Teräsvirta (2008). Univariate models are first fitted to each series by using the so called QuickShift methodology. Full information maximum likelihood (FIML) estimates of a bivariate system of temperature equations are then obtained. The system is then used to perform formal tests of co–system in the hemispheric series. The results show there is evidence of co–shifting in the temperature data, most notably since the early 1980s.

Keywords: Co–breaking; Co–shifting; Hemispheric surface temperatures; Vector nonlinear model; Structural change; Shifting–mean vector autoregression

JEL Classification Codes: C22; C32; C52; C53; Q54

---

* **Acknowledgement:** This research has been supported by a travel grant from the College of Commerce and Business Administration, University of Alabama and by the Danish National Research Foundation. Part of the work for this paper was carried out while the first author was visiting CREATES, Aarhus University, and the work continued during the second author's visit to the Department of Economics, Finance, and Legal Studies, University of Alabama. Kind hospitality at both institutions is gratefully acknowledged. Material from this paper has been presented at Aarhus University, Mississippi State University, Brigham Young University, and the University of Wisconsin, Madison. We wish to thank participants in these presentations for their comments. Errors and shortcomings in this work remain our own responsibility.

† Professor and Dwight Harrigan Endowed Faculty Fellow in Natural Resources Economics, Department of Economics, Finance & Legal Studies, University of Alabama, Box 870224, 248 Alston Hall, Tuscaloosa, AL 35487-0224, USA. Telephone: 205-348-8980. Fax: 205-348-0590. E-mail: mtholt@cba.ua.edu.

‡ CREATES, Department of Economics and Business, Aarhus University, Fuglesangs Allé 4, DK–8210 Aarhus V, Denmark. Telephone: +45 8716 5563. E-mail: tterasvirta@econ.au.dk

1 Introduction

In recent years there has been considerable debate in the climate change literature, interestingly enough, not about whether global warming can be detected in available time series data but rather the proper way to characterize this phenomenon in the modelling process. The essence of the debate is this: do global and hemispheric temperature data follow a unit root (difference stationary) process wherein shocks to the world's climate can be expected to have a permeant effect, perhaps with a shared stochastic trend (cointegration) between northern and southern hemispheric observations? Or do the observed data fluctuate around a deterministic, perhaps nonlinear trend wherein the mean (trend) of the series might exhibit occasional breaks or shifts? This recent debate is perhaps best characterized by four recent publications in the journal *Climatic Change*. These are: Gay-Garcia, Estrada and Sánchez (2009), Kaufmann, Kauppi and Stock (2010), Estrada, Gay and Sánchez (2010), and Mills (2010). Arguing in favor of trend stationarity with occasional but infrequent mean breaks, Gay-Garcia, Estrada and Sánchez (2009) build on prior work by Perron (1989), Perron (1990), Leybourne, Newbold and Vougas (1998), Harvey and Mills (2002), and others who have developed tests for unit roots against the alternative of trend stationarity with occasional mean breaks (shifts). Related work that has concluded that temperature data are best characterized by unit roots and, possibly, stochastic trends include Harvey and Mills (2001), Kaufmann and Stern (2002), Liu and Rodríguez (2005), Kaufmann, Kauppi and Stock (2006a,b), Johansen (2010), Breusch and Vahid (2011). Relevant studies that have assumed that temperature series are stationary but that they follow a deterministic and possibly breaking trend include Harvey and Mills (2001), Seidel and Lanzante (2004), Gil-Alana (2008a,b), Ivanov and Evtimov (2010), and Breusch and Vahid (2011).[1]

The approach taken to specifying and estimating statistical models of hemispheric temperature data and their concomitant trends will, of course, have important implications for conducting inferences and, as well, for making long–term predictions. Even so, it may well be the case that available sample sizes, at least for hemispheric data (approximately 160 annual observations), are not sufficient to allow an analyst to adequately distinguish between models which utilize first differencing along with a common stochastic trend or, alternatively, which rely on deterministic trends, possibly with infrequent breaks or shifts (see Mills, 2010, for a related discussion). We therefore do not presume to settle this issue here. What is the case, however, is that prior work on modelling temperature trends has not entertained the possibility of combin-

[1] Harvey and Mills (2001) report results for both stochastic trend models as well as models with deterministic trends that change (shift) in a potentially smooth manner. As well, Breusch and Vahid (2011) also explore the possibility of both stochastic and deterministic trends for available temperature data, with the latter being allowed to break at least once.

ing certain features of both approaches. There is a small but growing literature in time series econometrics on modelling and testing for deterministic co–trending (or co–breaking) amongst sets of two or more related variables. See, for example, Chapman and Ogaki (1993), Hendry and Mizon (1998), Bierens (2000), Camarero and Ordóñez (2006), Hendry and Massmann (2007), and Franchi and Ordóñez (2008). to out knowledge the concept of co–trending has not been tested previously in the context of climate change analysis.

Considering the aforementiond, the primary goal of the present paper is to build upon and exploit the prior literature on co–trending to explore the possibility that there is deterministic "co–shifting" in underlying hemispheric temperature means. As such, our analysis employs variants of the smooth transition autoregrssive model (1994), wherein time is used as the transition variable (see, e.g., Lin and Teräsvirta, 1994; González and Teräsvirta, 2008). In so doing we employ the so called Shifting–Mean Vector Autoregression (SM–VAR) framework. That is, unlike most prior studies in this area we allow for the possibility that shifts in the underlying means of the data generating process have occurred gradually over time, and that one or more of these shifts may be shared by both of the series in question. In terms of a modelling framework the present paper is most closely related to that of Harvey and Mills (2001). Even so, these authors did not consider the possibility of co–shifting between the northern and southern hemisphere in their analysis.

The plan of the paper is a follows. In Section 2 we present the univariate SM–AR model and in Section 3 its multivariate counterpart, the SM–VAR model. Section 4 is devoted to modelling and a discussion of the concept of co-shifting. The hemispheric temperature series are introduced in Section 5. Univariate SM-AR results using these series are considered in Section 6, whereas Section 7 contains the multivariate SM-VAR results. Finally, Section 8 concludes.

2 Univariate Modelling Framework

We begin with a brief review of the Shifting Mean Autoregressive (SM–AR) procedures referred to as QuickShift; more details regarding this methodology are provided by González and Teräsvirta (2008). QuickShift works by starting with a univariate autoregression for the temperature series in questions, y_t, defined as:

$$y_t = \delta(t) + \sum_{j=1}^{p} \phi_j y_{t-j} + \varepsilon_t, \tag{1}$$

where $\varepsilon_t \sim \text{iid}(0, \sigma^2)$ and where $\delta(t)$ is an intercept term that possibly varies with time. For example, the intercept term might be defined as:

$$\delta(t) = \delta_0 + \sum_{i=1}^{r} \delta_i G(\gamma(\eta_i), c_i, t^*), \tag{2}$$

and where

$$G(\gamma(\eta_i), c_i, t^*) = (1 + \exp\{-\gamma(\eta_i)(t^* - c_i)\})^{-1}, \quad \gamma(\eta_i) > 0, \tag{3}$$

a logistic function in time. As well, $t^* = t/T$, $t = 1, \ldots, T$, (T is the total number of observations). Following Goodwin, Holt and Prestemon (2011), for numerical reasons we specify $\gamma(\eta_i) = \exp(\eta_i)$.[2] Let L denote the usual lag operator: $\mathsf{L}^i x_t = x_{t-i}$. It then follows from (1) that, assuming the characteristic roots of $|1 - \sum_{j=1}^{p} \phi_j \mathsf{L}^j|$ lie outside the unit circle, the corresponding shifting mean for y_t at time t is given by

$$\mathsf{E}_t y_t = \left(1 - \sum_{j=1}^{p} \phi_j \mathsf{L}^j\right)^{-1} \delta(t).$$

The specification given by (1)–(3) provides considerable flexibility in modelling shifting means in time series data, depending on the number of shifts, r and the values taken by the parameters. For example, for large values of $\gamma(\eta_i)$ the underlying shift, which occurs at time c_i, can be rather abrupt. Alternatively, for small values of $\gamma(\eta_i)$ (and assuming for the moment that $r = 1$), the shift from δ_0 to $\delta_0 + \delta_1$ is smooth, and will occur over a number of periods.

3 Multivariate Modelling Framework

The primary objective of this paper is to evaluate possible relationships between hemispheric temperatures and there trends. In order to consider the two temperature series jointly we therefore define a nonlinear multivariate model that we call the Shifting Mean Vector Autoregressive (SM–VAR) model. It is a multivariate generalisation of the univariate Shifting Mean Autoregressive model introduced by González and Teräsvirta (2008) and described briefly in the previous section.

[2]This parametrization for γ_i automatically ensures that $\gamma_i > 0$ holds, which implies that the identification condition $\gamma_i > 0$ holds. Moreover, it facilitates a grid search wherein equal spacings between (large) γ_i values are not optimal.

The basic SM–VAR model is defined as follows:

$$\mathbf{y}_t \;=\; \boldsymbol{\delta}(t) + \sum_{j=1}^{p} \boldsymbol{\Phi}_j \mathbf{y}_{t-j} + \boldsymbol{\varepsilon}_t \tag{4}$$

where $\mathbf{y}_t = (y_{1t}, ..., y_{kt})'$ is a $k \times 1$ stochastic vector, $\boldsymbol{\varepsilon}_t \sim \text{iid}(\mathbf{0}, \boldsymbol{\Omega})$, where $\mathsf{E}\boldsymbol{\varepsilon}_t = \mathbf{0}$, and $\boldsymbol{\Omega}$ is a $k \times k$ positive definite covariance matrix, and $\boldsymbol{\Phi}_j$, $j = 1, ..., p$, are $k \times k$ parameter matrices. We assume that the roots of $|\mathbf{I} - \sum_{j=1}^{p} \boldsymbol{\Phi}_j z^j| = 0$ lie outside the unit circle. Furthermore, $\boldsymbol{\delta}(t) = (\delta_1(t), ..., \delta_k(t))'$ is a $k \times 1$ time–varying intercept vector comparable to (2), where

$$\delta_j(t) = \delta_{j0} + \sum_{i=1}^{r_j} \delta_{ji} G(\gamma(\eta_{ji}), c_{ji}, t^*) \tag{5}$$

and where again

$$G(\gamma(\eta_{ji}), c_{ji}, t^*) = (1 + \exp\{-\gamma(\eta_{ji})(t^* - c_{ji})\})^{-1}, \tag{6}$$

for all i and j, and where as before $t^* = t/T$ and $\gamma(\eta_{ji}) = \exp(\eta_{ji})$. From these assumptions it follows that (4) defines a nonstationary process, whereas $\{\mathbf{y}_t - \boldsymbol{\delta}(t)\}$ is a stationary sequence. Write (4) as

$$\left(\mathbf{I} - \sum_{j=1}^{p} \boldsymbol{\Phi}_j \mathsf{L}^j\right) \mathbf{y}_t \;=\; \boldsymbol{\delta}(t) + \boldsymbol{\varepsilon}_t. \tag{7}$$

Similar to (2), the shifting mean of $\mathbf{y}_t$ is obtained from the infinite–order vector moving average representation of (7):

$$\mathsf{E}_t \mathbf{y}_t = (\mathbf{I} - \sum_{j=1}^{p} \boldsymbol{\Phi}_j \mathsf{L}^j)^{-1} \boldsymbol{\delta}(t) \;=\; \sum_{j=0}^{\infty} \boldsymbol{\Psi}_j \boldsymbol{\delta}(t-j) \tag{8}$$

where $\boldsymbol{\Psi}_0 = \mathbf{I}$.

Assuming that $|\delta_{ji}| \leq M < \infty$ for all i and j, the vector $\boldsymbol{\delta}(t)$ is a vector of bounded elements, and so $\mathbf{y}_t$ is bounded in probability. In this respect the SM–VAR model differs from nonstationary VARs with stochastic as well as broken linear trends, that is, the SM–VAR is different than the hemispheric models considered by Kaufmann, Kauppi and Stock (2006a,b, 2010) and others. It is also different from the smooth transition trend model of Harvey and Mills (2002).

4 Modelling with the SM–VAR

4.1 Specification of the Model

In practice, the number of lags in (4) and the number of transitions (logistic functions) in $\delta_j(t)$ are unknown and have to be determined from the data. It may be noted that there is a potential identification problem due to the construction of (5). If $\gamma(\eta_{ji}) = 0$, that is, $\eta_{ji} \to -\infty$ in (6), δ_{ji} and c_{ji} are unidentified nuisance parameters. This is a common problem in many nonlinear time series models; see for example Teräsvirta, Tjøstheim and Granger (2010, Chapter 5) for discussion. As a result, specification of the number of transitions r_j has to be carried out from specific to general. In this work we first determine the number of transitions and the lag length p thereafter. Finding r_j can be done as follows:

1. Test linearity against (4). This can be done jointly or equation by equation. In the latter case, the test is a special case (only the intercept is nonconstant under the alternative) of the parameter constancy test in Lin and Teräsvirta (1994). The null hypothesis of the joint test is $\boldsymbol{\delta}(t) = \boldsymbol{\delta}$, which is formulated as $\gamma(\eta_{ji}) = 0$, for all i and j. Note that the test statistic is the same independent of the (positive) number of transitions in the alternative. Let α_0 denote the significance level of the test.

2. If the null hypothesis is rejected, estimate the SM–VAR model with a single transition equation by equation. Here it is done using QuickNet, see González, Hubrich and Teräsvirta (2009). The idea is to form a grid consisting of values of η_{j1} and c_{j1}, and estimate the parameters by minimizing the residual sum of squares with respect to η_{j1}, c_{j1}, and δ_{ji}, $i = 0, 1$. This amounts to carrying out a set of linear regressions, whose number equals the number of points in the grid.

3. More generally, suppose the jth equation contains q transitions. Then test q against $q+1$ transitions in (5) using the same test with the distinction that the null model now has a time–varying intercept (q transitions). The significance level is $\tau^q \alpha_0$, $0 < \tau < 1$. The significance level is lowered at each step to favour parsimony. If the null hypothesis is rejected, estimate the $q + 1$ regression coefficients of the intercept and the two nonlinear parameters in the $(q + 1)$th transition function using the grid. Continue until the first non–rejection of the null hypothesis. For further details, see González, Hubrich and Teräsvirta (2009).

In this work we determine the transitions assuming $p = 0$. This means that the test statistic has to be robustified against heteroskedasticity and autocorrelation. An alternative would be to first specify and estimate a VAR model with a constant intercept and determine the number of transitions thereafter. Since we use a reverse order the lag length p is determined after finding the number of transitions.

4.2 Estimation of Parameters

The accuracy of the linear estimates obtained by using the grid may already be sufficient for many practical purposes if the grid is sufficiently dense. If not, it is always possible to build a finer grid around the previous minimum and repeat the search. If this is not yet satisfactory, one can use the parameter values thus obtained as starting-values for nonlinear maximum likelihood estimation by a suitable algorithm; see for example Teräsvirta, Tjøstheim and Granger (2010, Chapter 12). In the application we have used the Broyden, Fletcher, Goldfarb and Shanno (BFGS) quasi–Newton method. Note, however, that setting the initial step–length plays a crucial role. Specifically, if it is too long, the BFGS algorithm may leap to a point that is too far from the local maximum nearest the starting–value in the sense that the maximum value of the log–likelihood ultimately found by the algorithm may even be smaller than the one obtained by the grid.

4.3 Co–shifting

In interpreting empirical results it may be interesting to know whether the shifts in the equations of our vector SM-AR model have anything in common. In order to do that, we make use of the concept of common features in Engle and Kozicki (1993). The shifting mean (or now the shifting intercept) is the feature, and if there is a linear combination of the elements of $\mathbf{y}_t$ such that the feature is eliminated, it is a common feature. To discuss the idea in our framework, consider a bivariate SM–AR model with two shifts:

$$\mathbf{y}_t = \boldsymbol{\delta}_0 + (\boldsymbol{\delta}_1 \odot \mathbf{G}_{1t}) + (\boldsymbol{\delta}_2 \odot \mathbf{G}_{2t}) + \sum_{j=1}^{p} \boldsymbol{\Phi}_j \mathbf{y}_{t-j} + \boldsymbol{\varepsilon}_t$$

where $\mathbf{G}_{it} = (G_{i1}(\gamma(\eta_{i1}), c_{i1}, t/T), G_{i2}(\gamma(\eta_{i2}), c_{i2}, t/T))'$, $i = 1, 2$, and $\boldsymbol{\delta}_i = (\delta_{i1}, \delta_{i2})'$. According to Engle and Kozicki (1993), the shift is a common feature if there is a 2–dimensional vector $\boldsymbol{\alpha}$ such that $\boldsymbol{\alpha}'\mathbf{y}_t$ is linear, that is, the two shifts have been eliminated. A necessary condition for this to happen is that $G_{ij}(\gamma(\eta_{ij}), c_{ij}, t/T) = G_i(\gamma(\eta_i), c_i, t/T)$, $j = 1, 2$, or in matrix notation:

$$\mathbf{G}_{it} = G_i(\gamma(\eta_i), c_i, t/T)\mathbf{1}_2, \; i = 1, 2 \tag{9}$$

where $\mathbf{1}_2 = (1,1)'$. This implies that the slope and location parameters in the two transition functions of both transitions are the same. If (9) is completed by the condition $\boldsymbol{\delta}_2 = \lambda\boldsymbol{\delta}_1$, $\lambda \neq 0$, there exists $\boldsymbol{\alpha} \neq \mathbf{0}$ that eliminates the shift.

These conditions are quite strong. Therefore, a partial or weak co–shifting can be of interest. For example, assume

$$\mathbf{G}_{1t} \;=\; G_1(\gamma(\eta_1), c_1, t/T)\mathbf{1}_2.$$

Then there exists a nonzero vector $\boldsymbol{\alpha}$ such that $\boldsymbol{\alpha}'\boldsymbol{\delta}_1 = 0$, which eliminates the first but not the second shift:

$$\boldsymbol{\alpha}'\mathbf{y}_t \;=\; \boldsymbol{\alpha}'\boldsymbol{\delta}_0 + \boldsymbol{\alpha}'(\boldsymbol{\delta}_2 \odot \mathbf{G}_{2t}) + \sum_{j=1}^{p} \boldsymbol{\alpha}'\boldsymbol{\Phi}_j\mathbf{y}_{t-j} + \boldsymbol{\alpha}'\boldsymbol{\varepsilon}_t.$$

If also

$$\mathbf{G}_{2t} \;=\; G_2(\gamma(\eta_{21}), c_{21}, t/T)\mathbf{1}_2$$

there exists another nonzero vector $\boldsymbol{\beta}$ such that $\boldsymbol{\beta}'\boldsymbol{\delta}_2 = 0$ and

$$\boldsymbol{\beta}'\mathbf{y}_t \;=\; \boldsymbol{\beta}'\boldsymbol{\delta}_0 + \boldsymbol{\beta}'(\boldsymbol{\delta}_1 \odot \mathbf{G}_{1t}) + \sum_{j=1}^{p} \boldsymbol{\beta}'\boldsymbol{\Phi}_j\mathbf{y}_{t-j} + \boldsymbol{\beta}'\boldsymbol{\varepsilon}_t.$$

As already discussed, a single vector $\boldsymbol{\alpha}$ can eliminate both shifts only if $\boldsymbol{\delta}_2 = \lambda\boldsymbol{\delta}_1$. Note that co–shifting does not mean that the shifts contribute to holding the two series together. This property of linear cointegration is not present in co–shifting. This is because the elements of $\boldsymbol{\delta}_1$ and $\boldsymbol{\delta}_2$ are not restricted to mimic this feature of cointegrated random variables.

The definition of weak co-shifting accords well with the definition of contemporaneous mean co-breaking, as stated in Hendry and Mizon (1998) and Hendry and Massmann (2007). These authors consider a location shift in the unconditional mean at time t: $\mathsf{E}(\mathbf{y}_t - \boldsymbol{\rho}_0) = \boldsymbol{\rho}_t$, If there is a time-point t such that $\boldsymbol{\rho}_t \neq \boldsymbol{\rho}_{t-1}$, this defines the co-break. Their definition, however, is only operational in connection with a single break. It works when

$$\begin{aligned}\boldsymbol{\rho}_t &= \boldsymbol{\rho}^{(0)} I(t/T < c) + \boldsymbol{\rho}^{(1)}\{1 - I(t/T < c)\} \\ &= \boldsymbol{\rho}^{(1)} + (\boldsymbol{\rho}^{(0)} - \boldsymbol{\rho}^{(1)}) I(t/T < c),\; \boldsymbol{\rho}^{(0)} \neq \boldsymbol{\rho}^{(1)}\end{aligned}$$

where $I(A)$ is an indicator function: $I(A) = 1$ when A holds, zero otherwise.

4.4 Testing Co–shifting Restrictions

Co–shifting may be viewed as a special case of common nonlinearity as defined by Anderson and Vahid (1998). They derived a general test of common nonlinearity as a test of overidentifying restrictions in the generalized method of moments framework. Our test of co–shifting is simply a test of parameter restrictions in the SM-VAR model. If we test against strong co–shifting, the null hypothesis is

$$\mathrm{H}_0\text{: } (\gamma(\eta_{i1}), c_{i1}) \;=\; \ldots \;=\; (\gamma(\eta_{ik}), c_{ik}) \,,\; i = 1, ..., r, \tag{10}$$

and $\boldsymbol{\delta}_j = \lambda_j \boldsymbol{\delta}_1$, $\lambda_j \neq 0$, $j = 2, ..., r$. This amounts to testing $k(3r-1)$ restrictions in (4). If only a subset of shifts are under test, the number of restrictions decreases. Note that the number of transitions need not be the same in all equations of (4).

To implement tests of co–shifting in the SM–VAR it is natural to consider a likelihood ratio test. The test statistic is defined as:

$$LR = T\left[\ln|\tilde{\Omega}| - \ln|\hat{\Omega}|\right], \tag{11}$$

where LR has an asymptotic $\chi^2_{q_0}$ distribution under the null hypothesis. In (11) $\hat{\Omega}$ denotes the maximum likelihood estimate of the residual covariance matrix for the general (unrestricted) model and $\tilde{\Omega}$ likewise denotes the corresponding estimates for the co–shifting (restricted) model that involves q_0 restrictions on the parameters of (4).

A well known problem with the test statistic in (11) is that its asymptotically $\chi^2_{q_0}$ distribution under the null is not a good approximation to its finite–sample null distribution when the dimension of the model and the null hypothesis are large compared to the length of the time series. In that case the test suffers from positive size distortion; see, for example, Candelon and Lütkepohl (2001) and Shukur and Edgerton (2002). A number of remedies have been proposed, but simulations almost invariably show that the best solution to the problem is to use Rao's F–test, see for example (Rao, 1973, p. 556). This does require, however, that the errors can be assumed independent and identically distributed. If the presence of conditional heteroskedasticity is suspected, one can generate a finite–sample null distribution of the test statistic by wild bootstrap; see, for example, Ahlgren and Catani (2012). This can be done even if there is no conditional or unconditional heteroskedasticity, but in that case computational ease speaks for Rao's F–test. Additional details on computing and using Rao's F in the context of an LM–based test for stationary vector autoregressive systems are provided by Teräsvirta, Tjøstheim and Granger (2010, pp. 100–102).

4.5 Evaluation

The estimated SM–VAR model has to be evaluated. One has to check whether or not the stability condition concerning the roots of $|\mathbf{I} - \sum_{j=1}^{p} \mathbf{\Phi}_j z^j| = 0$ holds. Misspecification tests also need to be carried out. They include the multivariate test of normality, see Lütkepohl and Krätzig (2004, p. 128), the multivariate test of no error autocorrelation adopted from Yang (2012), and the test of constancy of the error covariance matrix by Eklund and Teräsvirta (2007). Their framework can be used for testing against various alternatives. In this work the alternative is that the variances are changing smoothly over time whereas the correlations remain constant. It would also be possible to develop a test of linearity for the VAR component against smooth transition VAR. In our application, however, this component is rather minor, so we do not need such a test here.

5 Data

The data used in the empirical analysis are annual average hemispheric temperature data from 1850–2010, and are described in detail in Brohan et al. (2006). As mentioned previously, versions of these data have been used by various authors in recent years to empirically investigate feature of global warming, including Gay-Garcia, Estrada and Sánchez (2009), Kaufmann, Kauppi and Stock (2010), Estrada, Gay and Sánchez (2010), and Mills (2010). A time series plot of the basic data is reported in Figure 1. As illustrated there, temperatures in both hemispheres appear to have a slight downward trend between approximately 1880 and 1910, with temperatures in the southern hemisphere showing what appears to be a slightly steeper decline. Then, from about 1910 until approximately 1945 both series exhibit an upward trend, with temperatures in the northern hemisphere appearing to increase more rapidly than those in the southern. There is then a leveling off between the 1940s and approximately 1980, after which both series exhibit a rather steep upward trend and, moreover, appear to increase at approximately the same rate. Of course these observations are simply based on a casual inspection of the data and trends in Figure 1; they are not the result of any formal model specification, estimation, and testing strategy, to which we now turn.

6 Univariate Results

6.1 Stationary Shifting Means versus Stochastic Trends

Our working hypothesis is that each hemispheric temperature series may be adequately represented as being stationary around a deterministic albeit shifting mean. A plausible alternative, of course, is that each series is difference stationary, and therefore behaves in a manner consistent with possessing a unit root (i.e., a stochastic trend). In this sense the natural testing framework is consistent with that proposed originally by Kwiatkowski, Phillips, Schmidt and Shin (1992) (KPSS) wherein the null hypothesis is one of stationarity and the alternative is one of difference stationarity. The standard KPSS test is not appropriate, however, if the series in question contains several or more breaks or shifts over time. To this end Becker, Enders and Lee (2006) (BEL) propose a stationarity test similar to the KPSS test but where an unknown number of intercept shifts are accounted for by using a select number of low frequency terms from a Fourier approximation (Gallant, 1981). Because of the globally flexible properties of the Fourier approximation, the approach is useful even if the underlying shifts were, in fact, generated by some combination of logistic functions as defined by (2) and (3). See Becker et al. (2006) for additional discussion and simulation results regarding this point.

Following Becker, Enders and Lee (2006), we assume we can decompose the series y_t into four parts: (1) an intercept and a trend term; (2) leading terms from a Fourier approximation; (3) a random walk; and (4) a stationary error.[3] Specifically, consider the following model:

$$y_t = \delta_0 + \zeta t + \sum_{k=1}^{n} \{\alpha_k \sin(2\pi kt/T) + \beta_k \cos(2\pi kt/T)\} + r_t + \varepsilon_t \tag{12a}$$

$$r_t = r_{t-1} + u_t, \quad r_0 = 0 \tag{12b}$$

where ε_t are stationary errors and where $u_t \sim \text{iid}(0, \sigma_u^2)$. Likewise, in (12a) the sine and cosine terms are included as leading terms from a Fourier series used to approximate any unspecified shifts in the deterministic term. To that end, n denotes the number of (cumulative) frequencies used in the approximation ($n < T/2$), and k denotes a particular frequency. In practice it is likely the case that most shifts can be reasonably approximated by setting n to a rather low number, say, $n = 2$ or $n = 3$.

[3] Given the nature of the observed hemispheric temperature anomalies, we consider only the case where y_t is a trend–stationarity process. Another option is to consider the possibility that y_t is a level–stationary process, although such a specification hardly seems plausible for hemispheric temperature data.

The null hypothesis that the y_t series is stationary is then simply

$$\mathrm{H}_0 : \sigma_u^2 \;=\; 0 \tag{13}$$

so that the process defined in (12) is stationary around a smooth deterministic component.

The testing regression for the BEL test of stationarity is then:

$$y_t \;=\; \delta_0 + \zeta t + \sum_{k=1}^{n} \left\{ \alpha_k \sin\left(2\pi k t / T\right) + \beta_k \cos\left(2\pi k t / T\right) \right\} + e_t. \tag{14}$$

Let $\tilde{e}_t$, $t = 1, \ldots, T$, denote the residuals obtained from the estimate of (14). The BEL test statistic associated with testing (13) becomes

$$\tau_\tau\left(n\right) = \frac{1}{T^2} \frac{\sum_{t=1}^{T} \tilde{S}_t\left(n\right)^2}{\tilde{\sigma}^2}, \tag{15}$$

where $\tilde{S}_t\left(n\right) \;=\; \sum_{j=1}^{t} \tilde{e}_j$. The test statistic can be viewed as a comparison of the short–run variance to the long–run variance, $\tilde{\sigma}^2$. A nonparametric estimate of the long–run variance is typically obtained from:

$$\tilde{\sigma}^2 = \tilde{\gamma}_0 + 2 \sum_{j=1}^{\ell} w_j \tilde{\gamma}_j,$$

where $\tilde{\gamma}_j$ is the jth sample autocovariance of the residuals $\tilde{e}_t$ from (14), ℓ is the truncation lag and w_j are a set of weights. In implementing the test we use the Newey and West (1987) procedure, and set the truncation lag, ℓ, to the greatest integer less than or equal to $4\left(T/100\right)^{2/9}$ (see, e.g., Lee and Mossi, 1996).

The BEL stationarity test was carried out on the hemispheric temperature data. A key issue is the number of frequencies, n, to include in the auxiliary regression in (14). By considering all cumulative frequencies over the range $n = 1, \ldots, 5$, the Rissanen–Schwarz Bayesian Information Criterion (BIC) indicated that $n = 3$ was the preferred number of cumulative frequencies for both series. A truncation lag of $\ell = 4$ was used to estimate the long–run variance in the denominator of the $\tau_\tau\left(n\right)$ test statistic in (15). Test results are recorded in Table 1. As indicated there, in both instances the null hypothesis of stationarity (around a shifting mean) cannot be rejected at any usual significance level. Moreover, additional test results reported in Table 1–notably, a bootstrapped F test of the joint significance of the trigonometric Fourier terms–the so called F_{trig} test–indicate that movements of each series over time cannot be adequately characterized by a linear trend alone.

Results for the BEL stationarity test compare favorably with those reported previously by Harvey and Mills (2002) and Gay-Garcia, Estrada and Sánchez (2009), among others. The BEL test results therefore provide further support for the QuickShift and SM–VAR analysis, to which we now turn.

6.2 QuickShift Results

To implement QuickShift with the temperature data for the northern ($j = 1$) and southern ($j = 2$) hemispheres, we set $\alpha_0 = 0.20$ and $\tau = 0.25$ in controlling the significance level, $\alpha_q = \tau^q \alpha_0$, in the LM testing sequence. We set the upper limit for the number of mean shifts, $q_{\max}$, to ten; we restrict the search for c_i's to 100 equally spaced values in the $[0.05, 0.95]$ interval; and we search over 100 equally spaced values for η_i in the $[-1, 3.401]$ interval.[4] As well, the parameter $\gamma(\eta)$ is normalized by $\hat{\sigma}_{t^*}$, the 'standard deviation' of $t^* = t/T$, in order to render it unit free.

The results obtained by applying QuickShift with these settings are as follows:

Northern Hemisphere:

$$\begin{aligned} y_{1t} = & -\underset{(0.027)}{0.277} + \underset{(0.231)}{0.687}(1 + \exp\{-\underset{(-)}{1.823}[t^* - \underset{(-)}{0.950}]/0.289\})^{-1} \\ & + \underset{(0.056)}{0.253}(1 + \exp\{-\underset{(-)}{30}[t^* - \underset{(-)}{0.460}]/0.289\})^{-1} \\ & + \underset{(0.092)}{0.376}(1 + \exp\{-\underset{(-)}{30}[t^* - \underset{(-)}{0.905}]/0.289\})^{-1} \\ & - \underset{(0.040)}{0.117}(1 + \exp\{-\underset{(-)}{30}[t^* - \underset{(-)}{0.205}]/0.289\})^{-1} + \widehat{\varepsilon}_{1t} \end{aligned} \tag{16a}$$

$$R^2 = 0.777, \ \widehat{\sigma}_1 = 0.136,$$

Southern Hemisphere:

$$\begin{aligned} y_{2t} = & -\underset{(0.011)}{0.473} + \underset{(0.090)}{1.360}(1 + \exp\{-\underset{(-)}{1.744}[t^* - \underset{(-)}{0.932}]/0.289\})^{-1} \\ & - \underset{(0.045)}{0.115}(1 + \exp\{-\underset{(-)}{30}[t^* - \underset{(-)}{0.250}]/0.289\})^{-1} \\ & + \underset{(0.026)}{0.153}(1 + \exp\{-\underset{(-)}{30}[t^* - \underset{(-)}{0.096}]/0.289\})^{-1} + \widehat{\varepsilon}_{2t} \end{aligned} \tag{16b}$$

$$R^2 = 0.796, \ \widehat{\sigma}_2 = 0.117.$$

The estimated transitions in (16a) and (16b) appear in the order they are selected by QuickShift. All standard errors, which appear in parentheses below estimated parameters, are obtained from the heteroskedasticity–autocorrelation (HAC) consistent

[4]The corresponding grid for γ is $[0.368, 30]$.

covariance matrix, which in turn are obtained by applying the estimator developed by Newey and West (1987). Of course estimates for the γ_i and c_i parameters have no corresponding standard errors, as these values are determined by using the QuickShift grid search. Plots of the underlying transition functions (dash–dot lines), shifting means (dashed lines), and actual observations (solid lines) are reported in Panel A (northern hemisphere) and Panel B (southern hemisphere) of Figure 2. The dates that correspond to when each transition function for each series, $G(\hat{\gamma}_{ji}, \hat{c}_{ji}, t^*)$, $j = 1, 2$, $i = 1, \ldots, r_j$, equals, respectively, 0.1 (10%), 0.5 (Centre), and 0.9 (90%) are reported in the left–hand columns of Table 2.

The Quickshift procedure picks four transitions for the Northern hemisphere temperature series and three for the Southern. Of interest is that in both instances the first transition is rather slow, that is, the estimated γ_1 parameters, are qualitatively small (-1.823 and -1.744, respectively) and the corresponding c_1 parameters are qualitatively large (0.950 and 0.932, respectively). In Figure 2 these are the relatively long, smooth transitions. Moreover, given that both $\widehat{\delta}_{11}$ and that $\widehat{\delta}_{21}$ are positive, these transitions capture the general upward trend in northern and southern hemispheric temperatures over the past 160 years. As indicated in (16), the estimate $\widehat{\delta}_{21} = 1.360$ is almost twice as large as that of $\widehat{\delta}_{11} = 0.687$, which implies that the long–term increase in temperatures (i.e., over approximately the last 80 years) in the southern hemisphere has occurred at effectively twice the rate as in the northern. Finally, both of these transitions have an estimated c value that is at or close to the upper bound: 0.95. As recorded in Table 2, the implication is that these long–term upward trends do not near completion until well into the later part of the current century.

For both series the remaining shifts are rather abrupt–in each case the corresponding γ_i parameters are set at 30, the upper limit–and generally pick out various features of the data prior to 1940. See Table 2. The exception is the third shift identified for the northern hemisphere, which is abrupt, positive, and occurs toward the end of the sample period. Even so, the fourth shift for the northern series and the second for the southern seem to be rather sharp changes that occur at approximately the same time in the late 1800s. As indicated in Figure 2, both are associated with a rather abrupt but temporary downward movement in temperatures.

7 SM–VAR Results

The results of the previous section suggest that a shifting–mean vector autoregressive modelling approach is potentially feasible. As well, given the close correspondence between the long–term shifts (trends) for both hemispheres, the possibility of (weak) co–shifting also seems plausible. To this end we specified and estimated a bivariate SM–VAR for the hemispheric temperature data similar to that specified in (4).

The analysis begins by taking as given the number of shifts for each series, r_j, identified in the univariate analysis. We also use the parameter estimates obtained by QuickShift as starting values for the BFGS estimation of the bivariate system.[5] The system lag length, p, was determined by using a sequence of likelihood ratio tests; coincident with results reported by Harvey and Mills (2001), we find that $p = 3$ is adequate. The results of various system tests applied to the SM–VAR are reported in Table 3.

A question of primary interest is whether the two series in question exhibit any form of co–shifting. As already illustrated with the QuickShift results, there is reason to believe that the hemispheric temperature data have at least one logistic function intercept shift in common, and perhaps two. The possibilities include a long, relatively slow shift that started in the second half of the 20th century and another shift that occurred late in the 19th century. The results for a test of this co–shifting hypothesis are reported in Table 3.[6]

Based on either a standard Likelihood Ratio (LR) test or Rao's F, the null hypothesis of co–shifting cannot be rejected at any standard significance level. We therefore proceed by imposing the corresponding co–shifting restrictions. Following Kaufmann and Stern (1997) and Harvey and Mills (2001), we also perform Granger non–causality tests with respect to lags of northern temperatures in the southern equation. Results for this test, also reported in Table 3, strongly suggest that the implied exclusion restrictions cannot be rejected. Finally, by using the base SM–VAR model a test for remaining (excluded) mean shifts is performed. Specifically, this test is conducted by using a third–order Taylor approximation of the (omitted) logistic function in each equation in the system; again, the null hypothesis cannot be rejected at any plausible significance level. Based on these results we conclude that the SM–VAR with four mean shifts in the northern hemisphere equation; with three mean shifts in the southern hemisphere equation; with two sets of co–shifting restrictions imposed; and with temperatures in the southern hemisphere being treated as strongly exogenous is a reasonable representation of the hemispheric temperature data.

Estimation results for the preferred bivariate SM–VAR model are recorded in Table 4. As well, results for various system diagnostic tests are also reported at the bottom of Table 4. Importantly, the estimated SM–VAR with two co–shifting restrictions maintained provides a reasonable characterization of the hemispheric temperature data–there is no evidence of omitted of (vector) autoregressive errors; the estimated

[5] In obtaining ML estimates we follow established practice (see, e.g., van Dijk, Strikholm and Teräsvirta, 2003) and constrain values of $\gamma_i(\eta_i)$ to be bounded above, in this case at 50. Doing so helps avoid potential numerical problems. As with our implementation of QuickShift, we also constrain the values of c_i parameters to be bounded on the $[0.05, 0.95]$ interval.

[6] As a result of the way the shifts were re–ordered when obtaining the initial system estimates, the second candidate for co–shifting, that is, the shifts that occurred in the later part of the 19th century, are now associated with the fourth shift in the northern series and the third in the southern.

covariance matrix is seemingly stable over time; and the model's estimated residuals behave in a manner consistent with multivariate normality. Furthermore, the dominant root for the companion matrix associated with the VAR terms in the SM–VAR model is complex, but with a modulus of 0.584, implying the estimated model is dynamically stable. Plots of the implied shifting means along with the estimated transition functions are presented in Figure 3. As well, we again report dates that correspond to when each transition function for each series, $G(\hat{\gamma}_{ji}, \hat{c}_{ji}, t^*)$, $j = 1, 2$, $i = 1, \ldots, r_j$, equals, respectively, 0.1 (10%), 0.5 (Centre), and 0.9 (90%); these values are recorded in the right–hand columns of Table 2.

Results in Table 2 show that the estimated transition functions have some similarities to those obtained by using QuickShift, but there are some differences as well. For example, the leading logistic component for each equation (and the one associated with co–shifting) occurs more rapidly and is centered earlier (1994) than the comparable components obtained with QuickShift (2014 for the north and 2011 for the south). Of interest, and as indicated in Table 4, is that the corresponding estimates for $\hat{\delta}_{11}$ (1.017) and $\hat{\delta}_{21}$ (0.434), while continuing to be positive, have very different implications relative to the comparable QuickShift results. In other words, the SM–VAR results imply that over approximately the past 40 years temperatures in the north have been increasing at a rate that is over twice that of the south. This result seems reasonable in light of Figure 3. The other shared logistic function component captures a relatively slow shift in temperatures that occurred during the later part of the 19th and the early part of the 20th century (from approximately 1873 through 1920, and centered during 1897). See Table 2.

As identified in (8), the shifting mean for northern hemisphere temperatures in particular will be a function of all logistic function shifts included in the system, and notably those embedded in the equation for temperature anomalies in the southern hemisphere. Even so, because in the present case a form of co–shifting was found to hold, the shifting mean will be function of only five unique shifts. With the model's estimates in hand, it is possible to derive the shifting means for each series by using (8).[7] The results are reported in Figure 3 as the dashed line passing through the observed data points.

Focusing first on the shifting mean for the north in upper panel of Figure 3, results show that temperatures declined from the early 1880s until approximately 1912, at which point there is a slight but steady decrease until the early 1920s. Moreover, the mean shifts the temperatures in the northern hemisphere up until the early 1920s are, as indicated in upper panel of Figure 3, apparently due to the interactions of the fourth logistic function shift in the equation for the north and the third in the

[7] In approximating the moving average representation for the shifting means in (8), we truncate the lag order at six. Indeed, only minor changes to the imputed shifting means were observed after truncating the approximation at four lags.

equation for the south. Starting in the early 1920s mean temperatures in the north increase abruptly, with much of this shift being completed by the late 1920s. In turn, this shift is apparently due to the second logistic function in the equation for the north. See Table 2. The mean path for northern temperatures rises rather slowly from the early 1930s until the mid 1960s, with most of this increase attributable to the dominant first (and shared) logistic function shift. Mean temperatures for the north decline rapidly from the late 1960s until the early 1970s, due entirely to the third logistic function shift in the northern equation. From the mid 1970s on until the present, mean temperatures in the north increase rapidly, again due to the first logistic function shift.

The shifting mean for southern temperatures is plotted in the lower panel of Figure 3. As illustrated there, there is a notable increase in mean temperature beginning in 1850 continuing through 1885. This increase is attributed to third logistic function in the southern equation, a logistic function that is, moreover, co–shifting with the northern equation. Indeed, and as reported in Table 4, the estimate, $\hat{\delta}_{23}$, associated with this logistic function has the largest positive value among the three logistic function components included in the southern equation. Of additional interest is that no sustained increase in temperatures over the same period is observed for the northern hemisphere. The second logistic function shift in the southern equation is associated with a comparatively large and negative estimate of the corresponding δ_{22} parameter. The result, as indicated in Figure 3, is that the shifting mean for southern temperatures has a rather steep decline from 1885 until approximately 1905, at which point it begins to turn up again, and then increases rapidly from about 1910 until the early 1930s. The shifting mean for temperatures in the south continue to increase modestly from the late 1930s through the mid 1960s, at which point the effects of the first logistic function component in the southern equation begin to take hold (a component that is also co–shifting with the northern equation). Due entirely to this component (and, of course, the corresponding positive estimate for δ_{21}), the shifting mean for temperature anomalies in the south increases rapidly starting in the mid 1970s and continues to increase through the end of the sample period.

There are notable differences between the shifting means obtained by applying QuickShift and those obtained by FIML estimation. These differences are identified by contrasting Figure 2 with Figure 3. Compared with the QuickShift results, the FIML results that incorporate co–shifting apparently offer greater refinements to the estimates of shifting means. Importantly, the FIML results imply steeper increases in the shifting mean for northern temperatures over the past 35 years relative to the QuickShift results. What seems apparent is that QuickShift is a reasonable diagnostic tool, but that further insights may be obtained by using a system estimation framework where, moreover, appropriate co–shifting restrictions are incorporated.

Additionally, the model was simulated ahead for an additional 50 years. In conducting the simulations random errors were added to the model by drawing from the multivariate normal distribution with a covariance matrix equal to $\widehat{\Omega}$, the estimated covariance matrix for the SM–VAR model. In this manner the forward simulations were repeated 25,000 times. Actual values over the sample period (1860–2010) and simulated values (2011–2060) are illustrated in Figure 4. Approximate 95–percent confidence intervals are represented by the shaded areas during the forecasting period. The results show that temperatures in both hemispheres will continue to increase for at least the next twenty years, with temperatures in the south remaining somewhat below those in the north. Even so, the confidence intervals overlap, so that it is difficult to say if differences in future hemispheric temperatures will truly be significant.

As a final caveat, this analysis is based on the assumption that no further mean shifts in temperature occur. Given past experience, this assumption is probably too restrictive and, consequently, the confidence intervals presented in Figure 4 are likely too narrow. Constructing more realistic confidence intervals in the SM–VAR framework is an interesting research problem that is left for future work.

8 Conclusions

This paper builds on prior work that has examined co–trending among two or more variables. While there is a small but relevant literature on the topic of co–trending, to our knowledge this concept has not been used heretofore to examine co–movements in hemispheric temperatures. Nor, for that matter, has the logistic function framework (i.e., the TV–AR model of Lin and Teräsvirta, 1994) been formally developed as a useful construct for examining co–trending (co–shifting) amongst two or more related variables. This paper therefore accomplishes two important yet related goals. Firstly, we present a framework that defines the concept of and the testing for co–shifting in the context of an SM–VAR model. Secondly, we utilize the conceptual framework to examine the possibilities for co–shifting between surface temperatures in the northern and southern hemisphere.

In the empirical analysis the modelling sequence begins by using the QuickShift procedure developed by González and Teräsvirta (2008) to identify the number of relevant shifts in each series. In as much as QuickShift is a univariate procedure, and given that we wish to employ it in a bivariate setting, we find that it is expedient to use a relatively strict control sequence for the significance level cutoff applied in the tests conducted in each iteration of QuickShift. When we apply this approach to the hemispheric temperature data we find that four logistic function components are adequate to characterize the shifts in the mean of the northern series whilst only three are required for the southern series. These logistic function components and the parameter estimates embedded in them, as obtained by applying QuickShift, are then used as a starting point to estimate the parameters of the the bivariate SM–VAR. Subsequent testing reveal that two logistic function components are common to both equations, that is, the hemispheric temperatures data do, in fact, behave in a manner consistent with co–shifting. Moreover, one of these common shifts has occurred during the past 40 years, and, based on the estimate of the corresponding centrality parameter, is far from being complete.

Importantly, the estimated SM–VAR has been subjected to a battery of diagnostic and evaluative tests. In every instance the estimated model appears to be acceptable. We are therefore confident that the estimated SM–VAR model of hemispheric temperatures provides a reasonable representation of the data and, moreover, an accurate representation of the mean shifts that have occurred in these series over time.

References

Ahlgren, Niklas and Paul Catani, "Wild Bootstrap Tests for Autocorrelation in Vector Autoregressive Models," Hanken School of Economics Working Papers. No. 562, Hanken School of Economics, 2012.

Anderson, Heather M. and Farshid Vahid, "Testing Multiple Equation Systems for Common Nonlinear Components," *Journal of Econometrics*, 1998, *84*, 1–36.

Becker, Ralf, Walter Enders, and Junsoo Lee, "A Stationarity Test in the Presence of an Unknown Number of Smooth Breaks," *Journal of Time Series Analysis*, 2006, *27*, 381–409.

Bierens, Herman J, "Nonparametric Nonlinear Cotrending Analysis, with an Application to Interest and Inflation in the United States," *Journal of Business & Economic Statistics*, 2000, *18*, 323–337.

Breusch, Trevor and Farshid Vahid, "Global Temperature Trends," Monash Econometrics and Business Statistics, Working Paper 04–11, Monash University, Department of Econometrics and Business Statistics, 2011.

Brohan, P., J. J. Kennedy, I. Harris, S. F. B. Tett, and P. D. Jones, "Uncertainty Estimates in Regional and Global Observed Temperature Changes: A New Data Set from 1850," *Journal of Geophysical Research*, 2006, *111*, D12106.

Camarero, Mariam and Javier Ordóñez, "Is There a Nonlinear Co-Movement in the EU Countries' Unemployment?," *Economics Letters*, 2006, *93*, 157–162.

Candelon, Bertrand and Helmut Lütkepohl, "On the Reliability of Chow–type Tests for Parameter Constancy in Multivariate Dynamic Models," *Economics Letters*, 2001, *73*, 155–160.

Chapman, David A. and Masao Ogaki, "Cotrending and the Stationarity of the Real Interest Rate," *Economics Letters*, 1993, *42*, 133–138.

Eklund, Bruno and Timo Teräsvirta, "Testing Constancy of the Error Covariance Matrix in Vector Models," *Journal of Econometrics*, 2007, *140*, 753–780.

Engle, Robert F and Sharon Kozicki, "Testing for Common Features," *Journal of Business & Economic Statistics*, 1993, *11*, 369–380.

Estrada, Francisco, Carlos Gay, and Armando Sánchez, "A Reply to "Does Temperature Contain a Stochastic Trend? Evaluating Conflicting Statistical Results" by R. K. Kaufmann et al," *Climatic Change*, 2010, *101*, 407–414.

Franchi, Massimo and Javier Ordóñez, "Common Smooth Transition Trend–Stationarity in European Unemployment," *Economics Letters*, 2008, *101*, 106–109.

Gallant, A. Ronald, "On the Bias in Flexible Functional Forms and an Essentially Unbiased Form: The Fourier Flexible Form," *Journal of Econometrics*, 1981, *15*, 211–245.

Gay-Garcia, Carlos, Francisco Estrada, and Armando Sánchez, "Global and Hemispheric Temperatures Revisited," *Climatic Change*, 2009, *94*, 333–349.

Gil-Alana, L., "Time Trend Estimation with Breaks in Temperature Time Series," *Climatic Change*, 2008, *89*, 325–337. 10.1007/s10584-008-9407-z.

Gil-Alana, Luis A., "Warming Break Trends and Fractional Integration in the Northern, Southern, and Global Temperature Anomaly Series," *Journal of Atmospheric and Oceanic Technology*, 2008, *25*, 570–578.

González, Andrés and Timo Teräsvirta, "Modelling Autoregressive Processes with a Shifting Mean," *Studies in Nonlinear Dynamics & Econometrics*, 2008, *12*, No. 1, Article 1. Retrieved from: http://www.bepress.com/snde/vol12/iss1/art1.

González, Andrés, Kirstin Hubrich, and Timo Teräsvirta, "Forecasting Inflation with Gradual Regime Shifts and Exogenous Information," CREATES Research Papers 2009–03, School of Economics and Management, University of Aarhus, 2009.

Goodwin, Barry K., Matthew T. Holt, and Jeffrey P. Prestemon, "North American Oriented Strand Board Markets, Arbitrage Activity, and Market Price Dynamics: A Smooth Transition Approach," *American Journal of Agricultural Economics*, 2011, *93*, 993–1014.

Harvey, David I. and Terence C. Mills, "Modelling Global Temperature Trends Using Cointegration and Smooth Transitions," *Statistical Modelling*, 2001, *1*, 143–159.

_ and _ , "Unit Roots and Double Smooth Transitions," *Journal of Applied Statistics*, 2002, *29*, 675–683.

Hendry, David F. and Grayham E. Mizon, "Exogeneity, Causality, and Co–breaking in Economic Policy Analysis of a Small Econometric Model of Money in the UK," *Empirical Economics*, 1998, *23*, 267–294.

_ and Michael Massmann, "Co–Breaking: Recent Advances and a Synopsis of the Literature," *Journal of Business & Economic Statistics*, 2007, *25*, 33–51.

Ivanov, Martin A. and Stilian N. Evtimov, "1963: The Break Point of the Northern Hemisphere Temperature Trend During the Twentieth Century," *International Journal of Climatology*, 2010, *30*, 1738–1746.

Johansen, Søren, "The Analysis of Nonstationary Time Series Using Regression, Correlation and Cointegration with an Application to Annual Mean Temperature and Sea Level," CREATES Research Papers, School of Economics and Management, University of Aarhus, 2010.

Kaufmann, Robert, Heikki Kauppi, and James Stock, "Emissions, Concentrations, & Temperature: A Time Series Analysis," *Climatic Change*, 2006, *77*, 249–278. 10.1007/s10584-006-9062-1.

_ , _ , and _ , "The Relationship Between Radiative Forcing and Temperature: What Do Statistical Analyses of the Instrumental Temperature Record Measure?," *Climatic Change*, 2006, *77*, 279–289. 10.1007/s10584-006-9063-0.

_ , _ , and _ , "Does Temperature Contain a Stochastic Trend? Evaluating Conflicting Statistical Results," *Climatic Change*, 2010, *101*, 395–405. 10.1007/s10584-009-9711-2.

Kaufmann, Robert K. and David I. Stern, "Evidence for Human Influence on Climate from Hemispheric Temperature Relations," *Nature*, 1997, *388*, 39–44.

_ and _ , "Cointegration Analysis of Hemispheric Temperature Relations," *Journal of Geophysical Research*, 2002, *107(D2)*, 4012.

Kwiatkowski, Denis, Peter C. B. Phillips, Peter Schmidt, and Yongcheol Shin, "Testing the null hypothesis of stationarity against the alternative of a unit root : How sure are we that economic time series have a unit root?," *Journal of Econometrics*, 1992, *54*, 159–178.

Lee, Junsoo and Dante Mossi, "On improvements of Phillips-Perron unit root tests using optimal bandwidth estimates," *Applied Economics Letters*, 1996, *3*, 197–200.

Leybourne, Stephen, Paul Newbold, and Dimitrios Vougas, "Unit Roots and Smooth Transitions," *Journal of Time Series Analysis*, 1998, *19*, 83–97.

Lin, Chien-Fu Jeff and Timo Teräsvirta, "Testing the Constancy of Regression Parameters Against Continuous Structural Change," *Journal of Econometrics*, 1994, *62*, 211–228.

Liu, Hui and Gabriel Rodríguez, "Human Activities and Global Warming: a Cointegration Analysis," *Environmental Modelling & Software*, 2005, *20*, 761–773.

Lütkepohl, Helmut and Markus Krätzig, *Applied Time Series Econometrics* Cambridge Books, Cambridge: Cambridge University Press, 2004.

Mills, Terence C., "'Skinning a Cat': Alternative Models of Representing Temperature Trends," *Climatic Change*, 2010, *101*, 415–426.

Newey, Whitney K. and Kenneth D. West, "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 1987, *55*, 703–708.

Perron, Pierre, "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis," *Econometrica*, 1989, *57*, pp. 1361–1401.

_ , "Testing for a Unit Root in a Time Series with a Changing Mean," *Journal of Business & Economic Statistics*, 1990, *8*, 153–162.

Rao, C. Radhakrishna, *Linear Statistical Inference and its Applications*, New York: Wiley, 1973.

Seidel, Dian J. and John R. Lanzante, "An Assessment of Three Alternatives to Linear Trends for Characterizing Global Atmospheric Temperature Changes," *Journal of Geophysical Research*, 2004, *109*, D14108.

Shukur, Ghazi and David Edgerton, "The Small Sample Properties of the RESET Test as Applied to Systems of Equations," *Journal of Statistical Computation and Simulation*, 2002, *72*, 909–924.

Teräsvirta, Timo, "Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models," *Journal of the American Statistical Association*, 1994, *89*, 208–218.

_ , Dag Tjøstheim, and Clive W. J. Granger, *Modelling Nonlinear Economic Time Series*, Oxford: Oxford University Press, 2010.

van Dijk, Dick, Birgit Strikholm, and Timo Teräsvirta, "The Effects of Institutional and Technological Change and Business Cycle Fluctuations on Seasonal Patterns in Quarterly Industrial Production series," *Econometrics Journal*, 2003, *6*, 79–98.

Yang, Yukai, "Modelling Nonlinear Economic Time Series." PhD dissertation, Aarhus University 2012.

Table 1: Stationarity Test Results for Hemispheric Temperature Data.

	Northern Hemisphere	Southern Hemisphere
$\tau_\tau(n)$	0.0199	0.0172
F_{trig}	41.16 (0.001)	27.02 (0.001)
Critial Values:		
10%	0.0201	
5%	0.0222	
1%	0.0268	

Note: Test statistics are calculated by setting $\ell = 4$ for the truncation lag in computing the long–run variance, and by using $n = 3$ cumulative frequencies in the Fourier series approximation. The critical values are taken from Becker, Enders and Lee (2006), and correspond to a sample size of 100. F_{trig} is the standard F statistic associated with the null hypothesis that the corresponding trigonometric terms in the Fourier approximation should be excluded. Numbers in parentheses below the F_{trig} statistics are bootstrapped p–values based on 999 bootstrap draws of the null model without trigonometric terms. The F statistics as well as the empirical distributions of these statistics were constructed by using the Newey and West (1987) covariance estimator with truncation lag $\ell = 4$.

Table 2: Transitions and Shift Dates for QuickShift and the SM–VAR Model with Co–Shifting.

Panel A: Northern Hemisphere

	QuickShift					SM–VAR				
i	$\hat{\gamma}$	$\hat{c}$	10%	Centre	90%	$\hat{\gamma}$	$\hat{c}$	10%	Centre	90%
1	1.823	0.950	1958	2014	2070	3.890	0.900	1968	1994	2020
2	30	0.459	1932	1935	1938	29.971	0.461	1921	1925	1928
3	30	0.905	2003	2007	2010	49.459	0.743	1967	1969	1972
4	30	0.205	1891	1894	1897	4.288	0.284	1873	1897	1920

Panel B: Southern Hemisphere

	QuickShift					SM–VAR				
i	$\hat{\gamma}$	$\hat{c}$	10%	Centre	90%	$\hat{\gamma}$	$\hat{c}$	10%	Centre	90%
1	1.744	0.932	1952	2011	2070	3.890	0.900	1968	1994	2020
2	30	0.250	1898	1901	1905	7.027	0.275	1881	1896	1910
3	30	0.095	1873	1876	1880	4.288	0.284	1873	1897	1920

Note: Columns titled 10% (90%) denote the dates for which the relevant logistic function is associated with a value of 0.1 (0.9). Likewise, columns headed Centre denote dates for which $t^* = \hat{c}$ for the respective shift function.

Table 3: Results of Various System Tests for the Bivariate SM–VAR Model of Temperature Data in the Northern and Southern Hemispheres.

Hypothesis	$\ln L$	AIC	BIC	K	LR Test	Rao's F Test
Unrestricted‡	301.975	-9.055	-8.377	35	–	–
$H_0: \gamma_1 = \gamma_5, c_1 = c_5,$ $\gamma_4 = \gamma_7, c_4 = c_7$ Weak Co–shifting	298.536	-9.062	-8.461	31	$\chi^2_{(4)} = 6.878(0.142)$	$F_{(4,573)} = 1.567(0.181)$
$H_0: \Phi_{21} = \Phi_{22} = \Phi_{23} = 0$ $\vert\gamma_1 = \gamma_5, c_1 = c_5,$ $\gamma_4 = \gamma_7, c_4 = c_7$ Granger Non–Causality of North in the South	299.269†	-9.034†	-8.375†	34†	$\chi^2_{(3)} = 1.466(0.690)$	$F_{(3,567)} = 0.438(0.726)$
$H_0:$ No Remaining Mean Shifts $\vert\gamma_1 = \gamma_5, c_1 = c_5,$ $\gamma_4 = \gamma_7, c_4 = c_7$	301.506†	-9.024†	-8.307†	37†	$\chi^2_{(6)} = 5.940(0.430)$	$F_{(6,554)} = 0.065(0.999)$

‡ Results for the model with no co–shifting restrictions imposed but with strong exogeneity of temperatures in the Southern hemisphere maintained.

† Results are for the model estimated under the alternative.

Note: There are 158 usable observations. The unrestricted model contains 32 free parameters. $\ln L$ denotes the maximized likelihood function value. AIC and BIC denote, respectively, the system Akaike information criterion and the Rissanen–Schwarz Bayesian information criterion. The test statistic associated with χ^2_j is for a likelihood ratio test involving j parameter restrictions. Rao's F–test statistics are approximate system LM tests as described by Teräsvirta, Tjøstheim and Granger (2010). Values in parentheses are asymptotic p–values. The test for no remaining mean shifts is conducted by using a third–order Taylor approximation.

Table 4: System Estimates for the Hemispheric Temperature SM–VAR with Co–Shifting Restrictions Imposed.

Northern Hemisphere:[†]

$$y_t^n = -\underset{(0.043)}{0.147} + \underset{(0.257)}{1.017}(1+\exp\{-\underset{(0.685)}{3.890}[t/T - \underset{(0.036)}{0.900}]/0.289\})^{-1}$$

$$+\underset{(0.061)}{0.304}(1+\exp\{-\underset{(4.887)}{29.971}[t/T - \underset{(0.007)}{0.461}]/0.289\})^{-1}$$

$$-\underset{(0.064)}{0.296}(1+\exp\{-\underset{(11.982)}{49.459}[t/T - \underset{(0.005)}{0.743}]/0.289\})^{-1}$$

$$-\underset{(0.050)}{0.151}(1+\exp\{-\underset{(0.490)}{4.288}[t/T - \underset{(0.019)}{0.284}]/0.289\})^{-1}$$

$$+\underset{(0.070)}{0.144}\, y_{t-1}^n + \underset{(0.071)}{0.042}\, y_{t-2}^n - \underset{(0.071}{0.188}\, y_{t-3}^n + \underset{(0.097)}{0.392}\, y_{t-1}^s - \underset{(0.103)}{0.235}\, y_{t-2}^s + \underset{(0.099}{0.167}\, y_{t-3}^s + \widehat{\varepsilon}_t^n$$

$$R^2 = 0.846, \quad \widehat{\sigma}^n = 0.112, \quad Sk = 0.005, \quad Ek = 0.270, \quad LJB = 0.481(0.786)$$

Southern Hemisphere:[†]

$$y_t^s = -\underset{(0.037)}{0.299} + \underset{(0.111)}{0.434}(1+\exp\{-\underset{(0.685)}{3.890}[t/T - \underset{(0.036)}{0.900}]/0.289\})^{-1}$$

$$-\underset{(0.461)}{1.105}(1+\exp\{-\underset{(0.556)}{7.027}[t/T - \underset{(0.014)}{0.275}]/0.289\})^{-1}$$

$$+\underset{(0.443)}{1.292}(1+\exp\{-\underset{(0.490)}{4.288}[t/T - \underset{(0.019)}{0.284}]/0.289\})^{-1}$$

$$+\underset{(0.070)}{0.473}\, y_{t-1}^s - \underset{(0.080)}{0.182}\, y_{t-2}^s + \underset{(0.071)}{0.119}\, y_{t-3}^s + \widehat{\varepsilon}_t^s$$

$$R^2 = 0.858, \quad \widehat{\sigma}^s = 0.097 \quad Sk = 0.133, \quad Ek = -0.098, \quad LJB = 0.528(0.768)$$

$$\hat{\rho}_{ns} = 0.579$$

– –

System Statistics:

$$\ln L = 298.536, \ \text{AIC} = -9.062, \ \text{BIC} = -8.461, \ LM_{\Omega_t} = 3.196\times10^{-4}(0.999), \quad Sk_s = 0.681(0.711),$$

$$Ek_s = 0.351(0.839), \ LJB_s = 1.032(0.905), \ LM_{VAR(4)} = 0.260(0.999), \quad LM_{VAR(6)} = 0.378(0.997),$$

$$LM_{VAR(8)} = 0.517(0.988), \ LM_{VAR(10)} = 0.733(0.887), \ LM_{VAR(12)} = 1.027(0.428)$$

† See footnote 6.

Note: $\hat{\rho}_{ns}$ is the estimated correlation between the residuals. Sk denotes skewness and Ek excess kurtosis. LJB is the Lomnicki–Jarque–Bera test of normality of the residuals. These same statistics with a subscripted s are for the system, as described by Lütkepohl and Krätzig (2004). AIC and BIC denote, respectively, the system Akaike information criterion and the Rissamen–Schwarz Bayesian information criterion. LM_{Ω_t} denotes the system LM test of Eklund and Teräsvirta (2007) for a time–varying covariance matrix. $LM_{VAR(j)}$ denote system LM F–tests, based on Rao's F, for remaining vector autocorrelation at lags $j = 4, 6, 8, 10, 12$. Values in parentheses beside test statistics are p–values.

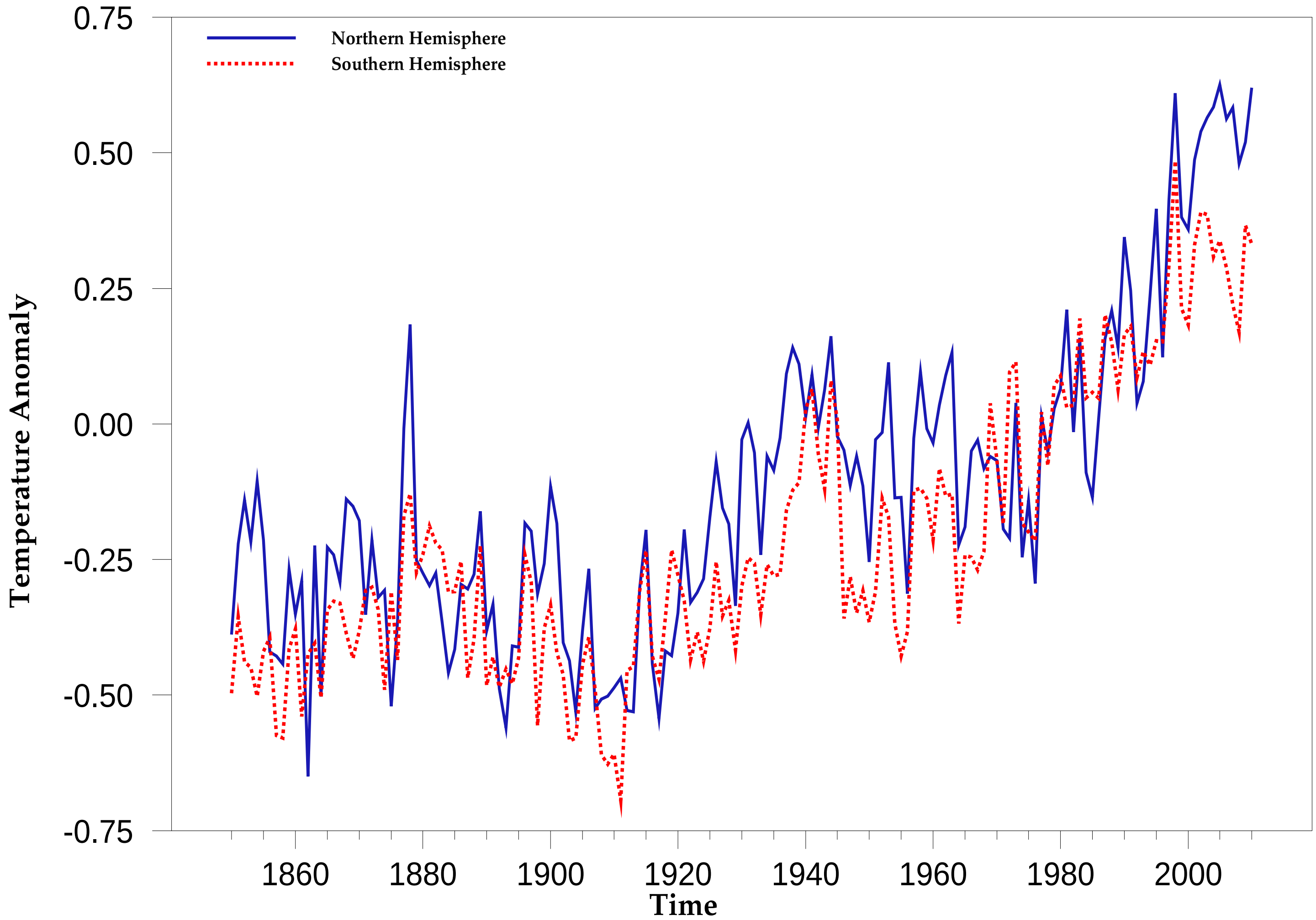

Figure 1: Temperature Anomalies in the Northern and Southern Hemisphere, 1850–2010.

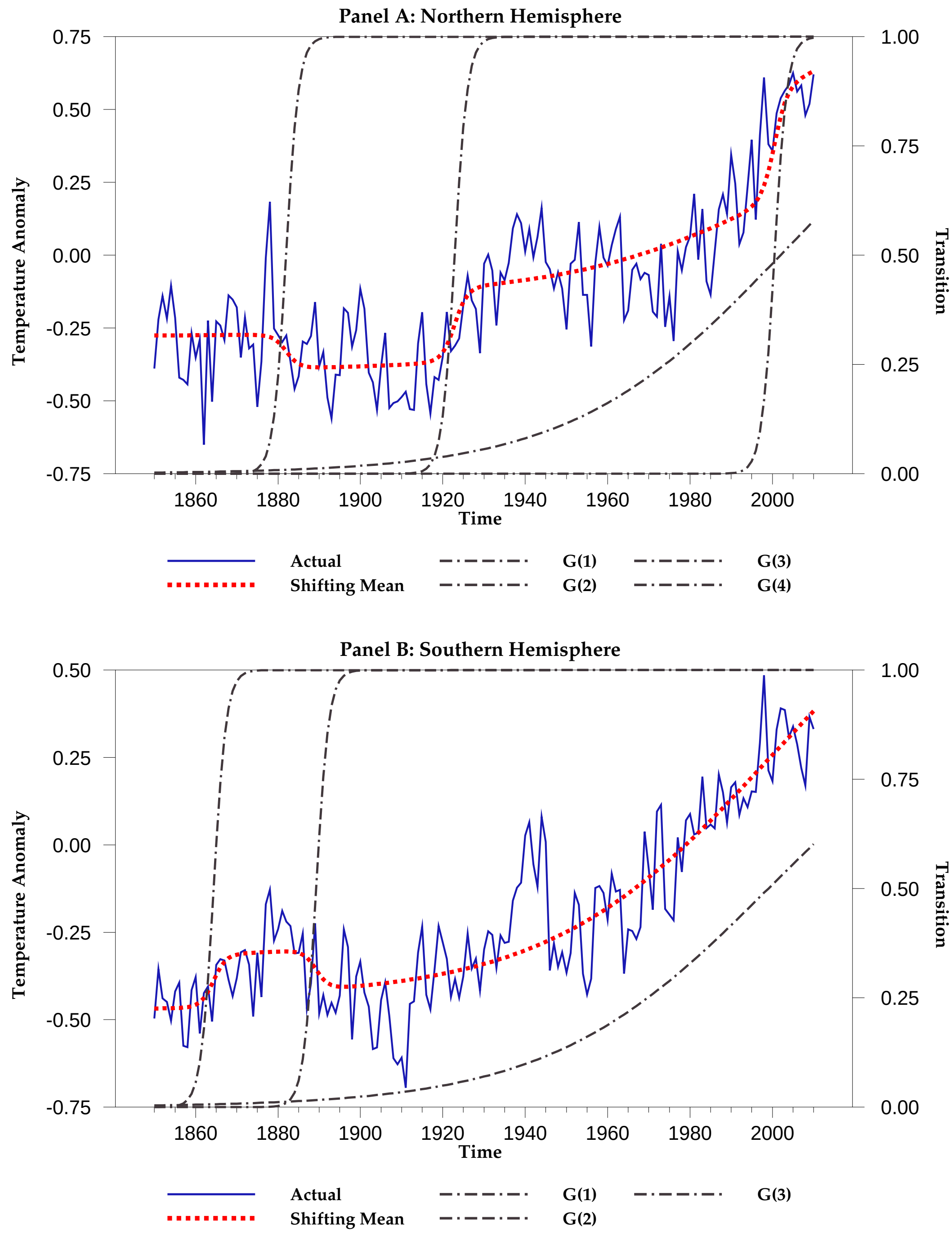

Figure 2: QuickShift Results for Temperature Anomalies for the Northern (Panel A) and Southern (Panel B) Hemispheres, 1850–2010. The dashed line indicates the shifting mean and the dash–dot lines indicate the estimated transition functions.

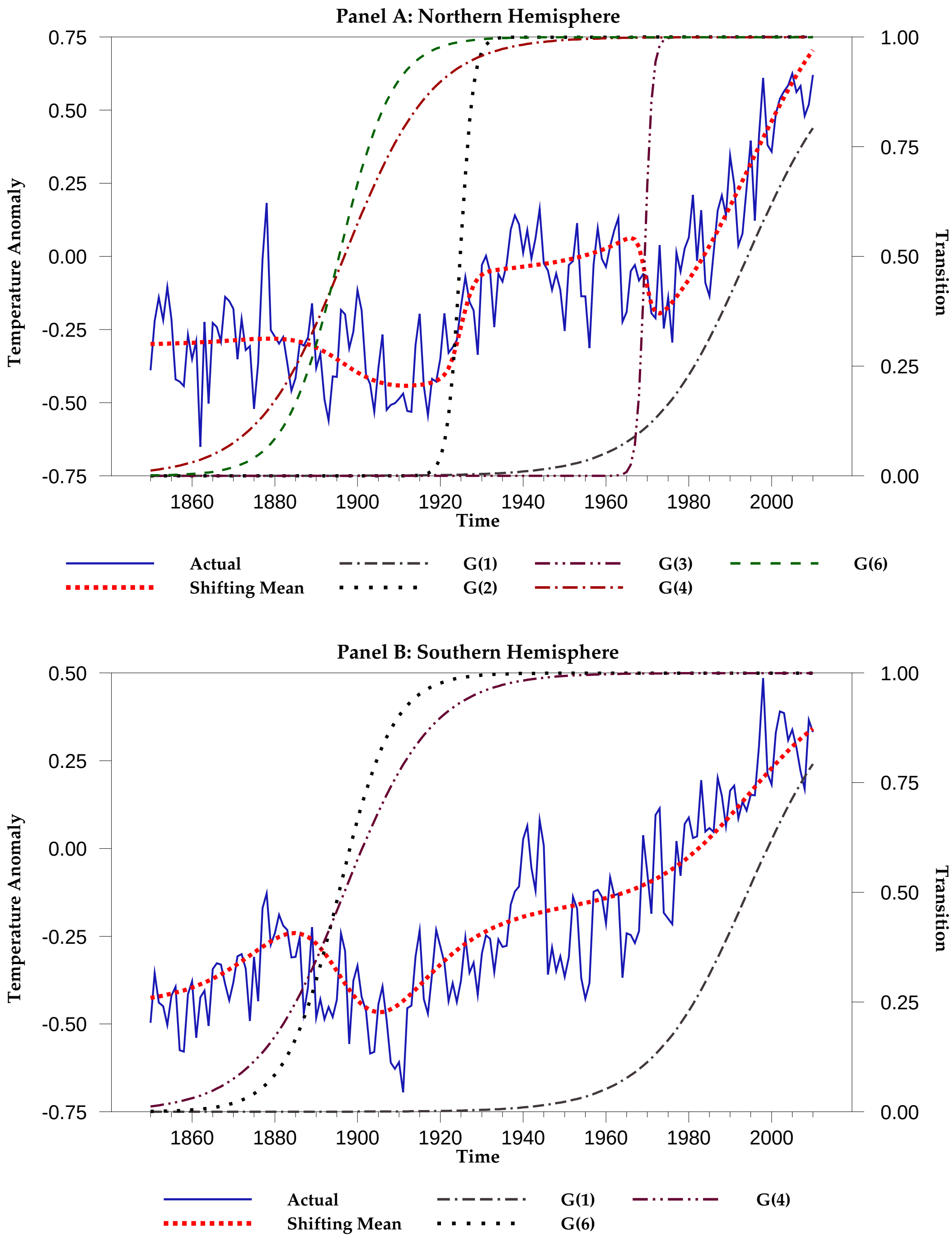

Figure 3: Bivariate VAR Results with Co–Trending Restrictions for Temperature Anomalies for the Northern (Panel A) and Southern (Panel B) Hemispheres, 1850–2010. The dashed line indicates the shifting mean and the dash–dot lines indicate the estimated transition functions.

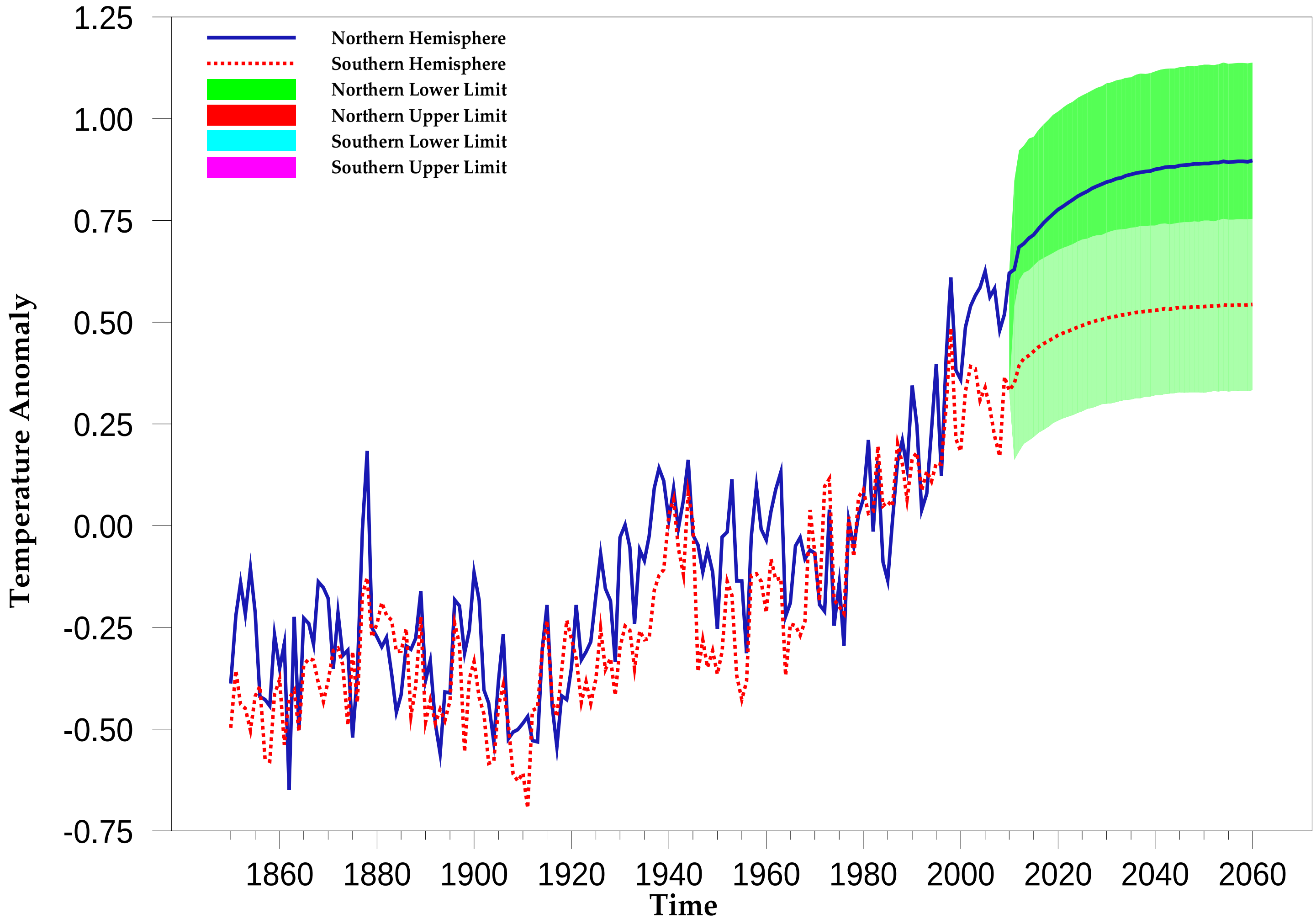

Figure 4: Temperature Anomalies in the Northern and Southern Hemisphere, Actual and Simulated, 1850–2060. Shaded areas denote 95–percent confidence intervals.

Research Papers 2012

2012-38: Anders Bredahl Kock and Laurent A.F. Callot: Oracle Efficient Estimation and Forecasting with the Adaptive LASSO and the Adaptive Group LASSO in Vector Autoregressions

2012-39: H. Peter Boswijk, Michael Jansson and Morten Ørregaard Nielsen: Improved Likelihood Ratio Tests for Cointegration Rank in the VAR Model

2012-40: Mark Podolskij, Christian Schmidt and Johanna Fasciati Ziegel: Limit theorems for non-degenerate U-statistics of continuous semimartingales

2012-41: Eric Hillebrand, Tae-Hwy Lee and Marcelo C. Medeiros: Let's Do It Again: Bagging Equity Premium Predictors

2012-42: Stig V. Møller and Jesper Rangvid: End-of-the-year economic growth and time-varying expected returns

2012-43: Peter Reinhard Hansen and Allan Timmermann: Choice of Sample Split in Out-of-Sample Forecast Evaluation

2012-44: Peter Reinhard Hansen and Zhuo Huang: Exponential GARCH Modeling with Realized Measures of Volatility Statistics

2012-45: Peter Reinhard Hansen and Allan Timmermann: Equivalence Between Out-of-Sample Forecast Comparisons and Wald

2012-46: Søren Johansen, Marco Riani and Anthony C. Atkinson: The Selection of ARIMA Models with or without Regressors

2012-47: Søren Johansen and Morten Ørregaard Nielsen: The role of initial values in nonstationary fractional time series models

2012-48: Peter Christoffersen, Vihang Errunza, Kris Jacobs and Hugues Langlois: Is the Potential for International Diversi…cation Disappearing? A Dynamic Copula Approach

2012-49: Peter Christoffersen, Christian Dorion , Kris Jacobs and Lotfi Karoui: Nonlinear Kalman Filtering in Affine Term Structure Models

2012-50: Peter Christoffersen, Kris Jacobs and Chayawat Ornthanalai: GARCH Option Valuation: Theory and Evidence

2012-51: Tim Bollerslev, Lai Xu and Hao Zhou: Stock Return and Cash Flow Predictability: The Role of Volatility Risk

2012-52: José Manuel Corcuera, Emil Hedevang, Mikko S. Pakkanen and Mark Podolskij: Asymptotic theory for Brownian semi-stationary processes with application to turbulence

2012-53: Rasmus Søndergaard Pedersen and Anders Rahbek: Multivariate Variance Targeting in the BEKK-GARCH Model

2012-54: Matthew T. Holt and Timo Teräsvirta: Global Hemispheric Temperature Trends and Co–Shifting: A Shifting Mean Vector Autoregressive Analysis