COMMON FACTORS IN CONDITIONAL DISTRIBUTIONS FOR BIVARIATE TIME SERIES

Clive W.J. Granger*
Department of Economics
University of California, San Diego
La Jolla, CA 92093-0508
USA
email: cgranger@ucsd.edu

Timo Teräsvirta
Department of Economic Statistics
Stockholm School of Economics
Box 6501, SE-113 83 Stockholm
Sweden
e-mail: sttimo@hhs.se

and

Andrew J. Patton
London School of Economics
London WC2A 2AE
United Kingdom
email: a.patton@lse.ac.uk

This version: 9 JUNE 2003

Abstract. A definition for a common factor for bivariate time series is suggested by considering the decomposition of the conditional density into the product of the marginals and the copula, with the conditioning variable being a common factor if it does not directly enter the copula. The links of this definition with a common factor being a dominant feature in standard linear representations is shown. An application using a business cycle indicator as the common factor in the relationship between U.S. income and consumption found that both series held the factor in their marginals but not in the copula.

Keywords. Common factor, dominant property, conditional distribution, copula

JEL Classification Codes: C32

Acknowledgements: Paper prepared for the conference "Common Features in Rio", Rio de Janeiro, July 2002. Part of the work was carried out when the second author was visiting Department of Economics, University of California, San Diego, whose kind hospitality he wishes to thank for. Material from the paper has also been presented in seminars at the Royal Institute of Technology, Stockholm, and the Swedish University of Agricultural Sciences, Uppsala. Comments from participants are gratefully acknowledged. The authors also wish to thank the referees for their constructive comments but retain the responsibility for any errors and shortcomings in this work.

---

*Corresponding Author

1. Introduction

This paper will initially consider common factors in a linear, bivariate framework and then ask if similar concepts can be extended for use with conditional distributions. For the start, it is important to have the idea of a "dominant property" (DP). Throughout, the dominant property will be thought of as being in a component of a process. If a series has several properties, it will be the DP that, in general will determine the relationship of the variable with others, and how it fits into models and equations. For the moment, we will consider only the case where there is a single dominant property and one (or more) dominated properties. In what follows, for a pair of random processes, X_t, Y_t, say, $X_t + Y_t$ is used as a convenient notation to denote the more general sum

$$X_t + AY_{t+m} \tag{1}$$

where *A, m* are some constants and $A \neq 0$. Some assumed properties are:

If X_t has DP and Y_t does not, then $X_t + Y_t$ will have the DP.

If X_t, Y_t both do not have a DP, then $X_t + Y_t$ will not have the DP.

Finally, it will generally be the case that if X_t and Y_t both have a common DP, then $X_t + Y_t$ has this DP.

Some of the usual examples of dominant properties or component processes are:

i. A trend in mean (either deterministic or stochastic)

ii. A strong seasonal component (either deterministic or stochastic);

iii. A strong business cycle component;

iv. Smooth transitions or distinct breaks in mean;

In (1) a lead of *m* (which may be negative) is allowed but for most of the dominant properties it is clear that taking *m* = 0 is little different in practice than any non-zero value. This is clear because there is little loss of information from knowing Y_{t+m} rather than Y_t unless *m* is quite large for all of the examples, except the last, near a break.

A persistent process (denoted I(1)) dominates a non-persistent process, denoted I(0). It has become the common practice to think of I(1) to be a unit root process, of a narrowly defined form, and I(0) to be a stationary linear process, such as an ARMA series, but again this is not necessary.

The relevance of a dominant property is clear in an explanatory model such as

$$X_t = a + bY_t + aW_t + e_t$$

as the two sides must balance, with the two sides having the same dominant properties. For example, if X_t has the DP whereas neither Y_t nor Z_t does, then the error term e_t must have the this property.

2. Dominant Common Factors

A particularly interesting case involving dominant properties and common factors is in the form

$$\left.\begin{aligned} X_t &= AW_t + Z_{1t} \\ Y_t &= W_t + Z_{2t} \end{aligned}\right\} \quad (2)$$

where W_t has the DP, Z_{1t}, Z_{2t} are independent of the DP, and $A \neq 0$ is some constant. From the rules given above, both X_t, Y_t will have the DP but $X_t - AY_t = Z_{1t} - AZ_{2t}$ will not have the DP. Thus, with this construction, a particular linear combination of two variables with a dominant property will not have the property.

If the DP is a trend, the variables are said to be "co-trending," if it is a break process, the variables are "co-breaking." From (2) it follows, however, that the breaks need not be simultaneous, as $m \neq 0$ is allowed. Furthermore, if W_t is a business cycle component, the variables are "co-cyclical," and if W_t has a strong seasonal, they can be thought of as being "co-seasonal." Finally when W_t is I(1) but the linear combination is I(0), they have been called "co-integrated." For a recent discussion of the co-cyclical literature, see Issler and Vahid (2001).

In this paper we concentrate on dominant common factors; that is a common factor that determines the major time series properties of two (or just a few) series. In general factor analysis, a common factor need not be dominant, but be present in largely unrelated processes. Such common factors can become dominant under cross-sectional aggregation (see Granger

(1987)). Sometimes a common factor can be important but not dominant such as the stock index in the Capital Asset Pricing Model in finance. Common factors may be either directly observed or derived from other series in the system, as in simple cointegration.

3. Conditional Distributions and Conditional Copula

The models considered in the previous section are relevant for the conditional expectation of a distribution, and are therefore somewhat limited in ambition. Similar examples can be constructed for the conditional variance. For a complete description of a relationship between random variables; however, one needs to consider a joint distribution. In our analysis of the joint distribution, we will employ a theorem of Sklar (1959), who showed that a bivariate density function can be decomposed into three parts: the two univariate marginal densities and a "copula" density. Suppose we concentrate just on the bivariate relationship between *X* and *Y*, conditional on *W*; then

$$f_{XY}(x,y\mid W)=f_X(x\mid W)f_Y(y\mid W)k(F_X(x\mid W),F_Y(y\mid W)\mid W) \tag{3}$$

where *k* is the conditional copula density function. As an example, when *X* and *Y* are conditionally independent given *W*, $k(x,y\mid W)\equiv 1$. In this special case, *k* is not dependent on *W*, although the marginals may still be dependent on *W*. Such situations will be of interest later on.

Equation (2) shows Sklar's theorem for density functions; the original theorem applied more generally to distribution functions:

$$F_{XY}(x,y\mid W)=C(F_X(x\mid W),F_Y(y\mid W)\mid W) \tag{4}$$

where F_{XY} is the joint conditional distribution function of *X*, *Y*, F_X is the conditional marginal distribution function of *X*, and similarly F_Y is the conditional marginal distribution function of *Y*. Sklar showed that there will always be a function *C*, called the copula distribution function, so that (4) holds. Differentiating (4) with respect to *x* and *y* gives (3). Function *C* itself is a cumulative distribution function, namely, a cumulative distribution function of two conditionally Uniform(0,1) distributed random variables. If *X* and *Y* are both continuous random variables, the copula is unique, and is the joint distribution conditional on *W*, of the random variables *u* and *v* which are defined as $u=F_X(x\mid W)$ and $v=F_Y(y\mid W)$.

The copula function represents the dependence between *X* and *Y* after taking out the effects of the marginals, which may be different, see Joe (1997) and Nelson (1999). What makes the copula important is that the marginal distributions and linear correlations determine the joint distribution of a set of random variables only if the latter are elliptically distributed, such as normally or *t*-distributed random variables. If this is not the case, the copula will take the place of the correlations. For discussion, see, for example, Embrechts, McNeil and Straumann (1999, 2001). Note, however, that the copula has a link to rank correlations. Kendall's τ for the dependence between *X* and *Y* is defined as

$$\tau(X,Y) = \Pr\{(X_i - X_j)(Y_i - Y_j) > 0\} - \Pr\{(X_i - X_j)(Y_i - Y_j) < 0\}$$

for $i \neq j$ where (X_i, Y_i) is a pair of observations from the joint distribution of *X* and *Y*. Now

$$\tau(X,Y) = 4\int_0^1\int_0^1 C(u,v)dC(u,v) - 1 = 4E[C(u,v)] - 1.$$

where *C* is the copula of the joint distribution *X* and *Y*. While the copula is a two-dimensional entity, Kendall's τ is a univariate measure of dependence between *X* and *Y*. It can similarly be shown that Spearman's rank correlation coefficient, ρ_S, is equal to

$$\rho_S(X,Y) = 12\int_0^1\int_0^1 uvdC(u,v) - 3 = 12E[uv] - 3.$$

The interpretation of the decomposition (3) is important for later developments in the paper. Let X_t and Y_t be defined as in (2) and assume that Z_{1t}, Z_{2t} are independent of $w_{t-j}, j \geq 0$. It should be noted that in (3) all of the univariate properties of the X_t process, such as the (conditional) mean, variance, higher moments, and so forth, are encapsulated in the conditional marginal distribution $f_Y(x \mid W)$. The copula involves none of these quantities and only contains measures relating to the extent that $X_t \mid W_t$ and $Y_t \mid W_t$ are interdependent. It is a bivariate function which generalizes the standard correlation coefficient, but which generally depends on the conditioning variable W_t. In general, the extent and manner in which X_t, Y_t are interrelated may change with the conditioning variable. However, in the system (3), without conditioning X_t and Y_t will be dependent largely through W_t, but on conditioning the dependence of X_t and Y_t will only depend on the joint distribution of Z_{1t} and Z_{2t}. It should be noted that in this example it is enough to condition on W_t as conditioning on W_t and its lags is not required.

Returning to the discussion in the previous section, if *W* has a dominant property, then the equivalent of equation (2) in distributions would be that both of the marginal densities $f_X(x|W), f_Y(y|W)$ are not independent of *W*. Thus, *W* does have an impact somewhere in the joint density. However, the equivalent of the linear common factor situation could be that the relationship between *X* and *Y* as expressed by the conditional copula density function does not depend on *W*. This will be discussed in Section 5, but one may already note the above-mentioned special case in which *X* and *Y* are conditionally independent given *W*.

4. Examples of Dominant Properties in Conditional Distribution

A process X_t can be said to have a seasonally varying distribution if it has a time-varying density $f_t(x)$ but, when measured monthly,

$$\|f_t(x) - f_{t+12}(x)\|$$

is small, using some suitable norm for densities. A plausible pseudo-norm is the Kullback-Leibler Criterion, see White (1994) for instance. X_t could be used as a conditioning variable in the common factor framework outlined above.

Similarly, a sequence of time-varying densities $f_t(x)$ could be called "trending" if $f_t(x)$ stochastically dominates (to order one) $f_s(x)$ for all $t > s$; i.e. $F_t(x) > F_s(x)$ for all x, $t > s$ where $F_t(x)$ is the distribution function corresponding to the density $f_t(x)$. If T_t is a random variable drawn from such a distribution, it might be called a trend and be a variable with a dominant property.

If $f_t(x)$ takes the form $f(x, \theta_t)$ where θ_t is some vector of parameters which are not necessarily constant, the densities can be called "breaking" if $\underline{\theta}_t = \underline{\theta}_0, t \le t_0$, $\underline{\theta}_t = \underline{\theta}_1 (\ne \underline{\theta}_0)$, $t > t_0$. There could be several breaks and they could be caused by other variables taking particular values. A variable W_t drawn from the distribution can be called a breaking process and used as a conditioning variable.

If B_t is a process that is closely linked with the business cycle, such as a coincident indicator, then it can be used directly as a candidate for a common factor in conditional distributions.

There are several ways that persistence can be defined. A useful way is to define a process W_t as being persistent if $F(W_t, W_{t+n}) \neq F(W_t)F(W_{t+n})$ as n becomes large. This can potentially be tested using some of the measures of dependence discussed in Joe (1997). If W_t is a persistent process, it can be used as a conditioning variable and it will have a dominant property.

The class of possible processes with dominant properties can be extended further to include "long-memory processes" for example, but these will not be considered here.

Tests for the existence or not of a particular dominant property will exist in some cases, such as for first-order stochastic dominance, but others will need to be developed.

Dominant factors need not be treated individually and a group of different trending variables, say, or a trend and a seasonal can be used jointly as conditioning variables. Further, other variables without dominant properties can also be included in the conditioning set. These extensions do complicate the picture and make analysis more difficult, although possibly more realistic. We leave such questions to be considered with the analysis of particular applications.

5. Common Factors in Distributions

We can now formally state our definition of common factors in distributions. The definition is adapted to time series situations where the observations are not independent but the present ones may depend on the previous ones.

Definition: *Let* X_t, Y_t *be a pair of processes and denote* $\overline{X}_t \equiv X_{t-j}, j \geq 0$ *and similarly* $\overline{Y}_t$ *is the present and the past of the Y's. A process* W_t *will be considered as one with a dominant common property, or a common factor in distribution, if the conditional marginals* $F_X(X_t \mid W_t, \overline{W}_{t-1}, \overline{X}_{t-1}, \overline{Y}_{t-1})$ *and* $F_Y(Y_t \mid W_t, \overline{W}_{t-1}, \overline{X}_{t-1}, \overline{Y}_{t-1})$ *do depend on* W_t *and possibly the lagged terms* $\overline{W}_{t-1}$*, but the conditional copula* $k(u_t, v_t \mid W_t, \overline{W}_{t-1}, \overline{X}_{t-1}, \overline{Y}_{t-1}) = k(u_t, v_t \mid \overline{X}_{t-1}, \overline{Y}_{t-1})$ *does not depend on* W_t*, either directly or through the lags.*

Thus, the effect of W_t on (x_t, y_t) is through the marginal distributions but not through their relationship. Although this could happen with any conditioning variable, it is particularly noteworthy for variables representing a dominant property. Thus, for example, a pair of variables could have marginals that vary seasonally, but their relationship, as characterized by the copula, does not vary seasonally. Similarly, a pair could have marginal distributions that change with the business cycle, not just in means but many quantities, yet the conditional copula density does not vary with the business cycle. Such possibilities lead to interesting interpretations for economic series. Again, suitable tests need development.

As an illustration of the definition, consider a two-factor model used to explore cointegration relationships. X_t, Y_t are a pair if I(1) series generated from W_t which is I(1) and Z_t which is I(0), by the equations

$$X_t = AW_t + C_1 Z_t + \varepsilon_{Xt} \tag{5}$$

$$Y_t = W_t + C_2 Z_t + \varepsilon_{Yt} \tag{6}$$

where $\varepsilon_{Xt}, \varepsilon_{Yt}$ are zero mean, independent series, independent of each other with the constraint that $C_1 - AC_2 = 1$. The two equations can be augmented by the addition of a finite number of weighted logs of $\Delta X_t, \Delta Y_t$. Such an augmentation merely complicates the algebra without greatly changing the important aspects of the model.

It follows directly that

$$Z_t = X_t - AY_t + e_{Zt} \tag{7}$$

where $e_{Zt} = e_{Xt} - Ae_{Yt}$; and

$$W_t = C_1 Y_t - C_2 X_t + e_{Wt} \tag{8}$$

where $e_{Wt} = C_1 e_{Xt} - C_2 e_{Xt}$.

To give a specific example, suppose that $\Delta W_t = \eta_t$ where η_t is zero mean iid and $Z_t = \rho Z_{t-1} + \theta_t$, $|\rho| < 1$, θ_t zero mean, iid. Taking changes in (5), (6) and using this example gives

$$\Delta X_t = C_1(\rho - 1) Z_{t-1} + \gamma_{1t} \tag{9}$$

$$\Delta Y_t = C_2(\rho - 1) Z_{t-1} + \gamma_{2t} \tag{10}$$

where

$$\gamma_{1t} = A\eta_t + C_1\theta_t + \Delta\varepsilon_{Xt} \tag{11}$$

$$\gamma_{2t} = \eta_t + C_2\theta_2 + \Delta\varepsilon_{Yt}. \tag{12}$$

In this example, X_t and Y_t are seen to have the dominant I(1) property because of the common I(1) factor W_t in equations (5)(6). However, there is a linear combination of X_t and Y_t which produces an estimate of Z_t with no W_t anywhere in its distribution. From (8) W_t can be estimated, but with an I(0) error.

Turning to the specific example and now concentrating on I(0) variables, (9) and (10) show that the relationship between ΔX_t and ΔW_t has become that between ΔX_t and η_t. Although η_t has mean zero, and so will not affect the conditional mean of ΔX_t, given $\Delta\overline{X}_{t-1}, \Delta\overline{Y}_{t-1}$, it will influence the conditional variance, and similarly for ΔY_t.

This example is standard for linear cointegration but does not quite fit into the examples discussed in this paper, as W_t is not directly observed in practice, only estimated from the raw data, X_t, Y_t. As η_t contains ΔX_t, there is simultaneity involved in distributional relationships. If W_t is separately observed, it can be conditioned on and used in our definition.

6. Application

As an empirical example for the ideas presented above we now present an analysis of the joint distribution of income and consumption, with a business cycle index variable as a possible common factor. Income and consumption are two of the most widely studied macroeconomic variables, and they both are known to vary individually over the business cycle, i.e., both consumption and income growth have cyclicality as a dominant property. The relationship between these variables has also been widely studied, though to our knowledge no stylized facts regarding the behavior of the conditional dependence between these variables over the business cycle are available. We will investigate whether the conditional dependence between these variables also exhibits cyclicality, by testing whether a business cycle index variable influences the conditional copula of income and consumption growth. If no evidence of influence is found, then the index variable is a common factor, and cyclicality is a dominant common property.

6.1 Data and Model

We used monthly data from January 1967 to November 2001 on U.S. real per capita disposable income (denoted Y_t) and U.S. real per capita consumption on nondurables (denoted C_t). The business cycle indicator used was the Stock and Watson experimental coincident index[1] (denoted B_t). As will be seen in the model, these variables appear in log-difference form.

We specified linear models for the conditional means of the two series and the autoregressive conditional heteroscedasticity (ARCH) model of Engle (1982) for the conditional variances. Our choice of specification for the marginal densities was guided by our desire to allow for conditional non-normality. Two of the most common deviations from normality are fat tails (excess kurtosis) and asymmetry or skewness. Two distributions that are commonly used to allow for excess kurtosis are the Student's *t* and the generalized error distribution (GED). Both of these distributions have been generalized to allow for skewness, and we selected the skewed Student's *t* of Hansen (1994) for its simplicity and its past success in modeling economic variables. The skewed *t* distribution has two parameters: one for skewness and one for tail thickness. The distribution is not generally elliptical and thus the conditional copula is the appropriate measure of conditional dependence between the two variables. The functional form of the skewed *t* density is given below.

$$\text{Skewed } t(y;\lambda,v) = \begin{cases} bc\left(1+\dfrac{1}{\upsilon+2}\left(\dfrac{by+a}{1-\lambda}\right)\right)^{-(\upsilon+1)/2} & for\ y \le -\dfrac{a}{b} \\ bc\left(1+\dfrac{1}{\upsilon+2}\left(\dfrac{by+a}{1+\lambda}\right)\right)^{-(\upsilon+1)/2} & for\ y > -\dfrac{a}{b} \end{cases} \tag{13}$$

where $a = 4\lambda c\left(\dfrac{\upsilon-2}{\upsilon-1}\right)$ with $c = \dfrac{\Gamma\left(\dfrac{\upsilon+1}{2}\right)}{\Gamma\left(\dfrac{\upsilon}{2}\right)\sqrt{\pi(\upsilon+2)}}$, and $b = \sqrt{1+3\lambda^2 - a^2}$.

[1] The data on consumption and income were taken from the St. Louis Federal Reserve web page, http://www.stls.frb.org/fred. The business cycle index series was taken from Jim Stock's web page, http://ksghome.harvard.edu/~.JStock.Academic.Ksg/xri/0201/xindex.asc.

Since the two marginal densities and the copula define a joint distribution, the natural estimation method is maximum likelihood. We employ the multi-stage maximum likelihood estimator presented in Patton (2001). Multi-stage estimation allows us to first estimate the marginal distributions separately, and then model the copula, which greatly simplifies the estimation of the model.

We used the Akaike Information criterion (AIC) and goodness-of-fit tests to find appropriate models for the each of the conditional moments of the two series. Lags of consumption, income and the business cycle variable were allowed to enter as explanatory variables for both dependent variables. In our particular example it happened that the best fitting models did not require lags of the "other" variable (ie, lags of Y_t in the model for C_t, and vice versa). This will not always be the case, and must be tested in each specific situation, as emphasized in Patton (2002, p10).

The final models and parameter estimates are presented below; standard errors are provided in parentheses, and parameter estimates significant at the 5% level are marked with an asterisk. We used the modified logistic transformation, Λ, to keep the skewness parameter, λ_t, in (-1,1) at all times.

$$\begin{aligned}
\Delta\log C_t &= \underset{(0.03)}{0.15^*} - \underset{(0.05)}{0.42^*}\,\Delta\log C_{t-1} - \underset{(0.05)}{0.20^*}\,\Delta\log C_{t-2} - \underset{(0.04)}{0.15^*}\,\Delta\log C_{t-12} - \underset{(0.04)}{0.20^*}\,\Delta\log C_{t-24} \\
&\quad + \underset{(0.03)}{0.32^*}\,\Delta\log B_t + \varepsilon_t, \ \text{ where } \frac{\varepsilon_t}{\sqrt{h_t^C}} \mid I_{t-1} \sim \textit{Skewed } t(\lambda_t^C, \upsilon_t^C) \\
h_t^C &= \underset{(0.03)}{0.31^*} - \underset{(0.02)}{0.01^*}\,\varepsilon_{t-1}^2 + \underset{(0.02)}{0.06^*}\,(\Delta\log B_t)^2 \\
\lambda_t^C &= \Lambda\left(-\underset{(0.17)}{0.01} + \underset{(0.10)}{0.36^*}(\Delta\log B_t)^2\right) \\
\upsilon_t^C &= \underset{(3.12)}{7.95^*} \\
\textit{where } \Lambda(a) &= \frac{1.998}{1+\exp\{-a\}} - 0.999
\end{aligned} \tag{14}$$

$$\Delta \log Y_t = \underset{(0.04)}{0.14}^* - \underset{(0.10)}{0.30}^* \Delta \log Y_{t-1} - \underset{(0.07)}{0.16}^* \Delta \log Y_{t-2} + \underset{(0.04)}{0.33}^* \Delta \log B_t + \eta_t,$$

$$\text{where } \frac{\eta_t}{\sqrt{h_t^Y}} \mid I_{t-1} \sim \textit{Skewed } t(\lambda_t^Y, \upsilon_t^Y)$$

$$h_t^Y = \underset{(0.11)}{0.26}^* + \underset{(0.18)}{0.46}^* \eta_{t-1}^2 + \underset{(0.03)}{0.03} \Delta \log B_t$$

$$\lambda_t^Y = \Lambda\left(\underset{(0.17)}{0.07} + \underset{(0.29)}{0.37} \eta_{t-1}^2 + \underset{(0.06)}{0.02} (\Delta \log B_t)^2 \right) \tag{15}$$

$$\upsilon_t^Y = 2.1 + \left(-\underset{(0.26)}{0.66}^* + \underset{(0.20)}{0.29} \eta_{t-1}^2 - \underset{(0.25)}{0.43} \Delta \log B_t \right)^2$$

$$\text{where } \Lambda(a) = \frac{1.998}{1 + \exp\{-a\}} - 0.999$$

No dynamics in the degrees of freedom parameter in the consumption density model were found, and so it was modeled as being constant. Many of the coefficients on the business cycle index variable in the conditional moment specifications were significant at conventional levels, confirming that both consumption and income vary over the business cycle. Although not all of the coefficients on the B_t terms are significant at the 5% level, these variables were needed for the model to pass the specification tests employed to check the adequacy of the proposed model. We conducted the specification tests presented in Patton (2002) to check the goodness-of-fit of the above specifications, and found no evidence that they are inadequate. The test results are presented in Appendix A.

In our search for the best specification of the conditional copula for these two variables, we considered eight alternative copula functional forms: normal, Clayton, rotated Clayton, Gumbel, rotated Gumbel, Plackett, Frank and the symmetrised Joe-Clayton. The first seven of these are presented in Joe (1997) and Nelsen (1999), while the eighth was introduced in Patton (2002). Each of these copulas implies a different type of dependence between the variables. For example, the Clayton copula would fit best if negative changes in consumption and income are more highly correlated than positive changes; the Gumbel and the rotated Clayton would fit best in the opposite situation. The Plackett and Frank copulas are symmetric, like the normal, but imply slightly different dependence structures. Without any economic theory to guide us on the choice of dependence structure, it becomes an empirical question to find the best fitting model.

We estimated constant versions of these copulas, and the Gumbel was found to provide the best fit in terms of the log-likelihood value[2]. We proceeded to use the Gumbel copula for the time-varying conditional copula specifications. The forms of the Gumbel copula cumulative distribution function and probability density function (C_{gumbel} and k_{gumbel} respectively) are given below.

$$
\begin{aligned}
C_{gumbel}(u,v;\kappa) &= \exp\left\{-\left((-\log u)^{\kappa} + (-\log v)^{\kappa}\right)^{1/\kappa}\right\} \\
k_{gumbel}(u,v;\kappa) &= \frac{C_{gumbel}(u,v;\kappa)\left((\log u)(\log v)\right)^{\kappa-1}}{uv\left((-\log u)^{\kappa} + (-\log v)^{\kappa}\right)^{2-1/\kappa}}\left(\left((-\log u)^{\kappa} + (-\log v)^{\kappa}\right)^{1/\kappa} + \kappa - 1\right)
\end{aligned} \tag{16}
$$

We allowed the parameter of the Gumbel copula, κ, to vary through time, setting it to be a function of the change and squared change in the business cycle index variable, and the average distance between the 'transformed' residuals, U_t and V_t. This average distance is a measure of the degree of dependence between the variables over the last six months[3], as under perfect positive dependence it always equals zero, under independence it is equal to one-third in expectation, and under perfect negative dependence it is equal to one-half in expectation.

$$
\begin{aligned}
\left(\frac{\Delta\log C_t - \mu_t^C}{\sqrt{h_t^C}}, \frac{\Delta\log Y_t - \mu_t^Y}{\sqrt{h_t^Y}}\right) &\sim F_{XY} = C(F_X, F_Y) \\
&= C_{\text{gumbel}}(\text{Skewed } t(\lambda_t^C, \upsilon_t^C), \text{Skewed } t(\lambda_t^Y, \upsilon_t^Y); \kappa_t)
\end{aligned} \tag{17}
$$
$$
\text{where } \kappa_t = 1 + \left(\gamma_0 + \gamma_1 \Delta\log B_t + \gamma_2 \Delta\log B_t^2 + \gamma_3 \sum_{j=1}^{6}\left|u_{t-j} - v_{t-j}\right|\right)^2
$$

The Gumbel copula parameter must be greater than or equal to one at all times, and we constrain the evolution equation for κ_t to ensure that this is the case.

We computed the covariance matrix of the parameter estimates of the joint distribution model, and present the results for the copula parameters in Table 6.1.

---

[2] As the Gumbel copula has a single parameter, and all the other copulas considered have either one or two parameters, selecting the Gumbel copula by the likelihood value is equivalent to selection by AIC or BIC.
[3] We also experimented with averaging over the preceding 12 and 24 months and found no significant improvement over using only 6 months.

Table 6.1: Copula parameter estimates and standard errors

	Coefficient	Standard error	*t*-statistic	Log-likelihood
Constant conditional copula				
Constant	1.0977	0.0361	2.7064*	7.9785
Time-varying conditional copula				
Constant (γ_0)	0. 2883	0.2106	1.3694	8.5526
$\Delta \log B_t$ (γ_1)	0. 0329	0.1040	0.3167	
$\Delta \log B_t^2$ (γ_2)	0. 0490	0.0490	0.9987	
$\Sigma\|u-v\|$ (γ_3)	-0. 0913	0.5870	-0.1555	

* This *t*-statistic is for the test of the null hypothesis that the parameter equals one (rather than zero), which corresponds to independence of the two variables.

As the results in the table show, none of the individual coefficient estimates of the variables used in the evolution equation for the conditional copula parameter are significant, and the joint hypothesis of no time variation in the conditional copula cannot be rejected either (a likelihood ratio test yields a *p*-value of 0.7655). This suggests that the conditional dependence between consumption and income is constant. We can, however, reject the hypothesis that the variables are independent at the 5% level. Most interestingly, our results suggest that the business cycle index variable is not important in describing the dependence between these two series, and thus may be a common factor in distribution for consumption and income.

In Figure 1 we present the time path of κ_t according to this model, along with the NBER recession periods. This figure confirms that the time variation in the conditional copula parameter is essentially unrelated to the business cycle.

It should be noted that for us to conclude with certainty that the dependence structure between these variables is independent of the business cycle we would need to try *all possible* functions of the business cycle index variable, not just the quadratic specification used above. It is of course possible that some other function of the business cycle index variable does influence the conditional dependence structure. Further, the results may be sensitive to the choice of B_t versus, say, B_{t-1}, or any other lag of B_t, or possibly the vector $[B_t, B_{t-1}, \ldots, B_{t-p}]$. While we found no evidence that B_t affected the conditional copula, in unreported results we did find some evidence that B_{t-1} was important for the conditional copula. Thus our conclusion is affected by the choice of lag on the business cycle index variable.

Overall, our preliminary results on this question give some support to the claim that the impact of the business cycle on the joint distribution of consumption and income is through the marginal distributions and not through their dependence structure, making it a "common factor in distribution" for consumption and income.

7.0 Conclusion

The paper proposes a definition for common factors in conditional distributions that is the analogy to that used in the linear context of the first and second moments. A wide variety of possible dominant factors are suggested and an application is presented concerning the income and consumption relationships over the business cycle. We find some evidence that a business cycle indicator variable is a common factor in the distribution of consumption and income. Many questions in this are remain unresolved, both concerning testing and also some properties of the common factor representation in particular cases. They are left for further research.

References

Clements, M.P., 2002, An Evaluation of the Survey of Professional Forecasters Probability Distributions of Expected Inflation and Output Growth, working paper, Department of Economics, University of Warwick.

Embrechts, P.A., A. McNeil, and D. Straumann (1999): Correlation: Pitfalls and Alternatives. *RISK* 12 (5), 11-21.

Embrechts, P.A., A. McNeil, and D. Straumann (2001): Correlation and Dependence Properties in Risk Management: Properties and Pitfalls, in M. Dempster ed., *Risk Management: Value at Risk and Beyond*, Cambridge University Press.

Engle, Robert F., 1982, Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflation, *Econometrica*, 50, 987-1007.

Granger, C.W.J., 1987, Implications of Aggregation with Common Factors, *Econometric Theory*, 3, 208-222.

Hansen, Bruce E., 1994, Autoregressive Conditional Density Estimation, *International Economic Review*, 35, 705-730.

Issler, J.V. and F. Vahid (2001): "Common Cycles and the Importance of Transitory Shocks to Macroeconomic Aggregates." *Journal of Monetary Economics* 47, 449-475.

Joe, H. (1997): Multivariate Models and Dependence Concepts. Chapman and Hall: London.

Nelson, R.B. (1999): An Introduction to Copulas. Springer: Berlin.

Patton, A. J. (2001): Estimation of Copula Models for Time Series of Possibly Different Lengths, Working Paper 2001-17, Department of Economics, University of California, San Diego.

Patton, A. J. (2002): Modelling Time-Varying Exchange Rate Dependence Using the Conditional Copula, Working Paper 2001-09, Department of Economics, University of California, San Diego. Revised July 2002.

Sklar, A., 1959, Fonctions de répartition à *n* dimensions et leurs marges, *Publ. Inst. Statist. Univ. Paris*, 8, 229-231.

Shao, J., 1999, Mathematical Statistics, Springer-Verlag, New York.

Wallis, K.F., 2001, Chi-squared Tests of Interval and Density Forecasts, and the Bank of England's Fan Charts, forthcoming, *International Journal of Forecasting*.

White, H. (1994): Estimation, Inference, and Specification Analysis. Cambridge University Press: New York.

Appendix A: Marginal distribution specification test results

Before proceeding to model the conditional copula, it is critical to test the goodness-of-fit of the models employed for the conditional marginal distributions. Mis-specification in the marginal densities implies that the probability integral transforms, denoted U_t and V_t above, will not be uniformly distributed on (0,1), and thus any copula will automatically be mis-specified. Mis-specification in the dynamics of the conditional marginal distribution models can lead to spurious findings for the dynamics of the conditional copula.

A simple test for a density specification (ignoring the impact of estimation error) is the Kolmogorov-Smirnov test, see Shao (1999). Applying this test to the series U_t and V_t we obtain test statistics (p-values) of 0.0228 (0.9834) and 0.0246 (0.9650), suggesting that both densities are well-specified.

To test jointly for the adequacy of the dynamics and the density specifications in the marginal distribution models we employ a test discussed in Patton (2002), variations of which were also presented in Clements (2002) and Wallis (2002). This test divides the support of the density into regions, R_i, and then applies interval forecast evaluation techniques to each region separately, and then all regions jointly. If the entire density is well-specified, then the derived interval forecast in each region should also be well-specified. We break the support of U and V into 5 regions: [0,0.1], (0.1,0.25], (0.25,0.75] , [0.75,0.9) and [0.9,1]. We construct 'hit' variables for each region, as $\text{Hit}_{i,t}^{U} = 1\{U_t \in R_i\}$ and $\text{Hit}_{i,t}^{V} = 1\{V_t \in R_i\}$, which take the value 1 if the realized value is in the region, and 0 otherwise. Under the null of correct specification, each of these Hit variables should be iid Bernoulli(U-L), where L and U are the lower and upper boundaries of the region.

To test individual regions we estimate a logistic regression of the hit variables on a constant and variables that should, if the model is well specified, have no influence on the hit variable. We used the first lag of the both hit variables for the same region (ie, both $\text{Hit}_{i,t-1}^{U}$ and $\text{Hit}_{i,t-1}^{V}$) to capture serial correlation, and the lagged business cycle index variable in levels and squares[4] to capture any information in this variable that may have been omitted from the model. Under the null hypothesis that the density models are well specified the test statistic is a χ_5^2 random variable.

To test all regions jointly we estimate a multinomial logit model, with the same specifications for each region as for the individual tests. The test statistic for the joint test is a χ_{20}^2 under the null hypothesis. The p-values for each test statistic are presented below.

Region	U_t	V_t
[0,0.1]	0.5935	0.9824
(0.1,0.25]	0.5320	0.3008
(0.25,0.75]	0.5343	0.5794
[0.75,0.9)	0.6833	0.4782
[0.9,1]	0.2264	0.1801
Joint test	0.6395	0.7116

This table shows that both specifications pass all of the individual region tests (p-values are all greater than 0.05) and the joint test. We thus conclude that these specifications are adequate representations of the conditional marginal distributions, and move on to modeling the copula.

[4] We also tried adding other variables, which counted the number of hits over the past 6 and/or 12 periods, to capture higher-order serial dependence. Further, we tried using only levels of the business cycle variable, and only using "own" lagged hits. None of these changes affected the final conclusion.

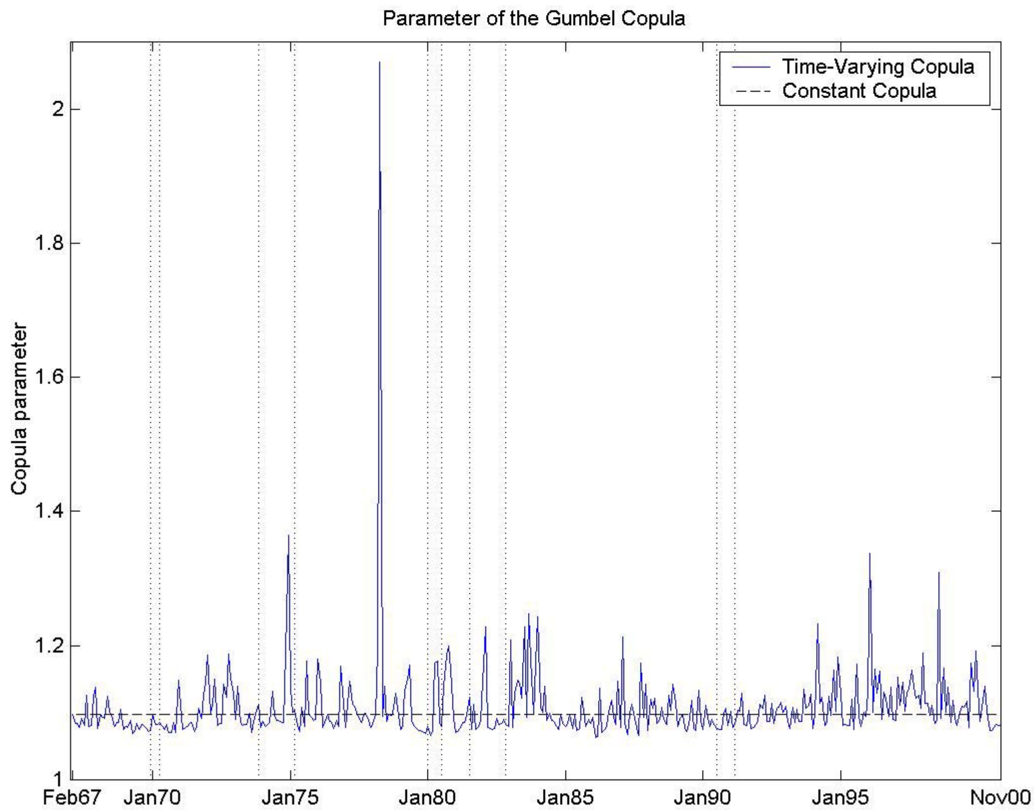

Figure 1: This figure shows the time path of the Gumbel copula parameter using the model in equation 17, along with its value in the constant conditional copula model. The vertical dotted lines are the NBER recession periods.