Time Series of Cross-Sectional Distributions with Common Stochastic Trends*

Yoosoon Chang†, Changsik Kim‡ and Joon Y. Park§

Abstract

We develop a new framework and methodology for the time series analysis of cross sectional distributions with stochastic trends. Often individual time series of cross sectional distributions have nonstationary persistent components that may be characterized effectively as functional unit roots. This paper shows how to model and analyze the presence of common trends in multiple time series of such cross sectional distributions. For an illustration, we use the CEX income and expenditure data to investigate the dynamic interactions between the household income and consumption distributions. Many interesting longrun and shortrun interactions between the household income and consumption distributions are found.

This Version: June 16, 2014

JEL Classification: C33, C12, C13

Key words and phrases: Time series of cross-sectional distributions, distributional unit roots and cointegration, income and consumption dynamics, income response functions to consumption.

*This version is prepared for the presentation at the 2014 Panel Conference in Japan by Yoosoon Chang. The authors are grateful to Minkyeong Baik for her excellent research assistance.

†Department of Economics, Indiana University

‡Department of Economics, Sungkyunkwan University

§Department of Economics, Indiana University and Sungkyunkwan University

1 Introduction

Many cross-sectional observations are available over time for economic analysis. Though some of them are made over time for the same set of individuals, for many others cross-sectional observations are collected for different groups of individuals as time changes. The former are called genuine panels, while the latter are often referred to as pseudo panels. For most panel studies in economics, genuine panels have been used. The use of pseudo panels has been rather limited. The studies relying on genuine panels typically use the data sets that include large dimensional cross-section observations with relatively much smaller time series dimensions. If we need to analyze observations over a long span, which is necessary to study persistent changes in any economic relationships over time, only pseudo panels are available in most cases. True, pseudo panels do not contain as much information as genuine panels. Obviously, however, they include much more information to be exploited than their cross-sectional aggregates used in the conventional time series analysis.

In the paper, we develop a new framework to analyze the longrun relationships between two time series of cross-sectional distributions, which have some persistent features. The persistent features of individual time series of cross-sectional distributions are characterized by distributional unit roots, and the longrun relationships between two time series of cross-sectional distributions having distributional unit roots are modeled as distributional cointegration. Our framework requires only pseudo panels, and therefore, it is widely applicable in practice. In our approach, we consider time series of probability densities representing cross-sectional distributions. The densities for cross-sectional distributions are estimated from cross-sectional observations, and we analyze them as time series of functional observations. Our analysis relies on the statistical theory that has been developed earlier by several authors including Bosq (2000), Park and Qian (2012) and Chang et al. (2012), among others.

The monograph by Bosq (2000) presents a basic idea and methodology on how to analyze the stationary time series of functional data. Park and Qian (2012) proposes a framework to analyze the time series of probability densities representing cross-sectional distributions and develops the relevant statistical theory, assuming the stationarity of the underlying probability densities over time. More recently, their framework has been extended by Chang et al. (2012) to allow for the unit root type nonstationarity. They demonstrate that time series of cross-sectional distributions such as the time series of income distributions may be nonstationary and have distributional unit roots, and develop the methodology to draw inference on the nonstationarity in time series of cross-sectional distributions. In the paper, we further extend their approach to multiple nonstationary time series of cross-sectional distributions individually having unit roots, and create a new framework to accommodate the presence of common stochastic trends in their time series that we call the distributional cointegration in the paper.

Our approach makes it possible to decompose the time series of cross-sectional distributions into stationary and nonstationary components. Accordingly, we may separately identify the shortrun and longrun relationships between two time series of cross-sectional distributions. The distributional cointegration between them provides their longrun relationships. Their shortrun relationships are specified and estimated by the stationary distributional regression of their stationary components. The longrun and shortrun relationships are measured and interpreted using what we define in the paper respectively as the longrun and shortrun response functions. Roughly speaking, response functions show how the other distribution is affected by one distribution at each of its different levels. If the response function of one distribution to the other has a peak at a certain level, it implies that the effect of a change in one distribution on the other is maximized if the change occurs in one distribution at that level.

Our new framework and methodology are applied to analyze the distributional cointegration between the time series of cross-sectional income and consumption distributions, using the U.S. households monthly income and consumption data from the Consumer expenditure (CE) survey series during the period from October 1979 to February 2013. We demonstrate that both the time series of income and consumption distributions have unit roots: The time series of income distributions has two unit roots, whereas the time series of consumption distributions has only one unit root. Furthermore, we find the presence of cointegration between the time series of income and consumption distributions: There is one cointegrating relationship between them. We slao obtain the longrun and shortrun response functions of income to consumption. They show how consumption is responded in the longrun and in the shortrun to income changes at each income level. Consumption is responded most in the longrun to changes in income of households with monthly earnings approximately \$2,000, but in the shortrun changes in income of household with monthly earnings approximately \$1,000 entails the biggest impact on consumption changes.

The rest of the paper is organized as follows. Section 2 presents the model and methodology. Time series of cross-sectional distributions are formally introduced with the basic framework to analyze their stationary and nonstationary relationships. In particular, the concepts of distributional unit roots and cointegration and the methodology to characterize them are developed. In Section 3, we present all statistical procedures required to do inference on our model. The methods of inference on both the stationary and nonstationary components of the model are provided. An empirical application on the study of interactive income-consumption dynamics follows in Section 4. The section summarizes all our findings on the characteristics of the time series of income and consumption, and on their interactions in the shortrun and in the longrun. Especially, we present the longrun and shortrun responses of consumption distributions to the changes in income distributions at each income level. Section 5 concludes the paper, and followed by Appendix that includes mathematical proofs.

2 Model and Methodology

In this section, we introduce a new framework to analyze the time series of densities representing cross-sectional distributions of some economic variables. We allow for the unit root type nonstationarity, possibly having some common stochastic trends, in the time series of distributions. Therefore, the notions of unit roots and cointegration in distributions naturally arise. Under such a general setup, we provide a methodology that is very useful to learn and interpret the longrun and shortrun relationships between two time series of cross-sectional distributions.

2.1 Distributional Time Series

Let (f_t) and (g_t) be two time series of densities representing cross-sectional distributions of some economic variables, which we call *distributional time series* for short. We regard the densities (f_t) and (g_t) as random elements taking values on the Hilbert space H of square integrable functions on $\mathbb{R}$. As usual, we define the inner product $\langle v, w\rangle = \int v(s)w(s)ds$ for any $v, w \in H$. For the main application in the paper, we designate (f_t) and (g_t) respectively to be the monthly time series of densities for income and consumption distributions. They are of course not directly observable and should be estimated using cross-sectional observations on household income and consumption. However, to present our framework and methodology more effectively, we tentatively assume that they are observable. In fact, the errors incurred in estimating cross-sectional densities are expected to be negligible and vanish asymptotically for many practical applications with much larger cross-sectional dimensions relative to their time series dimensions.[1]

For the time series of densities (f_t) and (g_t), we define

$$\big(\langle v, f_t\rangle\big) \quad \text{and} \quad \big(\langle w, g_t\rangle\big)$$

[1]This is because the estimation errors for cross-sectional densities decrease uniformly over any compact interval as the number of cross-sectional observations increases.

to be the *coordinate processes* of (f_t) and (g_t) respectively in the directions of v and w for any $v, w \in H$. The coordinate processes of (f_t) and (g_t) in the direction of ι_κ, where

$$\iota_\kappa(s) = s^\kappa, \tag{1}$$

are particularly importance, since we have

$$\langle \iota_\kappa, f_t \rangle = \int s^\kappa f_t(s) ds \quad \text{and} \quad \langle \iota_\kappa, g_t \rangle = \int s^\kappa g_t(s) ds,$$

which represent the κ-th moments of the distributions represented by f_t and g_t for each $t = 1, \ldots, T$. They will also be referred subsequently to as the κ-th *cross-sectional moments* of (f_t) and (g_t) respectively.

In what follows, we often consider, though not exclusively, the distributional regression

$$g_t = \mu + Af_t + e_t \tag{2}$$

for $t = 1, \ldots, T$, where the regressand and regressor are time series of densities for cross-sectional distributions, μ and A are respectively function and operator parameters, and (e_t) is a function-valued error process. The operator A generalizes the regression coefficient in finite-dimensional regression, and may be called the regression operator. In the distributional regression (2), we allow for nonstationarity in both (f_t) and (g_t). In particular, we let some of their coordinate processes $(\langle v, f_t \rangle)$ and $(\langle w, g_t \rangle)$ have unit roots and cointegration, which will be referred to as the *distributional unit roots* and *cointegration*. They will be discussed in detail and fully analyzed in the next subsections. Typically, we assume that the error process (e_t) is stationary and mean zero, i.e., $\mathbb{E}e_t = 0$ for all $t = 1, \ldots, T$. Furthermore, for some of our subsequent results, we need to impose some exogeneity condition for (f_t) and this will be introduced later.

The coordinate regression of (g_t) in any direction $w \in H$ can be readily obtained from our distributional regression. In fact, it follows directly from (2) that

$$\begin{aligned} \langle w, g_t \rangle &= \langle w, \mu \rangle + \langle w, Af_t \rangle + \langle w, e_t \rangle \\ &= \langle w, \mu \rangle + \langle A^* w, f_t \rangle + \langle w, e_t \rangle \end{aligned} \tag{3}$$

for any $w \in H$, where A^* is the adjoint operator of A and $t = 1, \ldots, T$. A coordinate regression represents a relationship between particular coordinate processes of (g_t) and (f_t). Clearly, the coordinate regression (3) may be interpreted as the usual bivariate regression of the coordinate process $(\langle w, g_t \rangle)$ of (g_t) on the coordinate process $(\langle v, f_t \rangle)$ of (f_t) with

$v = A^*w$ for any $w \in H$. The regression reveals the effect of the distribution represented by (f_t) on the coordinate process $(\langle w, g_t\rangle)$ for $w \in H$. The effect is summarized by the corresponding $v = A^*w = A^*\iota_\kappa$, which we call the *response function* of (f_t) to the coordinate process $(\langle w, g_t\rangle)$. If we set $w = \iota_\kappa$, the coordinate regression (3) reveals how the κ-th cross-sectional moment of (g_t) is affected by the distribution represented by (f_t), and the response function $v = A^*w = A^*\iota_\kappa$ measures the effect of (f_t) on the cross-sectional moments of (g_t). In the paper, we analyze the coordinate regression (3) separately for stationary and nonstationary components of (f_t) and (g_t).

We may also consider the distributional regression (2) in a demeaned form as

$$y_t = Ax_t + \varepsilon_t, \tag{4}$$

where

$$x_t = f_t - \frac{1}{T}\sum_{t=1}^{T} f_t, \qquad y_t = g_t - \frac{1}{T}\sum_{t=1}^{T} g_t \tag{5}$$

and $\varepsilon_t = e_t - T^{-1}\sum_{t=1}^{T} e_t$ for $t = 1, \ldots, T$. Of course, our definitions of (x_t), (y_t) and (ε_t) are all dependent upon T, and should be denoted more appropriately as, say, (x_t^T), (y_t^T) and (ε_t^T) with the superscript T. However, for the sake of simplicity in our notation, we suppress the superscript T in our subsequent discussions. Note that $\varepsilon_t \approx e_t - \mathbb{E}e_t = e_t$ for large T, since we assume that (e_t) is stationary and has mean zero. However, in general, (x_t) and (y_t) do not behave the same as $(f_t - \mathbb{E}f_t)$ and $(g_t - \mathbb{E}g_t)$ even asymptotically, since (f_t) and (g_t) are nonstationary.

In our statistical analysis, we mainly deal with the demeaned densities (x_t) and (y_t) defined in (5). To implement our methodology, we assume that the densities (f_t) and (g_t) all have supports included in a compact subset K of $\mathbb{R}$, for $t = 1, \ldots, T$. Under the assumption, the demeaned densities (x_t) and (y_t) take values in

$$L_0^2(K) = \left\{ w \in H \,\middle|\, \int_K w(s)ds = 0, \int_K w^2(s)ds < \infty \right\}, \tag{6}$$

which is a subspace of the Hilbert space $L^2(\mathbb{R})$ of square integrable functions on $\mathbb{R}$ endowed with the usual inner product. The moment functions ι_κ are redefined as

$$\iota_\kappa(s) = s^\kappa - \frac{1}{|K|}\int_K s^\kappa ds,$$

where $|K|$ denotes the length of K, so that they belong to $L_0^2(K)$. As is well known, the Hilbert space $L_0^2(K)$ is separable and has a countable basis. For all our actual computations, we use an approximate one-to-one correspondence between $L_0^2(K)$ and $\mathbb{R}^M$ for some large M using a Wavelet basis in $L_0^2(K)$.[2]

2.2 Distributional Unit Root and Cointegration

As discussed, we allow for nonstationarity in (f_t) and (g_t). More precisely, we assume that, in the directions of some v and w for $v, w \in H$, the coordinate processes $(\langle v, f_t \rangle)$ and $(\langle w, g_t \rangle)$ have unit roots. Following Chang et al. (2012), we define the subspaces F_S and G_S of H as

$$F_S = \left\{ v \in H \middle| \langle v, f_t \rangle \text{ is stationary} \right\}$$
$$G_S = \left\{ w \in H \middle| \langle w, g_t \rangle \text{ is stationary} \right\},$$

which are called respectively the *stationary subspaces* of (f_t) and (g_t), and let F_N and G_N be the orthogonal complements of F_S and G_S, called respectively the *nonstationary subspaces* of (f_t) and (g_t), so that $H = F_N \oplus F_S = G_N \oplus G_S$. We only consider the unit root type nonstationarity in (f_t) and (g_t), and therefore, it follows that the time series $(\langle v, f_t \rangle)$ and $(\langle w, g_t \rangle)$ are unit root processes for all $v \in F_N$ and $w \in G_N$. Throughout the paper, we assume that the nonstationarity subspaces F_N and G_N are finite-dimensional. Needless to say, the stationary subspaces F_S and G_S are infinite-dimensional. In what follows, we denote by P_N and Q_N the projections on the nonstationary subspaces F_N and G_N of (f_t) and (g_t), and similarly by P_S and Q_S the projections on the stationary subspaces F_S and G_S of (f_t) and (g_t), respectively.

Chang et al. (2012) show how we may consistently estimate the nonstationary subspaces F_N and G_N respectively from (f_t) and (g_t). For a given time series of densities, they propose to determine the dimension of its nonstationary subspace by recursively testing for the number of unit roots and estimate the nonstationary subspace itself, based on the functional principle component analysis. They also convincingly demonstrate that the time series of income distributions are nonstationary and have unit roots. The reader is referred to their paper for more details.

If (f_t) and (g_t) have the unit root type nonstationarity, it is natural to consider the possibility that some of their coordinate processes are cointegrated. That is, for some

[2] We may of course possibly use other bases such as trigonometric functions. However, we find that Wavelet bases work much better than other choices including trigonometric functions in dealing with demeaned densities.

$v \in F_N$ and $w \in G_N$, we may have

$$\langle w, g_t \rangle = \delta + \langle v, f_t \rangle + u_t \tag{7}$$

with some constant δ, where (u_t) is a general stationary process with mean zero. The relationship in (7) will be referred to as the *distributional cointegration* in the paper.

Now we assume more explicitly that F_N and G_N are p- and q-dimensional and there are p- and q-unit roots in (f_t) and (g_t), respectively. Therefore, we have $v_1, \ldots, v_p$ and $w_1, \ldots, w_q$, which are linearly independent and span F_N and G_N, such that $\langle v_i, f_t \rangle$ and $\langle w_j, g_t \rangle$ are unit root processes for $i = 1, \ldots, p$ and $j = 1, \ldots, q$.[3] If the $(p+q)$-dimensional process (z_t) defined as

$$z_t = \big(\langle v_1, f_t \rangle, \ldots, \langle v_p, f_t \rangle, \langle w_1, g_t \rangle, \ldots, \langle w_q, g_t \rangle\big)' \tag{8}$$

is cointegrated with the cointegrating vector

$$c = (-a_1, \ldots, -a_p, b_1, \ldots, b_q)', \tag{9}$$

then the distributional cointegration in (7) holds with

$$v = a_1 v_1 + \cdots + a_p v_p \quad \text{and} \quad w = b_1 w_1 + \cdots + b_q w_q. \tag{10}$$

In the paper, we call the pair of functions v and w defined in (10) the *distributional cointegrating functions* of two time series (f_t) and (g_t) of densities, and denote them by pair of functions v_C and w_C.

The distributional cointegrating function (v_C, w_C) of (f_t) and (g_t) measures the longrun response v_C of the time series of cross-sectional distribution represented by (f_t) on the time series $(\langle w_C, g_t \rangle)$. In particular, we define v_C to be the *longrun response function* of (f_t) on $(\langle w_C, g_t \rangle)$, which we may interpret as summarizing the longrun effect of (f_t) on the longrun movement of (g_t) in the direction of w_C.

Clearly, there are at most r-number of linearly independent distributional cointegrating relationships, $r \leq \min(p, q)$, between (f_t) and (g_t). This is because otherwise we would have a cointegrating vector c in (9) of the form $c = (-a_1, \ldots, -a_p, 0, \ldots, 0)'$ or $c = (0, \ldots, 0, b_1, \ldots, b_q)'$, which implies that there is a linear combination of $v_1, \ldots, v_p$ or $w_1, \ldots, w_q$ whose inner product with (f_t) or (g_t) becomes stationary, contradicting the as-

[3] Of course, they are not uniquely defined. However, our subsequent analysis does not require the individual identification of $v_1, \ldots, v_p$ and $w_1, \ldots, w_q$, and becomes invariant with respect to their choices as long as they span F_N and G_N respectively.

sumption that $v_1, \ldots, v_p$ and $w_1, \ldots, w_q$ are linearly independent functions that span F_N and G_N respectively. In case $r > 1$, we use the notations (v_k^C, w_k^C), $k = 1, \ldots, r$, for the distributional cointegrating functions of (f_t) and (g_t). However, in this case, the distributional cointegrating functions (v_k^C, w_k^C), $k = 1, \ldots, r$, of (f_t) and (g_t) are not individually identified, unless we impose some specific restrictions on their normalization. The subspaces of $F_N \times G_N$ spanned by them are nevertheless well identified, which we denote by $F_C \times G_C$ and call the *distributional cointegrating subspaces* of (f_t) and (g_t).

The distributional cointegration does not presume any distributional regression relationship like (2). However, for two time series of densities (f_t) and (g_t) that are given by the distributional regression model (2), we may easily deduce that

Lemma 2.1 Let (f_t) and (g_t) be given by the distributional regression model (2) with some stationary (e_t). Then for any $w \in G_N$ we have $A^*w \notin F_S$ and the distributional cointegration (7) holds with $v = P_N A^* w$.

If (f_t) and (g_t) are given by the distributional regression (2), then we have

$$G_C = G_N \quad \text{and} \quad r = q \leq p,$$

due to Lemma 2.1. Note that we still have $F_C \subset F_N$ in general. In this case, it follows that there exists a distributional cointegrating (v_C, w_C) function of (f_t) and (g_t) having

$$w_C = Q_N \iota_\kappa.$$

However, if we let $g_t^N = Q_N g_t$, then it follows that

$$\langle w_C, g_t \rangle = \langle Q_N \iota_\kappa, g_t \rangle = \langle \iota_\kappa, Q_N g_t \rangle = \langle \iota_\kappa, g_t^N \rangle,$$

and therefore, we may interpret the corresponding v_C as the *longrun response function* of (f_t) to the κ-th cross-sectional moment of (g_t^N), or the κ-th longrun cross-sectional moment of (g_t).[4] Recall that (g_t^N) is the nonstationary component of (g_t).

[4]It follows from Lemma 2.1 that v_C is given more explicitly as $v_C = P_N A^* Q_N \iota_\kappa$. However, we do not need to know A^* to find v_C. We may simply write w_C as a linear combination $w_C = b_1 w_1 + \cdots + b_q w_q$ and obtain the corresponding $v_C = a_1 v_1 + \cdots + a_p v_p$, once we estimate the cointegrating subspaces of (f_t) and (g_t) as in (10).

2.3 Stationary Distributional Regression

We let $f_t^S = P_S f_t$ and $g_t^S = Q_S g_t$, $t = 1, \ldots, T$, so that (f_t^S) and (g_t^S) are the stationary components of (f_t) and (g_t). Consider the stationary distributional regression

$$g_t^S = \nu + B f_t^S + e_t, \tag{11}$$

where ν is the constant parameter function and B is the regression operator, and (e_t) is a function-valued stationary error process with mean zero. To identify the regression operator B, we assume that (e_t) is uncorrelated with (f_t^S), i.e., $\mathbb{E} f_t^S \otimes e_t = 0$, in the stationary distributional regression (11).[5]

If (f_t) and (g_t) are given by the distributional regression model (2) with some stationary (e_t), it follows immediately that the stationary distributional regression (11) holds for (f_t^S) and (g_t^S). In fact, we have

Lemma 2.2 Let (f_t) and (g_t) be given by the distributional regression model (2) with some stationary (e_t). Then we have $Q_S A P_N = 0$ and the stationary distributional regression (11) holds with the regression operator $B = Q_S A$.

We may deduce from the stationary distributional regression (11) that

$$\begin{aligned} \langle w, g_t^S \rangle &= \langle w, \nu \rangle + \langle w, B f_t^S \rangle + \langle w, e_t \rangle \\ &= \langle w, \nu \rangle + \langle B^* w, f_t^S \rangle + \langle w, e_t \rangle, \end{aligned} \tag{12}$$

where B^* is the adjoint operator of B. If, in particular, we set $w = \iota_\kappa$, then $B^* w = B^* \iota_\kappa$ measures the response of the stationary component of (f_t) to the κ-th cross-sectional moment of the stationary component of (g_t), which we will simply refer to as the *shortrun response function* of (f_t) to the κ-th cross-sectional moment of (g_t).

In parallel with (4), we may write

$$y_t^S = B x_t^S + \varepsilon_t \tag{13}$$

in demeaned form, where (x_t^S) and (y_t^S) are defined from (f_t^S) and (g_t^S) exactly as (x_t) and (y_t) are defined from (f_t) and (g_t). As we will explain later, we use the demeaned stationary distributional regression (13) to estimate B. For the consistent estimation of regression operator B and its asymptotic theory, the reader is referred to Park and Qian

[5]For random elements x and y taking values in a Hilbert space H, we define their covariance as $\mathbb{E}(x - \mathbb{E}x) \otimes (y - \mathbb{E}y)$, which is an operator in H.

(2012). Clearly, under the assumption that (f_t) and (g_t) have supports contained in a compact subset K of $\mathbb{R}$, (x_t^S) and (y_t^S) belong to the Hilbert space $L_0^2(K)$, and therefore, can be represented as large dimensional vectors using a Wavelet basis in $L_0^2(K)$.

3 Statistical Procedure

In this section, we introduce the statistical procedures used to analyze our model and draw inferences on its various implications. It is shown how to interpret the distributional unit roots and cointegration, as well as how to estimate and test for them. We also demonstrate how to estimate and do inference on the stationary component of our model.

3.1 Inference on Distributional Unit Roots and Cointegration

Throughout this subsection, we let $(w_t) = (x_t)$ or (y_t), and denote $H_N = F_N$ or G_N and $H_S = F_S$ or G_S, and $\Pi_N = P_N$ or Q_N and $\Pi_S = P_S$ or Q_S, depending upon whether (w_t) is defined as $(w_t) = (x_t)$ or $(w_t) = (y_t)$.

Our test for unit roots in (w_t) is based on the sample variance operator

$$M^T = \sum_{t=1}^{T} w_t \otimes w_t, \tag{14}$$

whose quadratic form is given by

$$\langle v, M^T v \rangle = \sum_{t=1}^{T} \langle v, w_t \rangle^2 \tag{15}$$

for $v \in H$. The asymptotic behavior of the quadratic form (15) depends crucially on whether v is in H_N or in H_S. For $v \in H_S$, the coordinate process $(\langle v, w_t \rangle)$ becomes stationary and we expect that

$$T^{-1} \sum_{t=1}^{T} \langle v, w_t \rangle^2 \to_p \mathbb{E}\langle v, w_t \rangle^2 \tag{16}$$

as long as the expectation exists. On the other hand, if $v \in H_N$ and the coordinate process $(\langle v, w_t \rangle)$ is integrated, it follows under a very mild condition that

$$T^{-2} \sum_{t=1}^{T} \langle v, w_t \rangle^2 \to_d \int_0^1 V(r)^2 dr - \left(\int_0^1 V(r) dr \right)^2, \tag{17}$$

where V is a Brownian motion. This is well expected. Therefore, the quadratic form has different orders of magnitude, i.e., $O_p(T)$ and $O_p(T^2)$, depending upon whether the coordinate process $(\langle v, w_t\rangle)$ is stationary or integrated.

We let H_N be n-dimensional and denote by $v_1^T, v_2^T, \ldots$ the orthonormal eigenvectors of the sample variance operator M^T in (14). It is shown in Chang et al. (2012) that we have

$$v_i^T \to_p v_i \tag{18}$$

for $i = 1, 2, \ldots$, as $T \to \infty$, and

$$H_N = \bigvee_{i=1}^{n} v_i \quad \text{and} \quad H_S = \bigvee_{i=n+1}^{\infty} v_i,$$

where the symbol $\bigvee$ denotes span. However, if we define $\lambda_1^T \geq \lambda_2^T \geq \cdots$ to be the eigenvalues of M^T associated with the eigenvectors $v_1^T, v_2^T, \ldots$, then we have

$$\lambda_i^T = \langle v_i^T, M^T v_i^T\rangle = \sum_{t=1}^{T} \langle v_i^T, w_t\rangle^2$$

for $i = 1, 2, \ldots$. Therefore, it follows that

$$\lambda_i^T = \begin{cases} O_p(T^2) & \text{for } i = 1, \ldots, n \\ O_p(T) & \text{for } i = n+1, \ldots \end{cases},$$

due to (16), (17) and (18).

To determine the number of unit roots in (w_t), we consider the test of the null hypothesis

$$\mathrm{H}_0: \ \dim(H_N) = n \tag{19}$$

against the alternative hypothesis

$$\mathrm{H}_1: \ \dim(H_N) \leq n - 1 \tag{20}$$

successively. More precisely, we start testing the null hypothesis (19) against the alternative hypothesis (20) with $n = n_{\max}$, where $n_{\max}$ is large enough so that surely we have $\dim(H_N) \leq n_{\max}$, and continue with $n = n_{\max} - 1$ if the null hypothesis (19) is rejected in favor of the alternative hypothesis (20). Clearly, if, for any n, $\dim(H_N) \leq n$ and the null hypothesis (19) is not rejected, then we may conclude that $\dim(H_N) = n$. Therefore, we may estimate the number of unit roots in (w_t) by the smallest value of n for which we fail to reject the null hypothesis (19) in favor of the alternative hypothesis (20).

n	1	2	3	4	5
1%	0.0274	0.0175	0.0118	0.0103	0.0085
5%	0.0385	0.0223	0.0154	0.0127	0.0101
10%	0.0478	0.0267	0.0175	0.0139	0.0111

Table 1: **Critical Values for the Test Statistic τ_n^T.**

We expect that the eigenvalue λ_n^T would have a discriminatory power for the test of null hypothesis (19) against the alternative hypothesis (20), since it has different orders of stochastic magnitudes under the null and alternative hypothese. However, it cannot be used directly as a test statistic, since its limit distribution is dependent upon nuisance parameters. Therefore, we need to modify it appropriately to get rid of its nuisance parameter dependency problem.

To introduce our test, we let (z_t) be given by

$$z_t^T = (\langle v_1^T, w_t \rangle, \ldots, \langle v_n^T, w_t \rangle)' \tag{21}$$

for $t = 1, \ldots, T$. Moreover, we define the product sample moment $M_n^T = \sum_{t=1}^T z_t^T z_t^{T\prime}$, and the long-run variance estimator $\Omega_n^T = \sum_{|k| \leq \ell} \varpi_\ell(k) \Gamma_T(k)$ of (z_t^T), where ϖ_ℓ is the weight function with bandwidth parameter ℓ and Γ_T is the sample autocovariance function defined as $\Gamma_T(k) = T^{-1} \sum_t \Delta z_t^T \Delta z_{t-k}^{T\prime}$.[6] Our test statistic is defined as

$$\tau_n^T = T^{-2} \lambda_{\min}\left(M_n^T, \Omega_n^T\right), \tag{22}$$

where $\lambda_{\min}\left(M_n^T, \Omega_n^T\right)$ is the smallest generalized eigenvalue of M_n^T with respect to Ω_n^T. Under very general conditions, Chang et al. (2012) show that if the null hypothesis (19) holds, then we have

$$\tau_n^T \to_d \lambda_{\min}\left(\int_0^1 W_n(r) W_n(r)' dr - \int_0^1 W_n(r) dr \int_0^1 W_n(r)' dr \right) \tag{23}$$

as $T \to \infty$, where W_n is n-dimensional standard vector Brownian motion and $\lambda_{\min}(\cdot)$ denotes the smallest eigenvalue of its matrix argument. On the other hand, we have $\tau_n^T \to_p 0$ under the alternative hypothesis (20) as $T \to \infty$. Therefore, we reject the null hypothesis (19) in favor of the alternative hypothesis (20) if the test statistic τ_n^T takes small values. The critical values are obtained by Chang et al. (2012) and presented in Table 1 for easy reference.

[6]See, e.g., Andrews (1991) for more discussions on the estimation of longrun variances.

Once we determine n, we may estimate H_N by

$$H_N^T = \bigvee_{i=1}^{n} v_i^T,$$

i.e., the span of the n orthonormal eigenvectors of M^T associated with n largest eigenvalues of M_T in (14). Chang et al. (2012) establish the consistency of H_N^T for H_N.

As will be explained below, we may now find how much nonstationarity proportion exists in each cross-sectional moments. In what follows, we redefine ι_κ introduced in (1) as $\iota_\kappa - \int_K \iota_\kappa(s)ds$, so that we may regard it as an element in $L_0^2(K)$. We may decompose ι_κ as $\iota_\kappa = \Pi_N \iota_\kappa + \Pi_S \iota_\kappa$, from which it follows that

$$\|\iota_\kappa\|^2 = \|\Pi_N \iota_\kappa\|^2 + \|\Pi_S \iota_\kappa\|^2 = \sum_{i=1}^{n} \langle \iota_\kappa, v_i \rangle^2 + \sum_{i=n+1}^{\infty} \langle \iota_\kappa, v_i \rangle^2,$$

where (v_i), $i = 1, 2, \ldots$, is an orthonormal basis of $L_0^2(K)$ such that $(v_i)_{1 \leq i \leq n}$ and $(v_i)_{i \geq n+1}$ span H_N and H_S, respectively.

To measure the proportion of the component of ι_κ lying in H_N, we define

$$\pi_\kappa = \frac{\|\Pi_N \iota_\kappa\|}{\|\iota_\kappa\|} = \sqrt{\frac{\sum_{i=1}^{n} \langle \iota_\kappa, v_i \rangle^2}{\sum_{i=1}^{\infty} \langle \iota_\kappa, v_i \rangle^2}}. \tag{24}$$

We have $\pi_\kappa = 1$ and $\pi_\kappa = 0$, respectively, if ι_κ is entirely in H_N and H_S. Therefore, we may use π_κ to represent the proportion of nonstationary component in the κ-th cross-sectional moment of (w_t). The κ-th cross-sectional moment of (w_t) has more dominant unit root component as π_κ tends to unity, whereas it becomes more stationary as π_κ approaches to zero. Clearly, the κ-th cross-sectional moment of (w_t) becomes more difficult to predict if π_κ is closer to unity, and easier to predict if π_κ is small. Following Chang et al. (2012), π_κ is referred to as the *nonstationarity proportion* of the κ-th cross-sectional moment of (w_t).

The nonstationarity proportion π_κ of the κ-th cross-sectional moment defined in (24) is of course not directly applicable, since H_N and H_S are unknown. However, we may use its sample version

$$\pi_\kappa^T = \sqrt{\frac{\sum_{i=1}^{n} \langle \iota_\kappa, v_i^T \rangle^2}{\sum_{i=1}^{T} \langle \iota_\kappa, v_i^T \rangle^2}}. \tag{25}$$

The sample version π_κ^T in (25) of π_κ in (24) will be referred to as the *sample* nonstationarity proportion of the κ-th cross-sectional moment of (w_t). Chang et al. (2012) show that the sample nonstationarity proportion π_κ^T is a consistent estimator for the original nonstationarity proportion π_κ.

Now we assume that we find p and q, the numbers of unit roots in (f_t) and (g_t), and obtain consistent estimates (v_i^T) of (v_i) and (w_j^T) of (w_j), $i = 1, \ldots, p$ and $j = 1, \ldots, q$, which span the nonstationary subspaces F_N and G_N of (f_t) and (g_t). To test for distributional cointegration, we let (z_t^T) be defined as

$$z_t^T = \big(\langle v_1^T, x_t\rangle, \ldots, \langle v_p^T, x_t\rangle, \langle w_1^T, y_t\rangle, \ldots, \langle w_q^T, y_t\rangle\big)', \tag{26}$$

in place of (z_t^T) introduced in (21), and subsequently redefine τ_n^T in (22) from (z_t^T) in (26), exactly as it is defined from (z_t^T) in (21). Clearly, the newly defined statistic τ_n^T may be used to test for the number of unit roots in (z_t), $z_t = \big(\langle v_1, x_t\rangle, \ldots, \langle v_p, x_t\rangle, \langle w_1, y_t\rangle, \ldots, \langle w_q, y_t\rangle\big)'$, in (8). The critical values in Table 1 are applicable also for the newly defined statistic τ_n^T. The maximum number of unit roots for (z_t) in (8) is of course given by $p+q$, in which case we have no distributional cointegration in (f_t) and (g_t). If we find n-number of unit roots for (z_t) in (8), then it implies that we have r-number of cointegrating relationships with $r = (p+q) - n$. As discussed, we should have $r \leq \min(p, q)$.

3.2 Inference on Stationary Distributional Regression

Now we explain how to consistently estimate the regression operator B on F_S in the stationary distributional regression (11). Let

$$\begin{aligned} M_S &= \mathbb{E}\left[(f_t^S - \mathbb{E}f_t^S) \otimes (f_t^S - \mathbb{E}f_t^S)\right] \\ N_S &= \mathbb{E}\left[(g_t^S - \mathbb{E}g_t^S) \otimes (f_t^S - \mathbb{E}f_t^S)\right]. \end{aligned}$$

Then it follows from the orthogonality condition $\mathbb{E}f_t^S \otimes e_t = 0$ that

$$N_S = BM_S, \tag{27}$$

which we may use to estimate B. Unfortunately, however, it is generally impossible to use the relationship in (27) and define the regression operator B as $B = N_S M_S^{-1}$. This will be explained below.

We assume that M_S is a compact operator. Being compact and self-adjoint, M_S allows for the spectral representation

$$M_S = \sum_{i=1}^{\infty} \lambda_i (v_i \otimes v_i), \tag{28}$$

where (λ_i, v_i) are the pairs of eigenvalue and eigenvector of M_S. Even in the case $\lambda_i \neq 0$ for all i so that M_S^{-1} is well defined and given by $M_S^{-1} = \sum_{i=1}^{\infty} \lambda_i^{-1}(v_i \otimes v_i)$, M_S^{-1} is not defined on the entire domain of M_S. In fact, its domain is restricted to a proper subset of the domain of M_S given by $\left\{ w \,\middle|\, \sum_{i=1}^{\infty} \langle v_i, w \rangle^2 / \lambda_i^2 < \infty \right\}$. Therefore, we have $B = N_S M_S^{-1}$ only on the restricted domain. This problem is often referred to an ill-posed inverse problem.

The usual method to deal with this problem is to restrict the definition of M_S in a finite subset of its domain. Assuming $\lambda_1 > \lambda_2 > \cdots > 0$, we let F_{S_m} be the span of the m-eigenvectors $v_1, \ldots, v_m$ associated with the m-largest eigenvalues $\lambda_1, \ldots, \lambda_m$. Moreover, we denote by P_{S_m} the projection on F_{S_m}, and define $M_{S_m} = P_{S_m} M_S P_{S_m}$ and

$$M_{S_m}^{+} = \sum_{i=1}^{m} \frac{1}{\lambda_i} (v_i \otimes v_i), \tag{29}$$

i.e., the inverse of M_S on F_{S_m}. Subsequently, we let

$$B_m = N_S M_{S_m}^{+}, \tag{30}$$

which is the regression operator B restricted to the subspace F_{S_m} of F_S. Since (λ_i) decreases down to zero, we may well expect that B_m approximates B well if the dimension m of F_{S_m} increases. The reader is referred to Bosq (1998) for more detailed discussions.

The restricted regression operator B_m in (30) can be consistently estimated by its sample analogue. We define

$$M_S^T = \frac{1}{T} \sum_{t=1}^{T} x_t^S \otimes x_t^S \quad \text{and} \quad N_S^T = \frac{1}{T} \sum_{t=1}^{T} y_t^S \otimes x_t^S,$$

which are the sample analogue estimators of the operators M_S and N_S respectively, and denote by (λ_i^T, v_i^T) the pairs of eigenvalues and eigenvectors of M_S^T such that $\lambda_1^T > \lambda_2^T > \cdots$.

Then we define

$$M_{S_m}^{T+} = \sum_{i=1}^{m} \frac{1}{\lambda_i^T} (v_i^T \otimes v_i^T),$$

i.e., the sample analogue estimator of the operator $M_{S_m}^{+}$ in (29), and subsequently,

$$B_m^T = N_S^T M_{S_m}^{T+}, \tag{31}$$

which we use as an estimator for the regression operator B in (11). Park and Qian (2012) show that the estimator B_m^T is consistent for B under a very general set of conditions if we let $m \to \infty$ as $T \infty$ at a controlled rate.

4 Interactive Income-Consumption Dynamics

As an application of our model and methodology, we analyze the interactions between the income and consumption dynamics. For our analysis, we apply our theory developed thus far with (f_t) and (g_t) representing respectively the time series of household income and consumption distributions.

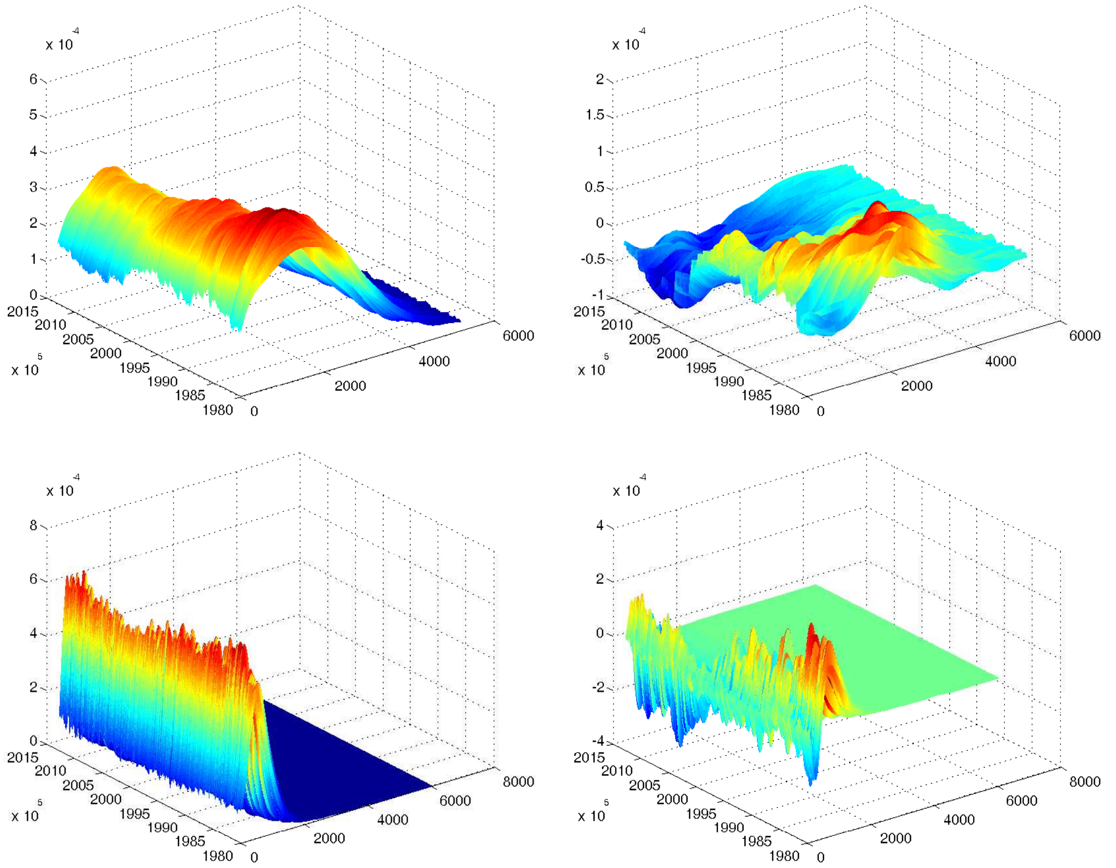
4.1 Data

We obtain the time-series of cross-sectional distributions of income and consumption using the U.S. households monthly income and consumption data from the Consumer expenditure (CE) survey series[7], which provides continuous flow of information on buying habits of the US customers. The survey is carried out by the U.S. Census Bureau under contract with the Bureau of Labor Statistics. We obtain the monthly data on income and consumption during the period from October 1979 to February 2013. During this sample period, each household included in the survey at most five times, and therefore the CE survey provides a pseudo panel data.

In order to construct monthly household income and consumption, we follow the definitions of income and consumption given by Krueger and Perri (2006), and aggregate the monthly values provided in Universal Classification Code (UCC) level for each month and year. See Krueger and Perri (2006) for detailed information. We then deflated the nominal income and consumption values by monthly CPI for all urban households with using a base year which varies among 1982, 1983 and 1984. The CPI used in here is provided by the Bureau of Labor Statistics.

[7]It is formerly called the Survey of Consumer Expenditures

Figure 1: Time Series of Income and Consumption Distributions



Notes: Time series of income and consumption distributions are presented at the upper and lower panels, without demeaning and with demeaning at the left and right panels, respectively.

The survey uses topcoding to change the values when the original data exceeds some prescribed critical values. The critical values may change annually and be applied at a different starting point. In our analysis we drop all top-coded values of household income and consumption. We correct the expenditure on food and impute services from vehicle and from primary residence, according to the regressions specified in Krueger and Perri (2006). Also, following the sample selection criteria given in Krueger and Perri (2006), we exclude observations with possible measurement error or inconsistency problem.

Figure 1 presents the time series of cross-sectional distributions for income and consumption with and without demeaning. Both the income and consumption distributions show some sign of nonstationary fluctuations evolving over time. In particular, it seems evident that the time series of their cross-sectional distributions do not randomly fluctuate around some fixed mean functions. This suggests the presence of nonstationarity in their time series.

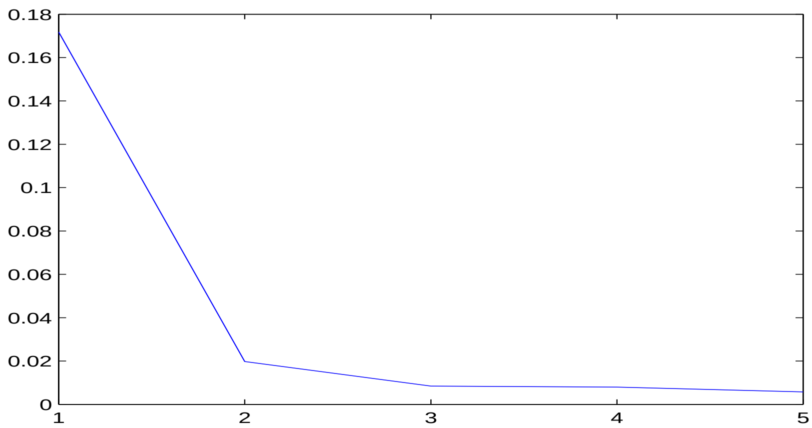
4.2 Nonstationarity in Income Dynamics

To determine the unit root dimension p in the time series of cross-sectional distributions of household incomes, use the test τ_n^T to test the null hypothesis $\mathrm{H}_0 : p = n$ against the alternative hypothesis $\mathrm{H}_1 : p \leq n-1$ with $n = 1, \ldots, 5$. The test results are given below.

n	1	2	3	4	5
τ_n^T	0.1077	0.0248	0.0101	0.0096	0.0083

Our test, strongly and unambiguously, rejects H_0 against H_1 successively for $n = 5, 4, 3$. Clearly, however, the test cannot reject H_0 in favor of H_1 for $n = 2$. Therefore, we conclude that there exists two-dimensional unit root, and set $p = 2$. Figure 2 shows that the leading principal component dominates all others, including the second principal component. However, it turns out that the second principal component is significantly larger than all other smaller components.

Figure 2: Scree Plot for Time Series of Income Distributions



Notes: Plotted are the five largest eigenvalues from the principal component analysis for the time series of income distributions.

We also compute the unit root portion estimates π_κ^T for the κ-th cross-sectional moments of household income distributions for $\kappa = 1, 2, 3$ and 4, as shown below.
The unit root proportions for the first four cross-sectional moments of household income distributions are all substantially large. In particular, the unit root proportions for the

π_1^T	π_2^T	π_3^T	π_4^T
0.6064	0.4137	0.2909	0.2154

first two cross-sectional moments are quite substantial. Needless to say, nonstationarity in the cross-sectional moments of household income would certainly make changes in the time series of income distributions more persistent.

4.3 Nonstationarity in Consumption Dynamics

To test for existence of unit root in time series of cross-sectional distributions of household consumptions, we also use the statistic τ_n^T to test the null hypothesis $\mathrm{H}_0 : q = n$ against the alternative hypothesis $\mathrm{H}_1 : q \leq n-1$ with $n = 1, \ldots, 5$. The test results are given below.

n	1	2	3	4	5
τ_n^T	0.0392	0.0143	0.0137	0.0074	0.0071

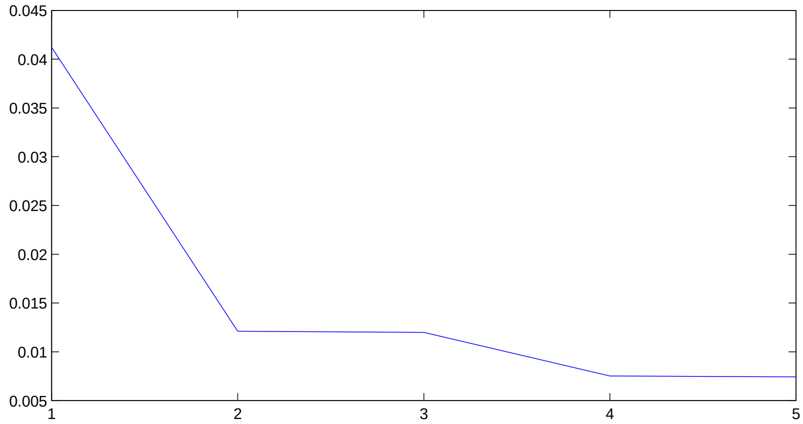
Our test successively rejects H_0 against H_1 for $n = 5, 4, 3, 2$. However, at 5% level, the test cannot reject H_0 in favor of H_1 for $n = 1$. Therefore, our test result implies $q = 1$. The scree plot for the time series of household consumption distributions is presented in Figure 3. As shown, there is one leading principal component, which supports our conclusion on the presence of unit root in the time series of consumption distributions. The second and third principal components are larger than other principal components of lower orders. However, the difference between them is tested to be insignificant.

Similarly as before, we compute the estimates π_κ^T of the unit root proportions π_κ for the first four cross-sectional moments of household consumption, assuming $q = 1$.

π_1^T	π_2^T	π_3^T	π_4^T
0.097	0.028	0.012	0.007

The unit root proportions are small for all of the first four moments, implying the nonstationarity in the cross-sectional distributions of household consumptions is not concentrated in the first four moments. However, the nonstationarity is relatively more concentrated in the first and the second moments, with the unit root proportion of the first moment being the largest.

Figure 3: Scree Plot for Time Series of Consumption Distributions



Notes: Plotted are the five largest eigenvalues from the principal component analysis for the time series of consumption distributions.

4.4 Interactive Dynamics of Income and Consumption

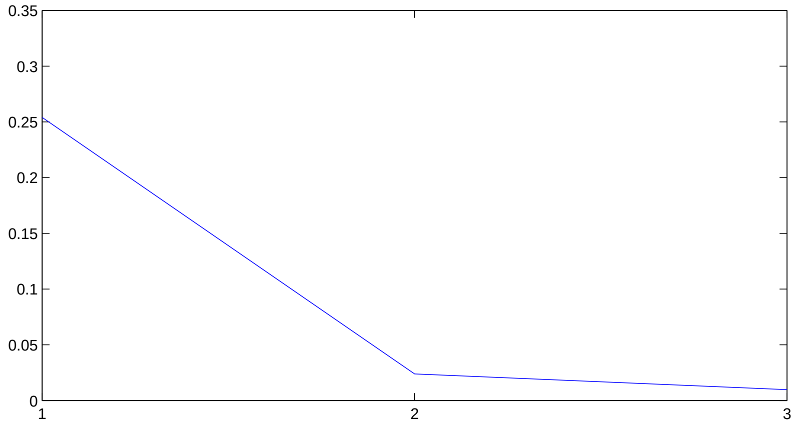
If the time series of income and consumption distributions have p- and q-number of unit roots and if they have r-number of cointegrating relationships, we would have $\big((p+q)-r\big)$-number of unit roots in their time series combined together. As discussed, we may use the test τ_n^T also in this case to find the number of unit roots in the combined time series of income and consumption distributions by testing the null hypothesis $\mathrm{H}_0 : (p+q)-r = n$ against the alternative hypothesis $\mathrm{H}_1 : (p+q)-r \leq n-1$. Given $p = 2$ and $q = 1$, we may have up to three unit roots in the time series of income and consumption distributions together. Therefore, we consider only $n = 1, 2$ and 3. The test results are summarized below.

n	1	2	3
τ_n^T	0.1362	0.0248	0.0116

Our test rejects H_0 against H_1 for $n = 3$. However, the test cannot reject H_0 in favor of H_1 for $n = 2$, giving $(p+q)-r = 2$. This implies $r = 1$, i.e., the presence of a single cointegrating relationship between household income and consumption distributions, since we have $p = 2$ and $q = 1$.

Let v_1 and v_2 be orthonormal functions that span the nonstationary subspace F_N of the time series (f_t) of income distributions, and let w be the normalized function generating

Figure 4: Scree Plot for Time Series of Income and Consumption Distributions



Notes: Plotted are the three eigenvalues from the principal component analysis for the time series of income and consumption distributions.

the nonstationary subspace G_N of the time series (g_t) of consumption distribution. We find the presence of cointegration in the time series of income and consumption distributions, and therefore, there exists constants a_1, a_2 and b such that

$$b\langle w, g_t\rangle = \delta + a_1\langle v_1, f_t\rangle + a_2\langle v_2, f_t\rangle + u_t$$

with some constant function δ and general stationary process with mean zero. In this case, we have

$$v_C = a_1 v_1 + a_2 v_2 \quad \text{and} \quad w_C = bw, \tag{32}$$

where (v_C, w_C) is the cointegrating function of (f_t) and (g_t).[8]

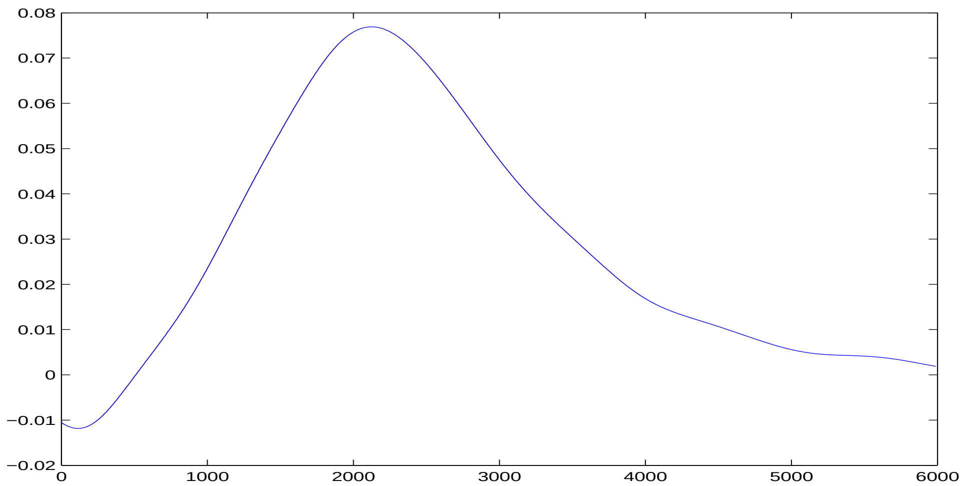
Using the procedures we introduce in the previous section, we may readily obtain estimates of v_C and w_C in (32), which we define as

$$v_C^T = a_1^T v_1^T + a_2^T v_2^T \quad \text{and} \quad w_C = b^T w^T, \tag{33}$$

from our estimates v_1^T, v_2^T and w^T of v_1, v_2 and w, and a_1^T, a_2^T and b^T of a_1, a_2 and b. We get the estimates v_1^T, v_2^T and w^T from our testing procedure for distributional unit roots, and the estimates a_1^T, a_2^T and b^T from our testing procedure for distributional cointegration, respectively in and between household income and consumption distributions. The estimated longrun response function of income distribution to consumption distribution is given by

[8]Obviously, we may set $b = 1$ without loss of generality and redefine a_1 and a_2 accordingly.

Figure 5: Longrun Income Response Function to Consumption



Notes: Presented is the longrun response function of income distribution to consumption distribution.

v_C^T defined in (33), and presented in Figure 5.

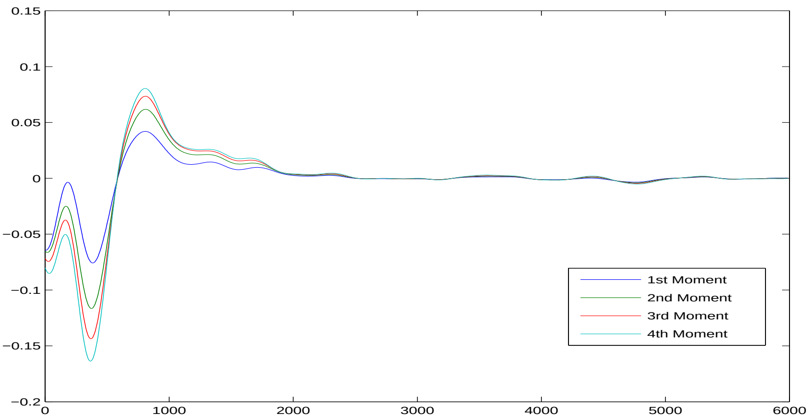
Our estimated longrun income response function to consumption reveals some interesting fact. For instance, it shows that the longrun trend in consumption is most affected by the income group with monthly earnings slightly over \$2,000. Roughly, all households with monthly earnings between \$1,000 and \$4,000 seem to play important roles in determining the persistent stochastic trend in consumption. As the level of monthly earning decreases below \$1,000, the longrun component of household's income has very little impact on the longrun consumption. The longrun component of household's income for the rich also does not have any major effect on the longrun consumption, though the magnitude of their effect decreases at a slower rate as their income increases than the rate it decreases as the income decreases for the poor. The income response to consumption is estimated to be negative for the household with monthly earnings less than approximately \$600, which we believe to be just an evidence of insignificant response.[9]

To analyze the shortrun response of income distribution to consumption, we compute and plot the shortrun response function of income distribution to the κ-th cross-sectional moments of consumption distribution, which is introduced in (12), for $\kappa = 1, \ldots, 4$. It is given by

$$B_m^{T*} \iota_\kappa,$$

[9]Observations for households with monthly earnings below approximately \$500 are scarce and irregular, so we do not expect to have any reliable results over very low income levels.

Figure 6: Shortrun Income Response Function to Consumption



Notes: Presented is the shortrun response function of income distribution to consumption distribution.

where B_m^{T*} is the adjoint operator of B_m^T defined in (31). The estimated shortrun response function of income distribution to the cross-sectional moments of consumption distribution is given in Figure 6.

Our estimated shortrun income response functions to consumption moments appear to provide some important clues on the shortrun relationship between income and consumption. All moments of consumption yield very similar income response functions. Except for the income group with monthly earnings less than approximately \$600, whose responses are negative and irregular, the income responses seem to be coherent and meaningful at all levels. As mentioned in our discussions on the longrun income response to consumption, we believe that our estimates for extreme low income levels are unreliable and ignorable. The shortrun income responses are maximized around the level a little below \$1,000 of monthly earnings for all moments of consumption. Needless to say, this means that the shortrun consumption is most affected by the transitory income of low income households. The shortrun income response decreases sharply as the level of income increases, at a much faster rate than the rate at which the longrun income response decreases, and becomes almost entirely negligible once the income level exceeds \$2,000 in monthly earnings.

5 Conclusion

This paper develops a new framework and methodology to analyze the relationships between two time series of cross-sectional distributions in the presence of distributional unit roots and cointegration. Their relationships, both in the longrun and in the shortrun, are revealed and summarized by what we call the response functions in the paper. Our analysis makes it possible to identify and estimate some important relationships between two time series of cross-sectional distributions, which we can never observe using the time series analysis relying on the aggregates. This is because the response functions provide us with the information about how one distribution affects the other distribution at each level of one distribution. Such information will never be revealed by the conventional time series analysis. We apply our approach to study the income and consumption dynamics, and find some interesting and important facts on the longrun and shortrun responses of consumption to income changes. Our findings have an immediate policy implication.

Mathematical Proofs

Proof of Lemma 2.1 If (g_t) and (f_t) are given by the distributional regression model (2), then it follows directly from the coordinate regression (3) that $A^*w \notin F_S$ for any $w \in G_N$. To see this, suppose on the contrary that there exists $w \in G_N$ such that $A^*w \in F_S$. Then for such w the time series $(\langle A^*w, f_t\rangle)$ becomes stationary, while the time series $(\langle w, g_t\rangle)$ is nonstationary. Clearly, this is a contradiction given that the error process (e_t) is stationary and the time series $(\langle w, e_t\rangle)$ is stationary for any $w \in H$.

We may easily deduce from (2) that

$$\begin{aligned} Q_N g_t &= Q_N\mu + Q_N A f_t + Q_N e_t \\ &= Q_N\mu + Q_N A P_N f_t + Q_N A P_S f_t + Q_N e_t. \end{aligned}$$

For any $w \in G_N$, it follows that

$$\langle w, Q_N g_t\rangle = \langle Q_N w, g_t\rangle = \langle w, g_t\rangle,$$

and

$$\langle w, Q_N A P_N f_t\rangle = \langle P_N A^* Q_N w, f_t\rangle = \langle P_N A^* w, f_t\rangle.$$

On the other hand, we have

$$\langle w, Q_N A P_S f_t\rangle = \langle A^* Q_N w, P_S f_t\rangle = \langle A^* w, P_S f_t\rangle,$$

which is stationary. Clearly, the time series $(\langle w, Q_N e_t\rangle) = (\langle Q_N w, e_t\rangle) = (\langle w, e_t\rangle)$ for any $w \in G_N$ is stationary. The proof is therefore complete. □

Proof of Lemma 2.2 We have

$$\begin{aligned} Q_S g_t &= Q_S \mu + Q_S A f_t + Q_S e_t \\ &= Q_S \mu + Q_S A P_S f_t + Q_S A P_N f_t + Q_S e_t, \end{aligned}$$

from which the stated results follow immediately. □

References

Andrews, D. W., 1991. Heteroskedasticity and autocorrelation consistent matrix estimation. Econometrica 59, 817–858.

Bosq, D., 1998. Nonparametric Statistics for Stochastic Processes, Lecture Notes in Statistics 110, 2nd Edition. Springer, New York.

Bosq, D., 2000. Linear Processes in Function Spaces, Lecture Notes in Statistics 149. Springer, New York.

Chang, Y., Kim, C., Park, J. Y., 2012. Nonstationarity in time series of state densities, Working paper, Indiana University.

Krueger, D., Perri, F., 2006. Does income inequality lead to consumption inequality? Evidence and theory. Review of Economic Studies 73, 163–193.

Park, J. Y., Qian, J., 2012. Functional regression of continuous state distributions. Journal of Econometrics 167, 397–412.