Time Series of Cross-Sectional Distributions with Common Stochastic Trends

Yoosoon Chang

Department of Economics
Indiana University

20th International Panel Data Conference
Hitotsubashi Hall
Tokyo, Japan
9-10 July 2014

Related Work

- "Nonstationarity in Time Series of State Densities," with Changsik Kim and Joon Park.
- "Time Series Analysis of Global Temperature Distributions: Identifying and Estimating Persistent Features in Temperature Anomalies," with Changsik Kim, J. Isaac Miller, Joon Y. Park and Sung-Keun Park.
- "Time Series of Cross-Sectional Distributions with Common Stochastic Trends," with Changsik Kim and Joon Y. Park.

Outline

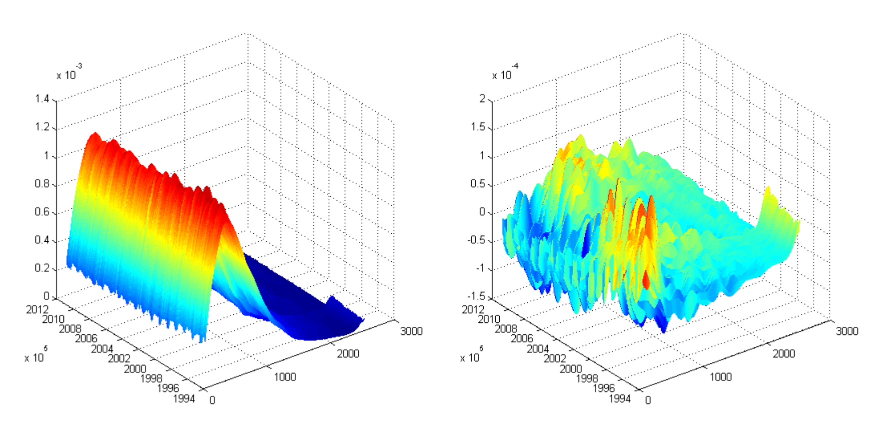
I. Basic Framework

Objective

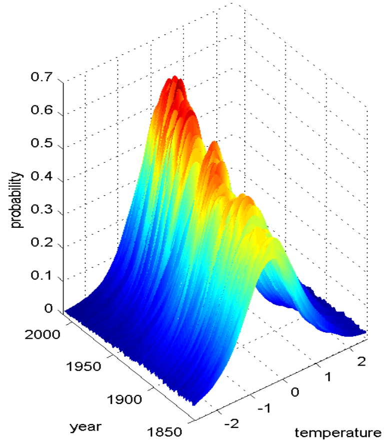
To analyze the time series of cross-sectional distributions such as

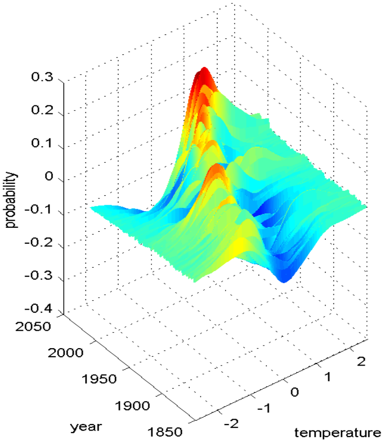
- individual earnings
- global temperatures
- household income
- household expenditures

Cross-sectional Distributions of Individual Earnings

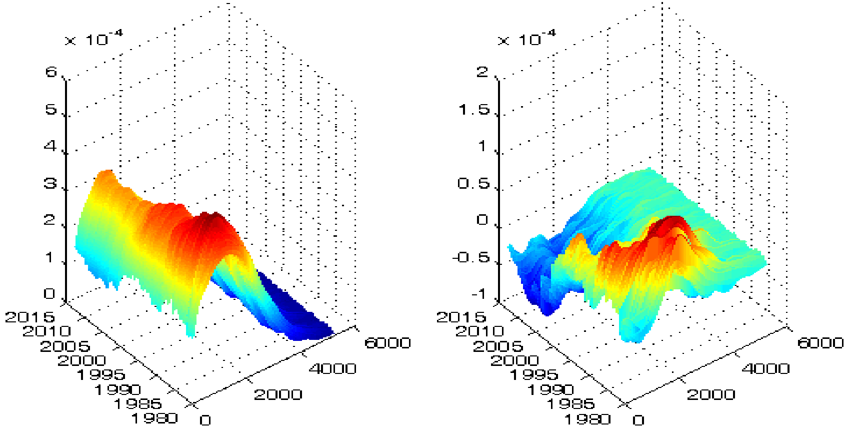


Global Temperature Distributions

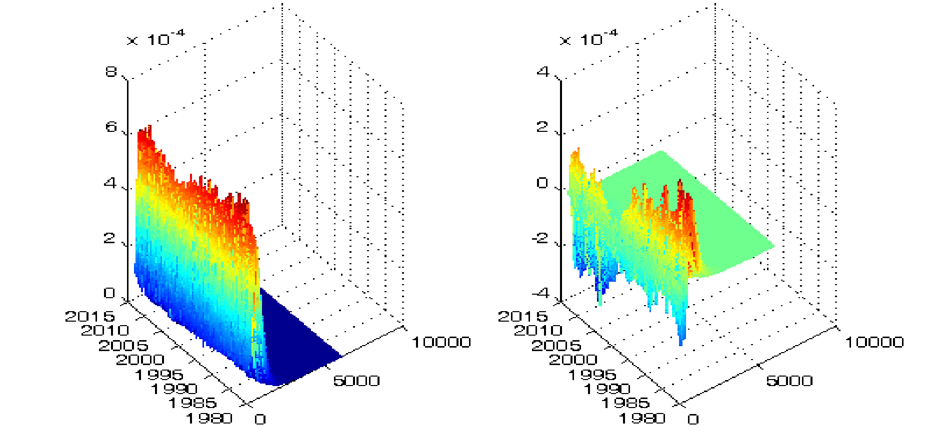




Cross-sectional Distributions of Household Income



Cross-sectional Distributions of Household Expenditures



Model for Persistent Functional Data

Model for Functional Data

- For each time $t = 1, 2, \ldots$, suppose there is a distribution represented by a probability density f_t, whose value at ordinate $s \in \mathbb{R}$ is denoted by $f_t(s)$.

- Denote by
$$w_t = f_t - \mathbb{E} f_t$$
a centered density function and treat w_t as functional data taking values in Hilbert space H.

- We define H to be the set of functions on a compact subset K of $\mathbb{R}$ that have vanishing integrals and are square integrable, i.e.,
$$H = \left\{ w \,\middle|\, \int_K w(s)ds = 0, \int_K w^2(s)ds < \infty \right\}$$
with inner product $\langle v, w \rangle = \int v(s)w(s)ds$ for $v, w \in H$.

Coordinate Process

- We assume that there exists an orthonormal basis (v_i) of H such that

$$w_t = \sum_{i=1}^{\infty} \langle v_i, w_t \rangle v_i$$

- And the i-th coordinate process

$$\langle v_i, w_t \rangle$$

  has a unit root for $i = 1, \ldots, n$, while it is stationary for all $i \geq n + 1$.

- By convention, we set $n = 0$ if all the coordinate processes are stationary.

Unit Root and Stationarity Subspaces

- Using the symbol $\bigvee$ to denote span, we let

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

so that $H = H_N \oplus H_S$. In what follows, H_N and H_S will respectively be referred to as the unit root and stationarity subspaces of H.

- We also let Π_N and Π_S be the projections on H_N and H_S, respectively. Moreover, we define

$$w_t^N = \Pi_N w_t \quad \text{and} \quad w_t^S = \Pi_S w_t$$

- Note that $\Pi_N + \Pi_S = 1$ (the identity operator on H), so in particular we have

$$w_t = w_t^N + w_t^S$$

Unit Root and Stationary Processes

- When $u_t = \triangle w_t = \Phi(L)\varepsilon_t$, it follows that

$$w_t^N = \Pi_N w_t = \Pi_N \Phi(1) \sum_{i=1}^{t} \varepsilon_i - \Pi_N \bar{u}_t$$

  and

$$w_t^S = \Pi_S w_t = -\Pi_S \bar{u}_t$$

  Clearly, (w_t^N) is an integrated process, while (w_t^S) is stationary.

- The unit root dimension n is unknown in practical applications.

- We will explain how to
  - Determine n statistically
  - Estimate the subspaces H_S and H_N

Functional Principal Component Analysis (FPCA)

- Our procedure to estimate H_N and test for its dimension M is based on the FPCA on the unnormalized sample variance operator of (w_t)

$$M_T = \sum_{t=1}^{T} w_t \otimes w_t$$

where T is the sample size.

- Denote the pairs of eigenvalues and eigenvectors of M_T by

$$(\lambda_i^T, v_i^T), \quad i = 1, \ldots, T$$

and order (λ_i^T) so that $\lambda_1^T \geq \cdots \geq \lambda_T^T$.

Sample Unit Root and Stationarity Subspaces

- Assuming $T > n$, we define sample unit root space as the subspace

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

spanned by the eigenvectors corresponding to n largest eigenvalues of M_T. Denote by Π_N^T the projection on H_N^T.

- The sample stationarity subspace is defined by $\Pi_S^T = 1 - \Pi_N^T$, so that we have $\Pi_N^T + \Pi_S^T = 1$ analogously as the relationship $\Pi_N + \Pi_S = 1$.

- We show that

$$\Pi_N^T = \Pi_N + O_p(T^{-1}) \quad \text{and} \quad \Pi_S^T = \Pi_S + O_p(T^{-1})$$

for all large T.

Decomposition of Sample Variance Operator

To develop our asymptotics, we decompose M^T as

$$M^T = T^2 M^T_{NN} + T M^T_{NS} + T M^T_{SN} + T M^T_{SS}$$

where

$$M^T_{NN} = \frac{1}{T^2}\Pi_N \left(\sum_{t=1}^T w_t \otimes w_t\right)\Pi_N = \frac{1}{T^2}\sum_{t=1}^T w_t^N \otimes w_t^N$$

$$M^T_{NS} = \frac{1}{T}\Pi_N \left(\sum_{t=1}^T w_t \otimes w_t\right)\Pi_S = \frac{1}{T}\sum_{t=1}^T w_t^N \otimes w_t^S$$

$$M^T_{SS} = \frac{1}{T}\Pi_S \left(\sum_{t=1}^T w_t \otimes w_t\right)\Pi_S = \frac{1}{T}\sum_{t=1}^T w_t^S \otimes w_t^S$$

and M^T_{SN} is the adjoint of M^T_{NS}, i.e., $M^T_{SN} = M^{T*}_{NS}$.

Asymptotics for Sample Variance Operators

Lemma Under some regularity conditions, we have

$$M_{NN}^T \ \to_d \ \ M_{NN} = \int_0^1 (W \otimes W)(r)dr$$

where W is Brownian motion on H_N with variance operator $\Pi_N \Phi(1) \Sigma \Phi(1)' \Pi_N$. Also, it follows that

$$M_{SS}^T \ \to_p \ \ M_{SS} = \Pi_S \left(\sum_{i=0}^{\infty} \bar{\Phi}_i \Sigma \bar{\Phi}_i' \right) \Pi_S$$

Moreover, we have

$$M_{NS}^T, M_{SN}^T = O_p(1)$$

for all large T.

Asymptotics for Eigenvalues and Eigenvectors

Theorem Under some regularity conditions, we have

$$\left(T^{-2}\lambda_i^T, v_i^T\right) \to_d \left(\lambda_i(M_{NN}), v_i(M_{NN})\right)$$

jointly for $i = 1, \ldots, n$, and

$$\left(T^{-1}\lambda_{n+i}^T, v_{n+i}^T\right) \to_p (\lambda_i, v_i)$$

for $i = 1, 2, \ldots$.

- In stationarity subspace H_S, eigenvectors and appropriately normalized eigenvalues of sample variance operator M_T of (w_t) converge in probability to their population counterparts.
- In unit root subspace H_N, they converge in distribution, and their distributional limits are given by the distributions of eigenvalues and eigenvectors of random operator M_{NN}.

II. Distributional Unit Roots

Sample Variance Operator

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

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

whose quadratic form is given by

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

for $v \in H$.

- Asymptotic behavior of the quadratic form of sample variance operator depends crucially on whether v is in H_N or in H_S.

Stationarity-Nonstationarity of Coordinate Processes

- For $v \in H_S$, the coordinate process $(\langle v, w_t \rangle)$ becomes stationary and we expect that

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

  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

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

  where V is a Brownian motion.

- 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.

Nonstationarity and Stationarity Subspaces

- 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.

- It is shown that

$$v_i^T \to_p v_i$$

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

Estimation of Nonstationarity Subspace

- Once we determine the number of unit roots n in (w_t), we may estimate the nonstationarity subspace H_N by

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

i.e., the span of the n orthonormal eigenvectors of the sample variance operator M^T associated with n largest eigenvalues of M_T.

- Recall

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

- We establish the consistency of H_N^T for H_N.

Functional Principal Component Analysis

- 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

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

Onto Testing for Distributional Unit Roots

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

$$\mathsf{H}_0: \ \dim(H_N) = n$$

against the alternative hypothesis

$$\mathsf{H}_1: \ \dim(H_N) \leq n - 1$$

successively downward.

- More precisely, we start testing the null with $n = n_{\max}$, where $n_{\max}$ is large enough so that $\dim(H_N) \leq n_{\max}$.
- Continue with $n = n_{\max} - 1$ if the null is rejected in favor of the alternative. If, for any n, $\dim(H_N) \leq n$ and the null 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.

Intuitive but Infeasible Test

- We expect that the eigenvalue λ_n^T would have a discriminatory power for the test of null against the alternative, since it has different orders of stochastic magnitudes under the null and alternative hypotheses.
- 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.

A Feasible Test for Unit Root Dimension

- To introduce our test, define (z_t^T) for $t = 1, \ldots, T$ by

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

- Also define the product sample moment $M_n^T = \sum_{t=1}^T z_t^T z_t^{T\prime}$ (sample variance in the unit root subspace), 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 $\Gamma_T(k) = T^{-1} \sum_t \Delta z_t^T \Delta z_{t-k}^{T\prime}$.

- Our test statistic is defined as

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

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.

Asymptotics for Distributional Unit Root Test

- Under very general conditions, we show that

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

  under the null, 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 as $T \to \infty$.

- Therefore, we reject the null in favor of the alternative if the test statistic τ_n^T takes small values.

Critical Values for Distributional Unit Root Test τ_n^T

- Critical values for the tests are obtained based on τ_n^T for $n = 1, \ldots, 5$, by simulations.
- For simulations, BM is approximated by standardized partial sum of mean zero i.i.d. normal random variates with sample size 10,000, and actual critical values are computed using 100,000 iterations.

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

Degree of Persistency in Cross-sectional Moments

- We may now find how much nonstationarity proportion exists in each cross-sectional moments.
- In what follows, we redefine ι_κ as $\iota_\kappa - \int_K \iota_\kappa(s)ds$, so that we may regard it as an element in H.
- 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 H such that $(v_i)_{1 \le i \le n}$ and $(v_i)_{i \ge n+1}$ span H_N and H_S, respectively.

Nonstationarity Proportion of Cross-sectional Moments

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

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

- $\pi_\kappa = 1$ 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.

Sample Nonstationarity Proportion

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

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

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

Empirical Illustrations

Overview

- We demonstrate how to define and estimate the state densities, and test for unit roots in the time series of densities representing cross-sectional distributions of economic variables.
- State densities are estimated using standard Gaussian kernel on cross-sectional observations, and their nonstationarities are analyzed using the test τ_T^n.
- Unit root dimension n of state densities is determined by applying τ_T^n successively downward starting from $n = n_{\max}$ with $n_{\max} = 5$.
- Unit root space H_N is then estimated and the unit root proportion (π_i) is computed for the first four moments. π_i provides the proportion of nonstationary fluctuation in the i-th moment of state distribution.

Representation of Functions as Numerical Vectors

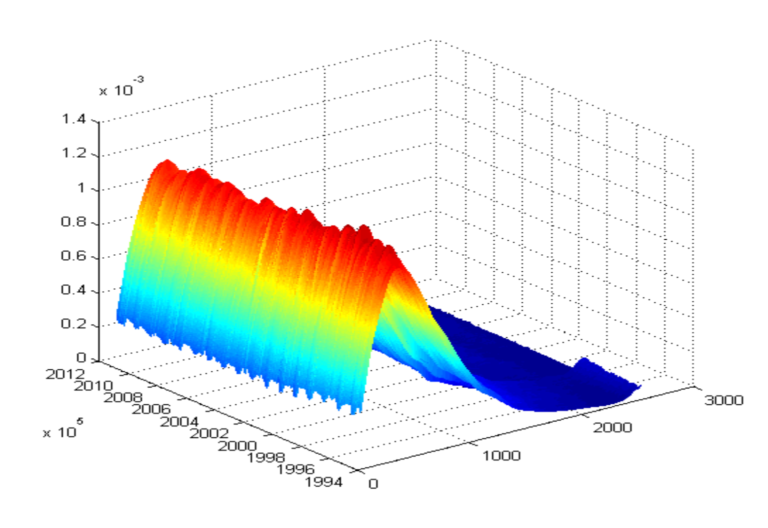
- For the representation of functions in Hilbert space as numerical vectors, we use a Daubechies wavelet basis.
- Wavelets are spatially varying orthonormal bases with two parameters, i.e., scale and translation, and hence they provide more flexibilities in fitting the state densities in our applications, some of which have severe asymmetry and time-varying support. The wavelet basis in general yields a much better fit than the trigonometric basis.
- The Daubechies wavelet is implemented with 1037 basis functions.

Cross-Sectional Distributions of Individual Earnings

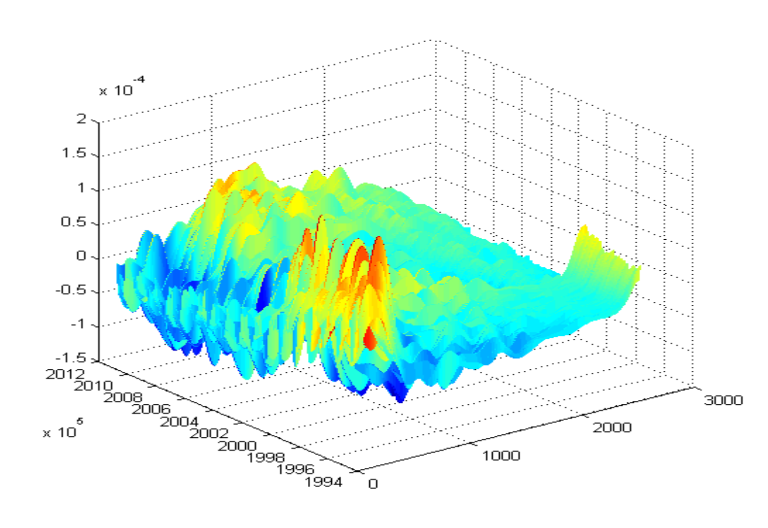
Cross-Sectional Distributions of Individual Earnings

- The cross-sectional observations of individual weekly earnings are obtained at monthly frequency from Current Population Survey (CPS) data set. The individual weekly earnings are deflated by consumer price index with base year 2005.
- The data set provides 204 time series observations spanning from January 1994 to December 2010, and the number of cross-sectional observations for each month ranges from 12,323 to 15,700.
- For confidentiality reasons, individual earnings are topcoded above a certain level. We drop all topcoded individual earnings as well as zero earnings as in Liu (2011) and Shin and Solon (2011).

Densities of Weekly Individual Earnings



Demeaned Densities of Weekly Individual Earnings



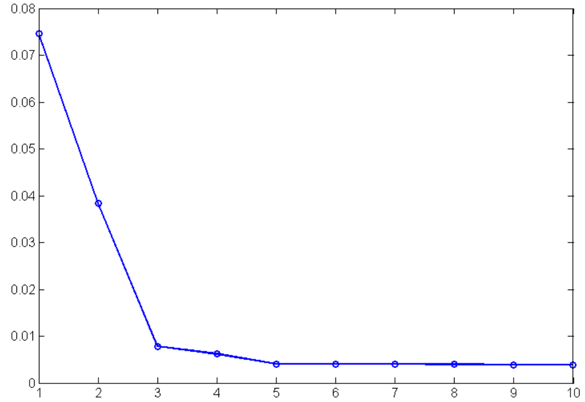
Testing for Unit Root Dimension

To determine the unit root dimension n in the time series of cross-sectional distributions of individual earnings, we use the feasible statistic τ_n^T to test for the null hypothesis $\mathsf{H}_0 : \dim(H_N) = n$ against the alternative $\mathsf{H}_1 : \dim(H_N) \leq n - 1$ with $n = 1, \ldots, 5$.

n	1	2	3	4	5
τ_n^T	0.0746	0.0383	0.0079	0.0062	0.0040

- 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 $n = 2$.
- We conclude that there exists two-dimensional unit root, and set $n = 2$.

Scree Plot of Eigenvalues - Individual Earnings



Unit Root Proportions in Moments

We compute the estimates π_i^T of the unit root proportions π_i with $\tau_n^T = 2$ for the first four moments.

π_1^T	π_2^T	π_3^T	π_4^T
0.5261	0.3420	0.2462	0.2013

- The unit root proportions for the first four moments are all nonnegligibly large. In particular, the unit root proportions for the first two moments are quite substantial.
- The presence of a substantial unit root proportion in the second moment explains the recent empirical findings on changes in volatilities of individual earnings. Dynan *et al* (2008) and others.
- Nonstationarity in time series of individual earnings distributions would certainly make their volatilities more persistent.

Global Temperature Distributions

Data

- Compiled by the Climatic Research Unit at the University of East Anglia and the Hadley Centre of the UK Met office.
- Global average of combined land and sea surface temperatures over widely dispersed locations, in a time series from 1850 to date (From 1,652 to 55,576 stations).
- Expressed as the deviation from the average of the period 1961-1990 and these deviations are called 'temperature anomalies'.
- Temperature anomalies on a 5° by 5° grid-box basis (number of Monthly grids : 36*72 = 2,592, number of Annual grids: 2,592*12 = 31,104)

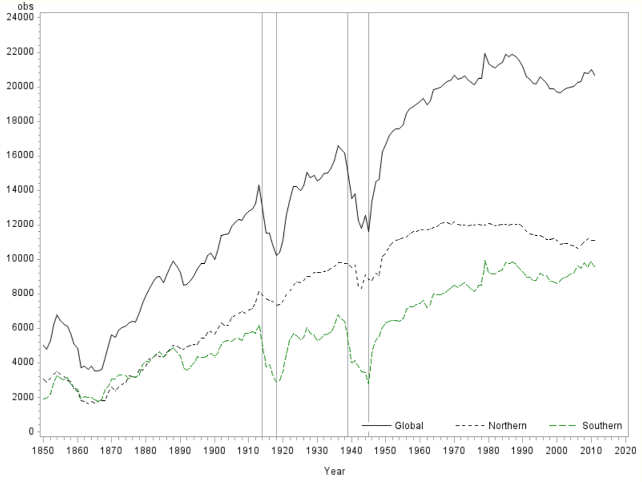


Figure: Number of observations for the globe, the northern hemisphere and the southern hemisphere; Total number of 5° by 5° grid-boxes is 31,104(=36*72*12) for the globe and 15,552(=18*72*12) for the northern and southern hemisphere.

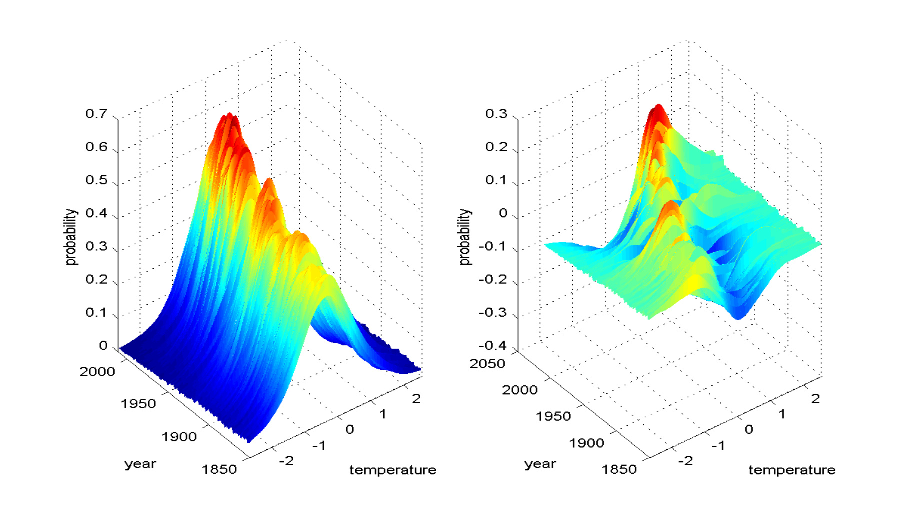


Figure: Annual temperature anomalies distributions (undemeaned and demeaned densities); Temperature anomalies on a 5° by 5° grid-box basis are used. A normal kernel with optimal fixed bandwidth is used for the estimation of the density functions.

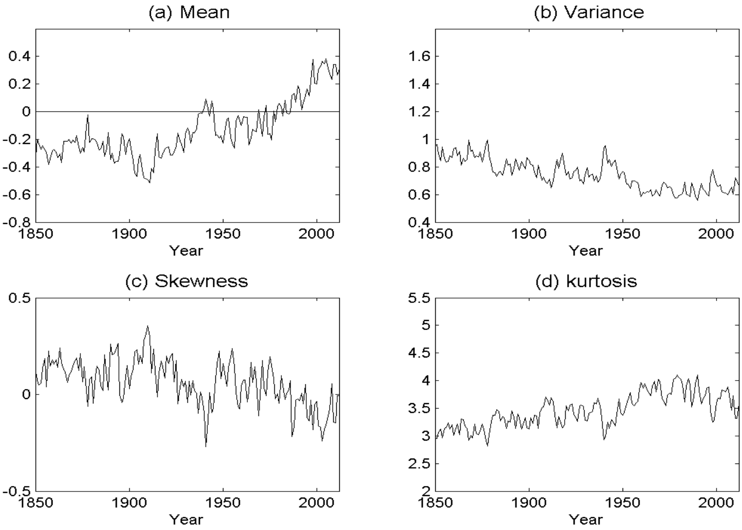


Figure: Mean, Variance, Skewness and Kurtosis for the globe

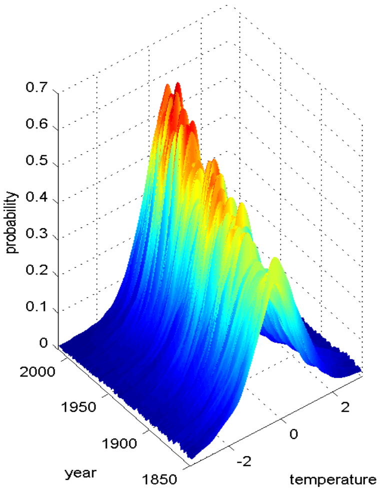


Figure: Annual temperature anomalies distributions for the northern hemisphere (undemeaned and demeaned densities)

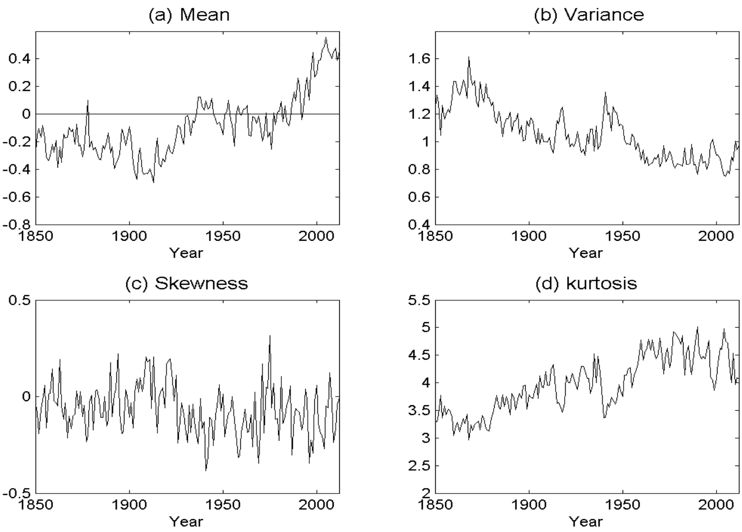


Figure: Mean, Variance, Skewness and Kurtosis for the northern hemisphere (Estimated annual temperature anomalies distributions are used.)

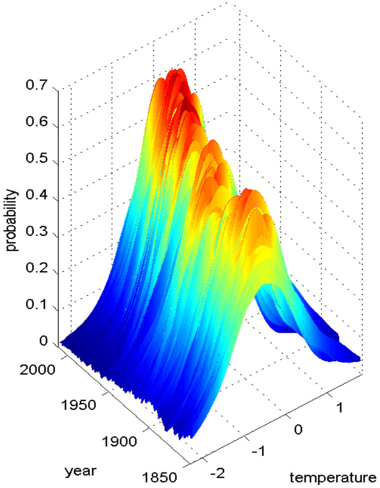


Figure: Annual temperature anomalies distributions for the southern hemisphere (undemeaned and demeaned densities)

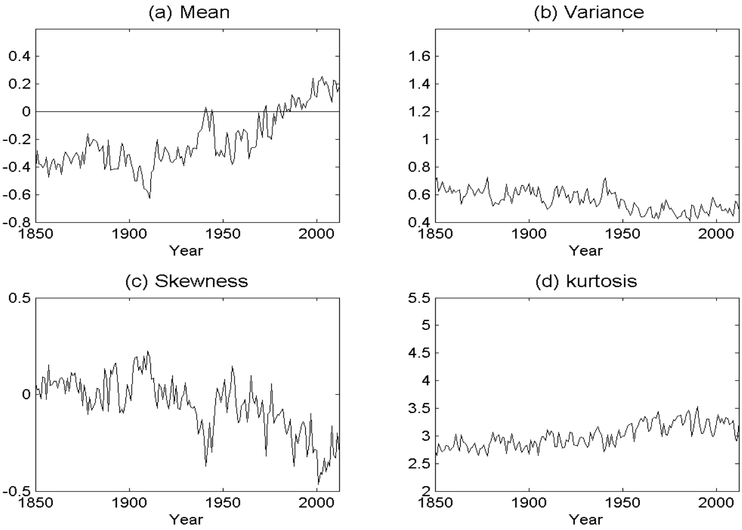


Figure: Mean, Variance, Skewness and Kurtosis for the southern hemisphere (Estimated annual temperature anomalies distributions are used.)

Findings

- We estimate annual temperature anomalies densities for the globe, northern and southern hemisphere, (f_t^g), (f_t^n) and (f_t^s).
- For the northern and the southern hemisphere, we choose the support that preserves 95% of the total probability mass of the northern$(\bar{f}^n)$ and the southern hemisphere$(\bar{f}^s)$.
- For the globe, we estimate total probability mass$(\bar{f}^g)$ as the average of $(\bar{f}^n)$ and $(\bar{f}^s)$ and find the support covering 95% of total probability mass for the globe$(\bar{f}^g)$.
- Densities for the globe, (f_t^g) is the average of the densities of northern and the southern, (f_t^n) and (f_t^s). This is so as not to give too much weight to the northern hemisphere, where there are more observations.

Table: Critical Values of the Test Statistics τ_n^T

$\tau_{n,1}^T$	n=1	n=2	n=3	n=4	n=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
$\tau_{n,2}^T$					
99%	0.7487	1.0073	1.2295	1.4078	1.5952
95%	0.4660	0.6787	0.8645	1.0336	1.1892
90%	0.3494	0.5399	0.7066	0.8574	1.0092

Table: Test Results for the Globe

(a) Values of Statistic for Testing n=m

	n=1	n=2	n=3	n=4		
$\tau_{n,1}^{T}$	0.0531	0.0289	0.0105	0.0097		
$\tau_{n,2}^{T}$	0.0531	0.0536				

(b) Critical Value

	n=1	n=2	n=3	n=4		
5%	0.0385	0.0223	0.0154	0.0127		
95%	0.4660	0.6787				

(c) Unit Root Proportions in First Seven Moments

$\hat{\pi}_1^T$	$\hat{\pi}_2^T$	$\hat{\pi}_3^T$	$\hat{\pi}_4^T$	$\hat{\pi}_5^T$	$\hat{\pi}_6^T$	$\hat{\pi}_7^T$
0.516	0.270	0.235	0.188	0.151	0.142	0.117

Table: Test Results for the Northern Hemisphere

(a) Values of Statistic for Testing n=m						
	n=1	n=2	n=3	n=4		
$\tau^T_{n,1}$	0.0387	0.0379	0.0119	0.0105		
$\tau^T_{n,2}$	0.0387	0.0407				
(b) Critical Value						
	n=1	n=2	n=3	n=4		
5%	0.0385	0.0223	0.0154	0.0127		
95%	0.4660	0.6787				
(c) Unit Root Proportions in First Seven Moments						
$\hat{\pi}^T_1$	$\hat{\pi}^T_2$	$\hat{\pi}^T_3$	$\hat{\pi}^T_4$	$\hat{\pi}^T_5$	$\hat{\pi}^T_6$	$\hat{\pi}^T_7$
0.409	0.205	0.160	0.129	0.101	0.094	0.078

Table: Test Results for the Southern Hemisphere

(a) Values of Statistic for Testing n=m

	n=1	n=2	n=3	n=4		
$\tau_{n,1}^T$	0.0611	0.0219	0.0097	0.0089		
$\tau_{n,2}^T$	0.0611					

(b) Critical Value

	n=1	n=2	n=3	n=4		
5%	0.0385	0.0223	0.0154	0.0127		
95%	0.4660					

(c) Unit Root Proportions in First Seven Moments

$\hat{\pi}_1^T$	$\hat{\pi}_2^T$	$\hat{\pi}_3^T$	$\hat{\pi}_4^T$	$\hat{\pi}_5^T$	$\hat{\pi}_6^T$	$\hat{\pi}_7^T$
0.633	0.199	0.331	0.168	0.212	0.140	0.157

Cross-sectional Distributions of Household Income and Consumption

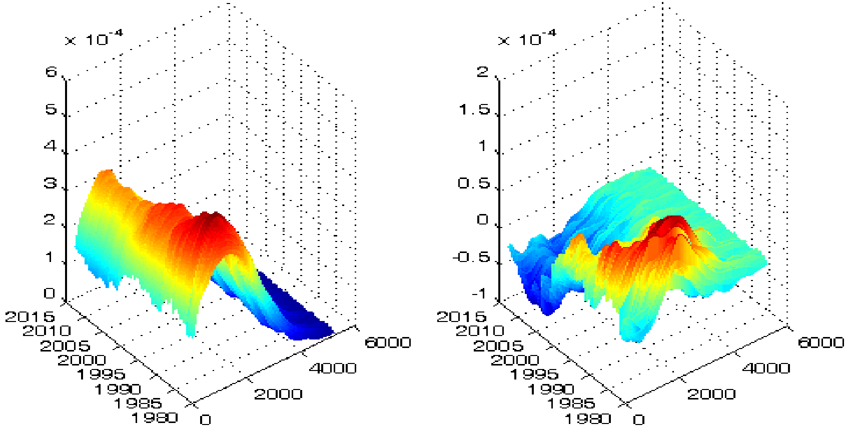
Data

- The cross-sectional observations of household income and consumptions are obtained at monthly frequency from Consumer Expenditure Survey (CES), collected for Bureau of Labor Statistics, US Census Bureau.
- CES consists of two surveys - Quarterly Interview Survey and Diary Survey, that provide information on buying habits, expenditures, income, and consumer unit (families and single consumers) characteristics. CES is the only Federal survey that provides the complete range of consumers' expenditures and incomes.
- CES data provide 400 time series obs from October 1979 to February 2013, with cross-sectional obs for each month ranging from 1,537 to 5,406.
- During this sample period, each household is included in the survey at most five times, and therefore the CE survey provides a pseudo panel data.

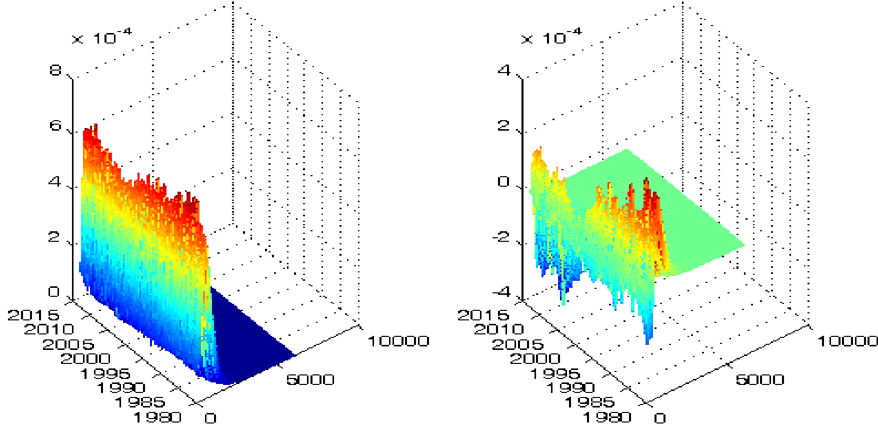
More on Data

- To construct monthly household income and consumption, we follow the definitions in Krueger and Perri (2006), and aggregate the monthly values provided in Universal Classification Code (UCC) level for each month and year.
- Nominal income and consumption values are deflated by monthly CPI provided by BLS for all urban households using a base year which varies among 1982, 1983 and 1984.
- The survey uses topcoding when original data exceeds some prescribed thresholds, which may change annually and be applied at a different starting point. We drop all top-coded values.
- We correct expenditure on food and impute services from vehicle and primary residence, according to the regressions specified in Krueger and Perri (2006). We also exclude obs with possible measurement error or inconsistency problem, following their sample selection criteria.

Cross-sectional Distributions of Household Income



Densities of Cross-sectional Distributions of Household Consumptions



Densities of Cross-sectional Distributions of Income and Consumption

- Figure 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.

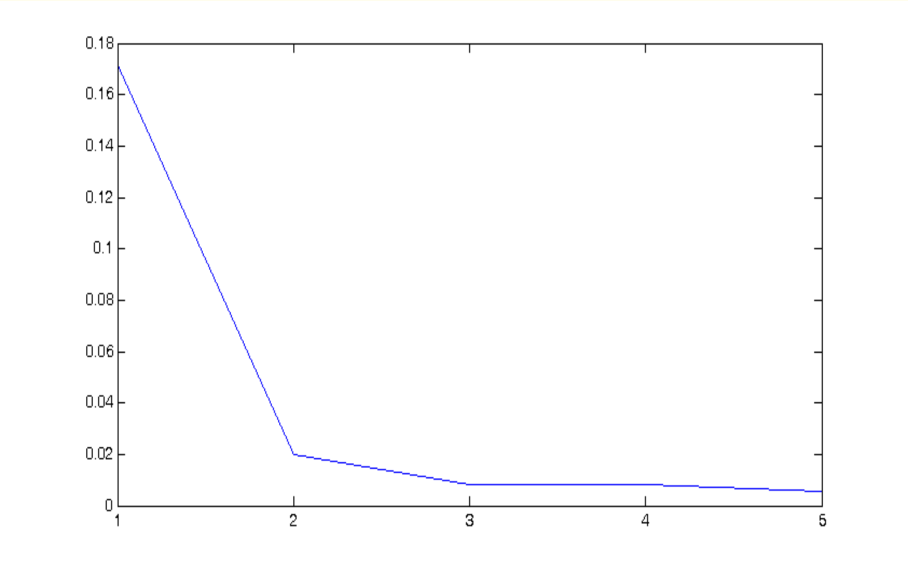
Nonstationarity in Income Dynamics

- 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 $n = 1, \ldots, 5$.

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 $p = 2$.
- Scree plot in the next slide 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.

Scree Plot for Time Series of Income Distributions



Unit Root Portion Estimates for Cross-sectional Moments of Household Income Distributions

- With $p = 2$, 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.

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

- 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 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.

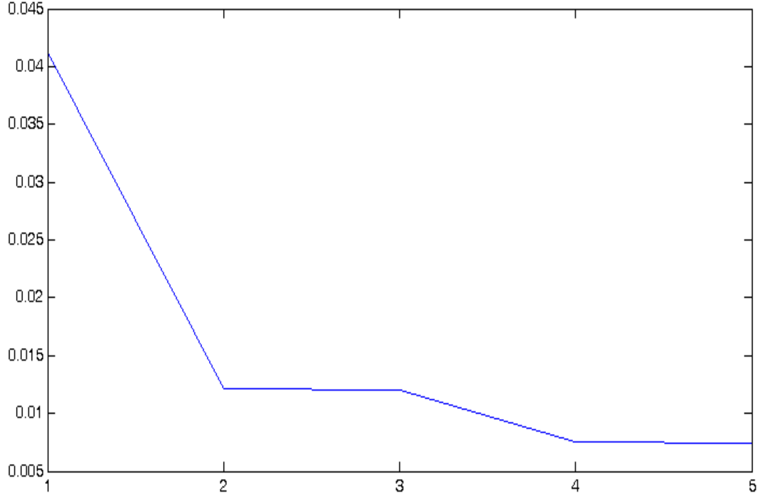
Nonstationarity in Consumption Dynamics

- 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 $\mathsf{H}_0 : q = n$ against the alternative hypothesis $\mathsf{H}_1 : q \leq n - 1$ with $n = 1, \ldots, 5$.

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 $q = 1$.
- Scree plot in the next slide shows that there is one leading principal component in the time series of household consumption distributions.

Scree Plot for Time Series of Consumption Distributions



Unit Root Portion Estimates for Cross-sectional Moments of Household Consumptions

- 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, with $q = 1$.

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.

III. Distributional Cointegration

Common Trends in Time Series of Cross-sectional Distributions

- Introduce a notion of distributional cointegration between two time series of densities representing cross-sectional distributions of some economic variables
- Explain how to estimate and test for such cointegrating relationships.

A New Framework

- To analyze time series of densities representing cross-sectional distributions allowing for unit root type of nonstationarity

- To analyze possible cointegration between cross-sectional distributions

- To learn and interpret both longrun and shortrun relationships between two time series of cross-sectional distributions

Model and Methodology

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 $\mathbb{R}$.
- 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.

Coordinate Processes

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

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

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

Cross-sectional Moments

- The coordinate processes of (f_t) and (g_t) in the direction of ι_κ, where

$$\iota_\kappa(s) = s^\kappa,$$

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 be referred subsequently to as the κ-th cross-sectional moments of (f_t) and (g_t) respectively.

Distributional Regression

- We consider the distributional regression

$$g_t = \mu + Af_t + e_t$$

  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.
- 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.
- We assume that (e_t) is stationary and mean zero, i.e., $\mathbb{E}e_t = 0$ for all $t = 1, \ldots, T$, and impose some exogeneity condition for (f_t) .

Coordinate Regression

- The coordinate regression of (g_t) in any direction $w \in H$ can be readily obtained from our distributional regression as

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

  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 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 $w \in H$.

More on Coordinate Regression

- The coordinate regression of (g_t) in any direction $w \in H$ is given as

$$\langle w, g_t \rangle = \langle w, \mu \rangle + \langle A^* w, f_t \rangle + \langle w, e_t \rangle$$

- The effect of the distribution represented by (f_t) on the coordinate process $(\langle w, g_t \rangle)$ 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 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 separately for stationary and nonstationary components of (f_t) and (g_t).

Regression in a Demeaned Form

- We may also consider the distributional regression in a demeaned form as

$$y_t = Ax_t + \varepsilon_t,$$

where

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

and $\varepsilon_t = e_t - T^{-1}\sum_{t=1}^{T} e_t$ for $t = 1, \ldots, T$.

- 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.
- We mainly deal with the demeaned densities (x_t) and (y_t) in our statistical analysis.

Demeaned Densities and Moment Functions

- We assume that the densities (f_t) and (g_t) all have supports included in a compact subset K of $\mathbb{R}$, for $t = 1, \ldots, T$.
- Then the demeaned densities (x_t) and (y_t) take values in

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

  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 $L_0^2(K)$.
- 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 $L_0^2(K)$.

Stationarity and Nonstationarity Subspaces

- 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.
- 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).
- Nonstationary subspaces F_N and G_N of (f_t) and (g_t) are defined as orthogonal complements of F_S and G_S, 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 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 $w \in G_N$.
- Nonstationarity subspaces F_N and G_N are assumed to be finite dimensional.

Estimation of Nonstationarity Subspace

- We 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.

Distributional Cointegration

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

$$\langle w, g_t \rangle = \pi + \langle v, f_t \rangle + u_t$$

with some constant π, where (u_t) is a general stationary process with mean zero.

Distributional Cointegrating Function

- 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 $j = 1, \ldots, q$. If the $(p+q)$-dimensional process (z_t) defined as

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

is cointegrated with the cointegrating vector

$$c = (-a_1, \ldots, -a_p, b_1, \ldots, b_q)',$$

then the distributional cointegration holds with

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

- The pair of functions v and w are called 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.

Longrun Response Function

- 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.

Possible Number of Cointegrating Relations

- Clearly, there are at most r-number of linearly independent distributional cointegrating relationships, $r \leq \min(p, q)$, between (f_t) and (g_t).
- Otherwise we would have a cointegrating vector c 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.
- This contradicts the assumption that $v_1, \ldots, v_p$ and $w_1, \ldots, w_q$ are linearly independent functions that span F_N and G_N respectively.

Distributional Cointegration

- The distributional cointegration does not presume any distributional regression relationship like $g_t = \mu + Af_t + e_t$. However, for two time series of densities (f_t) and (g_t) that are given by the above distributional regression model, we may easily deduce that

- We have

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

$$\langle w, g_t \rangle = \pi + \langle v, f_t \rangle + u_t$$

holds with $v = P_N A^* w$.

Longrun Response Function to Cross-sectional Moments

- If (f_t) and (g_t) are given by the distributional regression $g_t = \mu + Af_t + e_t$, then we have

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

- 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). Recall that (g_t^N) is the nonstationary component of (g_t).

Test for Distributional Cointegration

- 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

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

- Clearly, the test τ_n^T to determine the number of distributional unit roots may be used to test for the number of unit roots in (z_t), $z_t = \big(\langle v_1, x_t\rangle, \ldots, \langle v_p, x_t\rangle, \langle w_1, y_t\rangle, \ldots, \langle w_q, y_t\rangle\big)'$.
- The maximum number of unit roots for (z_t) is of course given by $p + q$ (no distributional cointegration in (f_t) and (g_t)).
- n-number of unit roots for (z_t) implies r-number of cointegrating relationships with $r = (p + q) - n$.

IV. Stationary Distributional Regression

Stationary Distributional Regression

- 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

$$g_t^S = \nu + B f_t^S + e_t,$$

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., $\mathbb{E} f_t^S \otimes e_t = 0$.

Stationary Distributional Regression

- If (f_t) and (g_t) are given by the distributional regression model $g_t = \mu + Af_t + e_t$ with some stationary (e_t), it follows immediately that the stationary distributional regression holds for (f_t^S) and (g_t^S).
- In fact, we have

Lemma Let (f_t) and (g_t) be given by the distributional regression model with some stationary (e_t). Then we have the stationary distributional regression

$$g_t^S = \nu + Bf_t^S + e_t$$

with the regression operator $B = Q_S A$.

Shortrun Response Function

- We may deduce from the stationary distributional regression that

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

  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).

- We will simply refer it as the shortrun response function of (f_t) to the κ-th cross-sectional moment of (g_t).

Demeaned Stationary Distributional Regression

- In parallel, we may write the stationary distributional regression

$$y_t^S = Bx_t^S + \varepsilon_t$$

  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).

- We use the demeaned stationary distributional regression to estimate B. For the consistent estimation of regression operator B and its asymptotic theory, see Park and Qian (2011).

- (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 $L_0^2(K)$, under the assumption that (f_t) and (g_t) have supports contained in a compact subset K of $\mathbb{R}$.

Inference on Stationary Distributional Regression

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

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

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

$$N_S = BM_S,$$

  which we may use to estimate B.

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

Ill-posed Inverse Problem

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

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

  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.

Restriction on Regression Operator

- 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

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

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

- We let

$$B_m = N_S M_{S_m}^{+},$$

  which is B restricted to F_{S_m} of F_S.

Estimation of Restricted Regression Operator

- The restricted regression operator B_m 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 $\lambda_1^T > \lambda_2^T > \cdots$.

Estimation of Restricted Regression Operator

- 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 $M_{S_m}^{+}$, and subsequently,

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

which we use as an estimator for the regression operator B.

- We show that the estimator B_m^T is consistent for B under very general conditions if we let $m \to \infty$ as $T\infty$ at a controlled rate.

Empirical IllustrationsInteractive Income-Consumption Dynamics

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.

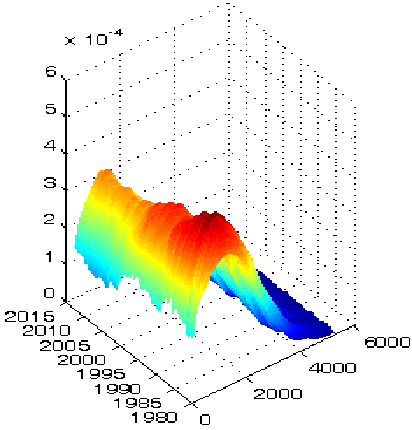
Data

- 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, 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 (BLS).
- 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.

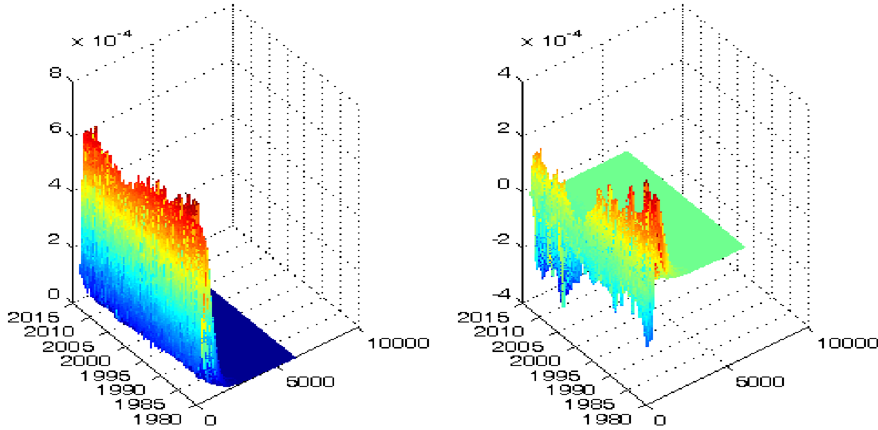
Data

- 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.
- We deflate the nominal income and consumption values by the monthly CPI provided by BLS for all urban households with using a base year which varies among 1982, 1983 and 1984.
- The survey uses topcoding to change the values when the original data exceeds some prescribed critical values, which may change annually and be applied at a different starting point. 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). We also exclude obs with possible measurement error or inconsistency problem.

Cross-sectional Distributions of Household Income



Densities of Cross-sectional Distributions of Household Consumptions



Interactive Dynamics of Income and Consumption

If

- the time series of income distributions has p unit roots
- the time series of consumption distributions has q unit roots
- there are r cointegrating relationships between them

Then, there are $(p+q)-r$ unit roots in their time series combined together.

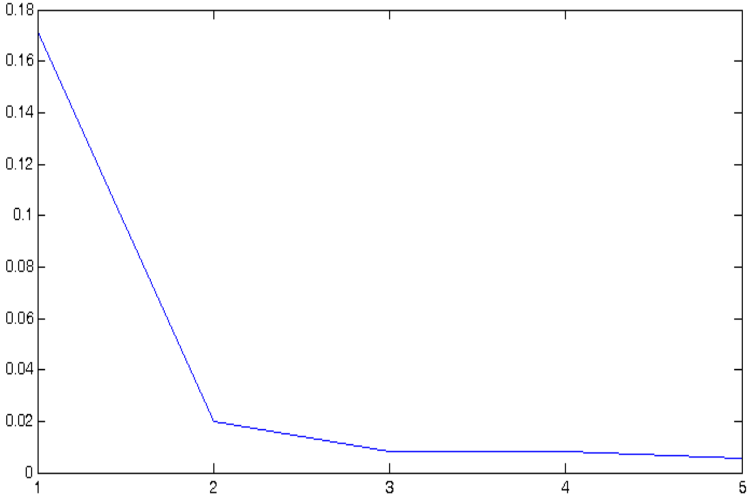
Test for Distributional Cointegration

- 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.

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 $q = 1$.

Scree Plot for Time Series of Income and Consumption Distributions



Cointegrating Function

- 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 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

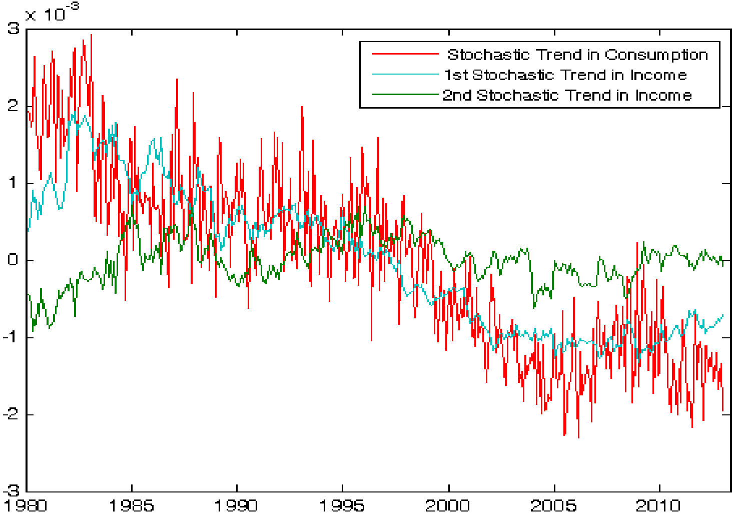
$$b\langle w, g_t\rangle = \delta + a_1\langle v_1, f_t\rangle + a_2\langle v_2, f_t\rangle + u_t$$

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

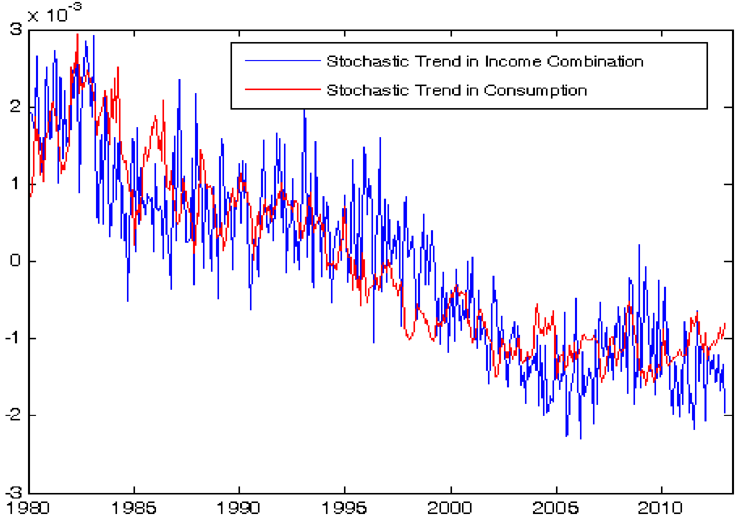
$$v_C = a_1 v_1 + a_2 v_2 \quad \text{and} \quad w_C = bw,$$

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

Integrated Coordinate Processes of Income and Consumption Distributions



Common Trends in Income and Consumption Distributions



Longrun Response Function of Income to Consumption

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

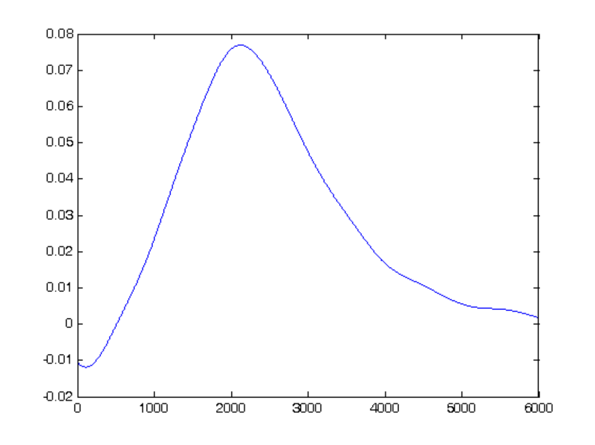
$$v_C^T = a_1^T v_1^T + a_2^T v_2^T \quad \text{and} \quad w_C = b^T w^T,$$

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.

- The estimates v_1^T, v_2^T and w^T are obtained 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 v_C^T.

Longrun Response Function of Income to Consumption



Longrun Response Function of Income to Consumption

- 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.

Note

- 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.
- 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.

Shortrun Response of Income Distribution to Consumption

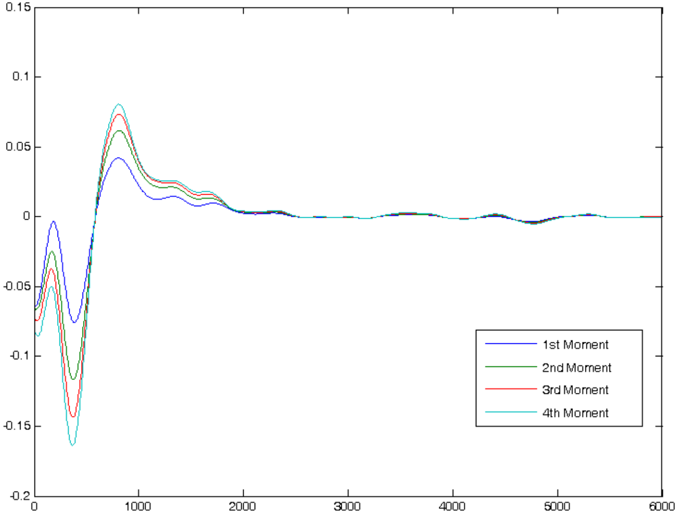
- To analyze the shortrun response of income distribution to consumption, we compute and plot the shortrun response function of income distribution to the κ-th cross-sectional moments of consumption distribution, which is introduced in (**??**), for $\kappa = 1, \ldots, 4$. It is given by

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

  where B_m^{T*} is the adjoint operator of B_m^T defined in (**??**).

- The estimated shortrun response function of income distribution to the cross-sectional moments of consumption distribution is given in the next slide.

Shortrun Income Response Function to Consumption



Shortrun Response of Income Distribution to Consumption

- 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.
- 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 income increases, and becomes almost entirely negligible once the income level exceeds $2,000 in monthly earnings.

Summary

- We develop 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, which provide the information about how one distribution affects the other.
- 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.
- 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.