

# Signal Extraction for Nonstationary Multivariate Time Series With An Illustration on Trend Inflation

Thomas M. Trimbur

Federal Reserve Board, Washington DC

with Tucker S. McElroy

Census Bureau, Washington DC

**Disclaimer** This report is released to inform interested parties of research and to encourage discussion. The views expressed on statistical issues are those of the author and not necessarily those of the U.S. Census Bureau or the Federal Reserve Board.

# 1 Introduction

- signal extraction widely used to focus on major movements in data (e.g., estimation of trend, long run component) and as input to policy-making
- multivariate case one of interest for signal extraction; for instance, central banks keep track of price movements in various sectors
- stochastic trends, nonstationarity, pervasive in economic data; also common trends, cointegration
- theory for single time series or for stationary multivariate case
- structure of filter design studied only for univariate case

## 2 Improved signal with multiple series

- Harvey and Trimbur (2003): Multivariate stochastic cycle model; empirical application to US business cycle, GDP-Investment
- Azevedo et. al (2006): applied MV higher order cycle model to Eurozone data
- Basistha and Startz (2008): Applied MV AR(2) [cyclical roots] model, common component, NAIRU
- Sinclair (2009) - Okun's Law pairing, GDP-URate model for US
- Kiley (2010): common trend (random walk) model for core and total inflation for US

### 3 Signal extraction research

Wiener (1949), Whittle (1963), Bell (1984): Wiener-Kolmogorov (WK) filters, bi-infinite benchmark, for multivariate stationary series or for nonstationary univariate series

McElroy (2008): Exact matrix formulas for nonstationary univariate case for actual finite-length series

Bi-infinite filters, generalization of WK, not yet derived for multivariate nonstationary case

Finite-series formulas not yet presented for the case of multivariate series, either stationary or nonstationary

## 4 Aims

- generalize the Wiener-Kolmogorov formula (bi-infinite series) to multiple nonstationary series, both under very general conditions and with the uniform structure usually used in the literature
- new formulas for actual data (finite series) for multivariate case (stationary or nonstationary), exact everywhere including near end of sample, simple and fast computation
- Treatment of Multivariate trends: related or common; examine some major generalized filters
- Illustrate with analysis of trend inflation using both core and total inflation data

## 5 Multivariate Filters

$y_t$  is  $N$  element vector series,  $t = -\infty, \dots, \infty$

Apply a filter  $F(L) = \sum_{j=-\infty}^{\infty} W_j L^j$  to  $y_t$  to estimate signal

$$\hat{z}_t = F(L)y_t = \sum_{j=-\infty}^{\infty} W_j L^j y_t = \sum_{j=-\infty}^{\infty} W_j y_{t-j}$$

$W_j$  is the  $N \times N$  matrix of coefficients for lag  $j$

$$z_t^{(I)} = \sum_{J=1}^N \sum_{j=-\infty}^{\infty} w_j^{(IJ)} y_{t-j}^{(J)}, \quad w_j^{(IJ)} \text{ applied to } J, \text{ at lag/lead } j.$$

## 6 Multivariate filters - frequency domain

Frequency Response - F.T. of  $W_j$ 's:  $\mathbf{FR}(\lambda) = \mathbf{F}(e^{-i\lambda})$

Gain function: how filter affects amplitudes at each frequency

Example: Low-pass filters for trends cut out high frequencies

$\mathbf{G}(\lambda) [N \times N]$ ,  $G_{IJ}(\lambda) = FR_{IJ}(\lambda)$  for  $\text{Im}(FR_{ij}(\lambda)) = 0$

For symmetric filters (past and present equally weighted for a given time separation), zero phase shift and  $\mathbf{G}(\lambda) = \mathbf{FR}(\lambda)$

## 7 Filters and Decomposition Models

Reason for filtering: series has a signal of interest,  $s_t$ , along with other, less informative, components - the noise,  $\mathbf{n}_t$

S-N decomposition model:  $\mathbf{y}_t = s_t + \mathbf{n}_t$

$s_t$  one component, e.g., trend, or combination such as trend plus cycle (regular movements)

$\mathbf{n}_t$  - remaining component(s), e.g., irregular or temporary part



## 8 Multivariate (MV) UC models

- intuitive form; flexible structure
- param. estimates, statistical fit and diagnostics
- close look at important individual series
- extracted components consistent with each other and with data; measure uncertainty in these estimates
- detailed correlation structure (components)

# 9 Theoretical Foundations for Signal Extraction (Optimal filters for doubly infinite Series)

long-term impact - abstracts from near-endpoint effects

approximately holds for most time points in a long enough series

uncorrelated components  $\rightarrow$  time-symmetric filters

compact expressions for filters and gain functions

essential properties and comparisons across filters

## 10 Existing Signal Extraction Formulas (I)

Matrix autocovariance generating function (ACGF)

$$\text{For stationary } \mathbf{x}_t: \Gamma_{\mathbf{x}}(L) = \sum_{j=-\infty}^{\infty} \Gamma_j L^j,$$

where  $\Gamma_j = E(\mathbf{x}_t \mathbf{x}_{t-j}')$  is auto-covariance matrix for lag  $j$ .

$$\text{Optimal filter: } F_{WK}(L) = \Gamma_s(L) [\Gamma_s(L) + \Gamma_n(L)]^{-1}$$

For Gaussian components, minimum mean-squared error (MMSE);

For MV white noise (WN) disturbances, best linear estimator

# 11 Pseudo-acgf

Univariate ARMA:  $\phi(L)x_t = \theta(L)\kappa_t$ ,  $\kappa_t \sim WN(0, \sigma_\kappa^2)$

$\phi(L)$ : AR polynomial,  $\theta(L)$ : MA polynomial, .

$$\text{ACGF: } \gamma_x(L) = \sigma_\kappa^2 \frac{\theta(L)\theta(L^{-1})}{\delta(L)\delta(L^{-1})\phi(L)\phi(L^{-1})},$$

Univariate ARIMA:  $\phi(L)\delta(L)x_t = \theta(L)\kappa_t$ , *eg*,  $\delta(L) = 1 - L$

$$\text{Pseudo-ACGF: } \gamma_x(L) = \frac{\theta(L)\theta'(L^{-1})}{\delta(L)\delta(L^{-1})\phi(L)\phi(L^{-1})}\sigma_\kappa^2,$$

## 12 Existing Signal Extraction Formulas (II)

For univariate nonstationary  $s_t$ ,  $n_t$  (with distinct nonstationary operators), use Bell's extension of WK formula:

$$\text{Optimal filter: } F_{WK}(L) = \gamma_s(L)[\gamma_s(L) + \gamma_n(L)]^{-1}$$

$$\text{Gain: } G(\lambda) = \gamma_s(e^{-i\lambda})[\gamma_s(e^{-i\lambda}) + \gamma_n(e^{-i\lambda})]^{-1}$$

Here  $\gamma_s(e^{-i\lambda})$  is a "pseudo-"spectrum - it doesn't exist at a nonstationary frequency, but a cancellation occurs for the gain function

## 13 Multivariate Nonstationary Case

$$s_t^{(j)}, n_t^{(j)}, y_t^{(j)} \quad j = 1, \dots, N$$

$$\{u_t^{(j)}\} = \{\delta_s^{(j)}(L)s_t^{(j)}\}, \quad \{v_t^{(j)}\} = \{\delta_n^{(j)}(L)n_t^{(j)}\},$$

$$\{w_t^{(j)}\} = \{\delta^{(j)}(L)y_t^{(j)}\}$$

Nonstationary operators,  $\delta_s^{(j)}(L) \neq \delta_n^{(j)}(L)$  for each  $j$

Let  $\{\mathbf{u}_t\} = \{u_t^{(1)}, \dots, u_t^{(N)}\}$ , likewise for  $\{\mathbf{v}_t\}$  and  $\{\mathbf{w}_t\}$ .

## 14 Multivariate Nonstationary Case

Let  $\delta_s^{(j)}(L), \delta_n^{(j)}(L)$  be uniform across  $j$  :  $\delta_s^{(j)}(L) = \delta_s(L)$  and  $\delta_n^{(j)}(L) = \delta_n(L)$  for all  $j$

Let  $\tilde{\delta}_s(L)$  be a diagonal matrix with elements  $\delta_s(L)$ .

Then,  $\tilde{\delta}_s(L)\mathbf{s}_t = \mathbf{u}_t$ , and for analogous  $\tilde{\delta}_n(L), \tilde{\delta}_n(L)\mathbf{n}_t = \mathbf{v}_t$ ,

Let  $\mathbf{F}_w$  denote the multivariate spectrum of  $\{\mathbf{w}_t\}$ . Likewise, define  $\mathbf{F}_u$  as the spectrum of  $\{\mathbf{u}_t\}$ , and  $\mathbf{F}_v$  as the spectrum of  $\mathbf{v}_t$ .

## 15 Theorem 1 - Assumptions

$\{u_t\}, \{v_t\}$  stationary and uncorrelated

Assumption  $M_\infty$ : Initial values  $y_*^{(j)}$  uncorrelated with  $\{u_t\}, \{v_t\}$

Assumption:  $F_w$  is invertible everywhere, except for a set of frequencies that has Lebesgue measure zero.



## 16 New Theorem for multivariate nonstationary signal extraction (bi-infinite series)

Define Pseudo-ACGF of  $\mathbf{s}_t$ :  $\Gamma_s(L) = \Gamma_u(L) [\tilde{\delta}_s(L)\tilde{\delta}_s(L^{-1})]^{-1}$

Define Pseudo-ACGF of  $\mathbf{n}_t$ :  $\Gamma_n(L) = \Gamma_v(L) [\tilde{\delta}_n(L)\tilde{\delta}_n(L^{-1})]^{-1}$

**Theorem 1** *The generalized optimal filter for nonstationary  $\mathbf{s}_t$  and  $\mathbf{n}_t$  defined above is given by*

$$F_{WK}(L) = \Gamma_s(L)[\Gamma_s(L) + \Gamma_n(L)]^{-1}$$

# 17 Optimal filters for multivariate nonstationary signal extraction

Gain:  $F_{WK}(e^{-i\lambda}) = \Gamma_s(e^{-i\lambda})[\Gamma_s(e^{-i\lambda}) + \Gamma_n(e^{-i\lambda})]^{-1}$

Depends on individual dynamic properties (variances of component disturbances) and on relationships (component cross-correlations)

Wold decomposition:  $\delta_s(L)s_t = \Xi(L)\zeta_t, \quad \zeta_t \sim WN(\mathbf{0}, \Sigma_\zeta)$

Covariance matrix  $\Sigma_\zeta$  ( $N \times N$ ) can have full (related signals) or reduced rank (common signals, cointegration)

## 18 Real Data: Finite-series setup

$y^{(I)} = [y_1^{(I)}, y_2^{(I)}, \dots, y_T^{(I)}]'$ , Likewise,  $s^{(I)}$  and  $n^{(I)}$

Seek estimator  $\hat{s}^{(I)}$  that minimizes the MSE criterion:

$$\hat{s}^{(I)} = \sum_{J=1}^N F^{IJ} y^{(J)} = E[s^{(I)} | y^{(1)}, y^{(2)}, \dots, y^{(N)}]$$

$F^{IJ}$  is  $T \times T$  dimensional,  $J \neq K$  cross filters,  $J = K$  own filters

Derive the set of  $F^{IJ}$ s from properties of  $s^{(I)}$  and  $n^{(I)}$

## 19 Stochastic trends

Many series are nonstationary and have a stochastic trend

Set  $\mathbf{s}_t = \boldsymbol{\mu}_t$ , signals long-run patterns (low-frequency movements)

Examples: trend inflation, potential output, NAIRU

Set  $\mathbf{n}_t = \boldsymbol{\varepsilon}_t$ , stationary irregular, e.g.,  $\boldsymbol{\varepsilon}_t \sim WN(\mathbf{0}, \boldsymbol{\Sigma}_\varepsilon)$

$$\mathbf{y}_t = \mathbf{s}_t + \mathbf{n}_t = \boldsymbol{\mu}_t + \boldsymbol{\varepsilon}_t,$$

## 20 Local Level Model

$$\mu_t = \mu_{t-1} + \eta_t, \quad \eta_t \sim WN(\mathbf{0}, \Sigma_\eta)$$

Random walk, simplest I(1) trend

$\text{Cov}(\eta_t, \varepsilon_t) = \mathbf{0} [N \times N]$ , trends driven by different types of factors than noise, leads to symmetric filters

Random, permanent shocks each period  $\rightarrow$  trend stays flat, on average, moving forward; frequent small changes in direction

## 21 Smooth trend

$$\mu_t = \mu_{t-1} + \beta_{t-1},$$

$$\beta_t = \beta_{t-1} + \zeta_t, \quad \zeta_t \sim WN(\mathbf{0}, \Sigma_\zeta), \quad \text{Cov}(\zeta_t, \varepsilon_t) = \mathbf{0} [N \times N]$$

time-varying slopes  $\beta_t$ , allow for permanent variation in growth rate of underlying signal

Usually, small diagonal values for  $\Sigma_\zeta \rightarrow$  estimated trends change direction slowly - "smooth"

## 22 Common Trends

$$\boldsymbol{\mu}_t = \boldsymbol{\mu}_{t-1} + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim WN(\mathbf{0}, \boldsymbol{\Sigma}_\eta)$$

Reduced rank  $\boldsymbol{\Sigma}_\eta$ :  $\boldsymbol{\mu}_t$  depends on  $K < N$  core trends, in  $\boldsymbol{\mu}_t^\dagger$

$$\boldsymbol{\mu}_t = \boldsymbol{\Theta} \boldsymbol{\mu}_t^\dagger + \boldsymbol{\mu}_0^\dagger$$

$$\boldsymbol{\mu}_t^\dagger = \boldsymbol{\mu}_{t-1}^\dagger + \boldsymbol{\eta}_t^\dagger$$

$\boldsymbol{\Theta}$  ( $N \times K$ ) contains the load factors, restrictions used for identification;  $\boldsymbol{\mu}_0^\dagger$  ( $K \times 1$ ) vector of constants

## 23 MV Butterworth low-pass filters

Filter, Time domain:  $\Psi(L) = G_\mu(L)[G_\mu(L) + G_\varepsilon(L)]^{-1}$

Gain based on spectra:  $\Psi(\lambda) = F_\mu(\lambda)[F_\mu(\lambda) + F_\varepsilon(\lambda)]^{-1}$

$$\Psi(\lambda) = \Sigma_\zeta (\Sigma_\zeta + (2 - 2 \cos \lambda)^m \Sigma_\varepsilon)^{-1}, \quad m = 1, 2, \dots$$

Full rank  $\Sigma_\zeta$  (irreducible related trends)  $\rightarrow \Psi(\lambda)$  well-defined and continuous everywhere for  $\lambda$  on  $[0, \pi]$

Separation of gains at lowest frequency:  $\Psi(0) = I_N$



## 24 Common Trends Low Pass

Reduced rank  $\Sigma_{\zeta}$  (common trends)  $\rightarrow \Psi(\lambda)$  well-defined and continuous everywhere on  $[0, \pi]$  except for  $\lambda = 0$

Common trends:  $\Sigma_{\zeta} = \Theta \Sigma_{\zeta}^{\dagger} (\Theta)'$

$$\Psi(0) = \lim_{\lambda \rightarrow 0} \Psi(\lambda) = \Theta \left( \Theta' \Sigma_{\epsilon}^{-1} \Theta \right)^{-1} \Theta' \Sigma_{\epsilon}^{-1}$$

Sharing of gains:  $\Theta = \iota$  (column vector of ones),  $\Sigma_{\epsilon} = \sigma_{\epsilon}^2 I_N \rightarrow$

$$\Psi(0) = \iota \iota' / N \text{ (equal weights on each input series)}$$

## 25 Application: Total Trend Inflation with Core

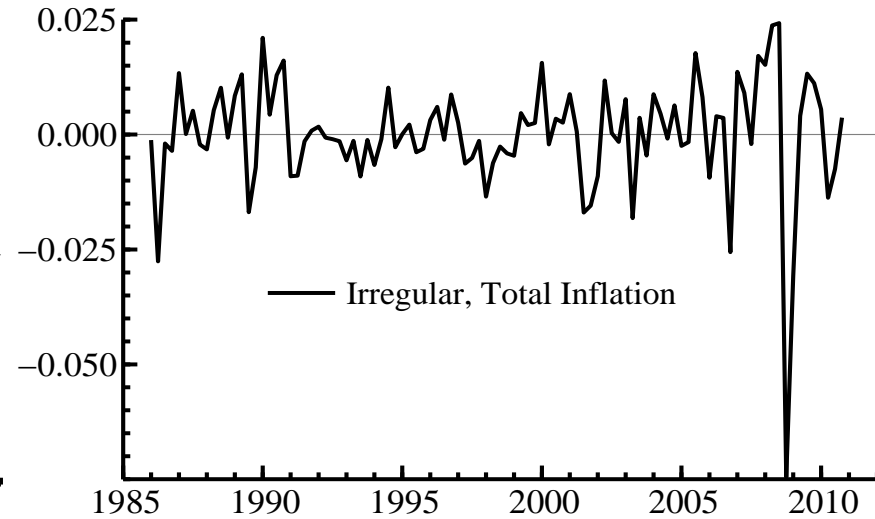
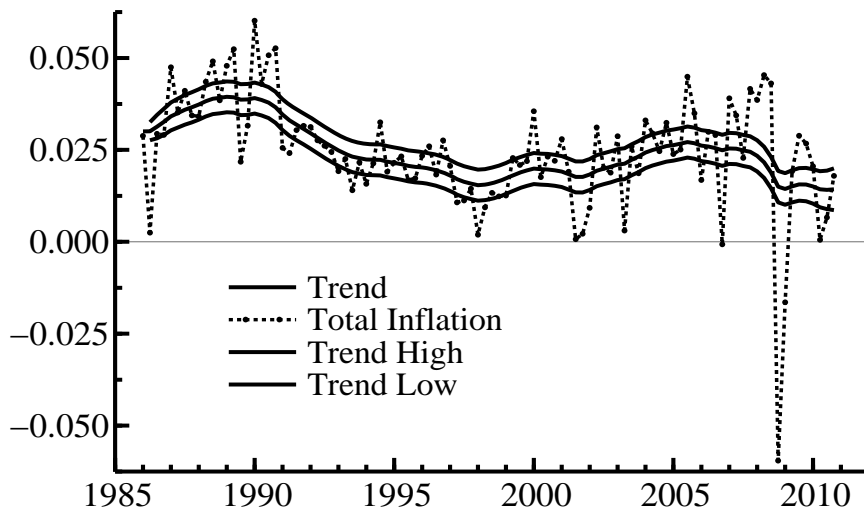
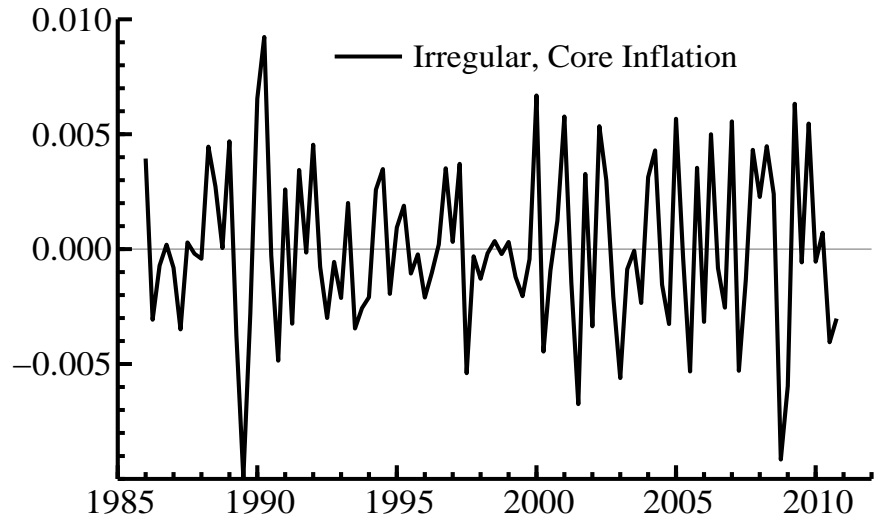
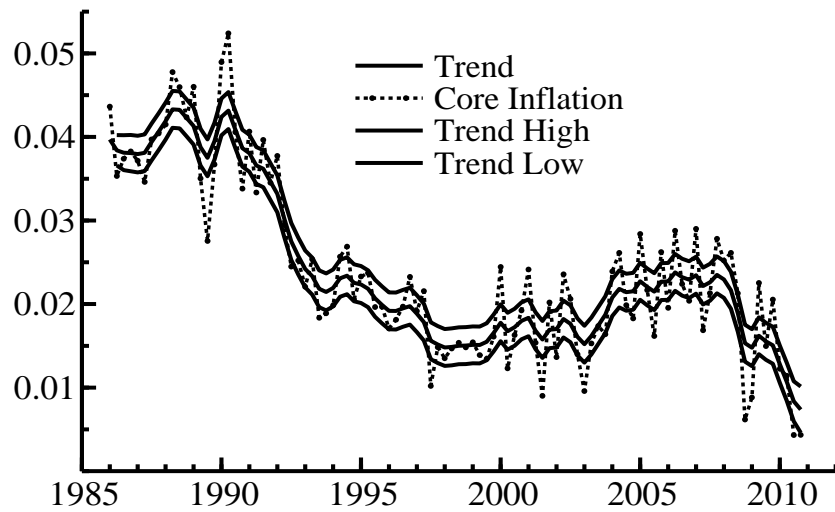
Example : trend in total inflation (additional information from core: has volatile food and energy components stripped out)

Data: Quarterly PCE inflation, total and core

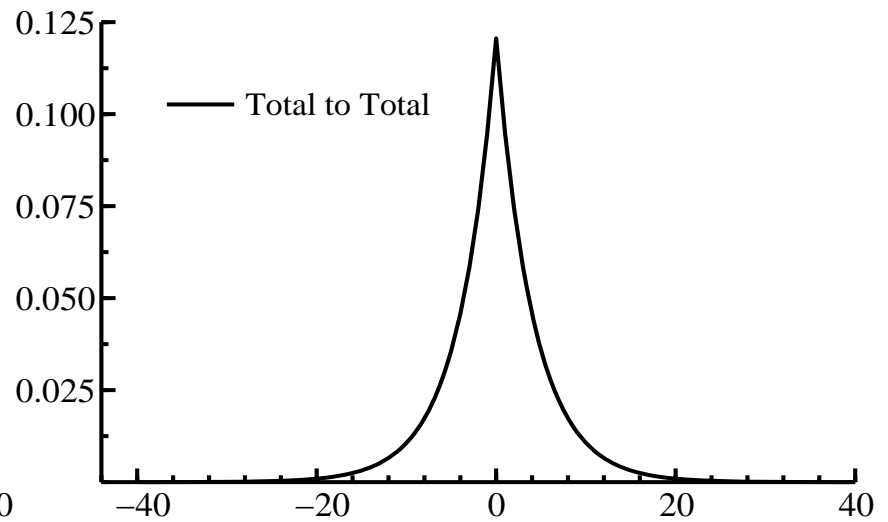
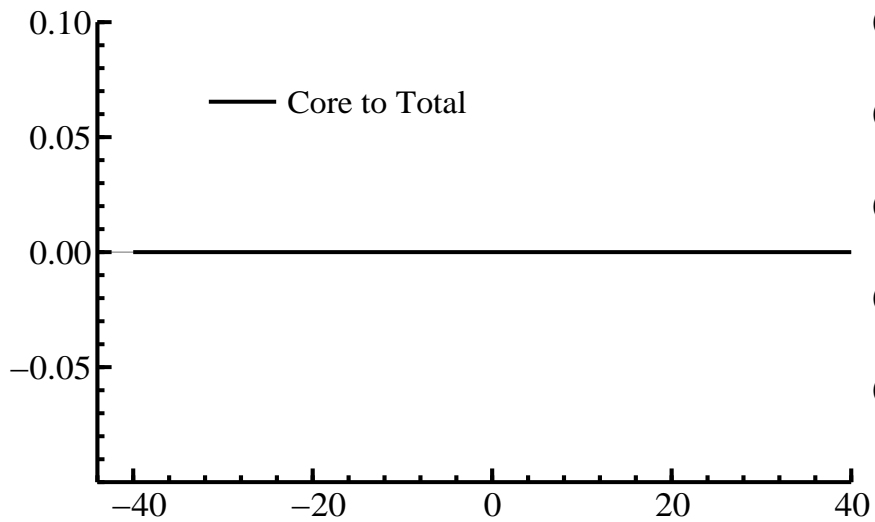
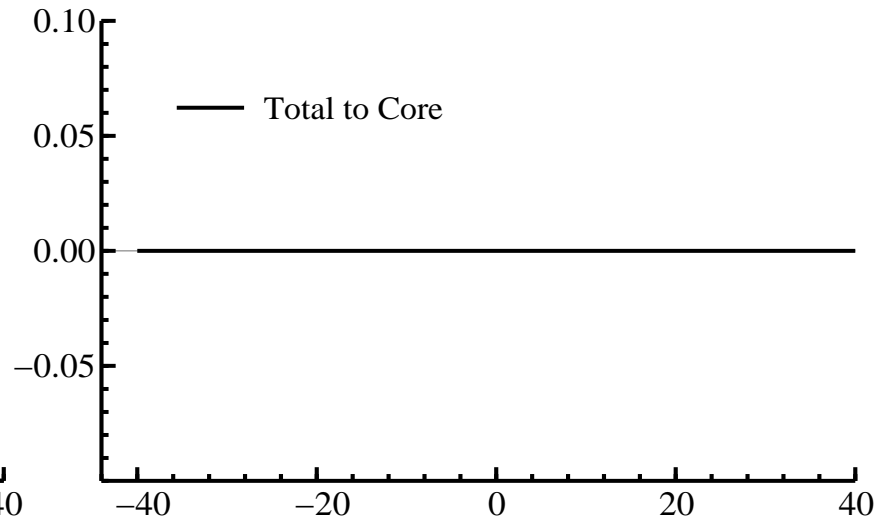
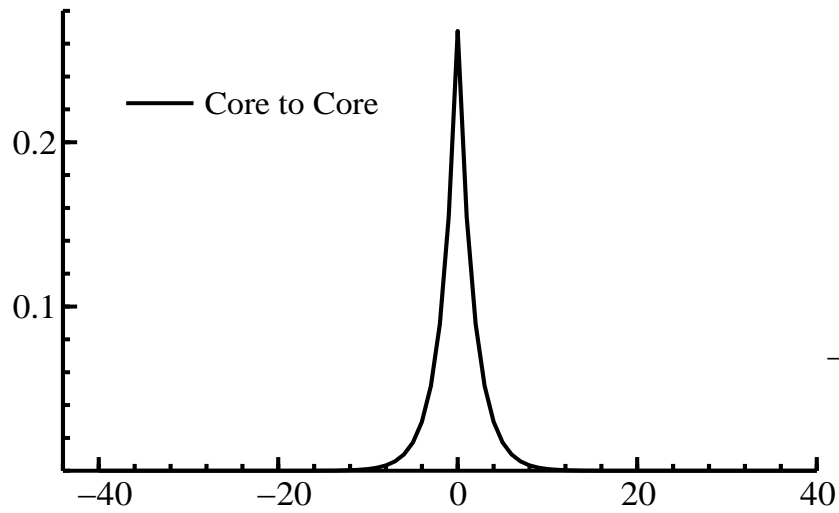
Sample: 1986:1 to 2010:4

Source: Bureau of Economic Analysis (Vintage: 2011Q1)

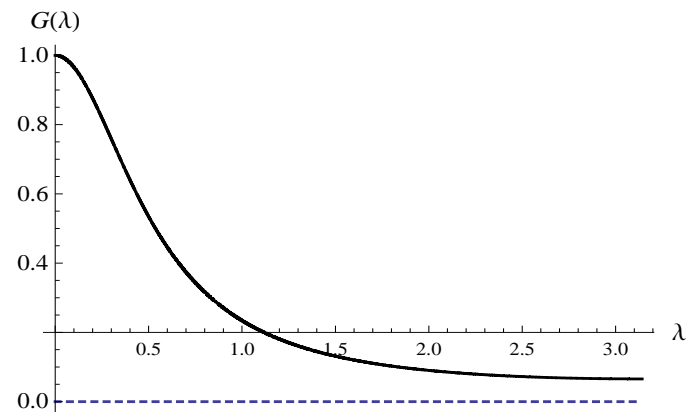
# Local Level Model - Univariate, Trends



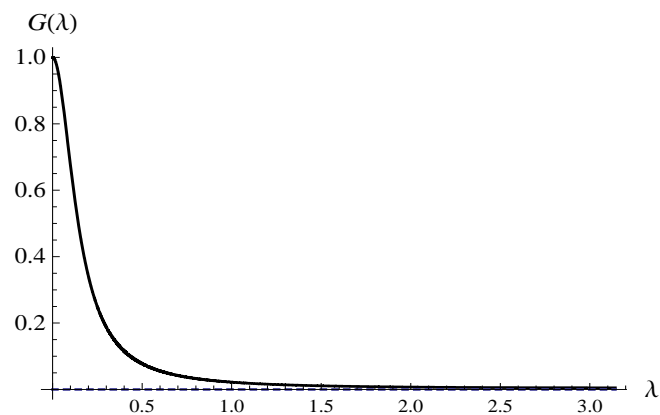
# Local Level Model - Univariate, Weights



# Local Level Model - Univariate, Gains

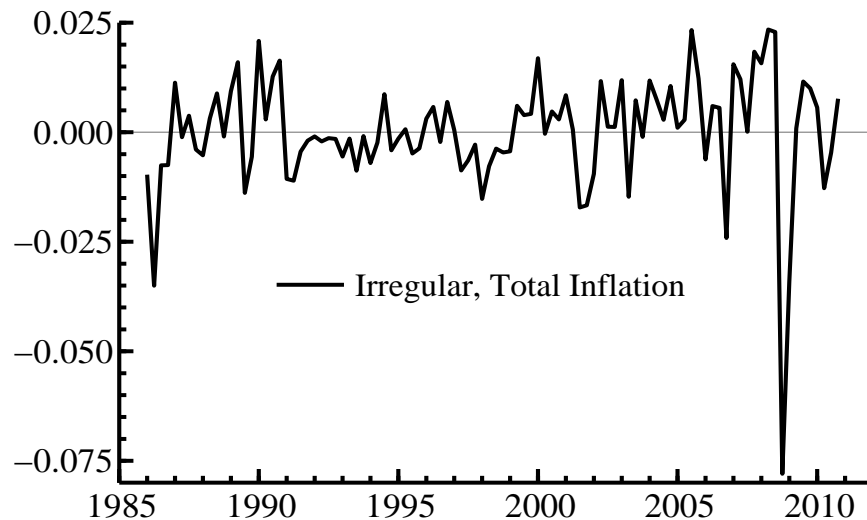
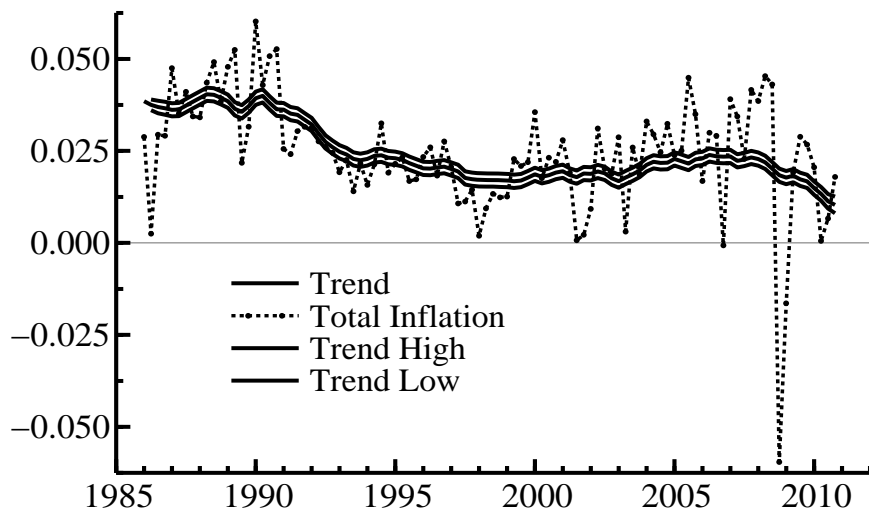
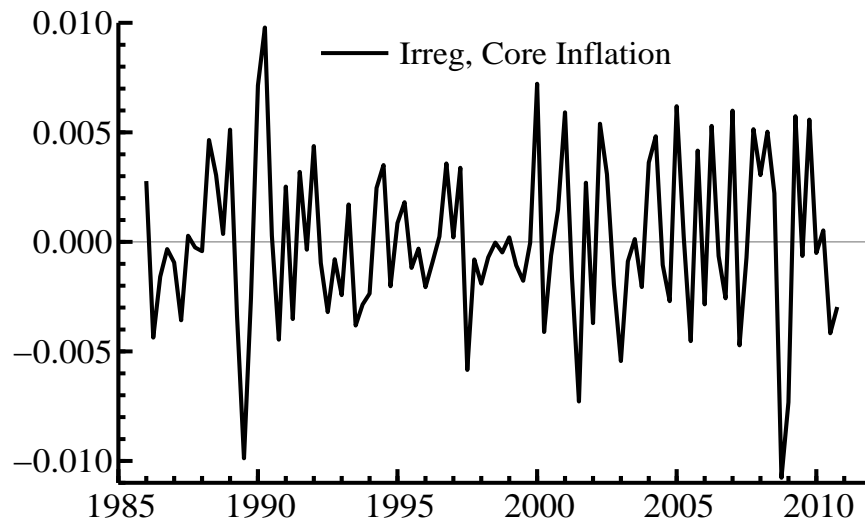
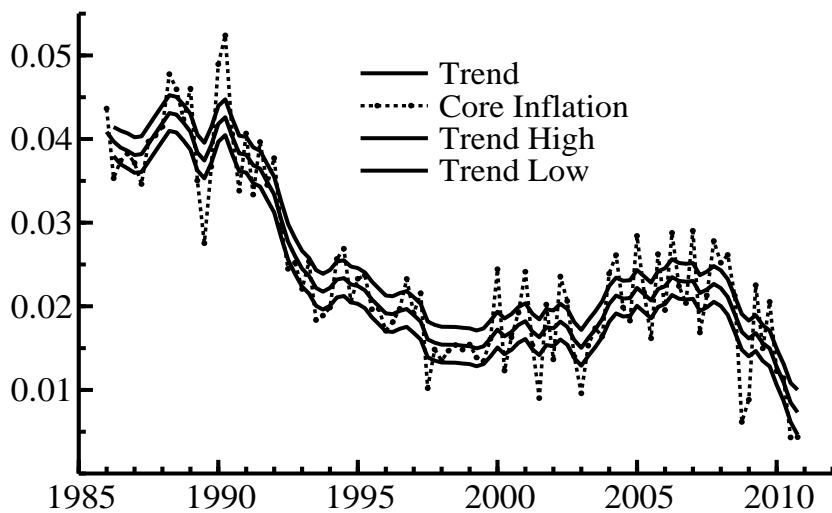


Solid: Core to Core. Dashed: Total to Core.

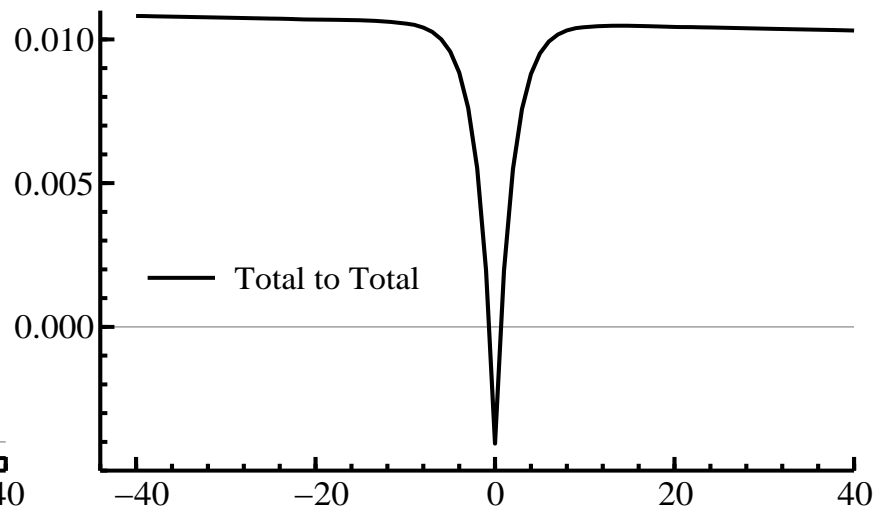
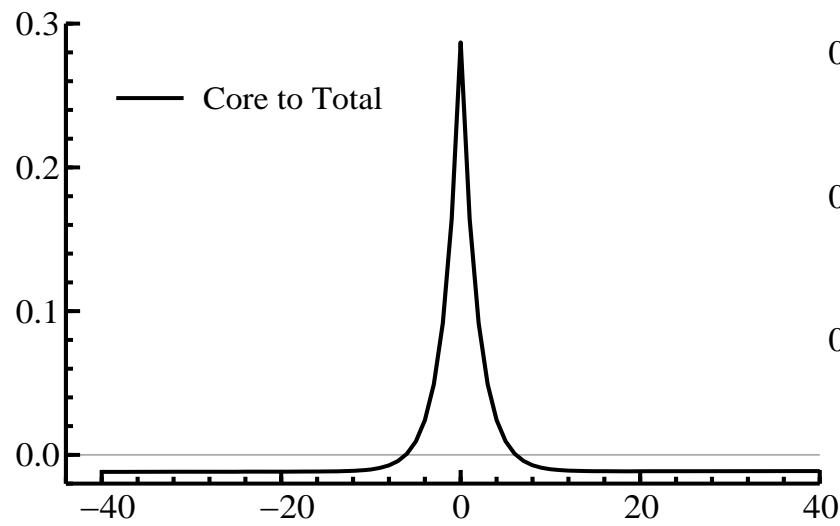
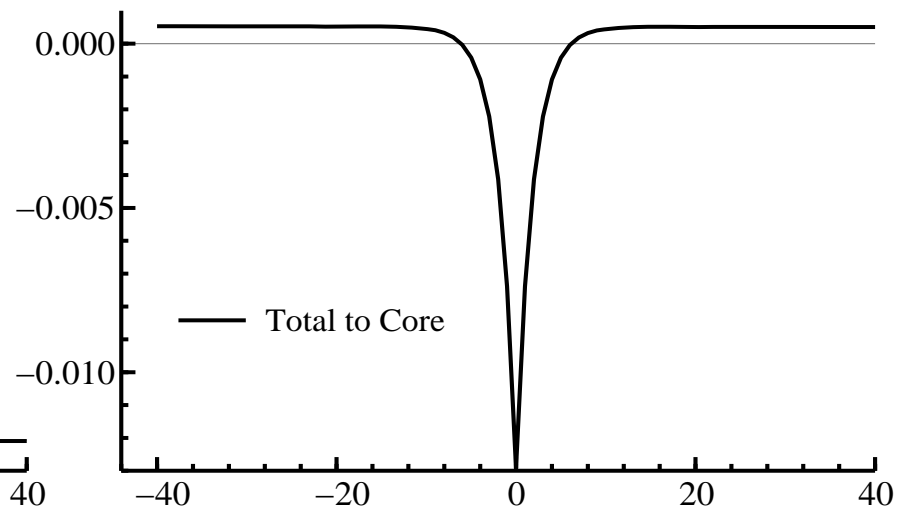
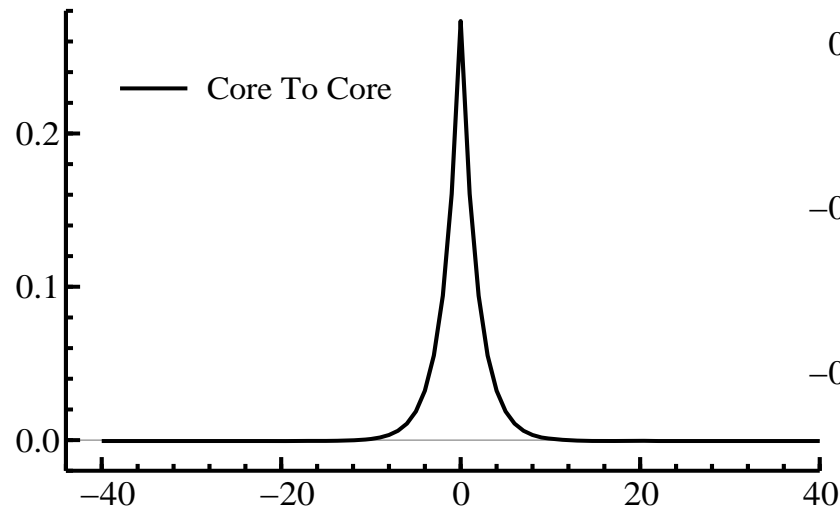


Solid: Total to Total. Dashed: Core to Total.

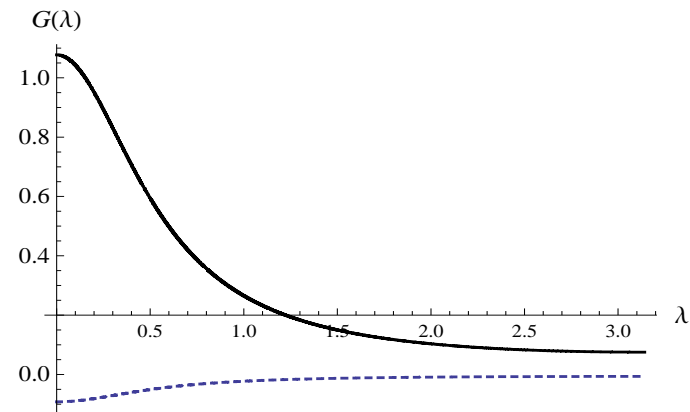
# Local Level Model - Bivariate, Trends



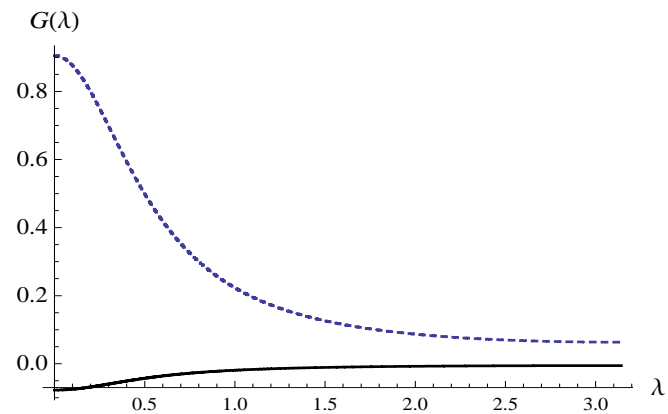
# Local Level Model - Bivariate, Weight patterns



## Local Level Model - Bivariate, Gains



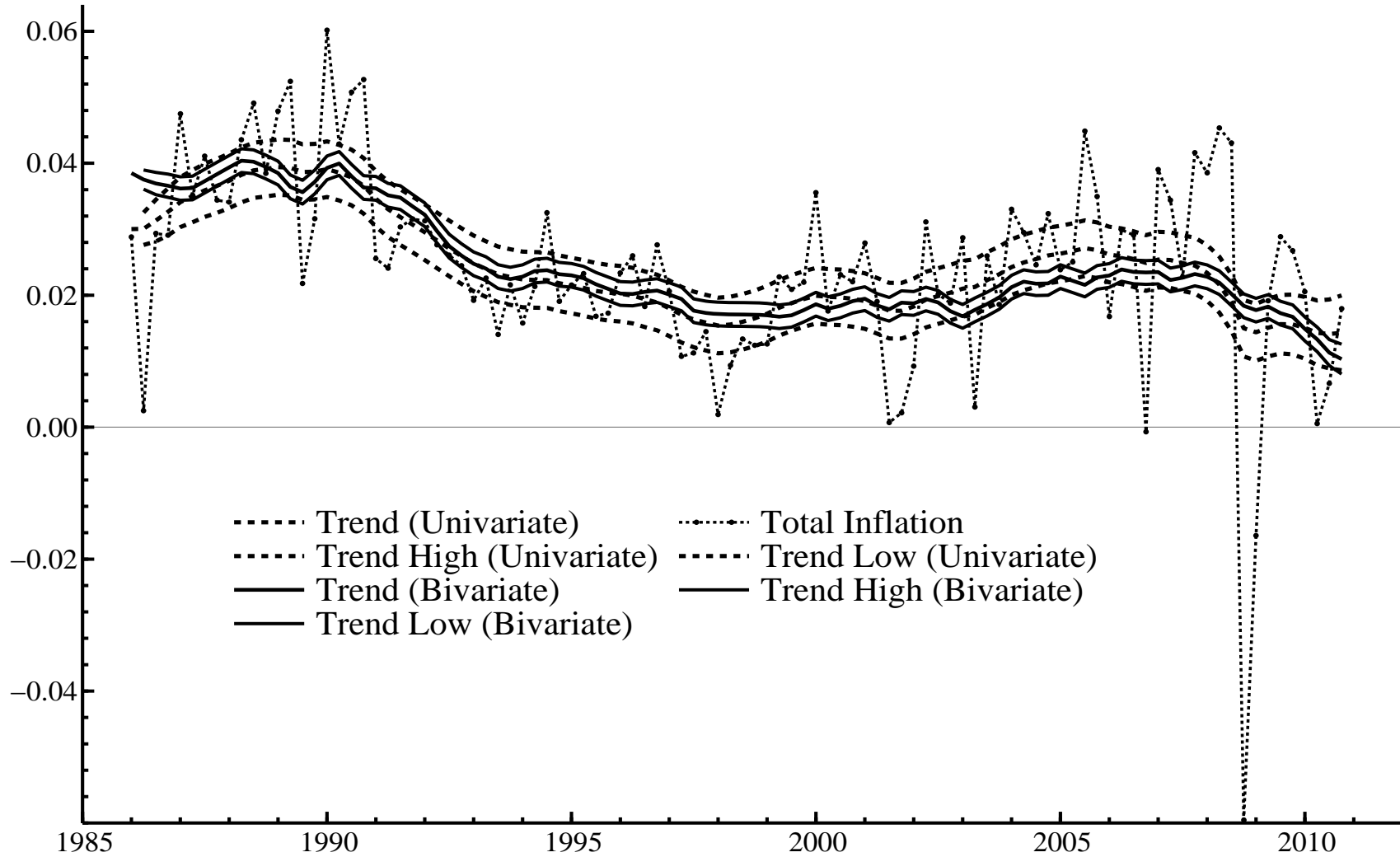
Solid: Core to Core. Dashed: Total to Core.



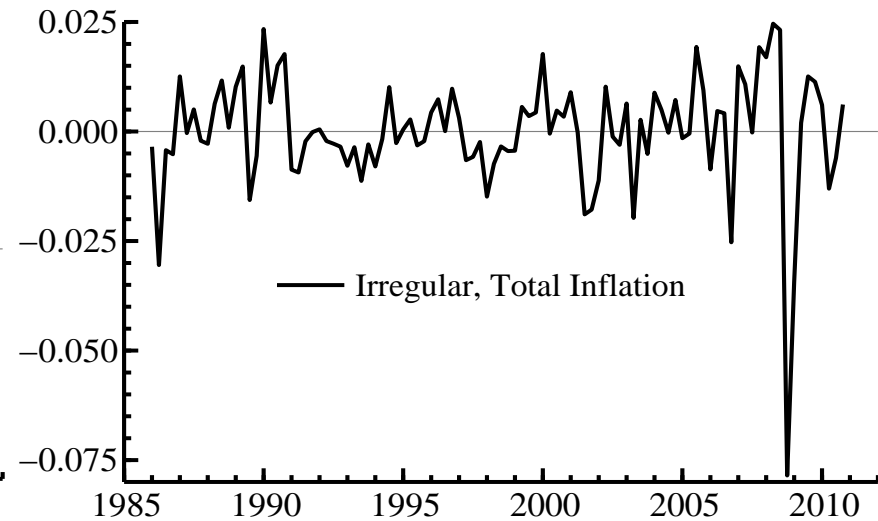
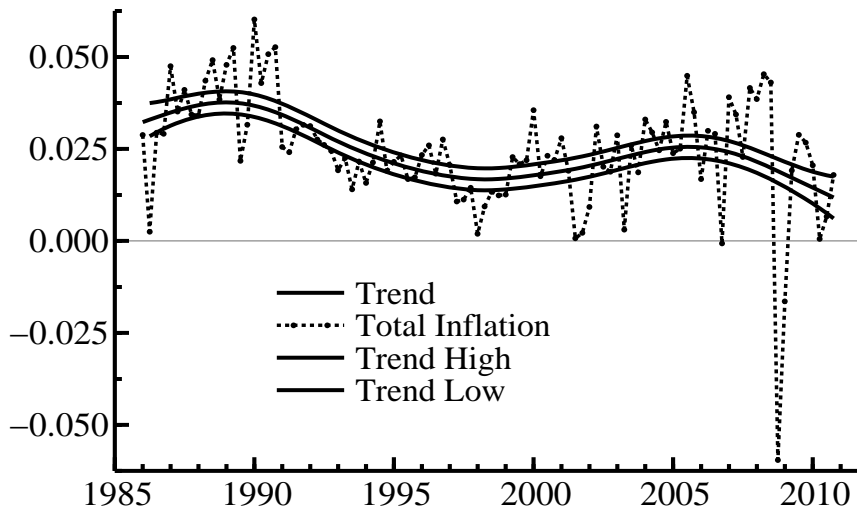
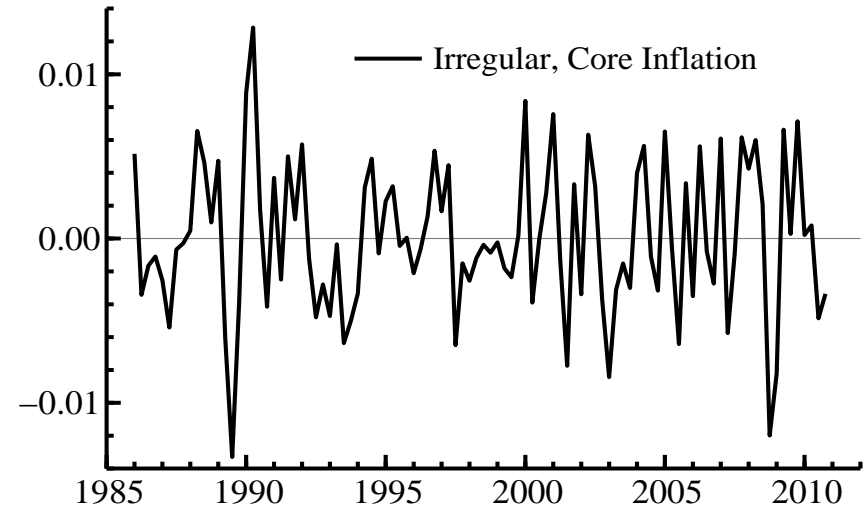
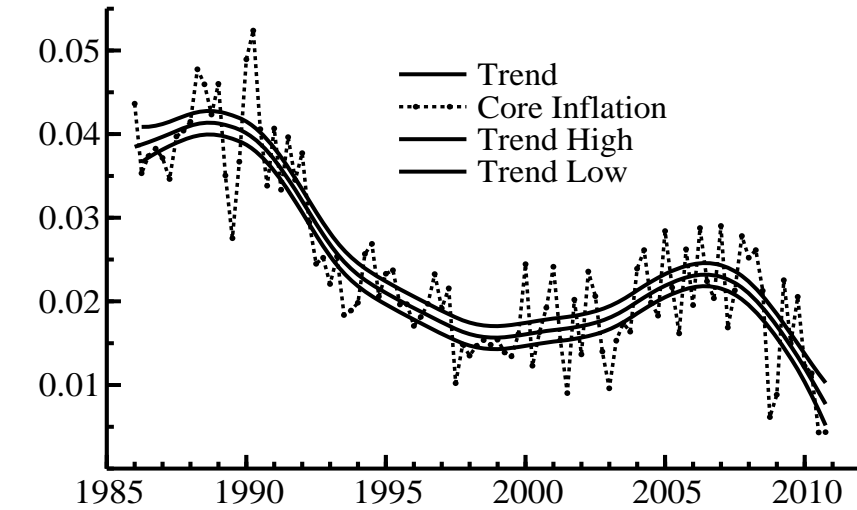
Solid: Total to Total. Dashed: Core to Total.



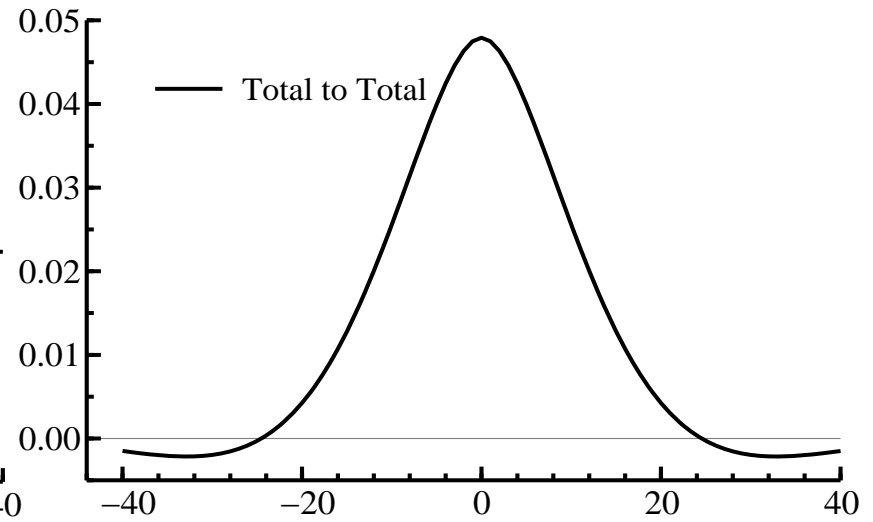
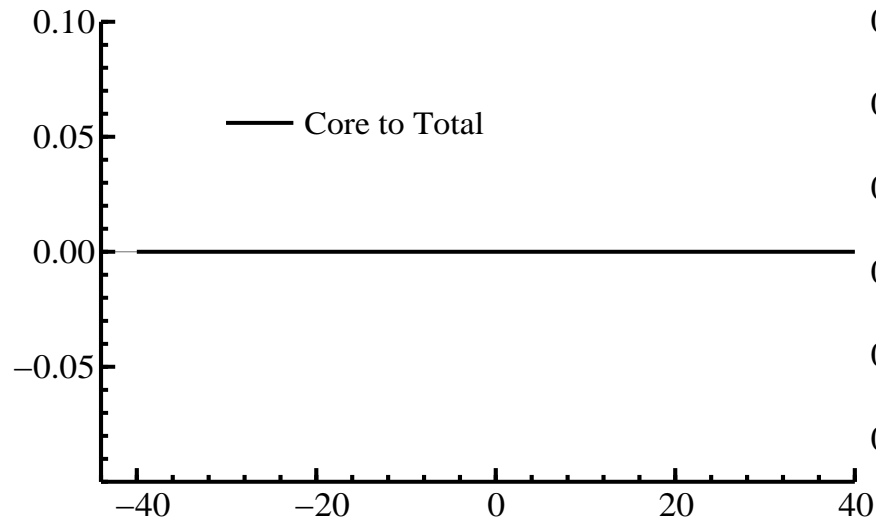
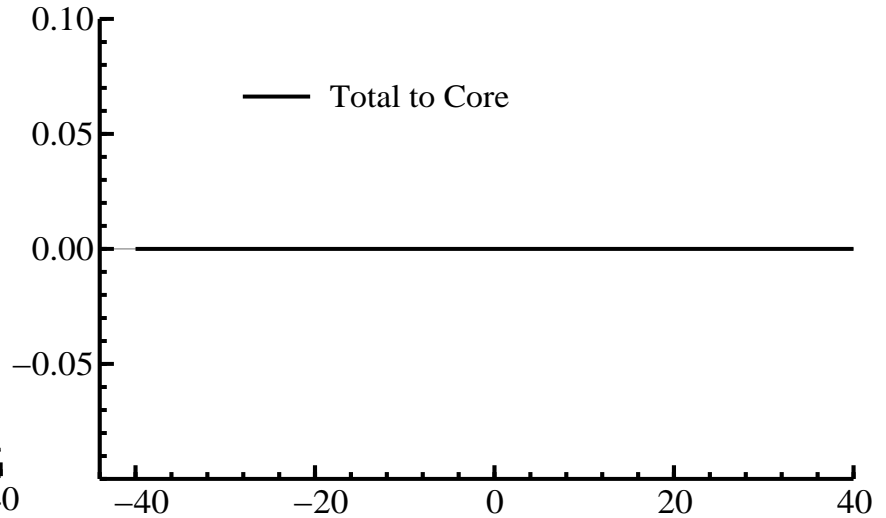
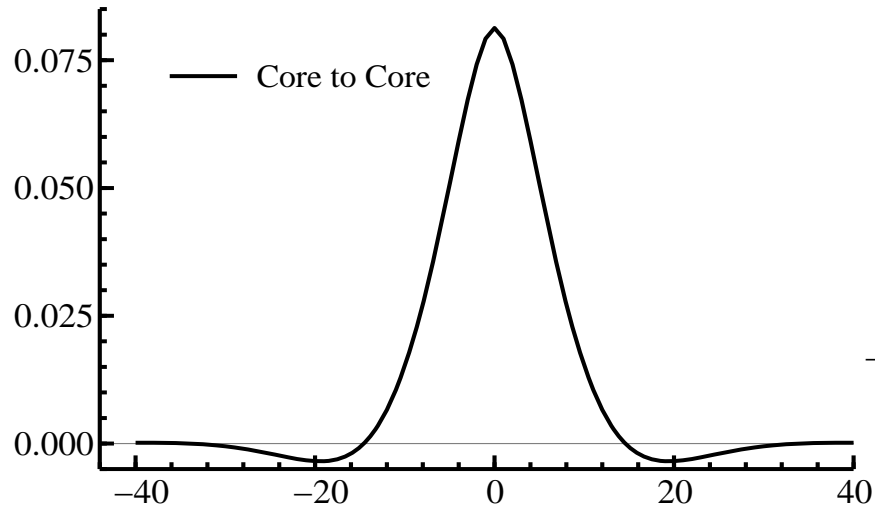
# Local Level Model, Univariate and Bivariate Trends



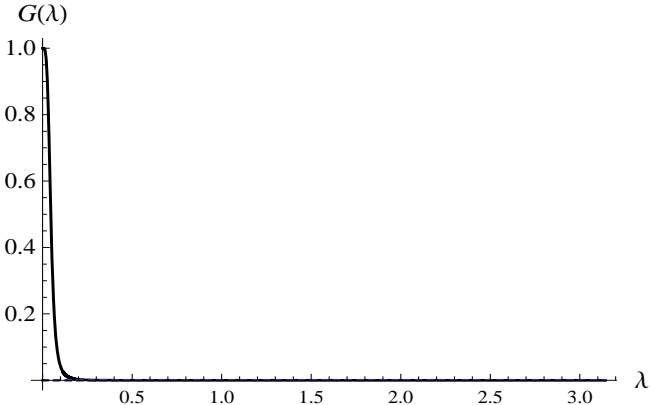
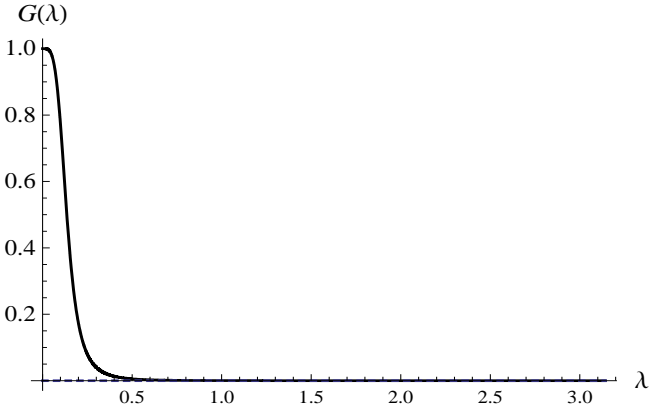
# Smooth Trend Model - Univariate, Trends



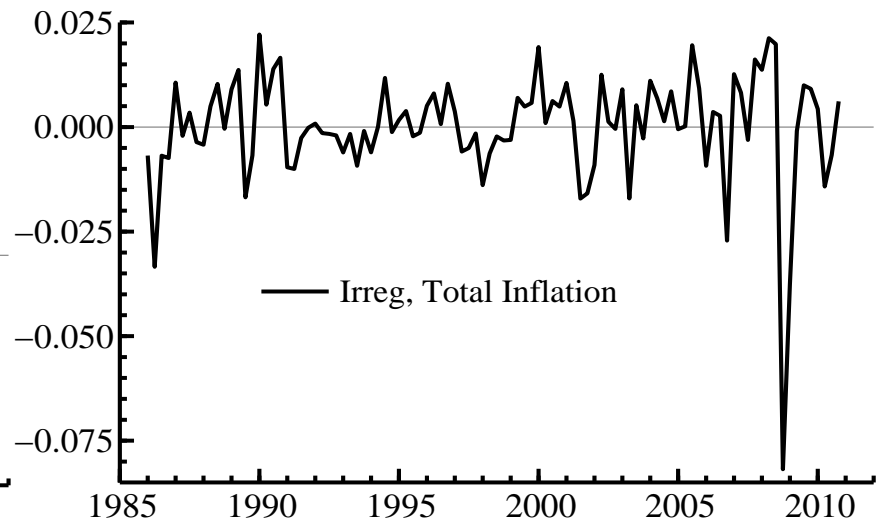
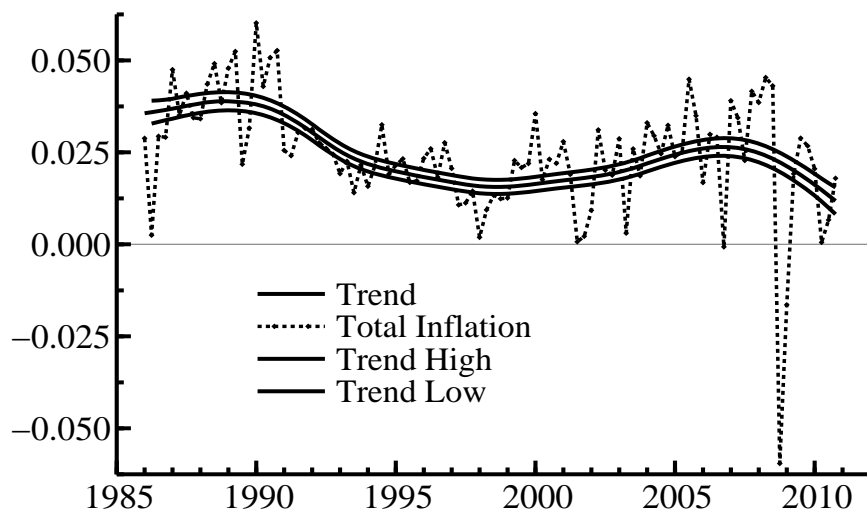
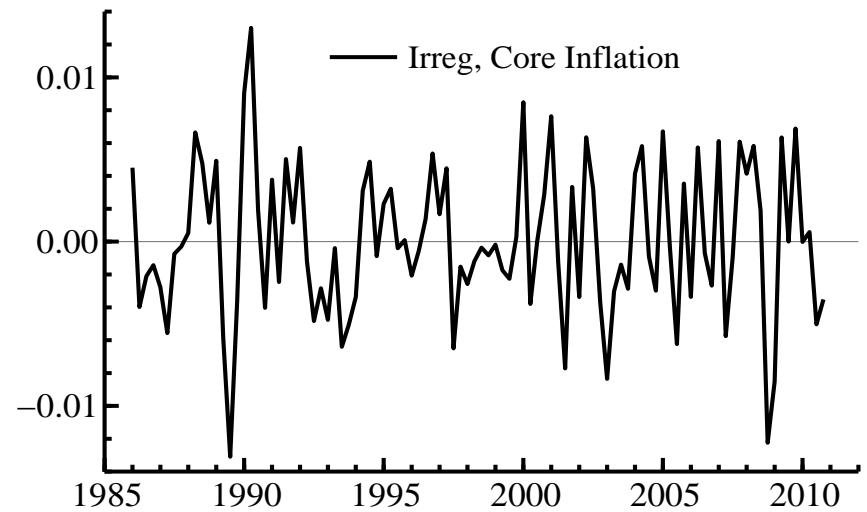
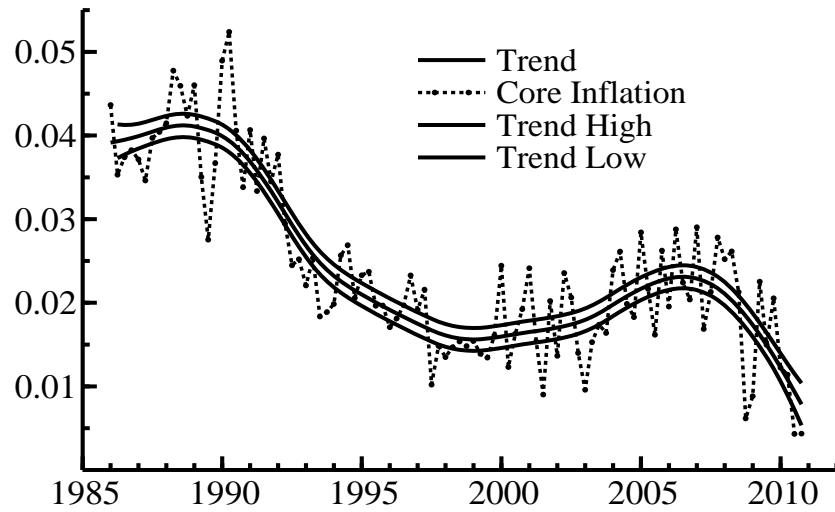
# Smooth Trend Model - Univariate, Weights



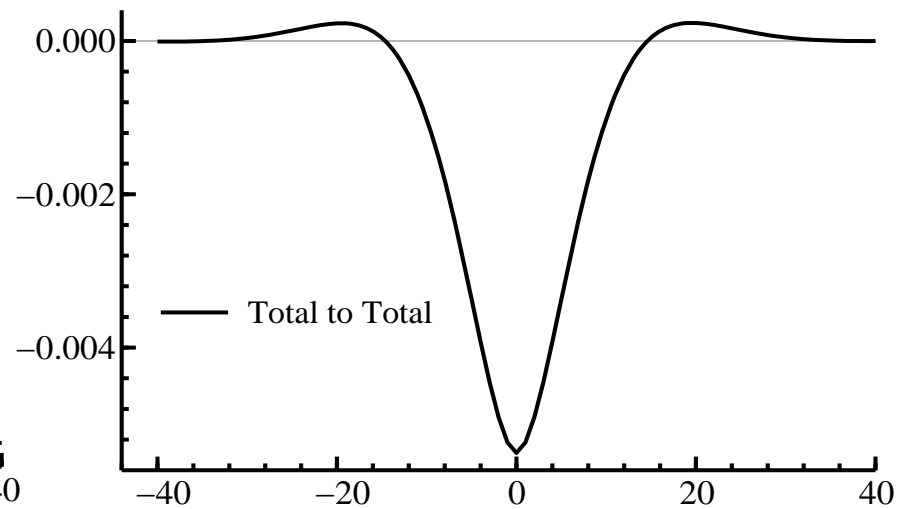
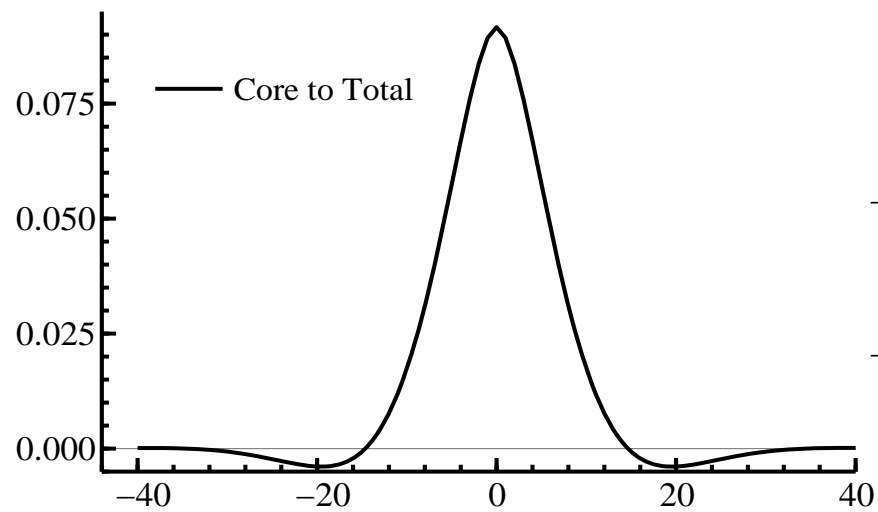
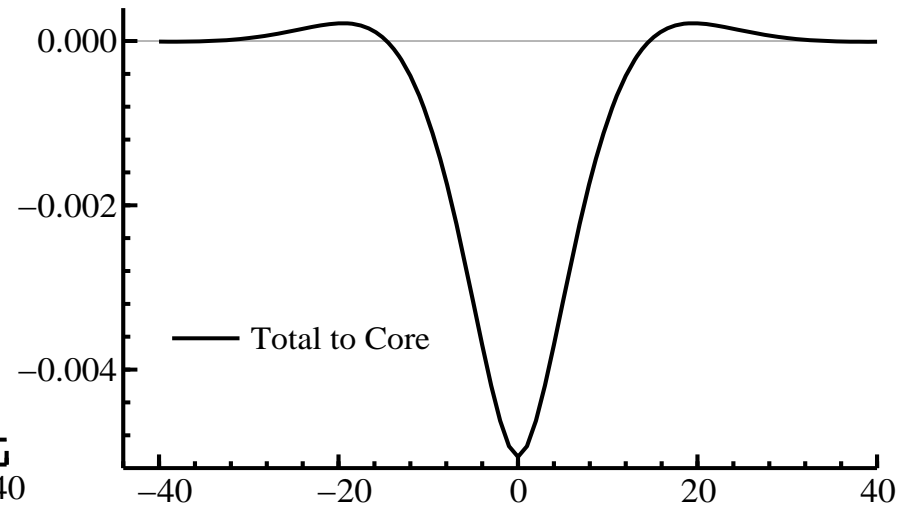
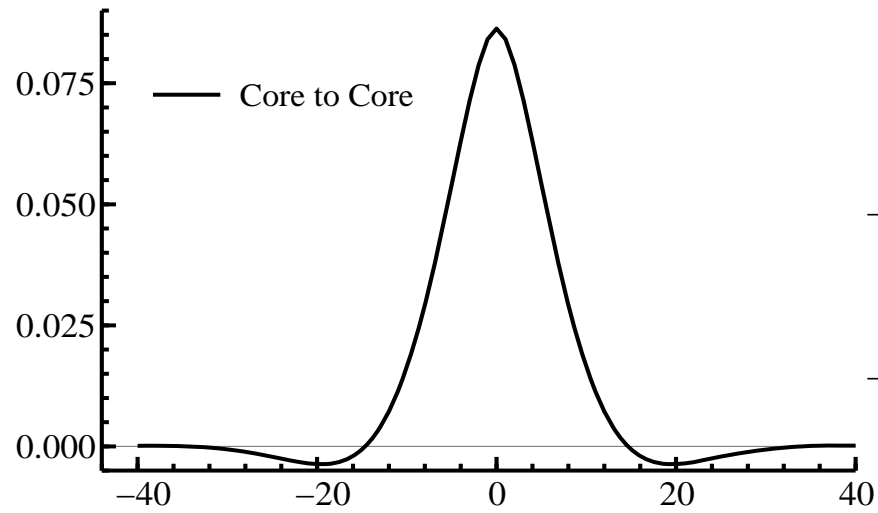
# Smooth Trend Model - Univariate, Gain functions



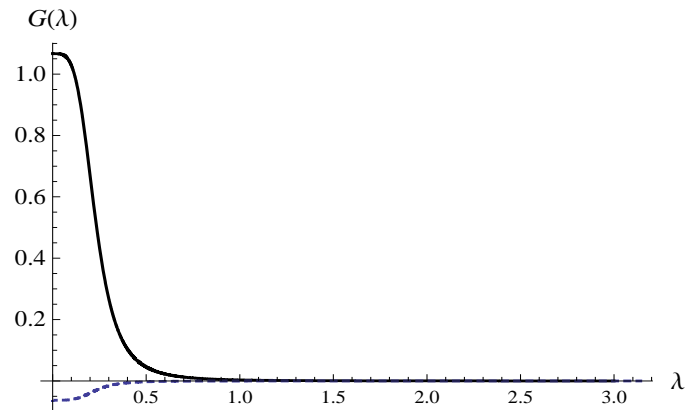
# Smooth Trend Model - Bivariate, Trends



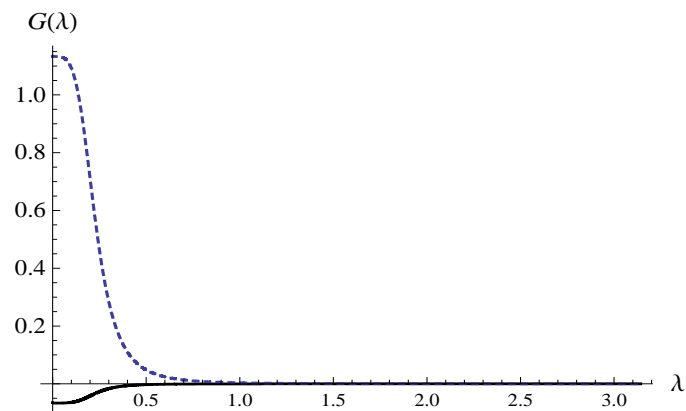
# Smooth Trend Model - Bivariate, Weight patterns



# Smooth Trend Model - Bivariate, Gain functions

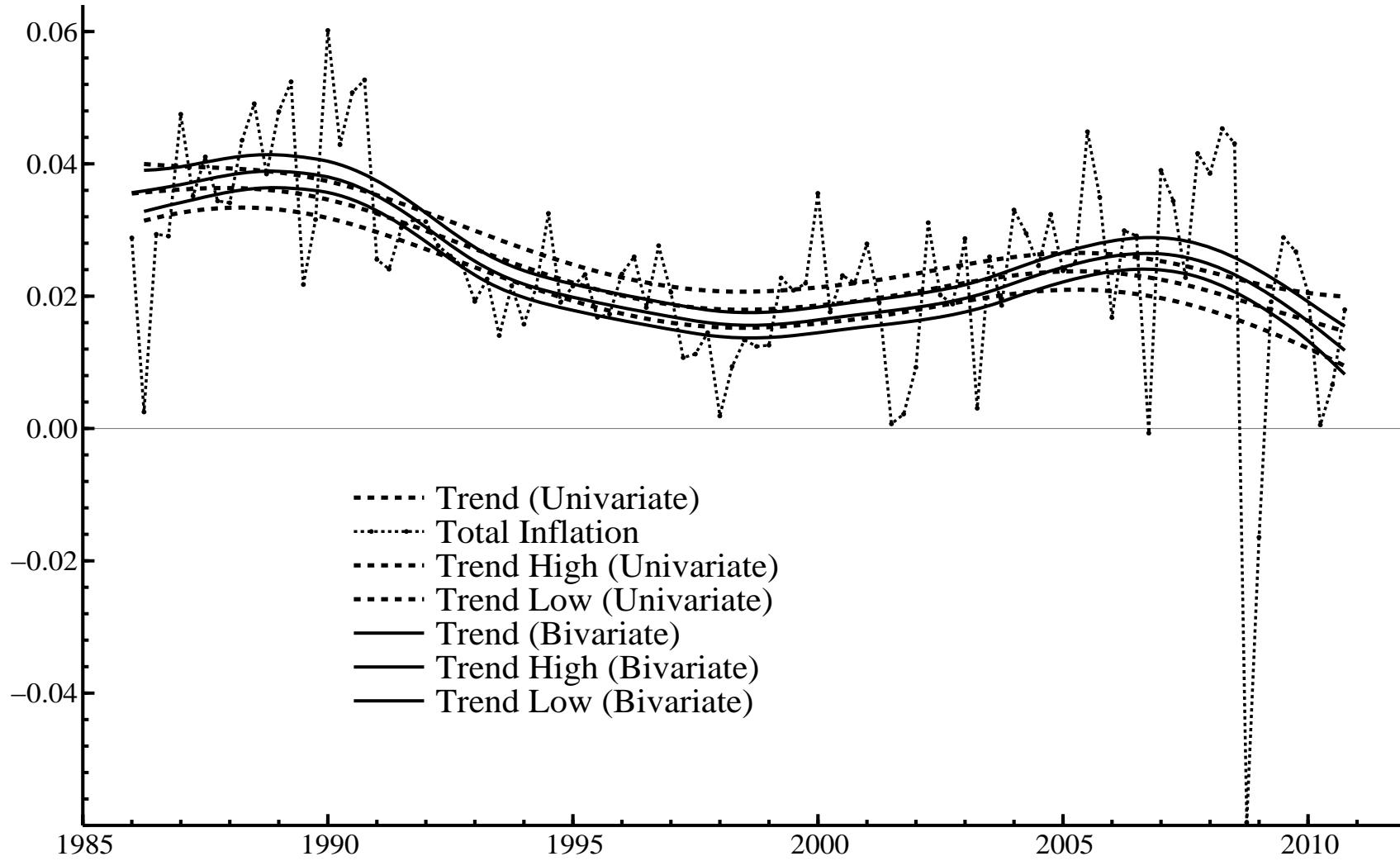


Dashed: Total to Core. Solid: Core to Core.



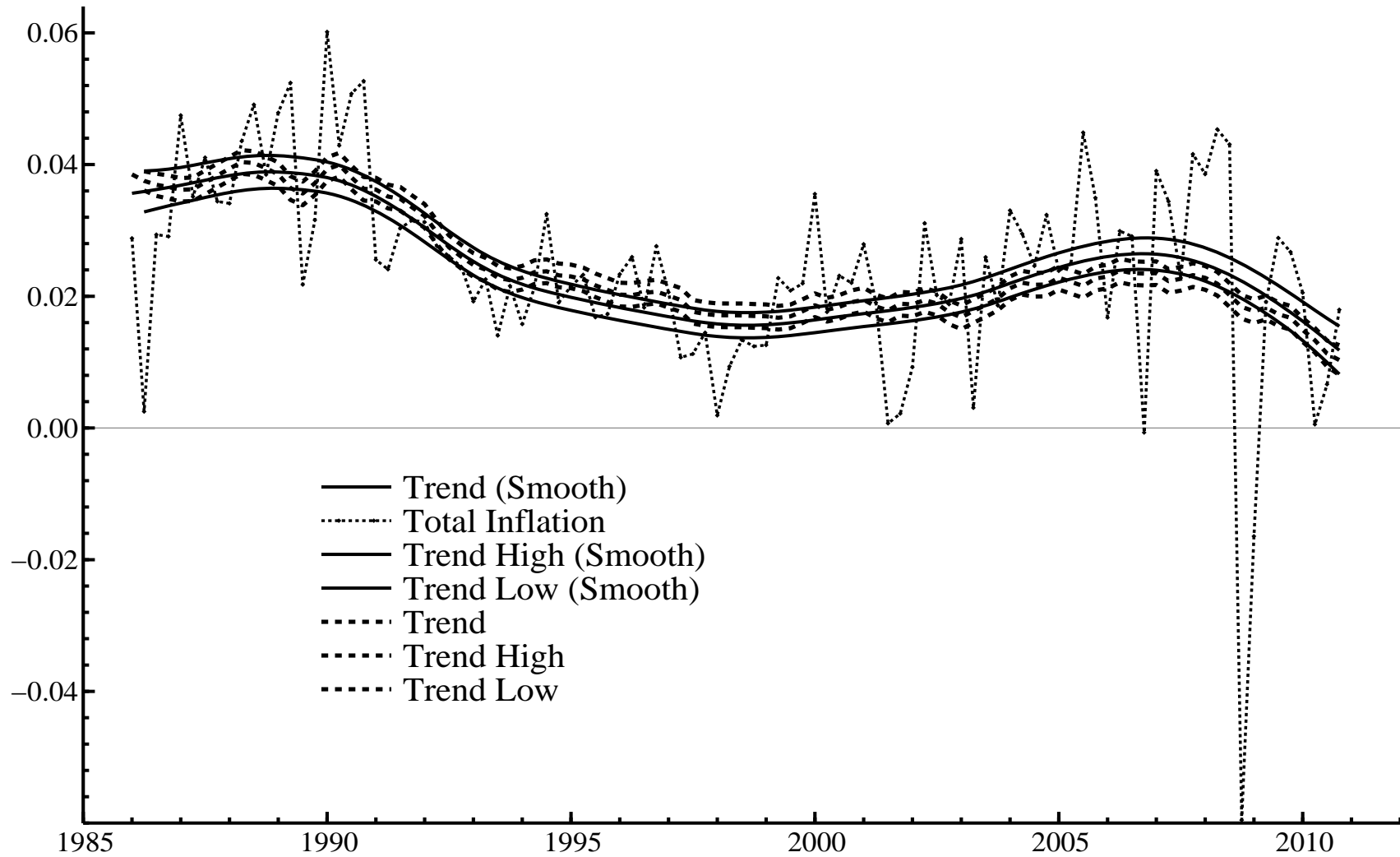
Dashed: Core to Total. Solid: Total to Total.

# Univariate, Bivariate Smooth trend models





# Bivariate Local level, Smooth trend models



## 26 Conclusions

new results on signal extraction for multiple series - generalized WK formula for nonstationary case and exact finite sample expressions

foundation for the extraction problem useful for nearly all economic applications of interest such as current analysis and policymaking

clarify the signal extraction architecture based on individual properties and relationships across series; otherwise, very difficult to determine how to combine information with cross-filters

flexibility in filter design; analytical expressions for multivariate gain functions - own-gains and cross-gains; consistency among filters and with set of series

Exact formulas for finite-sample weights

$$\mathbf{u}^{(J)} = \Delta_{\mathbf{s}}^{(J)} \mathbf{s}^{(J)} \text{ and } \mathbf{v}^{(J)} = \Delta_{\mathbf{n}}^{(J)} \mathbf{n}^{(J)}$$

matrices  $\Delta_{\mathbf{s}}^{(J)}$  and  $\Delta_{\mathbf{n}}^{(J)}$  contain differencing polynomials  $\delta_{\mathbf{s}}^{(J)}$  and  $\delta_{\mathbf{n}}^{(J)}$ . Let  $\mathbf{u}^{(J)} = \Delta_{\mathbf{s}}^{(J)} \mathbf{s}^{(J)}$  and  $\mathbf{v}^{(J)} = \Delta_{\mathbf{n}}^{(J)} \mathbf{n}^{(J)}$ , with cross-covariance matrices denoted  $\Sigma_{\mathbf{u}}^{IJ}$  and  $\Sigma_{\mathbf{v}}^{IJ}$ . Now assume there are no common roots among  $\delta_{\mathbf{s}}^{(J)}$  and  $\delta_{\mathbf{n}}^{(J)}$ , so that  $\delta^{(J)}(L) = \delta_{\mathbf{s}}^{(J)}(L)\delta_{\mathbf{n}}^{(J)}(L)$ . Then

$$\Delta^{(J)} = \underline{\Delta}_{\mathbf{n}}^{(J)} \Delta_{\mathbf{s}}^{(J)} = \underline{\Delta}_{\mathbf{s}}^{(J)} \Delta_{\mathbf{n}}^{(J)}, \quad (1)$$

where  $\underline{\Delta}_{\mathbf{n}}^{(J)}$  and  $\underline{\Delta}_{\mathbf{s}}^{(J)}$  are similar differencing matrices of reduced dimension, having  $T - d^J$  rows.

**Assumption  $M_T$ :** For each  $I = 1, 2, \dots, N$ , the initial values of  $\mathbf{y}^{(I)}$  are uncorrelated with  $\mathbf{u}$  and  $\mathbf{v}$ .

$$M^{II} = \Delta_{\mathbf{n}}^{(I)'} \Sigma_{\mathbf{v}}^{II-1} \Delta_{\mathbf{n}}^{(I)} + \Delta_{\mathbf{s}}^{(I)'} \Sigma_{\mathbf{u}}^{II-1} \Delta_{\mathbf{s}}^{(I)}.$$

With  $\hat{\mathbf{s}} = F\mathbf{y}$ , a compact matrix formula for  $F$  is given as follows. Define block-matrices  $A, B, C, D$  that have  $IJ$ th block matrix entries given by

$$A^{IJ} = \Delta_{\mathbf{s}}^{(I)'} \Sigma_{\mathbf{u}}^{II-1} \Sigma_{\mathbf{u}}^{IJ} \Sigma_{\mathbf{u}}^{JJ-1} \Delta_{\mathbf{s}}^{(J)}$$

$$B^{IJ} = \Delta_{\mathbf{n}}^{(I)'} \Sigma_{\mathbf{v}}^{II-1} \Sigma_{\mathbf{v}}^{IJ} \Sigma_{\mathbf{v}}^{JJ-1} \Delta_{\mathbf{n}}^{(J)}$$

$$C^{IJ} = \Delta_{\mathbf{s}}^{(I)'} \Sigma_{\mathbf{u}}^{II-1} \Sigma_{\mathbf{u}}^{IJ} \underline{\Delta}_{\mathbf{n}}^{(J)'}$$

$$D^{IJ} = \Delta_{\mathbf{n}}^{(I)'} \Sigma_{\mathbf{v}}^{II-1} \Sigma_{\mathbf{v}}^{IJ} \underline{\Delta}_{\mathbf{s}}^{(J)'}$$

let  $\widetilde{\Delta}$  block diagonal matrix with the matrix  $\Delta^{(I)}$  in the  $I$ th diagonal.

$$M = \tilde{A} + \tilde{B}$$

$$F = M^{-1} \left[ \tilde{B} + (C - D) \Sigma_{\mathbf{w}}^{-1} \tilde{\Delta} \right]$$

$$V = A + B + (C - D) \Sigma_{\mathbf{w}}^{-1} (C - D)',$$

# Optimal filter design

low-pass with varying curvature and location is optimal (MMSE)

curvature results from overlapping contributions (frequencies) of stochastic components

shape of low-pass varies across series which have different trend-noise relationships

which model generates a given filter, e.g., HP filter

model performance, model-filter selection, interpretation

the stochastic structure of signal and noise vectors - signal-noise ratios and correlations