How useful are no-arbitrage restrictions for forecasting the term structure of interest rates?

Raffaella Giacomini (with Andrea Carriero)

UCL(A) and Queen Mary

July 2008

Motivation

- Large literature on term structure modeling, focus on forecasting more recent

Motivation

- Large literature on term structure modeling, focus on forecasting more recent
- Ang and Piazzesi (2003): affine model building on assumption of no-arbitrage in bond markets $\Leftrightarrow$ VAR with parameter restrictions

Motivation

- Large literature on term structure modeling, focus on forecasting more recent
- Ang and Piazzesi (2003): affine model building on assumption of no-arbitrage in bond markets $\Leftrightarrow$ VAR with parameter restrictions
- Previous out-of-sample evaluation approaches (Ang and Piazzesi, 2003, Duffee, 2002, Diebold and Li, 2006, Carriero, 2007)

Motivation

- Large literature on term structure modeling, focus on forecasting more recent
- Ang and Piazzesi (2003): affine model building on assumption of no-arbitrage in bond markets $\Leftrightarrow$ VAR with parameter restrictions
- Previous out-of-sample evaluation approaches (Ang and Piazzesi, 2003, Duffee, 2002, Diebold and Li, 2006, Carriero, 2007)
    - Informal comparison of MSFE of no-arbitrage model, unrestricted VAR and random walk

Motivation

- Large literature on term structure modeling, focus on forecasting more recent
- Ang and Piazzesi (2003): affine model building on assumption of no-arbitrage in bond markets $\Leftrightarrow$ VAR with parameter restrictions
- Previous out-of-sample evaluation approaches (Ang and Piazzesi, 2003, Duffee, 2002, Diebold and Li, 2006, Carriero, 2007)
  - Informal comparison of MSFE of no-arbitrage model, unrestricted VAR and random walk
- Our goal: propose a formal framework for investigating the "usefulness" of no-arbitrage restrictions for forecasting

Questions we want to answer

- Are no-arbitrage restrictions only useful because they reduce estimation error?

Questions we want to answer

- Are no-arbitrage restrictions only useful because they reduce estimation error?
- Are the results different when considering economic vs. statistical measures of accuracy?

Questions we want to answer

- Are no-arbitrage restrictions only useful because they reduce estimation error?
- Are the results different when considering economic vs. statistical measures of accuracy?
- How important is the fact that no-arbitrage restrictions incorporate a time-varying risk premium?

Questions we want to answer

- Are no-arbitrage restrictions only useful because they reduce estimation error?
- Are the results different when considering economic vs. statistical measures of accuracy?
- How important is the fact that no-arbitrage restrictions incorporate a time-varying risk premium?
- Has the usefulness of no-arbitrage restrictions changed over time?

Limitations of existing approaches

- In-sample hypothesis testing not necessarily informative

Limitations of existing approaches

- In-sample hypothesis testing not necessarily informative
  - Restrictions may not be true but still be useful for forecasting (e.g., bias-variance tradeoff)

Limitations of existing approaches

- In-sample hypothesis testing not necessarily informative
  - Restrictions may not be true but still be useful for forecasting (e.g., bias-variance tradeoff)
  - Usefulness depends on the use to which one puts the model (e.g., constructing bond portfolios)

Limitations of existing approaches

- In-sample hypothesis testing not necessarily informative
  - Restrictions may not be true but still be useful for forecasting (e.g., bias-variance tradeoff)
  - Usefulness depends on the use to which one puts the model (e.g., constructing bond portfolios)
- Out-of-sample comparison tests (e.g., West, 1996, Clark and McCracken, 2001 etc.) not obviously applicable

Limitations of existing approaches

- In-sample hypothesis testing not necessarily informative
  - Restrictions may not be true but still be useful for forecasting (e.g., bias-variance tradeoff)
  - Usefulness depends on the use to which one puts the model (e.g., constructing bond portfolios)
- Out-of-sample comparison tests (e.g., West, 1996, Clark and McCracken, 2001 etc.) not obviously applicable
  - They compare performance of different *models*

Limitations of existing approaches

- In-sample hypothesis testing not necessarily informative
  - Restrictions may not be true but still be useful for forecasting (e.g., bias-variance tradeoff)
  - Usefulness depends on the use to which one puts the model (e.g., constructing bond portfolios)
- Out-of-sample comparison tests (e.g., West, 1996, Clark and McCracken, 2001 etc.) not obviously applicable
  - They compare performance of different *models*
  - We have one model and different restrictions on its parameters

Contributions

- Propose a measure of the usefulness of parameter restrictions for forecasting and show how to perform inference

Contributions

- Propose a measure of the usefulness of parameter restrictions for forecasting and show how to perform inference
- The measure is tailored to the forecaster's decision problem

Contributions

- Propose a measure of the usefulness of parameter restrictions for forecasting and show how to perform inference
- The measure is tailored to the forecaster's decision problem
- The measure can be time-varying

Environment

- One model (VAR for interest rates of different maturities)

Environment

- One model (VAR for interest rates of different maturities)
- Theoretical restrictions (Ang and Piazzesi's no-arbitrage restrictions)

Environment

- One model (VAR for interest rates of different maturities)
- Theoretical restrictions (Ang and Piazzesi's no-arbitrage restrictions)
- Atheoretical restrictions (random walk)

Environment

- One model (VAR for interest rates of different maturities)
- Theoretical restrictions (Ang and Piazzesi's no-arbitrage restrictions)
- Atheoretical restrictions (random walk)
- Statistical and economic loss functions (quadratic and utility of bond portfolio)

The idea in a nutshell

- Cast the problem in an out-of-sample forecast combination framework

The idea in a nutshell

- Cast the problem in an out-of-sample forecast combination framework
- Combine unrestricted (and random walk) forecast with no-arbitrage restricted forecast and estimate optimal weight for a general loss function

The idea in a nutshell

- Cast the problem in an out-of-sample forecast combination framework
- Combine unrestricted (and random walk) forecast with no-arbitrage restricted forecast and estimate optimal weight for a general loss function
- Optimal weight = measure of the usefulness of no-arbitrage restrictions

The idea in a nutshell

- Cast the problem in an out-of-sample forecast combination framework
- Combine unrestricted (and random walk) forecast with no-arbitrage restricted forecast and estimate optimal weight for a general loss function
- Optimal weight $=$ measure of the usefulness of no-arbitrage restrictions
- Time-varying weight $=$ time-varying measure of usefulness

Method. Rolling forecasts

- For each yield, produce three sequences of n out-of-sample forecasts using a rolling window scheme

Method. Rolling forecasts

- For each yield, produce three sequences of n out-of-sample forecasts using a rolling window scheme
  - unrestricted model $\Longrightarrow \left\{ f_t^U \right\}$

Method. Rolling forecasts

- For each yield, produce three sequences of n out-of-sample forecasts using a rolling window scheme
  - unrestricted model $\Longrightarrow \left\{ f_t^U \right\}$
  - model with no-arbitrage restrictions $\Longrightarrow \left\{ f_t^R \right\}$

Method. Rolling forecasts

- For each yield, produce three sequences of n out-of-sample forecasts using a rolling window scheme
  - unrestricted model $\Longrightarrow \left\{ f_t^U \right\}$
  - model with no-arbitrage restrictions $\Longrightarrow \left\{ f_t^R \right\}$
  - model with random walk restrictions $\Longrightarrow \left\{ f_t^{RW} \right\}$

Method. Rolling forecasts

- For each yield, produce three sequences of n out-of-sample forecasts using a rolling window scheme
  - unrestricted model $\Longrightarrow \left\{ f_t^U \right\}$
  - model with no-arbitrage restrictions $\Longrightarrow \left\{ f_t^R \right\}$
  - model with random walk restrictions $\Longrightarrow \left\{ f_t^{RW} \right\}$
- Two forecast combinations

$$
\begin{aligned}
f_t^c &= f_t^R + (1-\lambda)\left(f_t^U - f_t^R\right) \\
f_t^c &= f_t^R + (1-\lambda)\left(f_t^{RW} - f_t^R\right)
\end{aligned}
$$

Method. Global measure of usefulness

- For general loss $L(y_{t+1}, f_t)$:

$$\lambda^* = \arg\min_{\lambda \in R} E\left[L\left(y_{t+1}, f_t^R + (1-\lambda)(f_t^U - f_t^R)\right)\right]$$

Method. Global measure of usefulness

- For general loss $L\left(y_{t+1}, f_t\right)$:

$$\lambda^* = \arg\min_{\lambda \in R} E\left[L\left(y_{t+1}, f_t^R + (1-\lambda)(f_t^U - f_t^R)\right)\right]$$

- Estimate out-of-sample:

$$\widehat{\lambda} = \arg\min_{\lambda \in R} \frac{1}{n} \sum_t L\left(y_{t+1}, f_t^R + (1-\lambda)(f_t^U - f_t^R)\right).$$

Method. Global measure of usefulness

- For general loss $L(y_{t+1}, f_t)$:

$$\lambda^* = \arg\min_{\lambda \in R} E\left[L\left(y_{t+1}, f_t^R + (1-\lambda)(f_t^U - f_t^R)\right)\right]$$

- Estimate out-of-sample:

$$\widehat{\lambda} = \arg\min_{\lambda \in R} \frac{1}{n}\sum_t L\left(y_{t+1}, f_t^R + (1-\lambda)(f_t^U - f_t^R)\right).$$

- Large $\widehat{\lambda}$ = no-arbitrage restrictions are useful

Method. Global measure of usefulness

- For general loss $L\left(y_{t+1}, f_t\right)$:

$$\lambda^* = \arg\min_{\lambda \in R} E\left[L\left(y_{t+1}, f_t^R + (1-\lambda)(f_t^U - f_t^R)\right)\right]$$

- Estimate out-of-sample:

$$\widehat{\lambda} = \arg\min_{\lambda \in R} \frac{1}{n} \sum_t L\left(y_{t+1}, f_t^R + (1-\lambda)(f_t^U - f_t^R)\right).$$

- Large $\widehat{\lambda}$ = no-arbitrage restrictions are useful
- Small $\widehat{\lambda}$ = no-arbitrage restrictions are not useful

Method. Local measure of usefulness

- Usefulness of economic restrictions may vary over time (similar to Giacomini and Rossi, 2008 for testing problem)

Method. Local measure of usefulness

- Usefulness of economic restrictions may vary over time (similar to Giacomini and Rossi, 2008 for testing problem)
- Local measure $\Longleftrightarrow$ time-varying weights

$$\lambda_t^* = \arg\min_{\lambda_t \in R} E\left[L(y_{t+1}, f_t^R + (1-\lambda_t)\left(f_t^U - f_t^R\right)\right]$$

Method. Local measure of usefulness

- Usefulness of economic restrictions may vary over time (similar to Giacomini and Rossi, 2008 for testing problem)
- Local measure $\Longleftrightarrow$ time-varying weights

$$\lambda_t^* = \arg\min_{\lambda_t \in R} E\left[L(y_{t+1}, f_t^R + (1-\lambda_t)\left(f_t^U - f_t^R\right)\right]$$

- Estimate a "smoothed" version of λ_t^* over rolling windows of size d

$$\widehat{\lambda}_{t,d} = \arg\min_{\lambda \in R} \sum_{j=t-d+1}^{t} \left[L(y_{t+1}, f_t^R + (1-\lambda)\left(f_t^U - f_t^R\right)\right]$$

Method. Inference

- Global measure $\Longrightarrow$ test of $H_0 : \lambda = 0$ (restrictions are useless)

$$\text{Reject if } \left| \frac{\sqrt{n}\widehat{\lambda}}{\widehat{\sigma}} \right| > 1.96. \text{ We give valid } \widehat{\sigma}$$

Method. Inference

- Global measure $\Longrightarrow$ test of $H_0 : \lambda = 0$ (restrictions are useless)

$$\text{Reject if } \left| \frac{\sqrt{n}\widehat{\lambda}}{\widehat{\sigma}} \right| > 1.96. \text{ We give valid } \widehat{\sigma}$$

- Local measure $\Longrightarrow$ uniform confidence bands valid under $H_0 : \lambda_t^*$ constant $\forall t$

$$I = (\widehat{\lambda}_{t,d} - k_{\alpha,\pi}\frac{\widehat{\sigma}}{\sqrt{d}}, \widehat{\lambda}_{t,d} + k_{\alpha,\pi}\frac{\widehat{\sigma}}{\sqrt{d}})$$

We give $k_{\alpha,\pi}$ (same as Giacomini and Rossi, 2008)

Method. Inference

- Global measure $\Longrightarrow$ test of $H_0 : \lambda = 0$ (restrictions are useless)

$$\text{Reject if } \left| \frac{\sqrt{n}\widehat{\lambda}}{\widehat{\sigma}} \right| > 1.96. \text{ We give valid } \widehat{\sigma}$$

- Local measure $\Longrightarrow$ uniform confidence bands valid under $H_0 : \lambda_t^*$ constant $\forall t$

$$I = (\widehat{\lambda}_{t,d} - k_{\alpha,\pi}\frac{\widehat{\sigma}}{\sqrt{d}}, \widehat{\lambda}_{t,d} + k_{\alpha,\pi}\frac{\widehat{\sigma}}{\sqrt{d}})$$

  We give $k_{\alpha,\pi}$ (same as Giacomini and Rossi, 2008)

- If $0 \notin I$ at some $t \Longrightarrow$ reject that restrictions were useless $\forall t$

Statistical loss

- For a quadratic loss $L\left(y_{t+1}, f_t\right) = \left(y_{t+1} - f_t\right)^2$ measure of usefulness is

$$\lambda^* = \frac{E\left[\left(y_{t+1} - f_t^U\right)\left(f_t^R - f_t^U\right)\right]}{E\left[\left(f_t^R - f_t^U\right)^2\right]}$$

Statistical loss

- For a quadratic loss $L\left(y_{t+1}, f_t\right) = \left(y_{t+1} - f_t\right)^2$ measure of usefulness is

$$\lambda^* = \frac{E\left[\left(y_{t+1} - f_t^U\right)\left(f_t^R - f_t^U\right)\right]}{E\left[\left(f_t^R - f_t^U\right)^2\right]}$$

- Estimate by OLS

$$y_{t+1} - f_t^U = \lambda(f_t^R - f_t^U) + error$$

Statistical loss

- For a quadratic loss $L\left(y_{t+1}, f_t\right) = \left(y_{t+1} - f_t\right)^2$ measure of usefulness is

$$\lambda^* = \frac{E\left[\left(y_{t+1} - f_t^U\right)\left(f_t^R - f_t^U\right)\right]}{E\left[\left(f_t^R - f_t^U\right)^2\right]}$$

- Estimate by OLS

$$y_{t+1} - f_t^U = \lambda(f_t^R - f_t^U) + error$$

- Local measure obtained by estimating rolling regressions

Statistical loss

- For a quadratic loss $L\left(y_{t+1}, f_t\right) = \left(y_{t+1} - f_t\right)^2$ measure of usefulness is

$$\lambda^* = \frac{E\left[\left(y_{t+1} - f_t^U\right)\left(f_t^R - f_t^U\right)\right]}{E\left[\left(f_t^R - f_t^U\right)^2\right]}$$

- Estimate by OLS

$$y_{t+1} - f_t^U = \lambda(f_t^R - f_t^U) + error$$

- Local measure obtained by estimating rolling regressions
- $\widehat{\sigma}$ for the test and confidence bands is OLS standard error (HAC)

Illustrative example

- Suppose $y_t = \beta' x_t + \varepsilon_t$, $x_t \sim iid(0, \sigma_x^2 I_k)$, $\varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$

Illustrative example

- Suppose $y_t = \beta' x_t + \varepsilon_t$, $x_t \sim iid(0, \sigma_x^2 I_k)$, $\varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$
- Forecasts: $f_t^U = \widehat{\beta} x_{t+1}$ and $f_t^R = \widetilde{\beta} x_{t+1}$. If $\widehat{\beta}$ and $\widetilde{\beta}$ independent

$$\lambda^* = \frac{tr(Var(\widehat{\beta}))}{tr\left(Var\left(\widehat{\beta}\right) + Var(\widetilde{\beta}\Big) + \Big(bias(\widetilde{\beta})\Big)^2\right)}$$

Illustrative example

- Suppose $y_t = \beta' x_t + \varepsilon_t$, $x_t \sim iid(0, \sigma_x^2 I_k)$, $\varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$
- Forecasts: $f_t^U = \widehat{\beta} x_{t+1}$ and $f_t^R = \widetilde{\beta} x_{t+1}$. If $\widehat{\beta}$ and $\widetilde{\beta}$ independent

$$\lambda^* = \frac{tr(Var(\widehat{\beta}))}{tr\left(Var\left(\widehat{\beta}\right) + Var(\widetilde{\beta}\right) + \left(bias(\widetilde{\beta})\right)^2\right)}$$

- $\lambda^* \approx 0$ (restrictions not useful) if large bias of restricted estimator ($\Leftrightarrow$ restrictions too misspecified to be useful).

Illustrative example

- Suppose $y_t = \beta' x_t + \varepsilon_t$, $x_t \sim iid(0, \sigma_x^2 I_k)$, $\varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$
- Forecasts: $f_t^U = \widehat{\beta} x_{t+1}$ and $f_t^R = \widetilde{\beta} x_{t+1}$. If $\widehat{\beta}$ and $\widetilde{\beta}$ independent

$$\lambda^* = \frac{tr(Var(\widehat{\beta}))}{tr\left(Var\left(\widehat{\beta}\right) + Var(\widetilde{\beta}\right) + \left(bias(\widetilde{\beta})\right)^2\right)}$$

- $\lambda^* \approx 0$ (restrictions not useful) if large bias of restricted estimator ($\Leftrightarrow$ restrictions too misspecified to be useful).
- $\lambda^* \approx 1$ (restrictions useful) if large variance of unrestricted estimator ($\Leftrightarrow$ usefulness due to reduction in estimation uncertainty)

Limitations of statistical measures of accuracy

- Usefulness of no-arbitrage restrictions may be simply due to a reduction in dimensionality

Limitations of statistical measures of accuracy

- Usefulness of no-arbitrage restrictions may be simply due to a reduction in dimensionality
- Generally true for out-of-sample comparisons based on MSE

Economic loss function. Portfolio utility loss

- Similar to West, Edison and Cho (1993)

Economic loss function. Portfolio utility loss

- Similar to West, Edison and Cho (1993)
- Bond portfolio $w' y_{t+1}$. Optimal weights for quadratic utility

$$w^* = \arg\min_{w} \left\{ w' E_t \left[y_{t+1} \right] - \frac{\gamma}{2} w' \Sigma w \right\}$$

$\Sigma = Var(y_{t+1})$ and $\gamma = CRRA / \left(CRRA + 1 \right)$

Economic loss function. Portfolio utility loss

- Similar to West, Edison and Cho (1993)
- Bond portfolio $w' y_{t+1}$. Optimal weights for quadratic utility

$$w^* = \arg\min_w \left\{ w' E_t \left[y_{t+1} \right] - \frac{\gamma}{2} w' \Sigma w \right\}$$

$\Sigma = Var(y_{t+1})$ and $\gamma = CRRA / \left(CRRA + 1 \right)$

- Classical solution (Markowitz, 1952)

$$\begin{aligned} w^* &= C_1 + C_2 E_t \left[y_{t+1} \right], \\ C_1 &= \frac{\Sigma^{-1} \iota}{\iota' \Sigma^{-1} \iota}, \; C_2 = \frac{1}{\gamma} \left(\Sigma^{-1} - \frac{\Sigma^{-1} \iota \iota' \Sigma^{-1}}{\iota' \Sigma^{-1} \iota} \right) \end{aligned}$$

Economic loss function. Portfolio utility loss

- Different conditional mean forecasts $\Leftrightarrow$ different weights

Economic loss function. Portfolio utility loss

- Different conditional mean forecasts $\Leftrightarrow$ different weights
- For the combination forecast

$$w^*(\lambda) = C_1 + C_2 \left(f_t^R + (1 - \lambda)(f_t^U - f_t^R) \right)$$

Economic loss function. Portfolio utility loss

- Different conditional mean forecasts $\Leftrightarrow$ different weights
- For the combination forecast

$$w^*(\lambda) = C_1 + C_2 \left(f_t^R + (1-\lambda)(f_t^U - f_t^R) \right)$$

- Measure of usefulness

$$\begin{aligned} \lambda^* &= \arg\min_{\lambda \in R} \left[-w^*(\lambda)' E[y_{t+1}] + \frac{\gamma}{2} w^*(\lambda)' \Sigma w^*(\lambda) \right] \\ &= \frac{E\left[(f_t^R - f_t^U)' C_2' \left(y_{t+1} - \gamma \Sigma \left(C_1 + C_2 f_t^U \right) \right) \right]}{E\left[\gamma (f_t^R - f_t^U)' C_2' \Sigma C_2 (f_t^R - f_t^U) \right]} \end{aligned}$$

Economic loss function. Portfolio utility loss

- Different conditional mean forecasts $\Leftrightarrow$ different weights
- For the combination forecast

$$w^*(\lambda) = C_1 + C_2\left(f_t^R + (1-\lambda)(f_t^U - f_t^R)\right)$$

- Measure of usefulness

$$\begin{aligned}\lambda^* &= \arg\min_{\lambda \in R}\left[-w^*(\lambda)' E[y_{t+1}] + \frac{\gamma}{2} w^*(\lambda)' \Sigma w^*(\lambda)\right] \\ &= \frac{E\left[(f_t^R - f_t^U)' C_2' \left(y_{t+1} - \gamma\Sigma\left(C_1 + C_2 f_t^U\right)\right)\right]}{E\left[\gamma(f_t^R - f_t^U)' C_2' \Sigma C_2 (f_t^R - f_t^U)\right]}\end{aligned}$$

- Estimate by substituting $E[\cdot]$ with out-of-sample mean (global) or rolling out-of-sample means (local)

Empirical application. Ang and Piazzesi's model

- y_t = vector of yields of different maturities (1, 3, 12, 36, 60 months)

Empirical application. Ang and Piazzesi's model

- y_t = vector of yields of different maturities (1, 3, 12, 36, 60 months)
- State-space representation:

$$
\begin{aligned}
y_t &= A + BF_t + v_t \\
F_t &= \Psi F_{t-1} + \Omega \varepsilon_t
\end{aligned}
$$

Empirical application. Ang and Piazzesi's model

- y_t = vector of yields of different maturities (1, 3, 12, 36, 60 months)
- State-space representation:

$$
\begin{aligned}
y_t &= A + BF_t + v_t \\
F_t &= \Psi F_{t-1} + \Omega \varepsilon_t
\end{aligned}
$$

- F_t contains three latent factors ($\approx$ level, slope and curvature of term structure)

Empirical application. AP model as a restricted VAR

- No-arbitrage assumption restricts the elements of A and B

$$\bar{A}_{n+1} = \bar{A}_n + \bar{B}_n'(-\Lambda_0) + 0.5\bar{B}_n'\Omega\Omega'\bar{B}_n - \delta_0$$
$$\bar{B}_{n+1}' = \bar{B}_n'(\Psi - \Omega\Lambda_1) - \delta_1'$$

with $A_n = -\bar{A}_n/n$, $B_n = -\bar{B}_n/n$, $\bar{A}_1 = -\delta_0$ and $\bar{B}_1 = -\delta_1$

Empirical application. AP model as a restricted VAR

- No-arbitrage assumption restricts the elements of A and B

$$\begin{aligned}\bar{A}_{n+1} &= \bar{A}_n + \bar{B}_n'(-\Lambda_0) + 0.5\bar{B}_n'\Omega\Omega'\bar{B}_n - \delta_0 \\ \bar{B}_{n+1}' &= \bar{B}_n'(\Psi - \Omega\Lambda_1) - \delta_1'\end{aligned}$$

  with $A_n = -\bar{A}_n/n$, $B_n = -\bar{B}_n/n$, $\bar{A}_1 = -\delta_0$ and $\bar{B}_1 = -\delta_1$

- Λ_0 and Λ_1 are such that $\Lambda_t = \Lambda_0 + \Lambda_1 F_t$, with $\Lambda_t =$ market prices of risk

$$\begin{aligned}\Lambda_1 \neq 0 &\Longrightarrow \text{time-varying risk premium} \\ \Lambda_1 = 0 &\Longrightarrow \text{constant risk premium}\end{aligned}$$

Empirical application. AP model as a restricted VAR

- AP state-space model $\approx MA(\infty) \Longrightarrow$ approximate with VAR(3)

$$Y = X\Phi + U$$

Empirical application. AP model as a restricted VAR

- AP state-space model $\approx MA(\infty) \Longrightarrow$ approximate with VAR(3)

$$Y = X\Phi + U$$

- Impose the restrictions on the VAR by writing the sample likelihood of Y as

$$|\Sigma_u|^{-T/2} \exp\{-0.5tr[\Sigma_u^{-1}(\Gamma^*_{yy} - \Phi'\Gamma^*_{xy} - \Gamma^{*\prime}_{xy}\Phi + \Phi'\Gamma^*_{xx}\Phi)]\}$$

$\Gamma^*_{yy},\ \Gamma^*_{yx},\ \Gamma^*_{xx}$ = moments implied by state-space

Empirical application. AP model as a restricted VAR

- AP state-space model $\approx MA(\infty) \Longrightarrow$ approximate with VAR(3)

$$Y = X\Phi + U$$

- Impose the restrictions on the VAR by writing the sample likelihood of Y as

$$|\Sigma_u|^{-T/2} \exp\{-0.5 tr[\Sigma_u^{-1}(\Gamma^*_{yy} - \Phi'\Gamma^*_{xy} - \Gamma^{*\prime}_{xy}\Phi + \Phi'\Gamma^*_{xx}\Phi)]\}$$

$\Gamma^*_{yy}, \ \Gamma^*_{yx}, \ \Gamma^*_{xx}$ = moments implied by state-space

- Estimate by ML $\Longrightarrow$ no-arbitrage restricted estimator

Empirical application. AP model as a restricted VAR

- AP state-space model $\approx MA(\infty) \Longrightarrow$ approximate with VAR(3)

$$Y = X\Phi + U$$

- Impose the restrictions on the VAR by writing the sample likelihood of Y as

$$|\Sigma_u|^{-T/2} \exp\{-0.5tr[\Sigma_u^{-1}(\Gamma^*_{yy} - \Phi'\Gamma^*_{xy} - \Gamma^{*\prime}_{xy}\Phi + \Phi'\Gamma^*_{xx}\Phi)]\}$$

$\Gamma^*_{yy},\ \Gamma^*_{yx},\ \Gamma^*_{xx} =$ moments implied by state-space

- Estimate by ML $\Longrightarrow$ no-arbitrage restricted estimator
- The unrestricted estimator is the OLS estimator

Empirical results

VERY PRELIMINARY....

Quadratic loss. Global measure of usefulness

NA vs. unrestricted VAR

Yields	$\widehat{\lambda}$	Test of H_0 : NA restrictions useless
1-month	**0.427**	4.711*
3-month	**0.655**	3.547*
12-month	**0.979**	5.374*
36-month	**0.814**	4.823*
60-month	**0.826**	5.424*

Quadratic loss. Global measure of usefulness

NA with constant risk premium vs. unrestricted VAR

Yields	$\widehat{\lambda}$	Test of H_0 : NA restrictions useless
1-month	**0.337**	3.596*
3-month	**0.785**	4.559*
12-month	**0.796**	4.777*
36-month	**0.912**	5.894*
60-month	**0.742**	5.251*

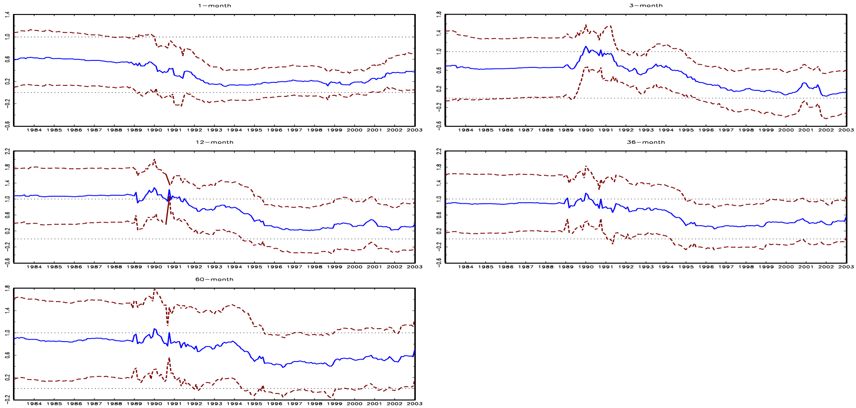
Quadratic loss. Global measure of usefulness

NA vs. random walk

Yields	$\widehat{\lambda}$	Test of H_0 : NA restrictions useless
1-month	**0.516**	4.649*
3-month	**-0.110**	-0.526
12-month	**0.276**	0.828
36-month	**0.099**	0.337
60-month	**0.274**	0.668

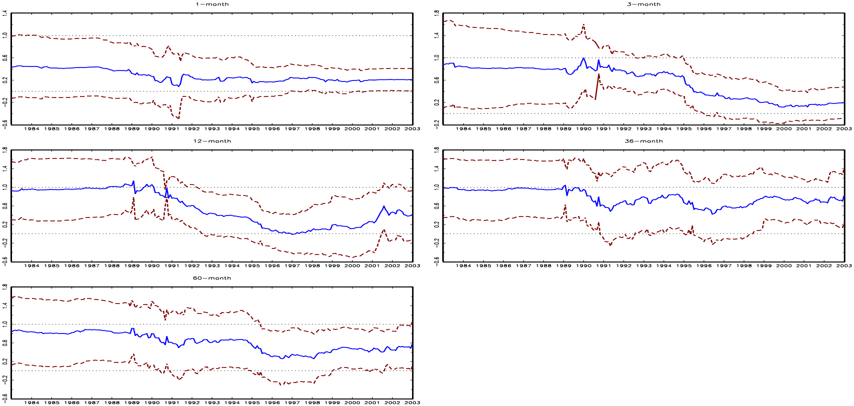
Quadratic loss. Local measure of usefulness

NA vs. unrestricted VAR



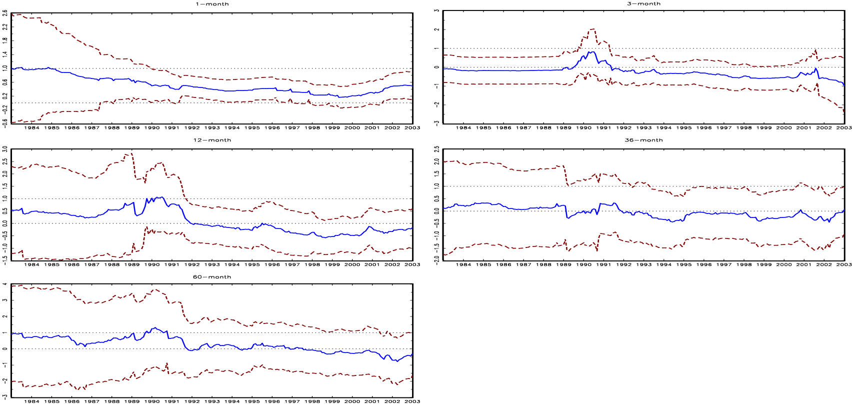
Quadratic loss. Local measure of usefulness

NA with constant risk premium vs. unrestricted VAR



Quadratic loss. Local measure of usefulness

NA vs. random walk



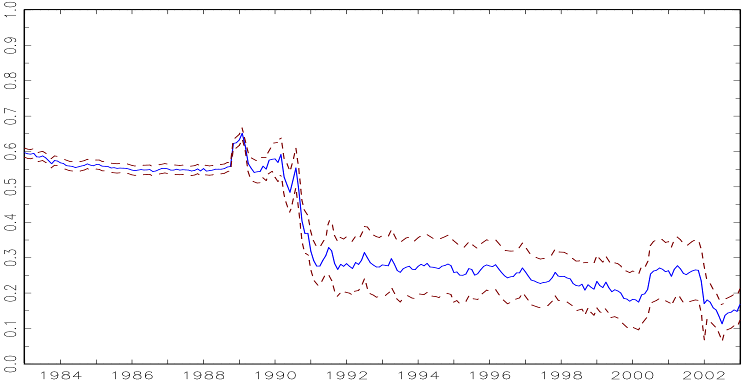
Portfolio utility loss. Global measure of usefulness

Results for CRRA = 1

	$\widehat{\lambda}$	H_0 : NA restr. useless
NA vs. unrestricted VAR	**0.416**	23.33^{*}
NA with const.risk premium vs. unrestricted VAR	**0.199**	26.80^{*}
NA vs. random walk	**0.512**	63.25^{*}

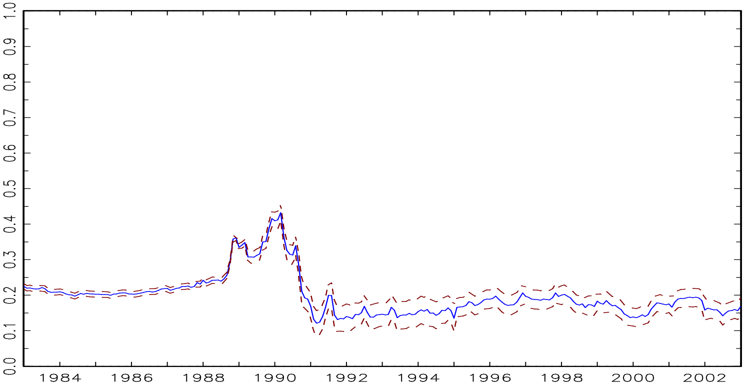
Portfolio utility loss. Local measure of usefulness

NA vs. unrestricted VAR



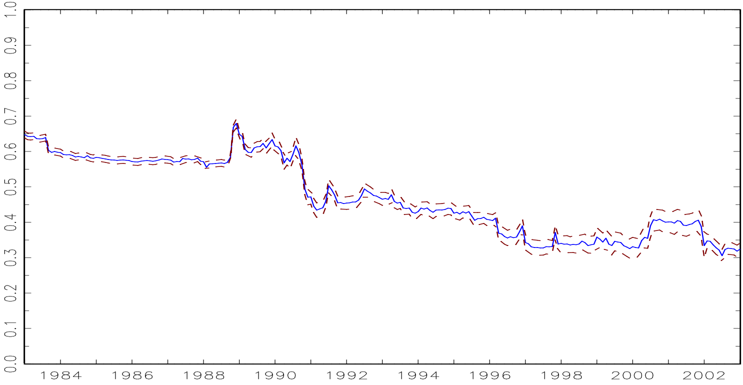
Portfolio utility loss. Local measure of usefulness

NA with constant risk premium vs. unrestricted VAR



Portfolio utility loss. Local measure of usefulness

NA vs. random walk



Conclusion. Theoretical framework

- No-arbitrage affine term structure model = VAR with theory-based parameter restrictions

Conclusion. Theoretical framework

- No-arbitrage affine term structure model = VAR with theory-based parameter restrictions
- Proposed a framework for measuring and performing inference about the usefulness of restrictions for forecasting

Conclusion. Theoretical framework

- No-arbitrage affine term structure model = VAR with theory-based parameter restrictions
- Proposed a framework for measuring and performing inference about the usefulness of restrictions for forecasting
- Usefulness depends on the forecaster's loss function

Conclusion. Theoretical framework

- No-arbitrage affine term structure model = VAR with theory-based parameter restrictions
- Proposed a framework for measuring and performing inference about the usefulness of restrictions for forecasting
- Usefulness depends on the forecaster's loss function
- Considered both global and local measures

Conclusion. Back to our initial questions

- Are no-arbitrage restrictions only useful because they reduce estimation error?

Conclusion. Back to our initial questions

- Are no-arbitrage restrictions only useful because they reduce estimation error?
  - Yes. In terms of statistical accuracy, they are no better than atheoretical ways to do the same

Conclusion. Back to our initial questions

- Are no-arbitrage restrictions only useful because they reduce estimation error?
  - Yes. In terms of statistical accuracy, they are no better than atheoretical ways to do the same
- Are the results different when considering economic vs. statistical measures of accuracy?

Conclusion. Back to our initial questions

- Are no-arbitrage restrictions only useful because they reduce estimation error?
  - Yes. In terms of statistical accuracy, they are no better than atheoretical ways to do the same
- Are the results different when considering economic vs. statistical measures of accuracy?
  - Yes. Dimension reduction not so important economically

Conclusion. Back to our initial questions

- Are no-arbitrage restrictions only useful because they reduce estimation error?
  - Yes. In terms of statistical accuracy, they are no better than atheoretical ways to do the same
- Are the results different when considering economic vs. statistical measures of accuracy?
  - Yes. Dimension reduction not so important economically
- How important is the fact that no-arbitrage restrictions incorporate a time-varying risk premium?

Conclusion. Back to our initial questions

- Are no-arbitrage restrictions only useful because they reduce estimation error?
  - Yes. In terms of statistical accuracy, they are no better than atheoretical ways to do the same
- Are the results different when considering economic vs. statistical measures of accuracy?
  - Yes. Dimension reduction not so important economically
- How important is the fact that no-arbitrage restrictions incorporate a time-varying risk premium?
  - Not important for statistical accuracy, but essential for constructing bond portfolios

Conclusion. Back to our initial questions

- Are no-arbitrage restrictions only useful because they reduce estimation error?
  - Yes. In terms of statistical accuracy, they are no better than atheoretical ways to do the same
- Are the results different when considering economic vs. statistical measures of accuracy?
  - Yes. Dimension reduction not so important economically
- How important is the fact that no-arbitrage restrictions incorporate a time-varying risk premium?
  - Not important for statistical accuracy, but essential for constructing bond portfolios
- Has the usefulness of the restrictions changed over time?

Conclusion. Back to our initial questions

- Are no-arbitrage restrictions only useful because they reduce estimation error?
  - Yes. In terms of statistical accuracy, they are no better than atheoretical ways to do the same
- Are the results different when considering economic vs. statistical measures of accuracy?
  - Yes. Dimension reduction not so important economically
- How important is the fact that no-arbitrage restrictions incorporate a time-varying risk premium?
  - Not important for statistical accuracy, but essential for constructing bond portfolios
- Has the usefulness of the restrictions changed over time?
  - Yes, they have become less useful