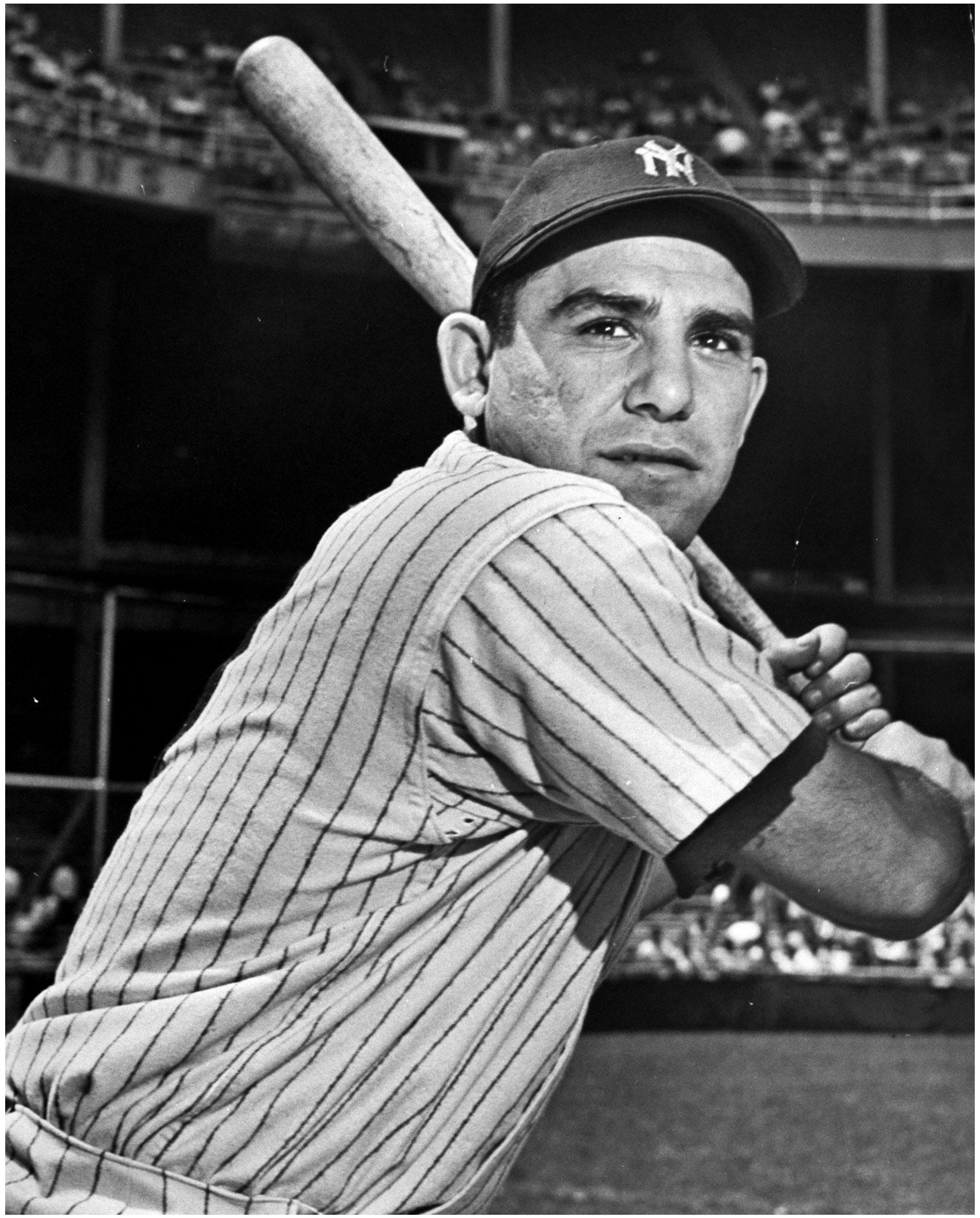
Expectations Formation and Forward Information

Nathan Goldstein
Bar-Ilan University

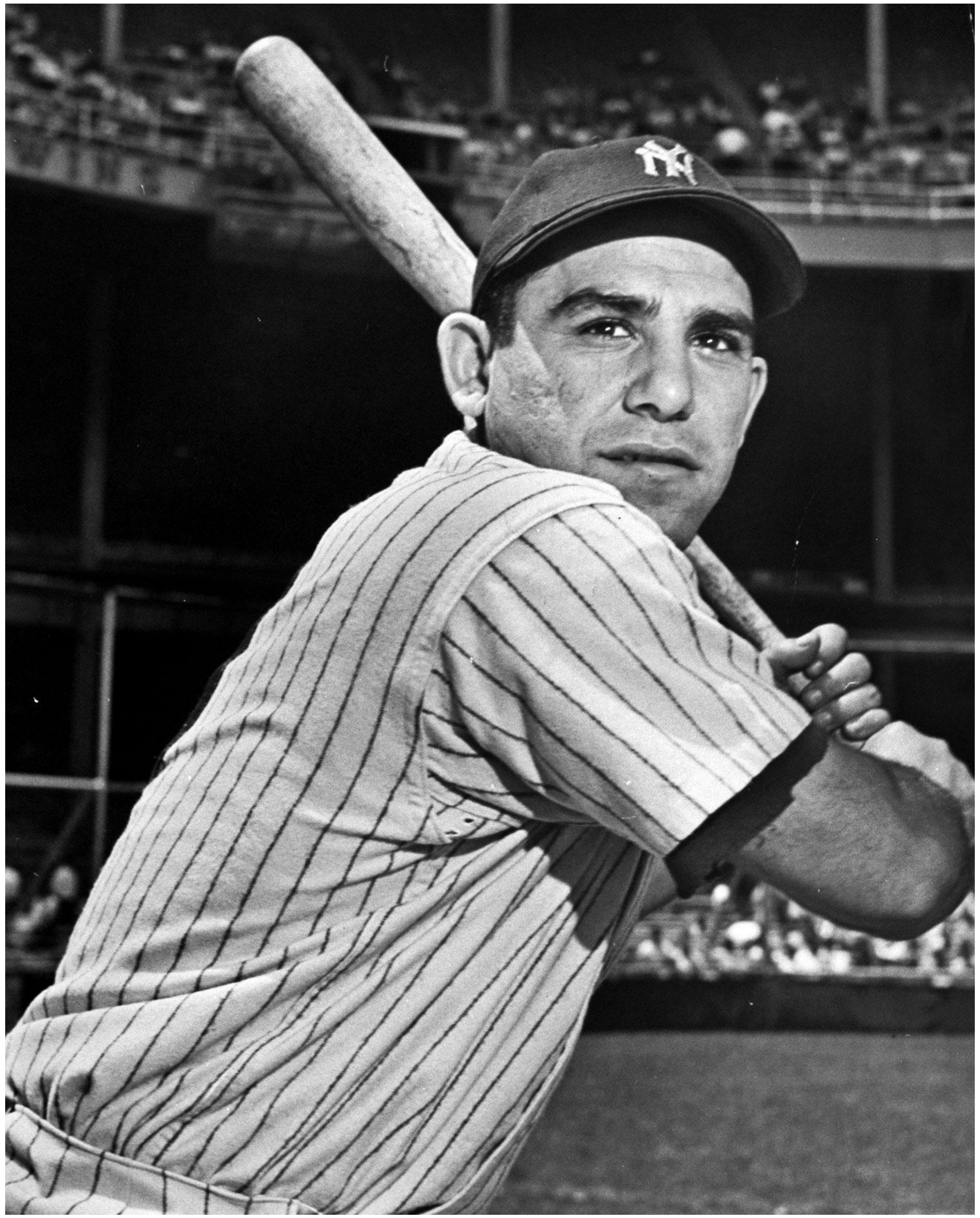
Yuriy Gorodnichenko
UC Berkeley &

CHALLENGES OF FORECASTING



Yogi Berra: “It’s tough to make predictions, especially about the future.”

CHALLENGES OF FORECASTING



Yogi Berra: “It’s tough to make predictions, especially about the future.”

But how do economic agents form expectations?

THE "STANDARD" (NOISY INFORMATION) MODEL

Fundamental (State):

$$x_t = \rho x_{t-1} + \omega_t$$

Signal (Measurement):

$$y_t^i = x_t + \nu_t^i$$

The "Standard" (Noisy Information) Model

Fundamental (State):

$$x_t = \rho x_{t-1} + \omega_t$$

Signal (Measurement):

$$y_t^i = x_t + \nu_t^i$$

Kalman filter:

$$E(x_t|Information_t) \equiv x_{t|t}^i = x_{t|t-1}^i + G\left(y_t^i - x_{t|t-1}^i\right)$$

THE "STANDARD" (NOISY INFORMATION) MODEL

Fundamental (State):

$$x_t = \rho x_{t-1} + \omega_t$$

Signal (Measurement):

$$y_t^i = x_t + \nu_t^i$$

Kalman filter:

$$E(x_t | Information_t) \equiv x_{t|t}^i = x_{t|t-1}^i + G\left(y_t^i - x_{t|t-1}^i\right)$$

Prediction:

$$x_{t+h|t}^i = \rho x_{t+h-1|t}^i = \rho^h x_{t|t}^i$$

The "Standard" (Noisy Information) Model

Fundamental (State):

$$x_t = \rho x_{t-1} + \omega_t$$

Signal (Measurement):

$$y_t^i = x_t + \nu_t^i$$

Kalman filter:

$$E(x_t | Information_t) \equiv x_{t|t}^i = x_{t|t-1}^i + G\left(y_t^i - x_{t|t-1}^i\right)$$

Prediction:

$$x_{t+h|t}^i = \rho x_{t+h-1|t}^i = \rho^h x_{t|t}^i$$

Practice:

$$x_{t+h|t}^i = \overbrace{\rho^h x_{t|t}^i}^{\text{"Model"}} + \{add\ factor\}_{t,h}$$

Add factor is subjective adjustment.

INFLATION FORECASTS

Prediction in the standard noisy-information model:

$$x^i_{t+h|t} = \rho x^i_{t+h-1|t} = \rho^h x^i_{t|t}$$

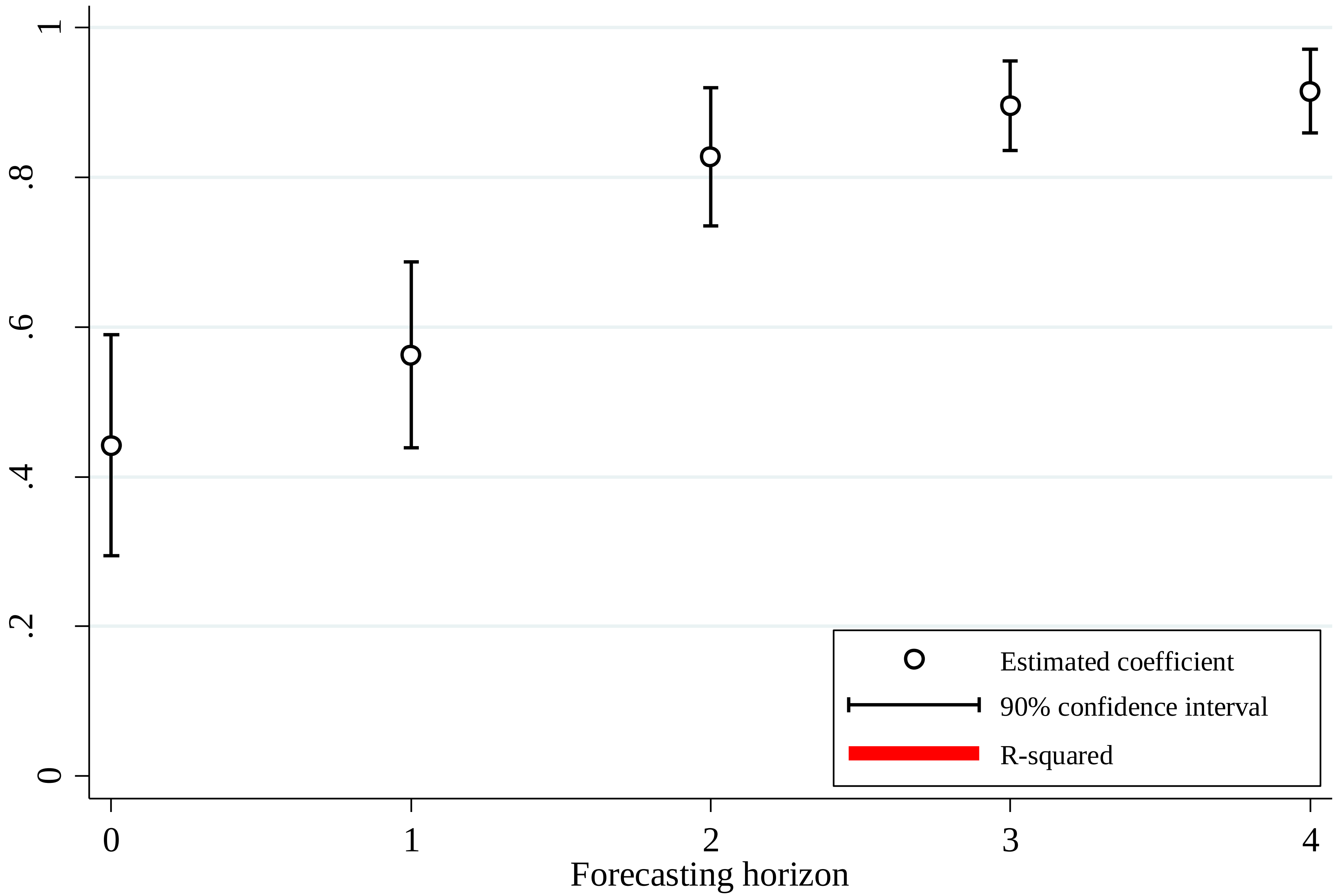
When estimating a regression of the forecast $x^i_{t+h|t}$ on the forecast $x^i_{t+h-1|t}$:

I. *The fit of the regression should be perfect ($R^2 = 1$) for any $h \geq 1$.*

II. *The regression coefficient recovers ρ, the persistence parameter, for any $h \geq 1$.*

SPF INFLATION FORECASTS

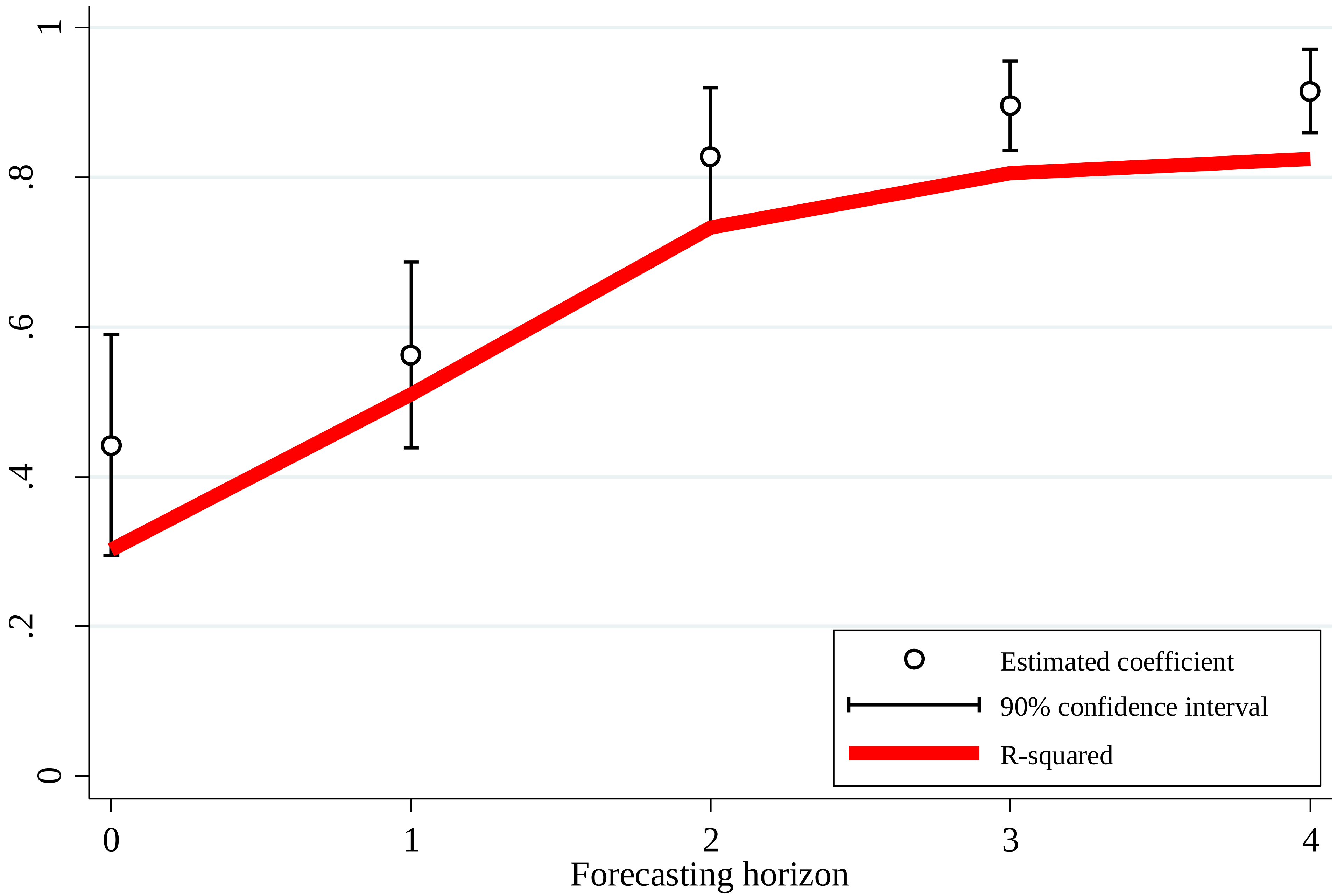
Regression: $x^i_{t+h|t} = \rho x^i_{t+h-1|t} + error$



Note: p-value (equality of estimated coefficients) <0.01

SPF INFLATION FORECASTS

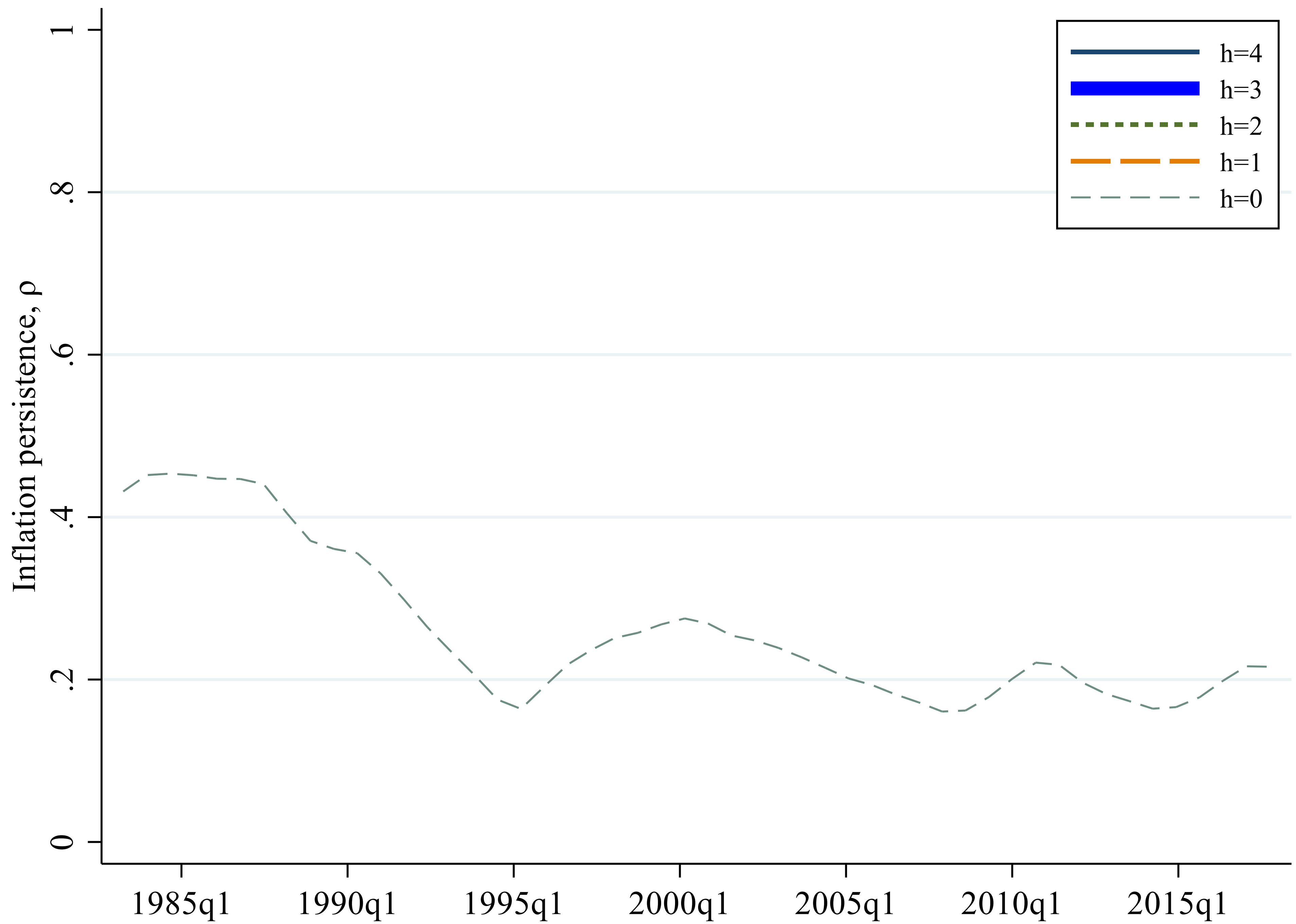
Regression: $x^i_{t+h|t} = \rho x^i_{t+h-1|t} + error$



Note: p-value (equality of estimated coefficients) <0.01

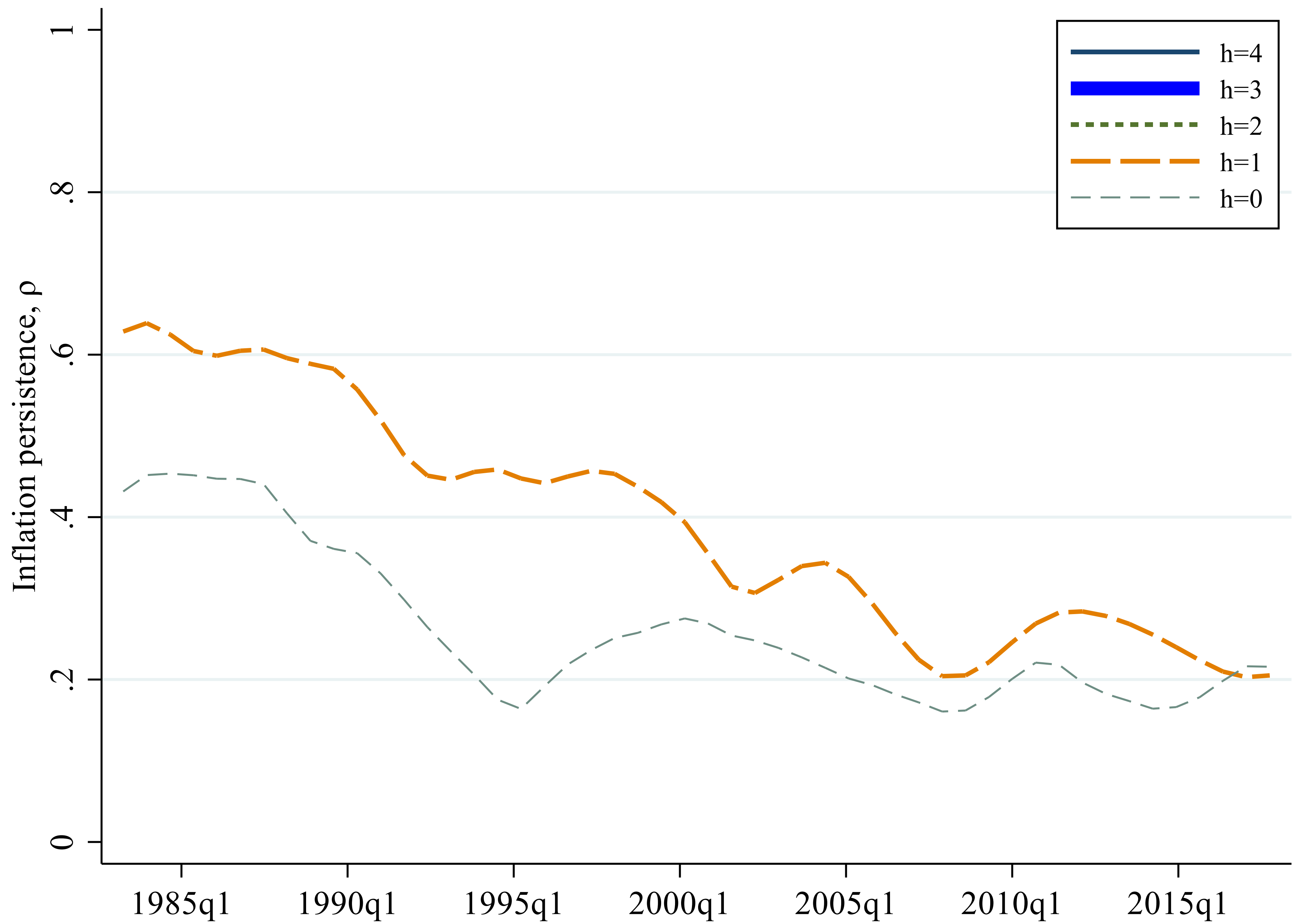
SPF INFLATION FORECASTS

Regression: $x^i_{t+h|t} = \rho x^i_{t+h-1|t} + error$



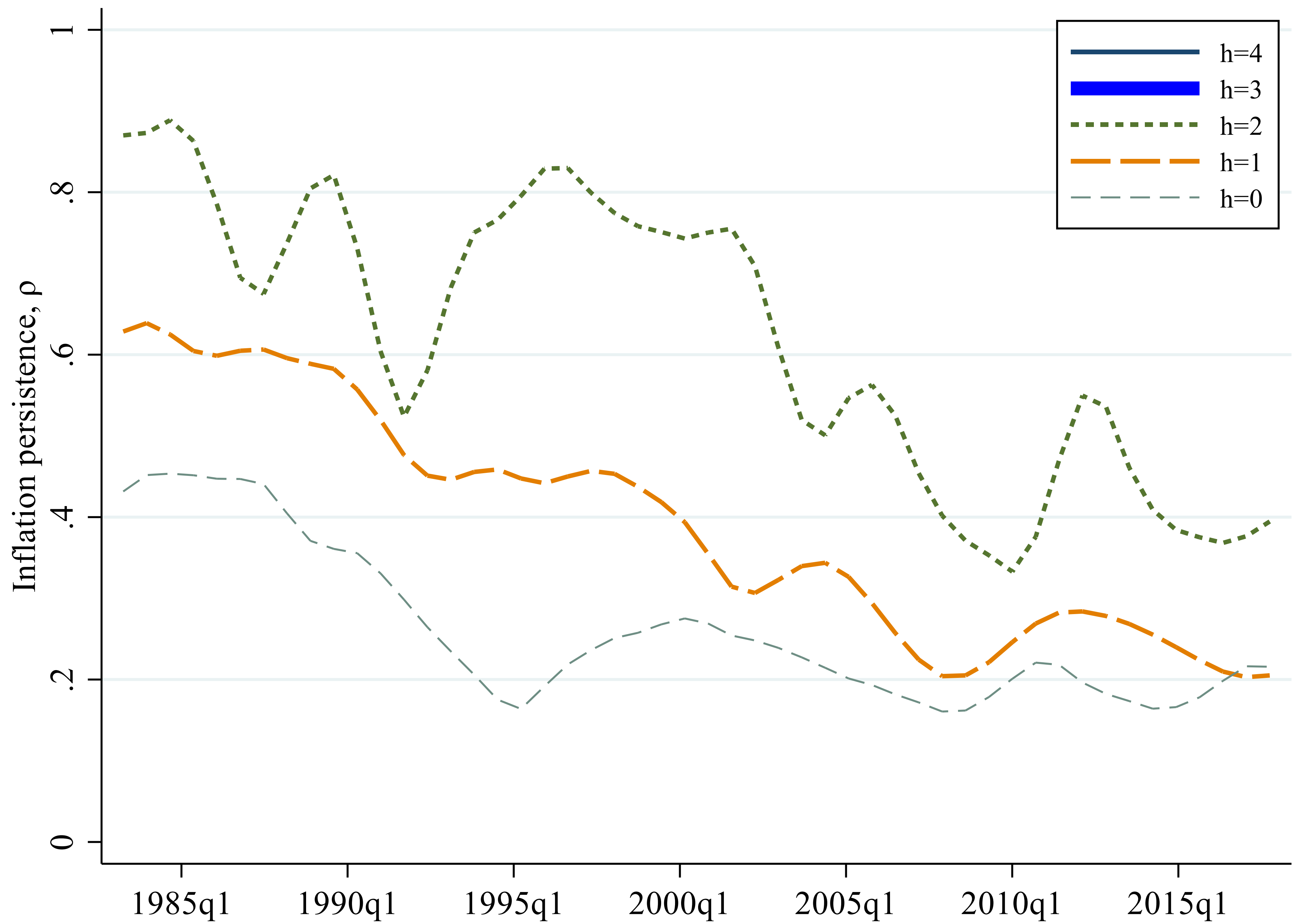
SPF INFLATION FORECASTS

Regression: $x^i_{t+h|t} = \rho x^i_{t+h-1|t} + error$



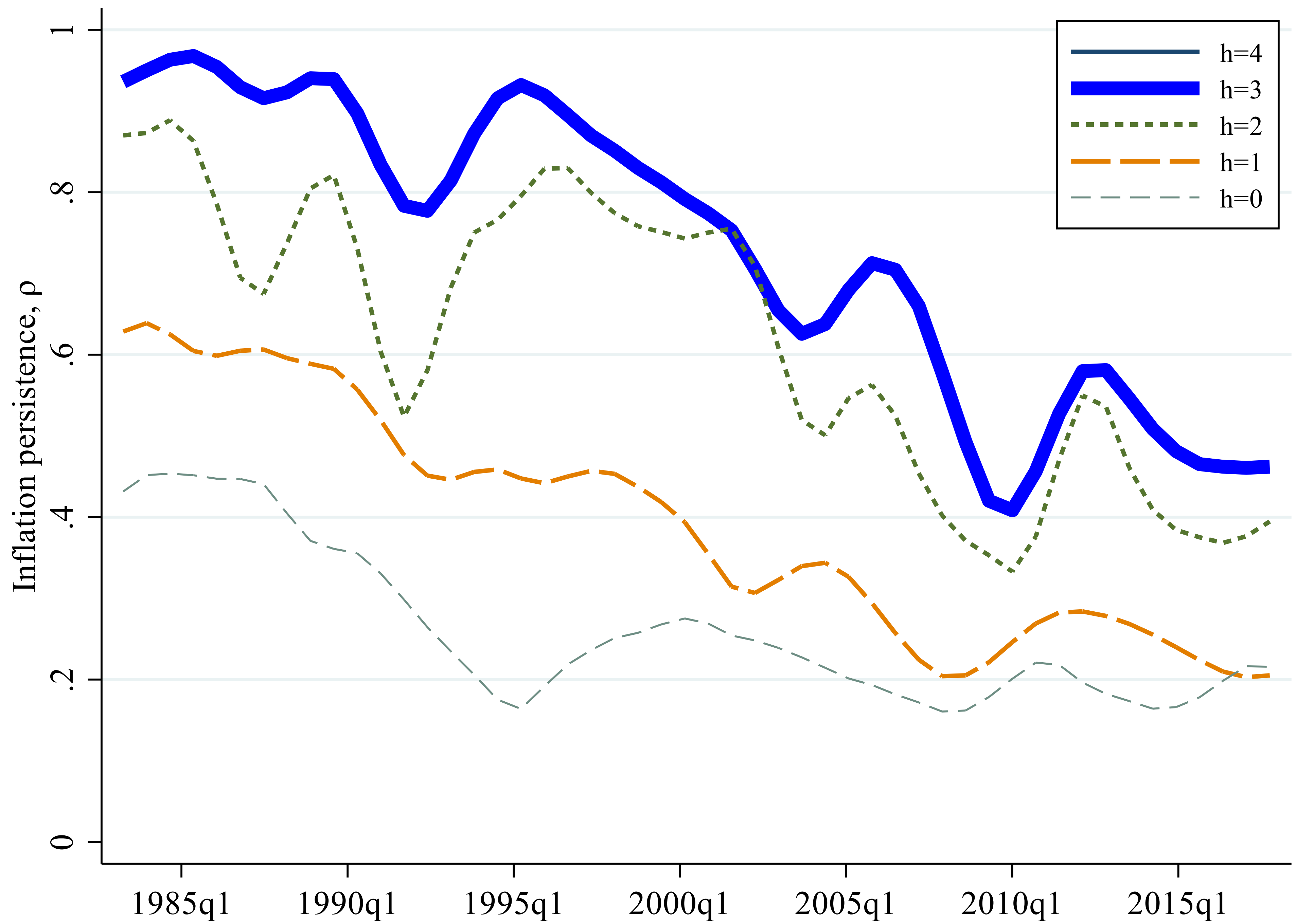
SPF INFLATION FORECASTS

Regression: $x^i_{t+h|t} = \rho x^i_{t+h-1|t} + error$



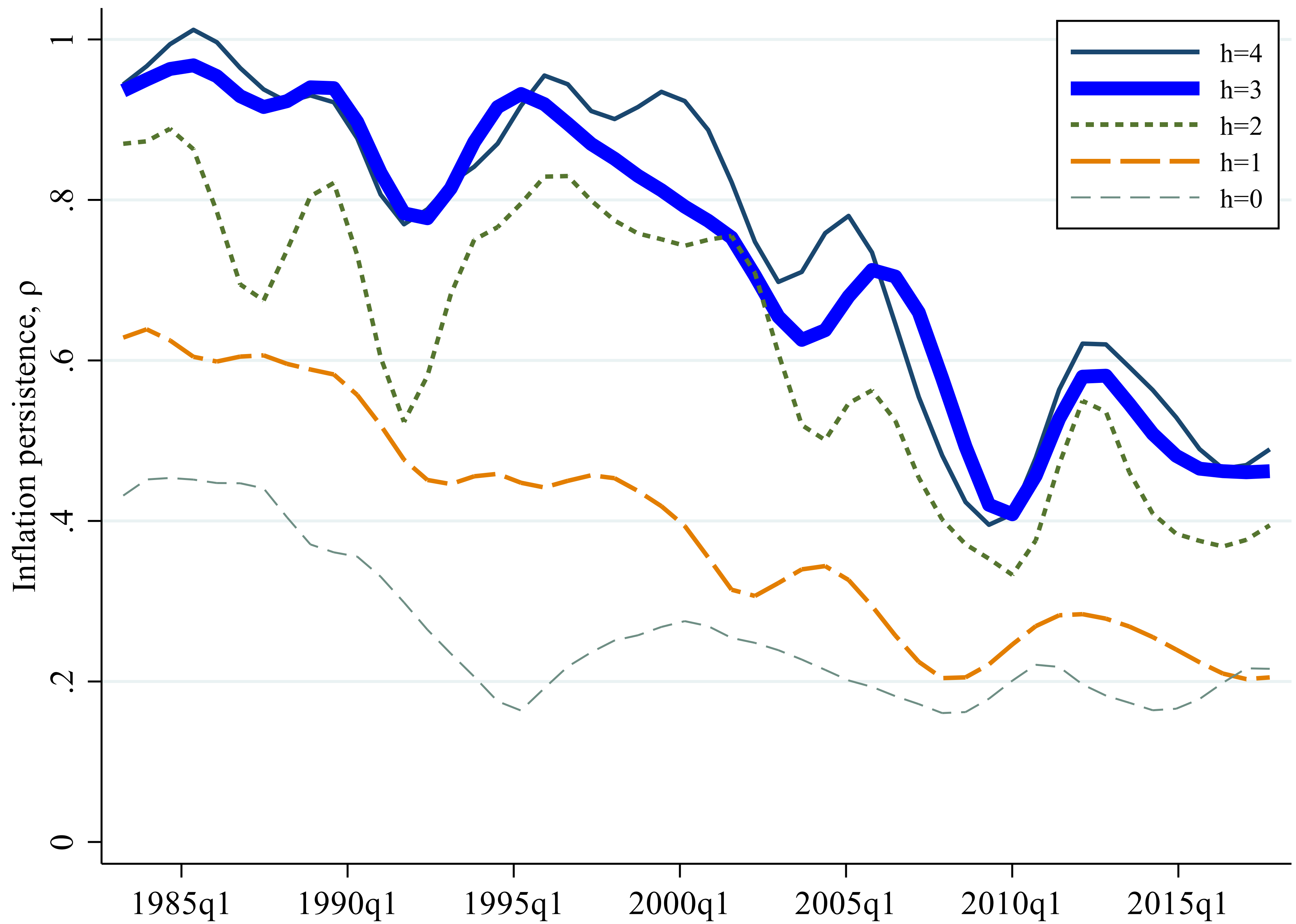
SPF INFLATION FORECASTS

Regression: $x^i_{t+h|t} = \rho x^i_{t+h-1|t} + error$



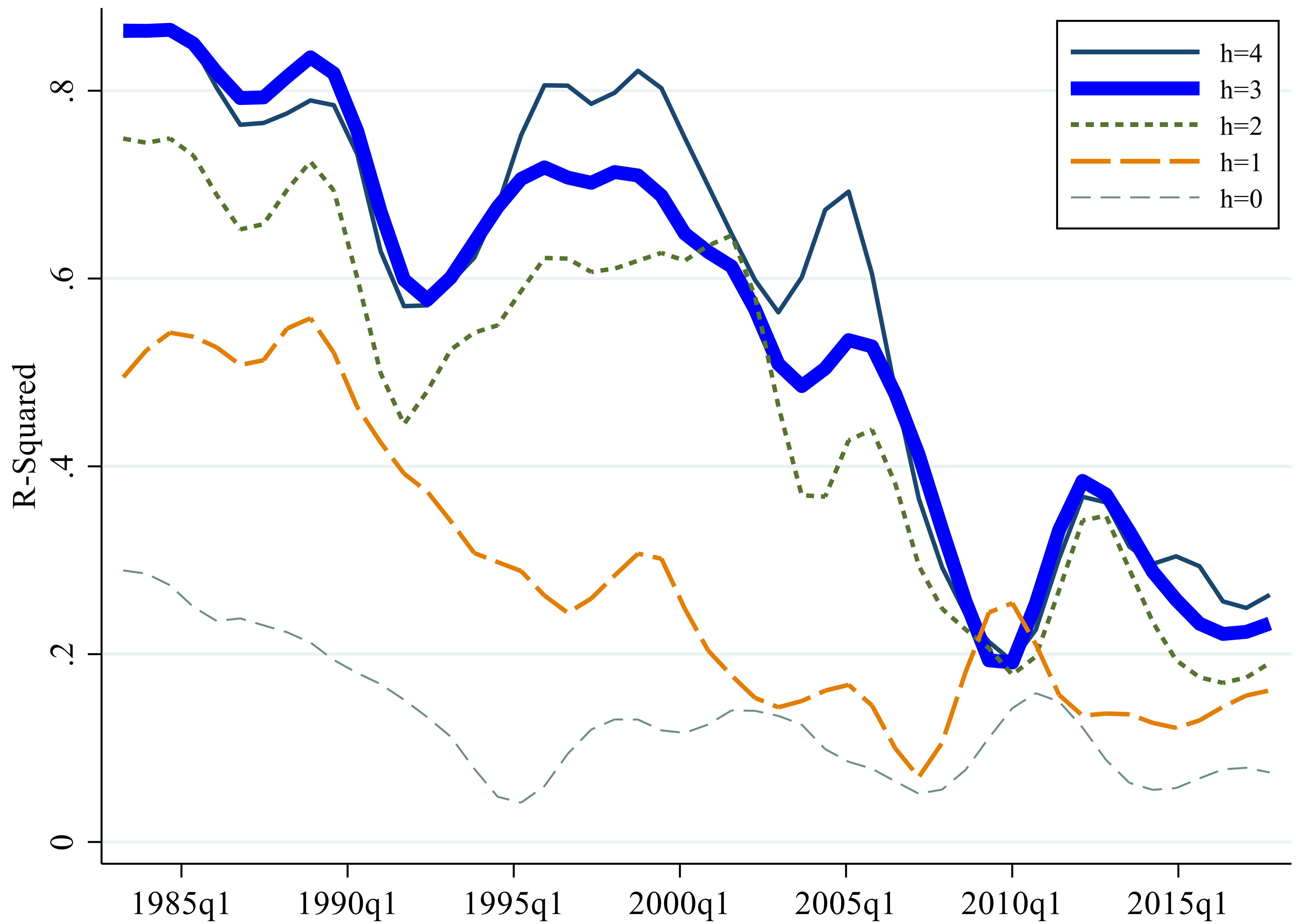
SPF INFLATION FORECASTS

Regression: $x^i_{t+h|t} = \rho x^i_{t+h-1|t} + error$



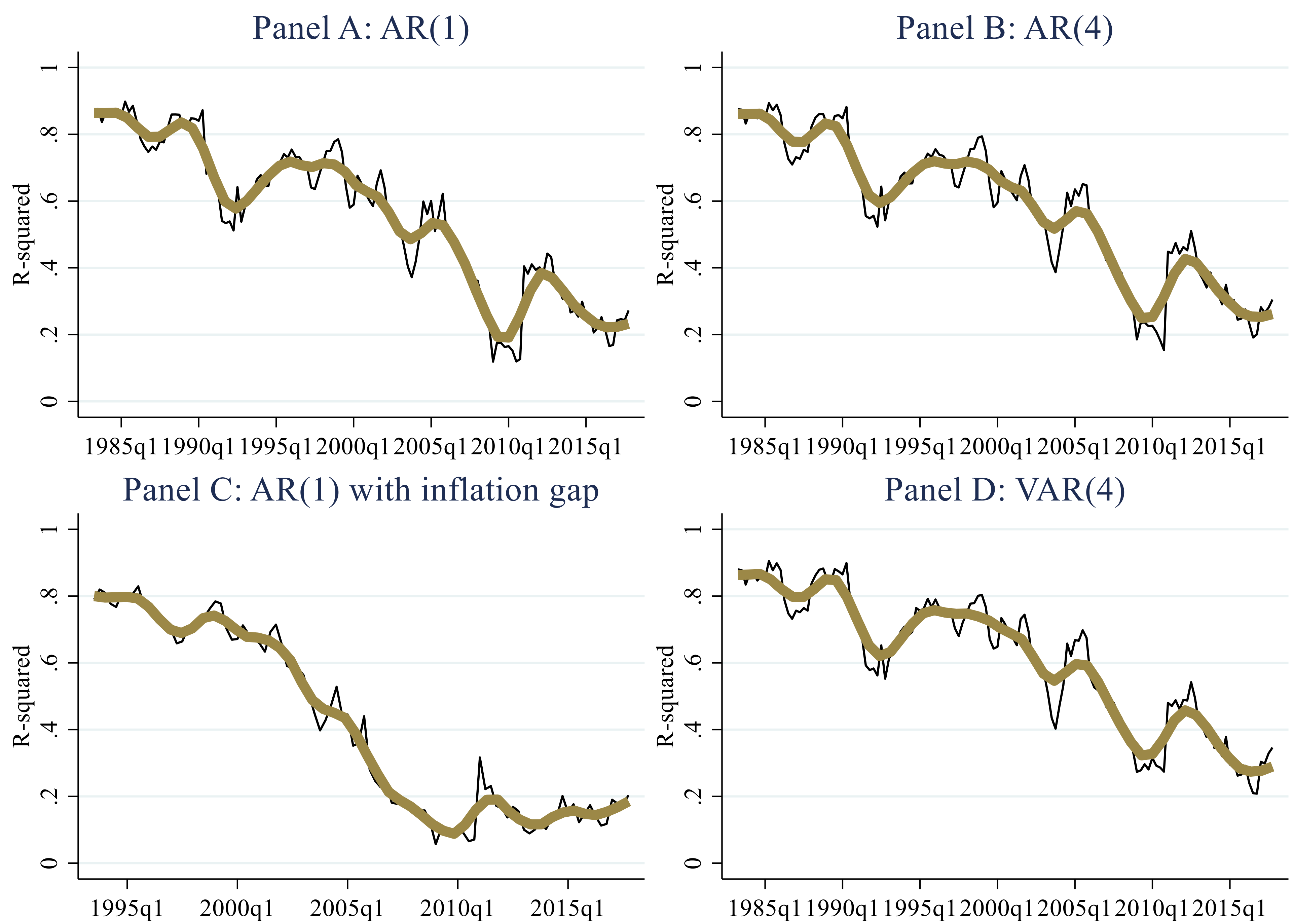
SPF INFLATION FORECASTS

Regression: $x^{i}_{t+h|t} = \rho x^{i}_{t+h-1|t} + error$



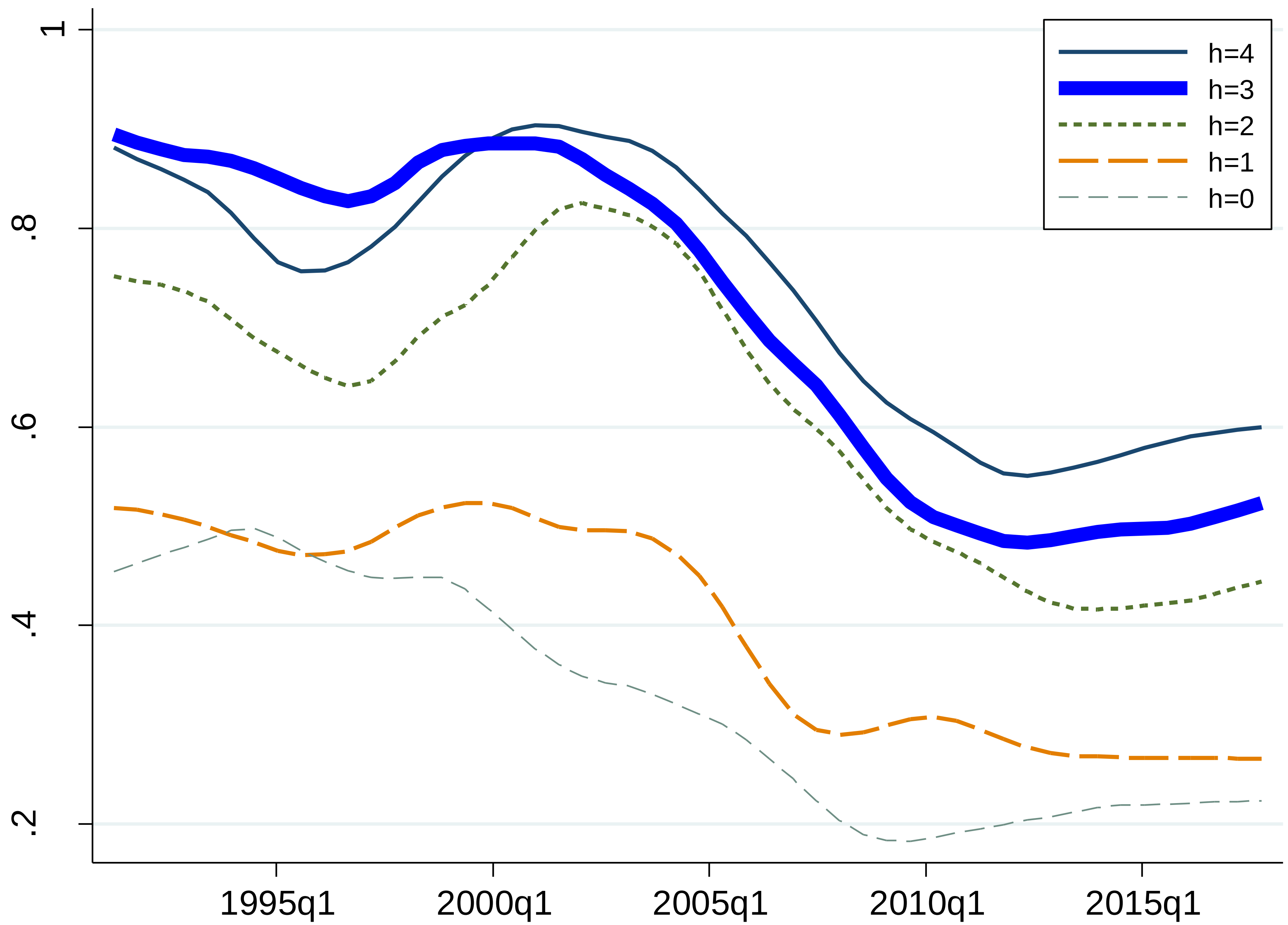
ROBUSTNESS 1: AR(1) CAPTURES WELL THE "MODEL"

Regressions: $x^i_{t+h|t} = \sum_p \boldsymbol{\rho}'_p \boldsymbol{z}^i_{t+h-p|t} + error$



ROBUSTNESS 2: DISAGREEMENT ABOUT THE "MODEL"

Regression: $x^i_{t+h|t} = \rho^i x^i_{t+h-1|t} + error$



Recap

Regression:

$$x^i_{t+h|t} = \rho^h x^i_{t|t} + error_{t,h}$$

Practice (SPF):

$$x^i_{t+h|t} = \rho^h x^i_{t|t} + \{add\ factor\}_{t,h}$$

Recap

Regression:

$$x^i_{t+h|t} = \rho^h x^i_{t|t} + error_{t,h}$$

Practice (SPF): (80% of forecasters do this according to a special survey of SPF)

$$x^i_{t+h|t} = \rho^h x^i_{t|t} + \{add\ factor\}_{t,h}$$

RECAP

Regression:

$$x^i_{t+h|t} = \rho^h x^i_{t|t} + error_{t,h}$$

Practice (SPF): (80% of forecasters do this according to a special survey of SPF)

$$x^i_{t+h|t} = \rho^h x^i_{t|t} + \{add\ factor\}_{t,h}$$

OECD: An add-factor is the adjustment made to equation-based projection over the forecasting period. For example, if an equation has under-predicted a variable in recent periods, then an "add factor" may be added to the equation if it is judged that the equation will under-predict over the forecast period as well. In short, add factors are equation-residuals applied over the forecast period.

Larry Klein: "After the preparation of preliminary predictions from the … Wharton-EFU Model, there is a discussion of the assumptions and properties of the prediction with business and government specialists. A priori information on impending labor disputes, hedge purchasing, production bottlenecks, major economic decisions and similar phenomena are then suggested for further modification of parameter or residual values, and a revised forecast in prepared."

Add factor is information about the future ("forward information", "news", etc.)

NOISY FORWARD INFORMATION

Fundamental (State):

$$x_t = \rho x_{t-1} + \omega_t$$

Signal (Measurement):

$$y^i_{t,t+h} = x_{t+h} + \nu^i_{t,t+h}$$

NOISY FORWARD INFORMATION

Fundamental (State):

$$\boldsymbol{x}_t \equiv \begin{bmatrix} x_{t+H} \\ x_{t+H-1} \\ \vdots \\ x_t \end{bmatrix} = \begin{bmatrix} \rho & 0 & \cdots & 0 \\ 1 & 0 & 0 & 0 \\ \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & 1 & 0 \end{bmatrix} \boldsymbol{x}_{t-1} + S'\omega_{t+H} = \mathrm{P}\boldsymbol{x}_{t-1} + S'\omega_{t+H}$$

Signal (Measurement):

$$\boldsymbol{y}_t^i \equiv \begin{bmatrix} y_{t,t+H}^i \\ y_{t,t+H-1}^i \\ \vdots \\ y_{t,t}^i \end{bmatrix} = \begin{bmatrix} x_{t+H} \\ x_{t+H-1} \\ \vdots \\ x_t \end{bmatrix} + \begin{bmatrix} \nu_{t,t+H}^i \\ \nu_{t,t+H-1}^i \\ \vdots \\ \nu_{t,t}^i \end{bmatrix} = \boldsymbol{x}_t + \boldsymbol{\nu}_t^i$$

NOISY FORWARD INFORMATION

Fundamental (State):

$$\boldsymbol{x}_t \equiv \begin{bmatrix} x_{t+H} \\ x_{t+H-1} \\ \vdots \\ x_t \end{bmatrix} = \begin{bmatrix} \rho & 0 & \cdots & 0 \\ 1 & 0 & 0 & 0 \\ \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & 1 & 0 \end{bmatrix} \boldsymbol{x}_{t-1} + S'\omega_{t+H} = \mathrm{P}\boldsymbol{x}_{t-1} + S'\omega_{t+H}$$

Signal (Measurement):

$$\boldsymbol{y}_t^i \equiv \begin{bmatrix} y_{t,t+H}^i \\ y_{t,t+H-1}^i \\ \vdots \\ y_{t,t}^i \end{bmatrix} = \begin{bmatrix} x_{t+H} \\ x_{t+H-1} \\ \vdots \\ x_t \end{bmatrix} + \begin{bmatrix} \nu_{t,t+H}^i \\ \nu_{t,t+H-1}^i \\ \vdots \\ \nu_{t,t}^i \end{bmatrix} = \boldsymbol{x}_t + \boldsymbol{\nu}_t^i$$

Kalman filter:

$$\boldsymbol{x}_{t|t}^i = \boldsymbol{x}_{t|t-1}^i + \mathrm{G}\left(\boldsymbol{y}_t^i - \boldsymbol{x}_{t|t-1}^i\right)$$

NOISY FORWARD INFORMATION: FORECASTS

$$x^i_{t+h|t} =$$

$$\rho x^i_{t+h-1|t}$$

$$+\left(x^i_{t+h|t-1} - \rho x^i_{t+h-1|t-1}\right)$$

$$+\left(\mathbf{G}_j - \rho \mathbf{G}_{j+1}\right)\left(\boldsymbol{x}_t - \boldsymbol{x}^i_{t|t-1}\right)$$

$$+\left(\mathbf{G}_j - \rho \mathbf{G}_{j+1}\right)\boldsymbol{\nu}^i_t$$

NOISY FORWARD INFORMATION: FORECASTS

$$x^i_{t+h|t} =$$

$$\rho x^i_{t+h-1|t}$$ standard

$$+\left(x^i_{t+h|t-1} - \rho x^i_{t+h-1|t-1}\right)$$

$$+\left(\mathbf{G}_j - \rho \mathbf{G}_{j+1}\right)\left(\boldsymbol{x}_t - \boldsymbol{x}^i_{t|t-1}\right)$$

$$+\left(\mathbf{G}_j - \rho \mathbf{G}_{j+1}\right)\boldsymbol{\nu}^i_t$$

NOISY FORWARD INFORMATION: FORECASTS

$$x^i_{t+h|t} =$$

$$\rho x^i_{t+h-1|t}$$ standard

$$+\left(x^i_{t+h|t-1} - \rho x^i_{t+h-1|t-1}\right)$$ revisions in weights on past signals

$$+\left(\mathbf{G}_j - \rho \mathbf{G}_{j+1}\right)\left(\boldsymbol{x}_t - \boldsymbol{x}^i_{t|t-1}\right)$$

$$+\left(\mathbf{G}_j - \rho \mathbf{G}_{j+1}\right)\boldsymbol{\nu}^i_t$$

NOISY FORWARD INFORMATION: FORECASTS

$x^i_{t+h|t} =$

$\rho x^i_{t+h-1\|t}$	standard
$+\left(x^i_{t+h\|t-1} - \rho x^i_{t+h-1\|t-1}\right)$	revisions in weights on past signals
$+\left(\mathbf{G}_j - \rho \mathbf{G}_{j+1}\right)\left(\boldsymbol{x}_t - \boldsymbol{x}^i_{t\|t-1}\right)$	information in new forward signals
$+\left(\mathbf{G}_j - \rho \mathbf{G}_{j+1}\right)\boldsymbol{\nu}^i_t$	

Noisy forward information: forecasts

$x^i_{t+h|t} =$

$\rho x^i_{t+h-1|t}$	standard
$+\left(x^i_{t+h|t-1} - \rho x^i_{t+h-1|t-1}\right)$	revisions in weights on past signals
$+\left(\mathbf{G}_j - \rho \mathbf{G}_{j+1}\right)\left(\boldsymbol{x}_t - \boldsymbol{x}^i_{t|t-1}\right)$	information in new forward signals
$+\left(\mathbf{G}_j - \rho \mathbf{G}_{j+1}\right)\boldsymbol{\nu}^i_t$	idiosyncratic noise

NOISY FORWARD INFORMATION: FORECASTS

$x^i_{t+h|t} =$

$\rho x^i_{t+h-1|t}$	standard
$+\left(x^i_{t+h|t-1} - \rho x^i_{t+h-1|t-1}\right)$	revisions in weights on past signals
$+\left(\mathbf{G}_j - \rho \mathbf{G}_{j+1}\right)\left(\boldsymbol{x}_t - \boldsymbol{x}^i_{t|t-1}\right)$	information in new forward signals
$+\left(\mathbf{G}_j - \rho \mathbf{G}_{j+1}\right)\boldsymbol{\nu}^i_t$	idiosyncratic noise

$$x^i_{t+h|t} = \rho x^i_{t+h-1|t} + \{\textbf{\textit{new terms}}\}$$

NOISY FORWARD INFORMATION: FORECASTS

$x^i_{t+h|t} =$

$\rho x^i_{t+h-1|t}$	standard
$+\left(x^i_{t+h|t-1} - \rho x^i_{t+h-1|t-1}\right)$	revisions in weights on past signals
$+\left(\mathbf{G}_j - \rho \mathbf{G}_{j+1}\right)\left(\boldsymbol{x}_t - \boldsymbol{x}^i_{t|t-1}\right)$	information in new forward signals
$+\left(\mathbf{G}_j - \rho \mathbf{G}_{j+1}\right)\boldsymbol{\nu}^i_t$	idiosyncratic noise

$$x^i_{t+h|t} = \rho x^i_{t+h-1|t} + \{\boldsymbol{new\ terms}\} \Rightarrow \begin{cases} \hat{\rho} \neq \rho \\ R^2 < 1 \end{cases}$$

NOISY FORWARD INFORMATION: FORECASTS

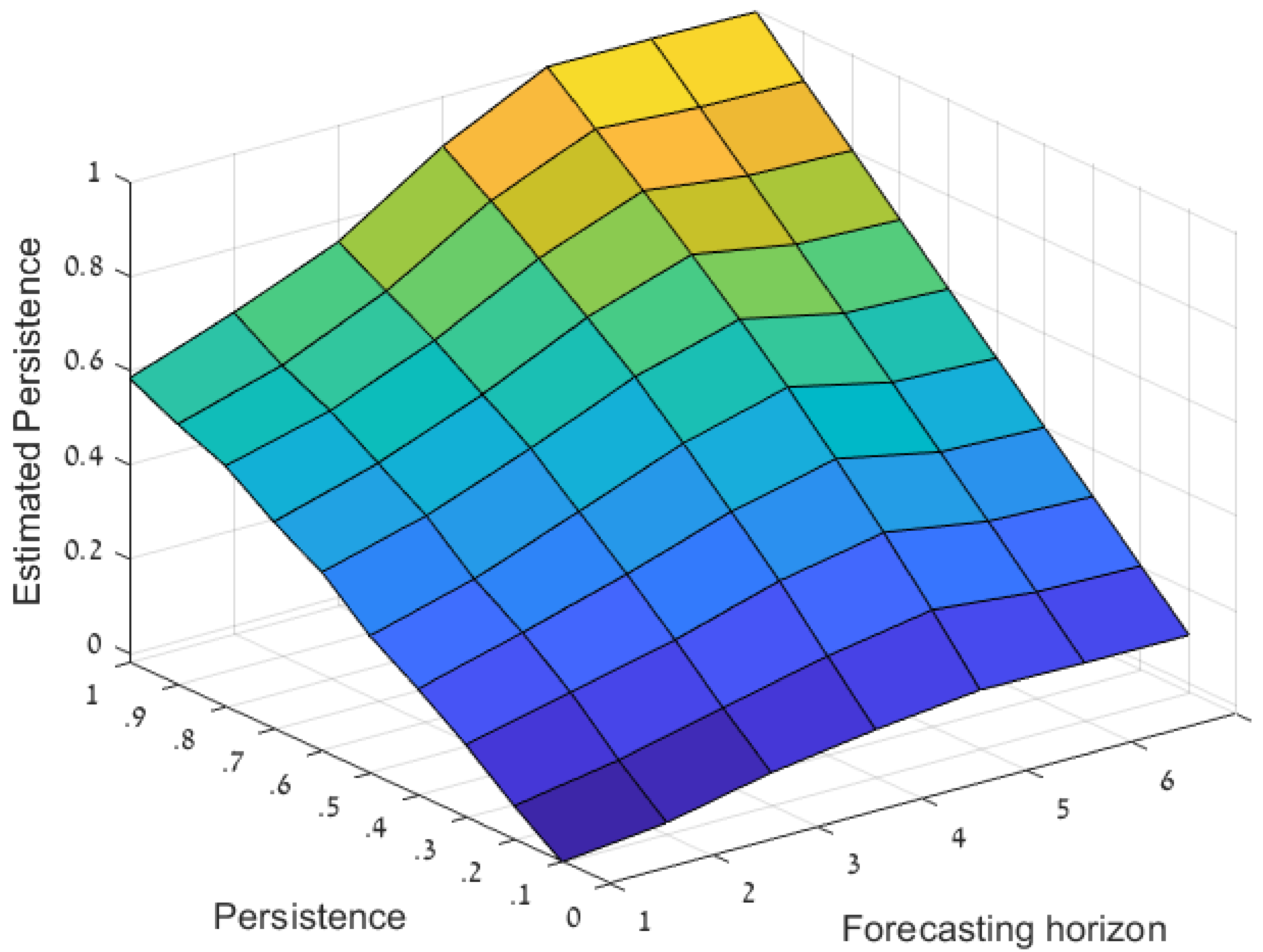
$x^i_{t+h|t} =$

$\rho x^i_{t+h-1|t}$	standard
$+\left(x^i_{t+h|t-1} - \rho x^i_{t+h-1|t-1}\right)$	revisions in weights on past signals
$+\left(\mathbf{G}_j - \rho \mathbf{G}_{j+1}\right)\left(\boldsymbol{x}_t - \boldsymbol{x}^i_{t|t-1}\right)$	information in new forward signals
$+\left(\mathbf{G}_j - \rho \mathbf{G}_{j+1}\right)\boldsymbol{\nu}^i_t$	idiosyncratic noise

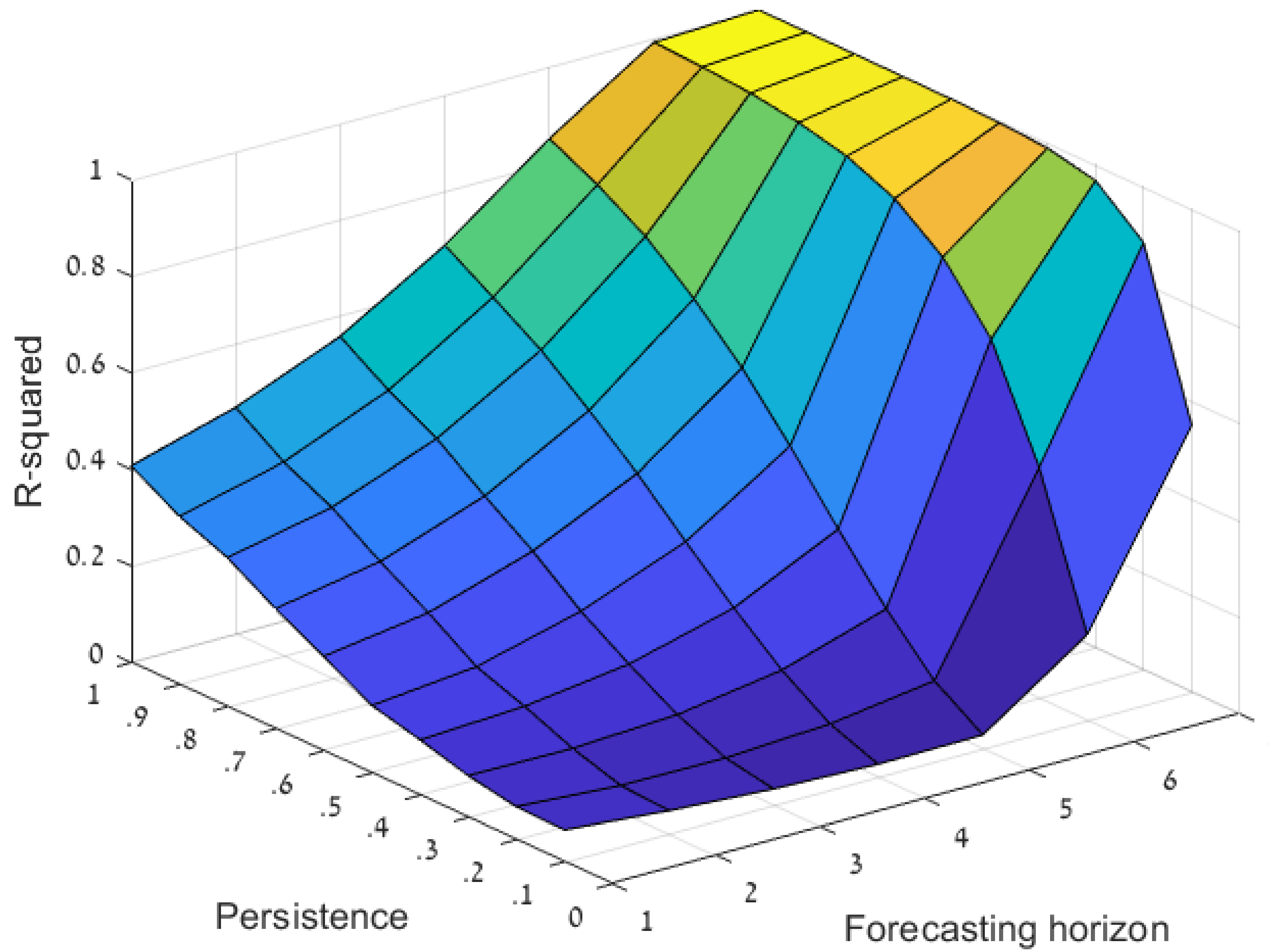
$$x^i_{t+h|t} = \rho x^i_{t+h-1|t} + \{\boldsymbol{new\ terms}\} \Rightarrow \begin{cases} \hat{\rho} \neq \rho \ (and\ \hat{\rho}_h < \hat{\rho}_{h+1}) \\ R^2 < 1 \ (and\ R^2_h < R^2_{h+1}) \end{cases}$$

Intuition: as $h \uparrow$, signals become less precise $\Rightarrow$ for some H we get $x^i_{t+H|t} \approx \rho x^i_{t+H-1|t}$

NOISY FORWARD INFORMATION: SIMULATIONS



NOISY FORWARD INFORMATION: SIMULATIONS



How can one make it useful?

How can one make it useful?

An insight from Goldstein (2021):

How can one make it useful?

An insight from Goldstein (2021):

Model:

$$\boldsymbol{x}_{t|t}^{i} = \boldsymbol{x}_{t|t-1}^{i} + \mathrm{G}\left(\boldsymbol{y}_{t}^{i} - \boldsymbol{x}_{t|t-1}^{i}\right) = \boldsymbol{x}_{t|t-1}^{i} + \mathrm{G}\left(\boldsymbol{x}_{t} + \boldsymbol{v}_{t}^{i} - \boldsymbol{x}_{t|t-1}^{i}\right) = (I - G)\boldsymbol{x}_{t|t-1}^{i} + G\left(\boldsymbol{x}_{t} + \boldsymbol{v}_{t}^{i}\right)$$

How can one make it useful?

An insight from Goldstein (2021):

Model:

$$\boldsymbol{x}_{t|t}^{i} = \boldsymbol{x}_{t|t-1}^{i} + \mathrm{G}\left(\boldsymbol{y}_{t}^{i} - \boldsymbol{x}_{t|t-1}^{i}\right) = \boldsymbol{x}_{t|t-1}^{i} + \mathrm{G}\left(\boldsymbol{x}_{t} + \boldsymbol{\nu}_{t}^{i} - \boldsymbol{x}_{t|t-1}^{i}\right) = (I - G)\boldsymbol{x}_{t|t-1}^{i} + G\left(\boldsymbol{x}_{t} + \boldsymbol{\nu}_{t}^{i}\right)$$

Take the average across individuals (hence we drop superscript i) and obtain:

$$\boldsymbol{x}_{t|t} = (I - G)\boldsymbol{x}_{t|t-1} + G\boldsymbol{x}_{t}$$

How can one make it useful?

An insight from Goldstein (2021):

Model:

$$\boldsymbol{x}_{t|t}^{i} = \boldsymbol{x}_{t|t-1}^{i} + \mathrm{G}\left(\boldsymbol{y}_{t}^{i} - \boldsymbol{x}_{t|t-1}^{i}\right) = \boldsymbol{x}_{t|t-1}^{i} + \mathrm{G}\left(\boldsymbol{x}_{t} + \boldsymbol{v}_{t}^{i} - \boldsymbol{x}_{t|t-1}^{i}\right) = (I - G)\boldsymbol{x}_{t|t-1}^{i} + G\left(\boldsymbol{x}_{t} + \boldsymbol{v}_{t}^{i}\right)$$

Take the average across individuals (hence we drop superscript i) and obtain:

$$\boldsymbol{x}_{t|t} = (I - G)\boldsymbol{x}_{t|t-1} + G\boldsymbol{x}_{t}$$

Subtract the bottom equation from top equation:

$$\boldsymbol{x}_{t|t}^{i} - \boldsymbol{x}_{t|t} = (I - G)\left(\boldsymbol{x}_{t|t-1}^{i} - \boldsymbol{x}_{t|t-1}\right) + G\boldsymbol{v}_{t}^{i}$$

How can one make it useful?

An insight from Goldstein (2021):

Model:

$$\boldsymbol{x}_{t|t}^{i} = \boldsymbol{x}_{t|t-1}^{i} + \mathrm{G}\left(\boldsymbol{y}_{t}^{i} - \boldsymbol{x}_{t|t-1}^{i}\right) = \boldsymbol{x}_{t|t-1}^{i} + \mathrm{G}\left(\boldsymbol{x}_{t} + \boldsymbol{\nu}_{t}^{i} - \boldsymbol{x}_{t|t-1}^{i}\right) = (I - G)\boldsymbol{x}_{t|t-1}^{i} + G\left(\boldsymbol{x}_{t} + \boldsymbol{\nu}_{t}^{i}\right)$$

Take the average across individuals (hence we drop superscript i) and obtain:

$$\boldsymbol{x}_{t|t} = (I - G)\boldsymbol{x}_{t|t-1} + G\boldsymbol{x}_{t}$$

Subtract the bottom equation from top equation:

$$\boldsymbol{x}_{t|t}^{i} - \boldsymbol{x}_{t|t} = (I - G)\left(\boldsymbol{x}_{t|t-1}^{i} - \boldsymbol{x}_{t|t-1}\right) + G\boldsymbol{\nu}_{t}^{i}$$

We can estimate this equation-by-equation with OLS and recover $I - G$

$$x_{t+h|t}^{i} - x_{t+h|t} = \beta_0\left(x_{t+H|t-1}^{i} - x_{t+H|t-1}\right) + \beta_1\left(x_{t+H-1|t-1}^{i} - x_{t+H-1|t-1}\right) +$$

$$\ldots + \beta_H\left(x_{t|t-1}^{i} - x_{t|t-1}\right) + error_t$$

How can one make it useful?

An insight from Goldstein (2021):

Model:

$$\boldsymbol{x}_{t|t}^{i} = \boldsymbol{x}_{t|t-1}^{i} + \mathrm{G}\left(\boldsymbol{y}_{t}^{i} - \boldsymbol{x}_{t|t-1}^{i}\right) = \boldsymbol{x}_{t|t-1}^{i} + \mathrm{G}\left(\boldsymbol{x}_{t} + \boldsymbol{\nu}_{t}^{i} - \boldsymbol{x}_{t|t-1}^{i}\right) = (I - G)\boldsymbol{x}_{t|t-1}^{i} + G\left(\boldsymbol{x}_{t} + \boldsymbol{\nu}_{t}^{i}\right)$$

Take the average across individuals (hence we drop superscript i) and obtain:

$$\boldsymbol{x}_{t|t} = (I - G)\boldsymbol{x}_{t|t-1} + G\boldsymbol{x}_{t}$$

Subtract the bottom equation from top equation:

$$\boldsymbol{x}_{t|t}^{i} - \boldsymbol{x}_{t|t} = (I - G)\left(\boldsymbol{x}_{t|t-1}^{i} - \boldsymbol{x}_{t|t-1}\right) + G\boldsymbol{\nu}_{t}^{i}$$

We can estimate this equation-by-equation with OLS and recover $I - G$

$$x_{t+h|t}^{i} - x_{t+h|t} = \beta_0\left(x_{t+H|t-1}^{i} - x_{t+H|t-1}\right) + \beta_1\left(x_{t+H-1|t-1}^{i} - x_{t+H-1|t-1}\right) +$$

$$\ldots + \beta_H\left(x_{t|t-1}^{i} - x_{t|t-1}\right) + error_t$$

Standard noisy-info model: $x_{t+h|t}^{i} - x_{t+h|t} = \beta\left(x_{t+h|t-1}^{i} - x_{t+h|t-1}\right) + error_t$

A TEST OF FORWARD INFORMATION

Dependent variable:	$x^i_{t+1\|t} - x_{t+1\|t}$	$x^i_{t+2\|t} - x_{t+2\|t}$	$x^i_{t+3\|t} - x_{t+4\|t}$	$x^i_{t+4\|t} - x_{t+4\|t}$
$x^i_{t\|t-1} - x_{t\|t-1}$	-0.013	-0.013	-0.063***	-0.062***
	(0.025)	(0.021)	(0.012)	(0.020)
$x^i_{t+1\|t-1} - x_{t+1\|t-1}$	0.220***	0.003	0.021	0.032
	(0.052)	(0.040)	(0.044)	(0.034)
$x^i_{t+2\|t-1} - x_{t+2\|t-1}$	0.130***	0.458***	-0.095**	-0.056*
	(0.050)	(0.057)	(0.046)	(0.032)
$x^i_{t+3\|t-1} - x_{t+3\|t-1}$	-0.126**	-0.120**	0.486***	0.103*
	(0.061)	(0.057)	(0.069)	(0.059)
$x^i_{t+4\|t-1} - x_{t+4\|t-1}$	0.071	0.056	0.037	0.362***
	(0.066)	(0.043)	(0.038)	(0.044)
Constant	-0.008	-0.001	0.005	0.007
	(0.009)	(0.006)	(0.008)	(0.008)
Obs.	3,854	3,856	3,855	3,853
R^2	0.053	0.146	0.213	0.178
BIC	10,515	8,565	7,434	7,323
BIC for standard noisy info	10,822	8,763	7,635	7,484

ANOTHER TEST OF FORWARD INFORMATION

Dependent variable: $x^i_{t\|t} - x_{t\|t}$ (backcasts)	Full Sample	1980s	1990s	2000s	2010s
$x^i_{t\|t-1} - x_{t\|t-1}$	0.018	0.118	0.017**	-0.000	0.002**
	(0.014)	(0.082)	(0.008)	(0.000)	(0.001)
$x^i_{t+1\|t-1} - x_{t+1\|t-1}$	0.009	0.014	0.032**	0.000	0.002
	(0.016)	(0.084)	(0.015)	(0.000)	(0.001)
$x^i_{t+2\|t-1} - x_{t+2\|t-1}$	-0.029	-0.202**	-0.034***	0.001	0.000
	(0.018)	(0.085)	(0.013)	(0.001)	(0.003)
$x^i_{t+3\|t-1} - x_{t+3\|t-1}$	0.009	0.078	0.013	-0.000	0.002
	(0.018)	(0.123)	(0.015)	(0.001)	(0.002)
$x^i_{t+4\|t-1} - x_{t+4\|t-1}$	0.001	-0.007	-0.009	-0.000	0.002**
	(0.014)	(0.067)	(0.009)	(0.001)	(0.001)
Constant	0.004	0.027	0.002	0.000	0.001
	(0.005)	(0.037)	(0.002)	(0.000)	(0.002)
Obs.	3,849	559	1,068	1,272	950
R^2	0.004	0.023	0.023	0.003	0.002

Intuition of the test: when all forecasters observe realized inflation in the same way, the deviation of their backcasts from the mean should not be persistent.

How to extract forward information?

Recall $x^i_{t+h|t} = \rho x^i_{t+h-1|t} + residual$

The residual measures forward information:

$$\boldsymbol{residual} \equiv FI^i_{t+h|t} = \left(x^i_{t+h|t-1} - \rho x^i_{t+h-1|t-1}\right) + (\mathbf{G}_{h+1} - \rho \mathbf{G}_h)\left(\boldsymbol{x}_t - \boldsymbol{x}^i_{t|t-1}\right) + (\mathbf{G}_{h+1} - \rho \mathbf{G}_h)\boldsymbol{\nu}^i_t$$

How to extract forward information?

Recall $x^i_{t+h|t} = \rho x^i_{t+h-1|t} + residual$

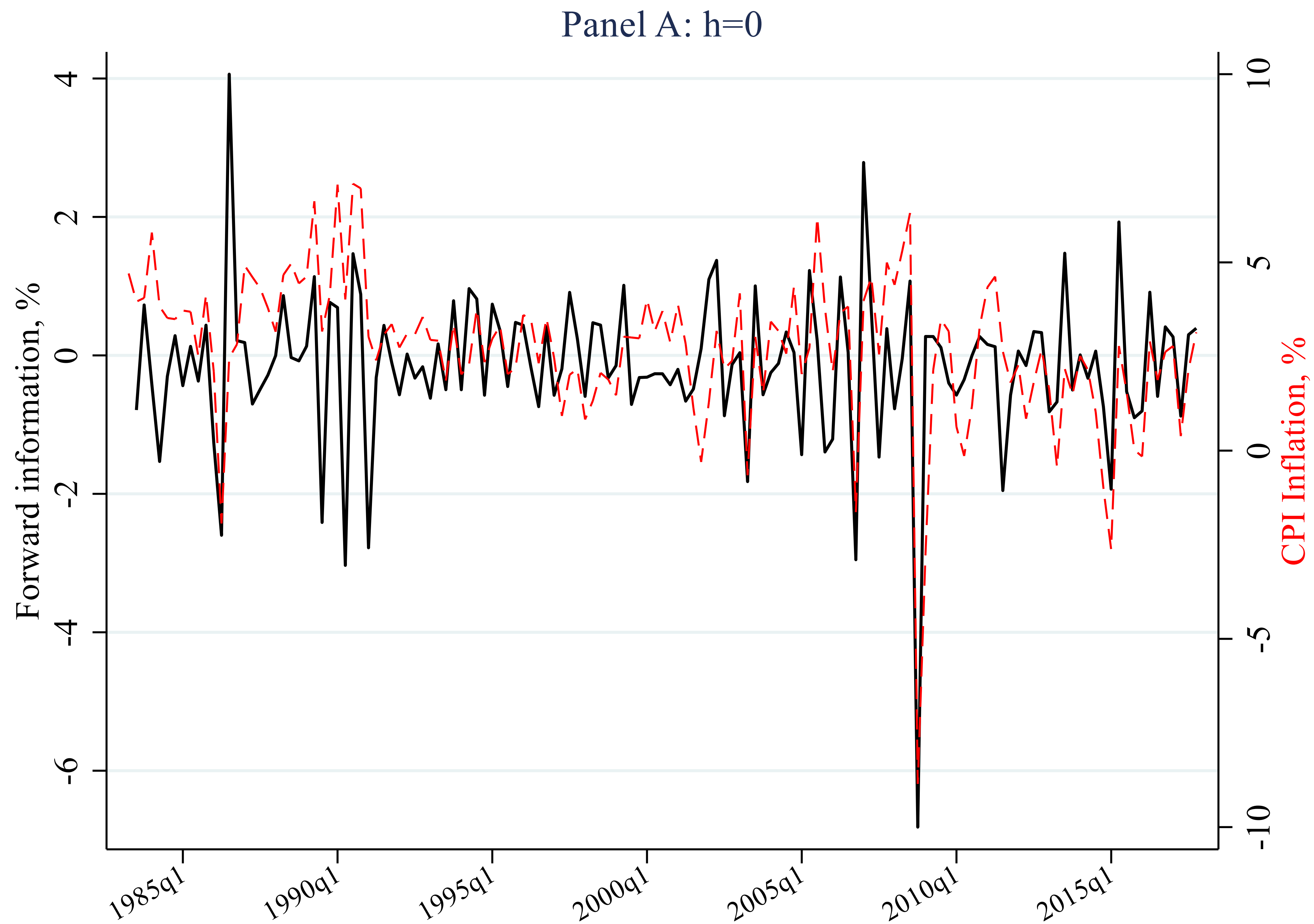
The residual measures forward information:

$$\boldsymbol{residual} \equiv FI^i_{t+h|t} = \left(x^i_{t+h|t-1} - \rho x^i_{t+h-1|t-1}\right) + (\mathbf{G}_{h+1} - \rho \mathbf{G}_h)\left(\boldsymbol{x}_t - \boldsymbol{x}^i_{t|t-1}\right) + (\mathbf{G}_{h+1} - \rho \mathbf{G}_h)\boldsymbol{\nu}^i_t$$

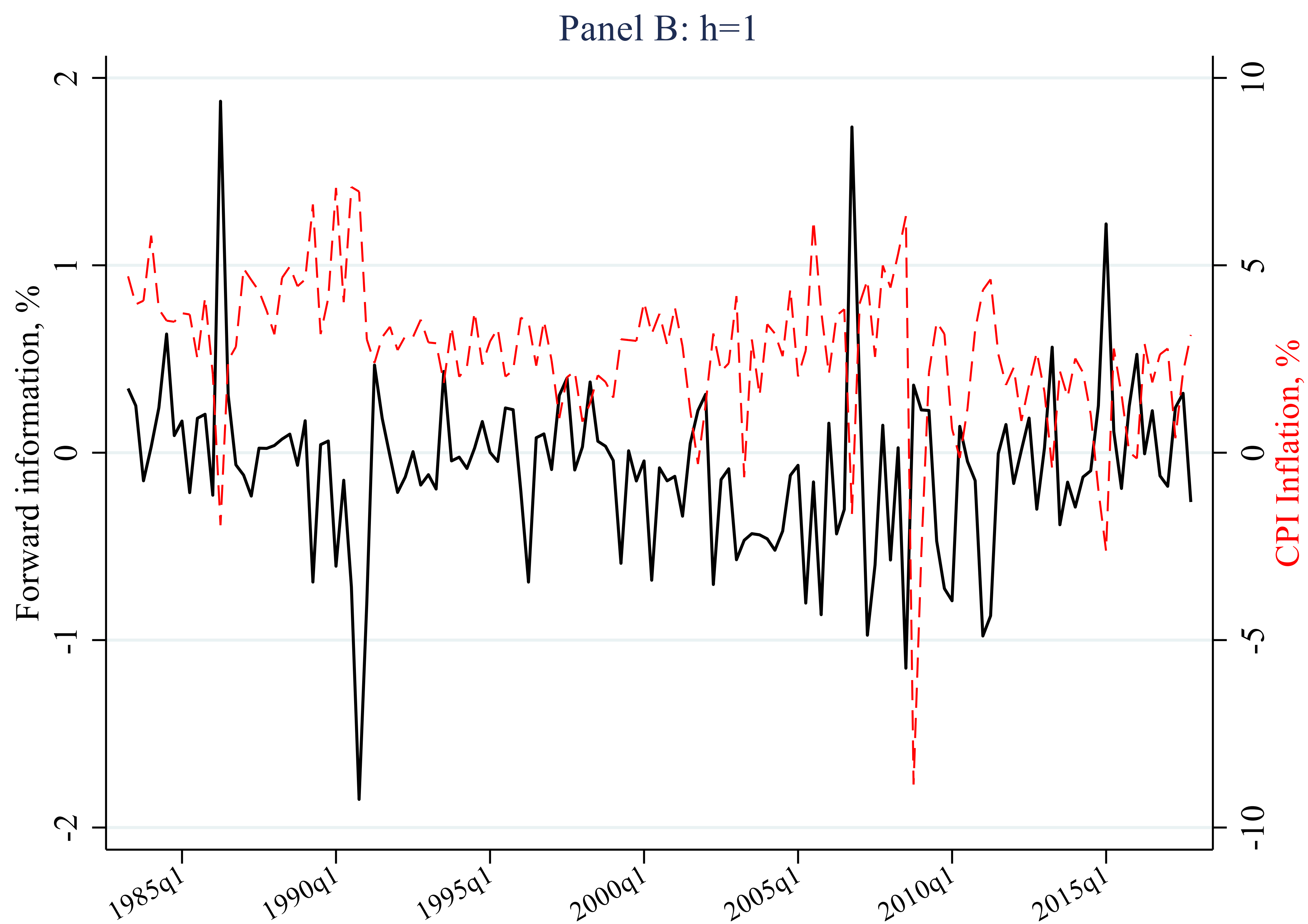
In practice: $FI^i_{t+h|t} = x^i_{t+h|t} - \left(\hat{c}_t + \hat{\rho}_t x^i_{t+h-1|t}\right)$

- Allow time-varying intercept (e.g. changes in trend inflation)
- Estimate ρ on a long forecast horizon h (as $h \uparrow$, forward info is less precise and the bias is smaller). Allow ρ to vary over time as well.

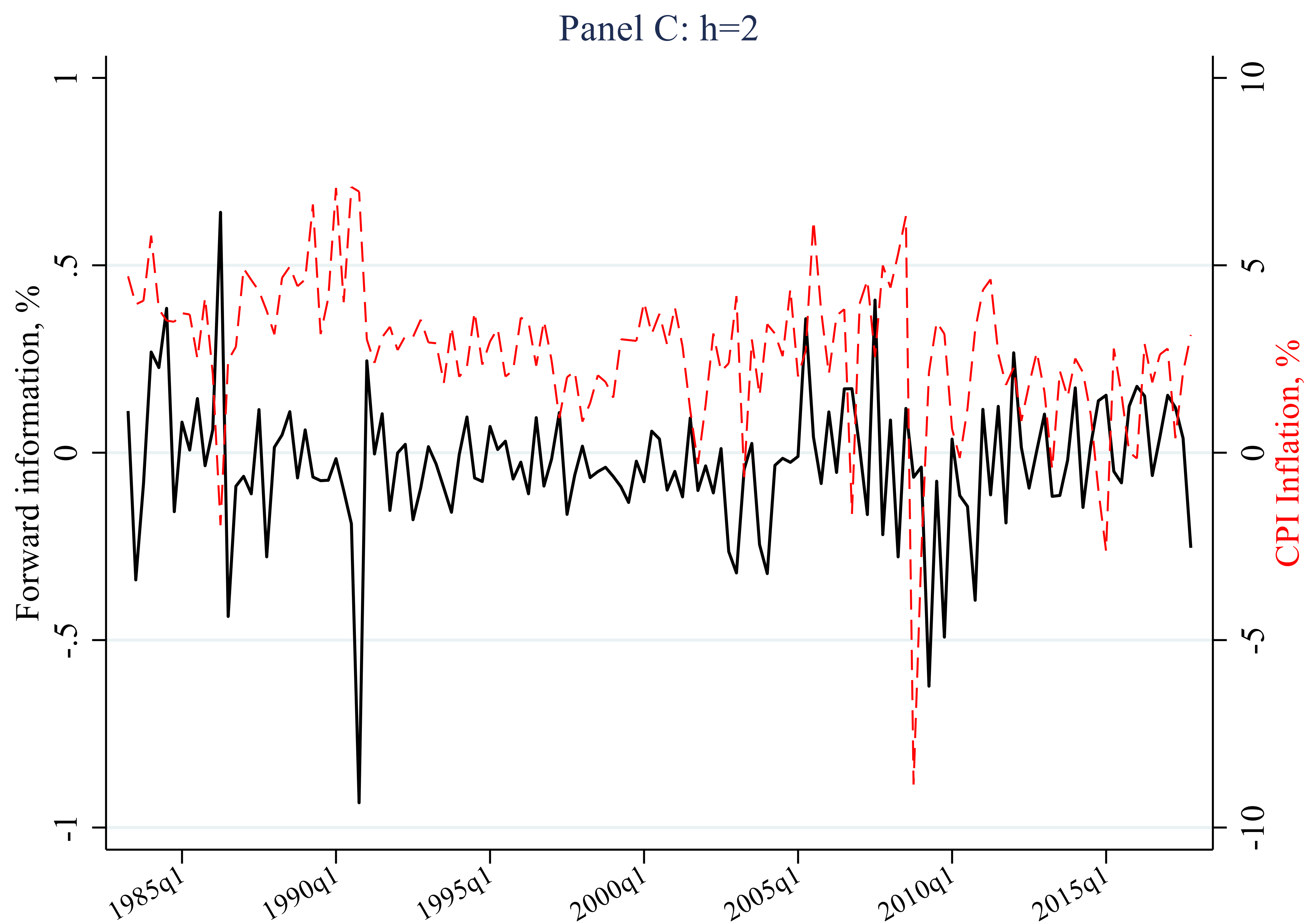
FORWARD INFORMATION



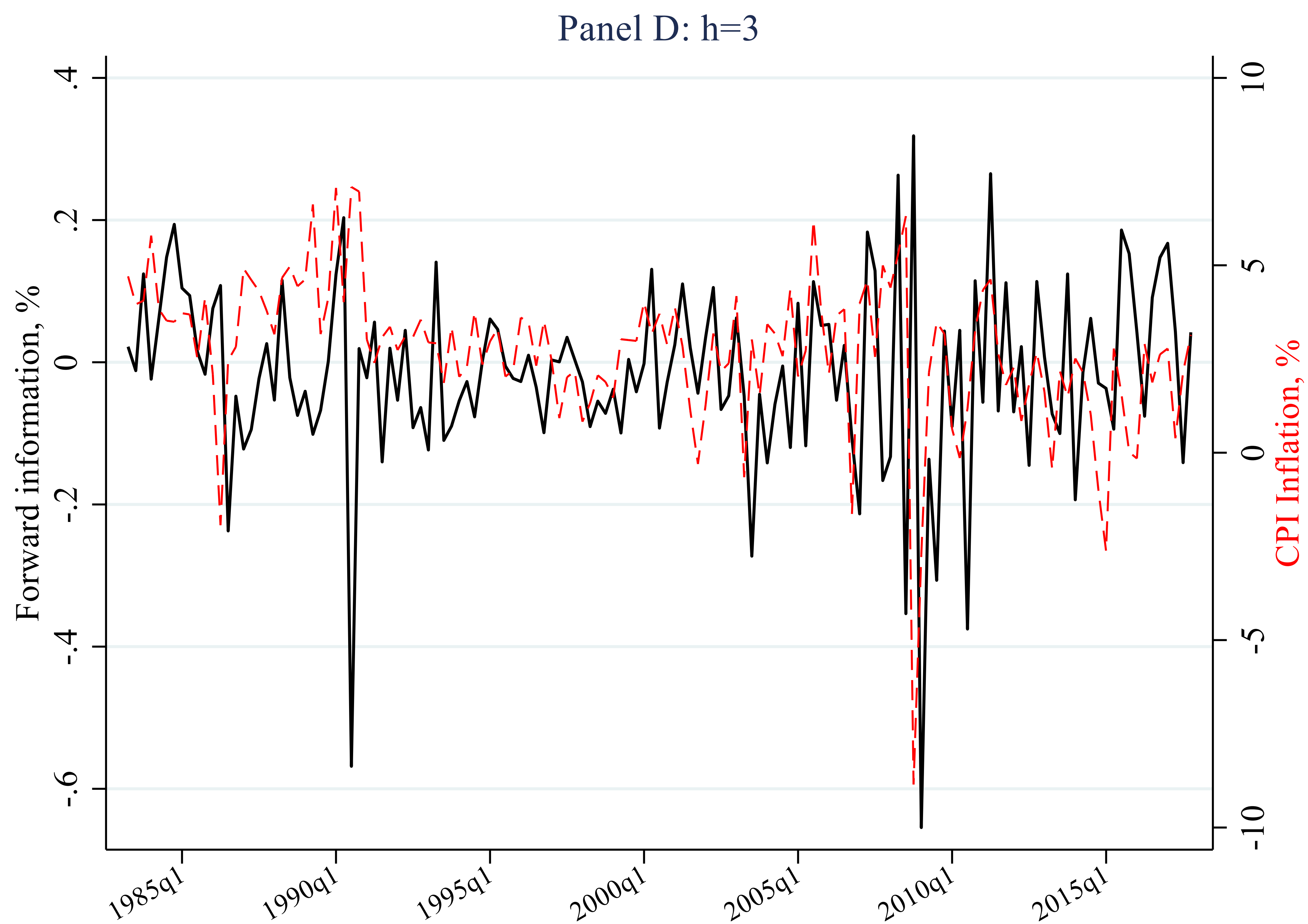
FORWARD INFORMATION



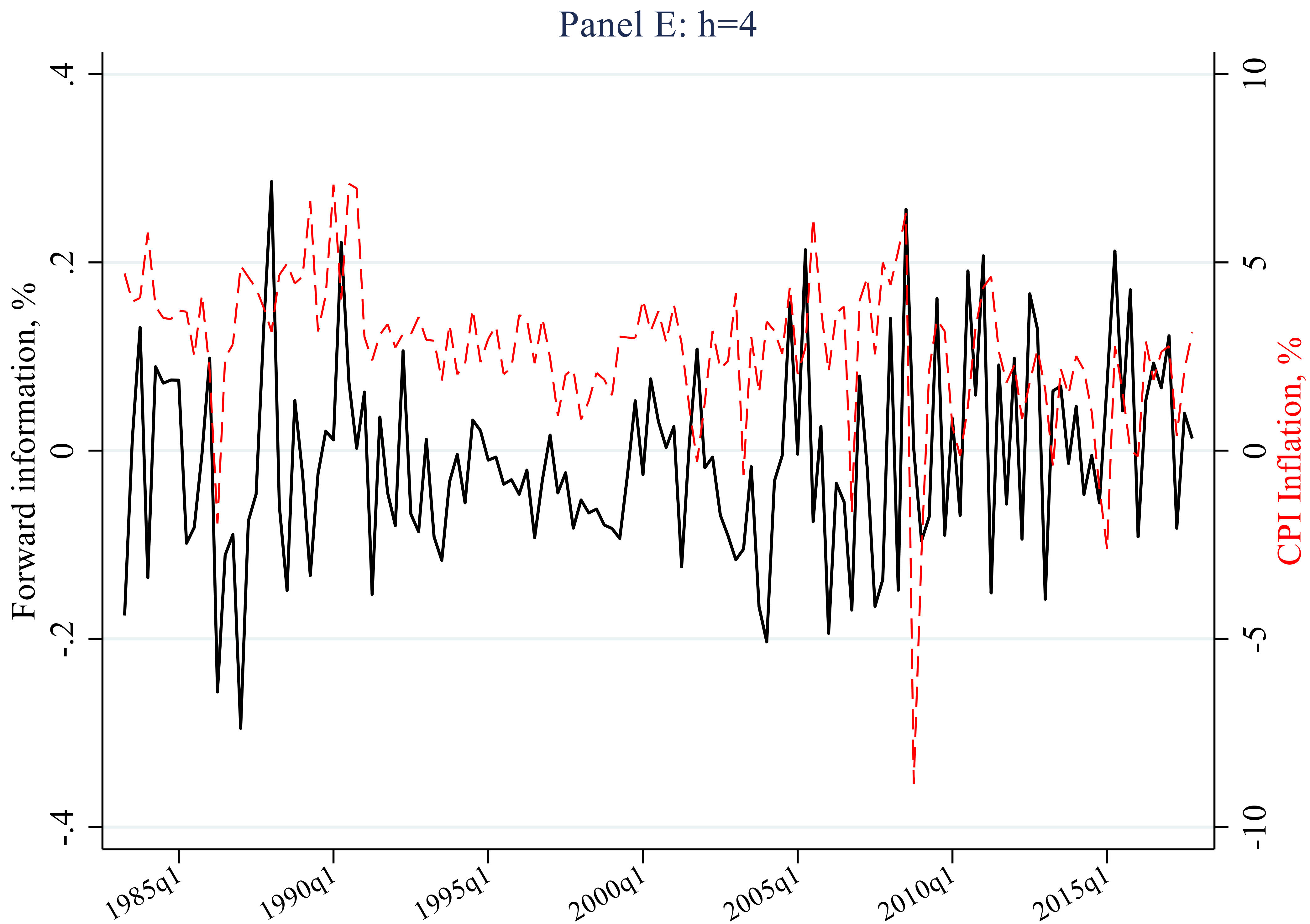
FORWARD INFORMATION



FORWARD INFORMATION



FORWARD INFORMATION



Forward information

Variable:	Mean	Standard deviation
$FI_{t\|t}$	-0.150	1.122
$FI_{t+1\|t}$	-0.078	0.454
$FI_{t+2\|t}$	-0.029	0.182
$FI_{t+3\|t}$	-0.017	0.134
$FI_{t+4\|t}$	-0.010	0.104
Actual inflation	2.704	1.992

Properties:

- The variation of forward information over time decreases in the horizon. This pattern is in line with diminishing information in forward signals for longer horizons. It is also driven by the decay in the change of weights on signals across horizons.

Forward information

Variable:	Mean	Standard deviation	Serial correlation
$FI_{t\|t}$	-0.150	1.122	-0.275
$FI_{t+1\|t}$	-0.078	0.454	0.204
$FI_{t+2\|t}$	-0.029	0.182	-0.073
$FI_{t+3\|t}$	-0.017	0.134	-0.209
$FI_{t+4\|t}$	-0.010	0.104	0.011
Actual inflation	2.704	1.992	0.350

Properties:

- The variation of forward information over time decreases in the horizon. This pattern is in line with diminishing information in forward signals for longer horizons. It is also driven by the decay in the change of weights on signals across horizons.
- Series for forward information should be serially correlated due to the overlap in forward signals over time. That is, previous forward signals which look beyond time t are still useful for the forecast made at time t.

Forward information

Variable:	Mean	Standard deviation	Serial correlation	Correlation between horizons
$FI_{t\|t}$	-0.150	1.122	-0.275	
$FI_{t+1\|t}$	-0.078	0.454	0.204	-0.258
$FI_{t+2\|t}$	-0.029	0.182	-0.073	0.324
$FI_{t+3\|t}$	-0.017	0.134	-0.209	0.069
$FI_{t+4\|t}$	-0.010	0.104	0.011	-0.094
Actual inflation	2.704	1.992	0.350	

Properties:

- The variation of forward information over time decreases in the horizon. This pattern is in line with diminishing information in forward signals for longer horizons. It is also driven by the decay in the change of weights on signals across horizons.
- Series for forward information should be serially correlated due to the overlap in forward signals over time. That is, previous forward signals which look beyond time t are still useful for the forecast made at time t.
- The series for forward information are correlated across horizons because the same signals are applied at each horizon. The correlation should eventually decay due to the diminishing variation.

FORWARD INFORMATION

Cross-sectional variation in inflation forecasts and forward information

	$var(x^i_{t+h\|t})$	$var(\rho_i x^i_{t+h-1\|t})$	$var\left(FI^i_{t+h\|t}\right)$	$\sqrt{\frac{var\left(FI^i_{t+h\|t}\right)}{var\left(x^i_{t+h\|t}\right)}}$
$h = 0$	0.889	0.398	1.143	1.134
$h = 1$	0.610	0.533	0.700	1.071
$h = 2$	0.498	0.336	0.324	0.807
$h = 3$	0.481	0.308	0.279	0.762
$h = 4$	0.459	0.304	0.178	0.623

Forward information (forecasters have different news about the future) accounts for a large share of forecast disagreement.

DOES FORWARD INFORMATION MATTER?

Policy rule: $r_t = c + \gamma \pi_{t|t}^{GB} + \theta_1 gap_{t|t}^{GB} + \theta_2 gr_{t|t}^{GB} + \rho_1^r r_{t-1} + \rho_2^r r_{t-2} + \varepsilon_t$

	(1)
$\pi_{t\|t}^{GB}$	0.051**
	(0.021)
$gap_{t\|t}^{GB}$	0.025*
	(0.014)
$gr_{t\|t}^{GB}$	0.149***
	(0.039)
r_{t-1}	1.134***
	(0.099)
r_{t-2}	-0.184**
	(0.089)
R^2	0.982

DOES FORWARD INFORMATION MATTER?

Policy rule: $r_t = c + \gamma \pi^{GB}_{t|t} + \theta_1 gap^{GB}_{t|t} + \theta_2 gr^{GB}_{t|t} + \rho^r_1 r_{t-1} + \rho^r_2 r_{t-2} + \varepsilon_t$

	(1)	(2)
$\pi^{GB}_{t\|t}$	0.051**	
	(0.021)	
$(\hat{c}_{t-1} + \hat{\rho}_{t-1}\pi_{t-1})$		0.031
		(0.053)
$FI^{GB}_{t\|t}$		0.058***
		(0.018)
$gap^{GB}_{t\|t}$	0.025*	0.024
	(0.014)	(0.015)
$gr^{GB}_{t\|t}$	0.149***	0.149***
	(0.039)	(0.039)
r_{t-1}	1.134***	1.151***
	(0.099)	(0.111)
r_{t-2}	-0.184**	-0.196**
	(0.089)	(0.097)
R^2	0.982	0.982

DOES FORWARD INFORMATION MATTER?

Policy rule: $r_t = c + \gamma\pi_{t|t}^{GB} + \theta_1 gap_{t|t}^{GB} + \theta_2 gr_{t|t}^{GB} + \rho_1^r r_{t-1} + \rho_2^r r_{t-2} + \varepsilon_t$

	(1)	(2)	(3)
$\pi_{t\|t}^{GB}$	0.051**		
	(0.021)		
$(\hat{c}_{t-1} + \hat{\rho}_{t-1}\pi_{t-1})$		0.031	0.124
		(0.053)	(0.083)
$FI_{t\|t}^{GB}$		0.058***	0.151***
		(0.018)	(0.039)
$FI_{t+1\|t}^{GB}$			0.176***
			(0.059)
$FI_{t+2\|t}^{GB}$			0.069
			(0.087)
$FI_{t+3\|t}^{GB}$			0.134
			(0.191)
$FI_{t+4\|t}^{GB}$			0.205
			(0.276)
$gap_{t\|t}^{GB}$	0.025*	0.024	0.027
	(0.014)	(0.015)	(0.017)
$gr_{t\|t}^{GB}$	0.149***	0.149***	0.143***
	(0.039)	(0.039)	(0.038)
r_{t-1}	1.134***	1.151***	1.089***
	(0.099)	(0.111)	(0.128)
r_{t-2}	-0.184**	-0.196**	-0.185*
	(0.089)	(0.097)	(0.103)
R^2	0.982	0.982	0.984

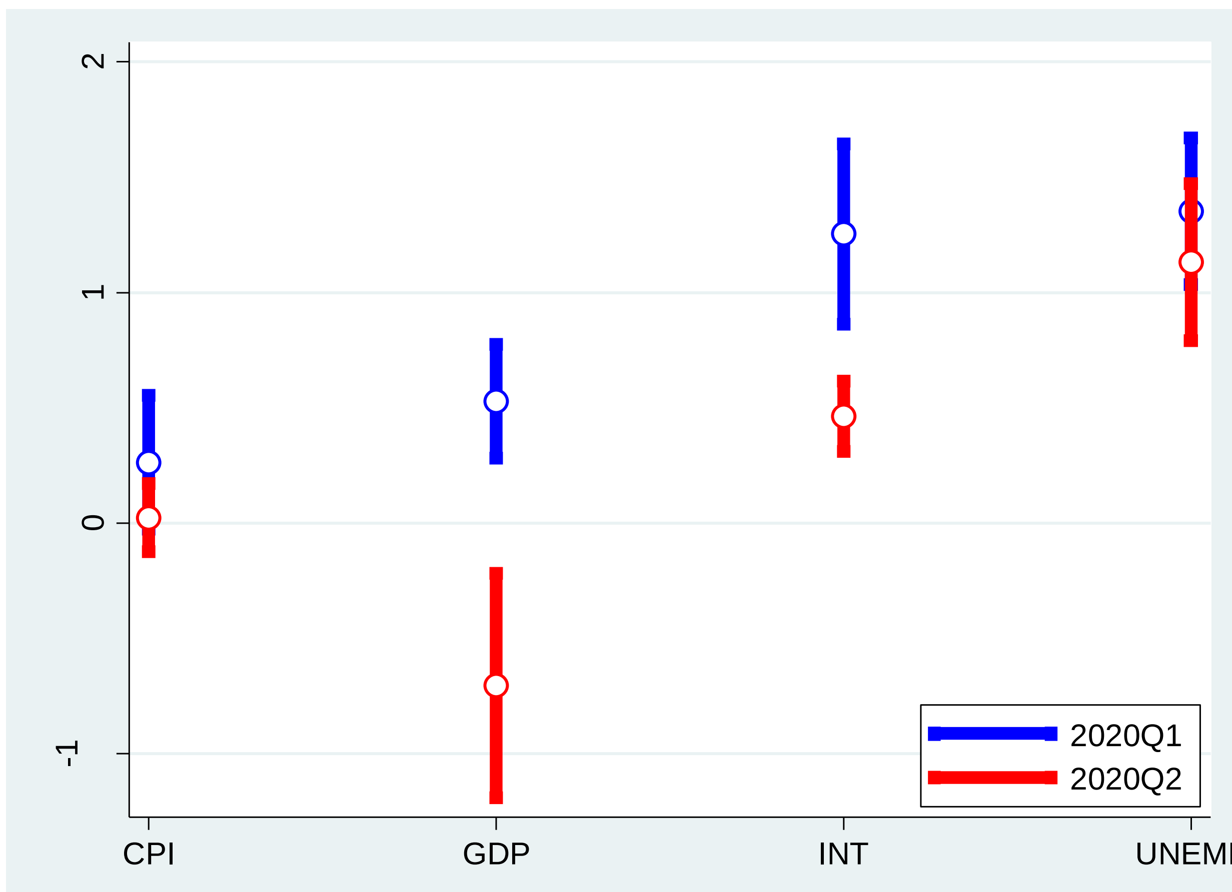
COVID: Forward information booming

Estimate $x^i_{t+h|t} = \rho x^i_{t+h-1|t} + error$, before and after the outbreak

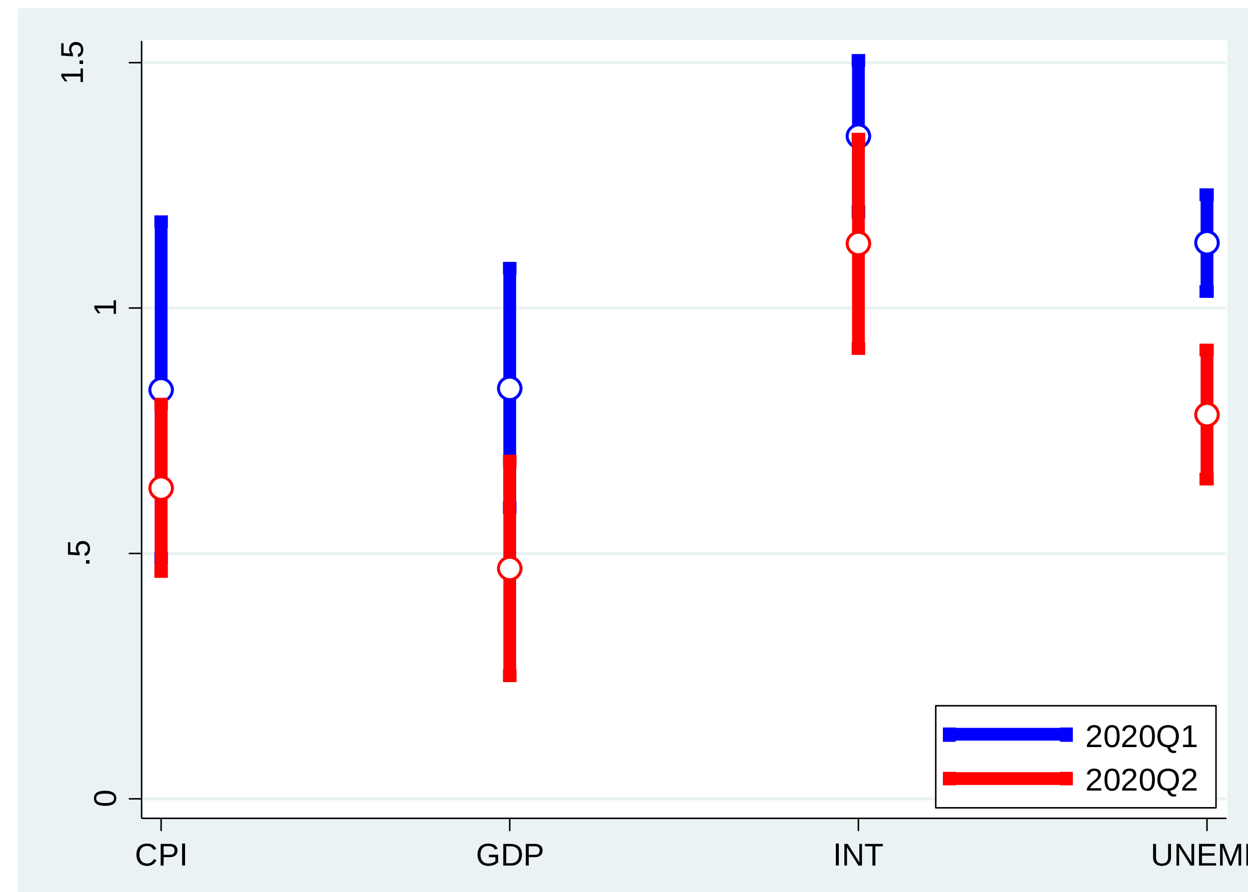
COVID: FORWARD INFORMATION BOOMING

Estimate $x^i_{t+h|t} = \rho x^i_{t+h-1|t} + error$, before and after the outbreak

$\underline{h = 1}$



$\underline{h = 3}$



The term structure of persistence is **shifted downward** in response to a big event.

CONCLUDING REMARKS

- Some key questions:
  - How do people form expectations?
  - Why do we see disagreement in forecasts?
  - What is the role of news?

CONCLUDING REMARKS

- Some key questions:
    - How do people form expectations?
    - Why do we see disagreement in forecasts?
    - What is the role of news?

- We offer a simple framework to measure forward information (private signals about future "fundamentals").

- Forward information appears to be an important force.

CONCLUDING REMARKS

- Some key questions:
  - How do people form expectations?
  - Why do we see disagreement in forecasts?
  - What is the role of news?

- We offer a simple framework to measure forward information (private signals about future “fundamentals”).

- Forward information appears to be an important force.

  Key insights:
  - Forward information accounts for the practice of "add-factoring" (forecast adjustment).
  - Information varies not only across agents, but also across **horizons** (information about the **past** could be homogenous!).