High-Frequency Expectations from Asset Prices: A Machine Learning Approach

Aditya Chaudhry Sangmin S. Oh

Modelling with Big Data & Machine Learning 2020

Motivation

Much work in finance and macro seeks to answer:

- How do economic agents form expectations?
- What real effects do expectations updates induce?
- How to identify causal effects on/of expectations updates at low frequency?
  - E.g. Why did real GDP growth expectations change between March 1 and April 1, 2020?

Common empirical tool: Low-frequency surveys of macro expectations

- E.g. Monthly Blue Chip survey, quarterly Survey of Professional Forecasters
- Unfortunately, many events occur between survey dates

Solution: Move to a higher frequency

Motivation

Much work in finance and macro seeks to answer:

- How do economic agents form expectations?
- What real effects do expectations updates induce?
- How to identify causal effects on/of expectations updates at low frequency?
  - E.g. Why did real GDP growth expectations change between March 1 and April 1, 2020?

Common empirical tool: Low-frequency surveys of macro expectations

- E.g. Monthly Blue Chip survey, quarterly Survey of Professional Forecasters
- Unfortunately, many events occur between survey dates

Solution: Move to a higher frequency

Motivation

Much work in finance and macro seeks to answer:

- How do economic agents form expectations?
- What real effects do expectations updates induce?
- How to identify causal effects on/of expectations updates at low frequency?
  - E.g. Why did real GDP growth expectations change between March 1 and April 1, 2020?

Common empirical tool: Low-frequency surveys of macro expectations

- E.g. Monthly Blue Chip survey, quarterly Survey of Professional Forecasters
- Unfortunately, many events occur between survey dates

Solution: Move to a higher frequency

This Paper

Goal: Construct a daily measure of aggregate growth expectations

- High-frequency series would enable clean identification in event studies.

Approach: Reinforcement Learning + Asset Prices

- Given: Quarterly cross-section of real GDP growth expectations from SPF
- Our task: Recover the daily series of expectations between two quarterly survey release dates

Application: Testing the "Fed Information Effect"

- Hawkish monetary policy surprises correspond to *increases* in surveyed real GDP growth expectations
- News between pre-FOMC survey and FOMC announcement may be omitted variable (Bauer and Swanson, 2020)

This Paper

Goal: Construct a daily measure of aggregate growth expectations

- High-frequency series would enable clean identification in event studies.

Approach: Reinforcement Learning + Asset Prices

- Given: Quarterly cross-section of real GDP growth expectations from SPF
- Our task: Recover the daily series of expectations between two quarterly survey release dates

Application: Testing the "Fed Information Effect"

- Hawkish monetary policy surprises correspond to *increases* in surveyed real GDP growth expectations
- News between pre-FOMC survey and FOMC announcement may be omitted variable (Bauer and Swanson, 2020)

This Paper

Goal: Construct a daily measure of aggregate growth expectations

- High-frequency series would enable clean identification in event studies.

Approach: Reinforcement Learning + Asset Prices

- Given: Quarterly cross-section of real GDP growth expectations from SPF
- Our task: Recover the daily series of expectations between two quarterly survey release dates

Application: Testing the “Fed Information Effect”

- Hawkish monetary policy surprises correspond to *increases* in surveyed real GDP growth expectations
- News between pre-FOMC survey and FOMC announcement may be omitted variable (Bauer and Swanson, 2020)

Key Findings

Measurement:

1. The RL approach successfully filters growth expectations from asset prices
   - R^2 of constructed daily series vs. observed quarterly series: 82.3%
   - Benchmarks: 64.7% for Naive, 2.3% for KF, 39.2% for MIDAS
2. Expectation updates correspond to salient macroeconomic events.

Application:

- No evidence to support the existence of Fed Information effect.
- Hawkish surprises correspond to *decreases* in real GDP growth expectations.

Key Findings

Measurement:

1. The RL approach successfully filters growth expectations from asset prices
   - R^2 of constructed daily series vs. observed quarterly series: 82.3%
   - Benchmarks: 64.7% for Naive, 2.3% for KF, 39.2% for MIDAS
2. Expectation updates correspond to salient macroeconomic events.

Application:

- No evidence to support the existence of Fed Information effect.
- Hawkish surprises correspond to *decreases* in real GDP growth expectations.

Key Findings

Measurement:

1. The RL approach successfully filters growth expectations from asset prices
   - R^2 of constructed daily series vs. observed quarterly series: 82.3%
   - Benchmarks: 64.7% for Naive, 2.3% for KF, 39.2% for MIDAS
2. Expectation updates correspond to salient macroeconomic events.

Application:

- No evidence to support the existence of Fed Information effect.
- Hawkish surprises correspond to *decreases* in real GDP growth expectations.

Literature

1. **Measurement of latent economic and financial variables in time-series**
   - Popular approaches: balanced panel regressions, state-space models, latent VARs
   - Stock and Watson (1989), Bernanke et al. (1997), Evans (2005), Van Binsbergen and Koijen (2010), Brandt and Kang (2001)
   - **Our Approach**: Estimate expectations of variables using a variant of the state-space approach.
2. **Application of machine learning methods in finance**
   - Popular approaches: shrinkage and selection, neural networks, and tree-based models for prediction
   - Rapach et al. (2013), Kelly et al. (2017), Giglio and Xiu (2018), Kozak et al. (2019), Moritz and Zimmermann (2016)
   - **Our Approach**: We show reinforcement learning can outperform the traditional filtering approach.

Roadmap

Empirical Framework

Empirical Performance of RL

Testing the “Fed Information Effect”

Conclusion

Empirical Framework

Data

Asset prices: equities and fixed income

- Baseline: CRSP value-weighted portfolio and CRSP U.S. Treasury five-year fixed-term index

Growth Expectations: cross-sectional mean of quarterly SPF surveys ▸ Summary Statistics

- We focus on one-quarter ahead real GDP growth forecasts
- E.g. Survey conducted in mid 2018:Q3 / Expectation pertains to growth in 2018:Q4

Time Period: 1990:Q3 – 2018:Q4

Data

Asset prices: equities and fixed income

- Baseline: CRSP value-weighted portfolio and CRSP U.S. Treasury five-year fixed-term index

Growth Expectations: cross-sectional mean of quarterly SPF surveys Summary Statistics

- We focus on one-quarter ahead real GDP growth forecasts
- E.g. Survey conducted in mid 2018:Q3 / Expectation pertains to growth in 2018:Q4

Time Period: 1990:Q3 – 2018:Q4

Data

Asset prices: equities and fixed income

- Baseline: CRSP value-weighted portfolio and CRSP U.S. Treasury five-year fixed-term index

Growth Expectations: cross-sectional mean of quarterly SPF surveys ▸ Summary Statistics

- We focus on one-quarter ahead real GDP growth forecasts
- E.g. Survey conducted in mid 2018:Q3 / Expectation pertains to growth in 2018:Q4

Time Period: 1990:Q3 – 2018:Q4

Framework: The Environment

Expected returns and dividend growth have a factor structure: GDP growth and some latent factor

$$\boldsymbol{\theta_{t+1}} = \mu + \delta \boldsymbol{\theta_t} + \epsilon_{t+1}, \quad \epsilon_{t+1} \sim N\left(0, \sigma_\epsilon^2\right) \qquad \text{(GDP Growth)}$$

$$\boldsymbol{\zeta_{t+1}} = \tau + \psi \boldsymbol{\zeta_t} + \xi_{t+1}, \quad \xi_{t+1} \sim N\left(0, \sigma_\xi^2\right) \qquad \text{(Latent Factor)}$$

$$\forall i, d_{t+1}^i - d_t^i = \gamma + \beta^i \boldsymbol{\theta_{t+1}} + \nu_{t+1}^i, \quad \nu_{t+1}^i \sim N\left(0, \sigma_\nu^2\right) \qquad \text{(Dividend Growth)}$$

$$\forall i, \mathbb{E}_t\left[r_{t+1}^i\right] = \alpha + \phi^i \boldsymbol{\zeta_t} \qquad \text{(Conditional Expected Returns)}$$

$$\text{Corr}\left(\epsilon_t, \xi_{t+1}\right) = \pi$$

Applying Campbell-Shiller (1988) decomposition yields the following where $\rho = 1/\left(1 + \exp\left(\overline{d-p}\right)\right)$:

$$\forall i, r_{t+1}^i = \gamma + \left(\beta^i + \frac{\delta\beta^i}{1-\rho\delta}\right)\boldsymbol{\theta_{t+1}} - \frac{\delta\beta^i}{1-\rho\delta}\boldsymbol{\theta_t} - \frac{\phi^i}{1-\rho\psi}\left(\boldsymbol{\zeta_{t+1}} - \boldsymbol{\zeta_t}\right) + \nu_{t+1} \qquad \text{(Realized Returns)}$$

Implication: Can filter estimates of latent growth rate θ_t from multiple asset returns r_t^i

Framework: The Environment

Expected returns and dividend growth have a factor structure: GDP growth and some latent factor

$$\boldsymbol{\theta_{t+1}} = \mu + \delta \boldsymbol{\theta_t} + \epsilon_{t+1}, \quad \epsilon_{t+1} \sim N\left(0, \sigma_\epsilon^2\right) \qquad \text{(GDP Growth)}$$

$$\boldsymbol{\zeta_{t+1}} = \tau + \psi \boldsymbol{\zeta_t} + \xi_{t+1}, \quad \xi_{t+1} \sim N\left(0, \sigma_\xi^2\right) \qquad \text{(Latent Factor)}$$

$$\forall i, d_{t+1}^i - d_t^i = \gamma + \beta^i \boldsymbol{\theta_{t+1}} + \nu_{t+1}^i, \quad \nu_{t+1}^i \sim N\left(0, \sigma_\nu^2\right) \qquad \text{(Dividend Growth)}$$

$$\forall i, \mathbb{E}_t\left[r_{t+1}^i\right] = \alpha + \phi^i \boldsymbol{\zeta_t} \qquad \text{(Conditional Expected Returns)}$$

$$\text{Corr}\left(\epsilon_t, \xi_{t+1}\right) = \pi$$

Applying Campbell-Shiller (1988) decomposition yields the following where $\rho = 1/\left(1 + \exp\left(\overline{d - p}\right)\right)$:

$$\forall i, r_{t+1}^i = \gamma + \left(\beta^i + \frac{\delta \beta^i}{1 - \rho\delta}\right) \boldsymbol{\theta_{t+1}} - \frac{\delta \beta^i}{1 - \rho\delta} \boldsymbol{\theta_t} - \frac{\phi^i}{1 - \rho\psi} \left(\boldsymbol{\zeta_{t+1}} - \boldsymbol{\zeta_t}\right) + \nu_{t+1} \qquad \text{(Realized Returns)}$$

Implication: Can filter estimates of latent growth rate θ_t from multiple asset returns r_t^i

Framework: Incorporating Bayesian Agents

SPF forecasters as Bayesians:

- Observes only returns (r_t^i), not underlyng GDP growth rate (θ_t)
- Estimates θ_t using the Kalman Filter (KF)

Cross-sectional disagreement:

- Cross-sectional disagreement is an important feature of survey data.
- Introduce heterogeneity among agents in *prior-mean* and *learning*.

1. Prior-mean heterogeneity
   Mean of each agent's prior belief regarding θ_t at the start of the quarter: drawn from a normal distribution.
2. Learning heterogeneity
   Each agent's parameter in the state and observation equations: drawn from a normal distribution centered at the true parameter value.
   - Parametrized by signal-to-noise ratio (s) that determines the parameter distribution

Implication: Forecasters differ in starting points and learning rules

Framework: Incorporating Bayesian Agents

SPF forecasters as Bayesians:

- Observes only returns (r_t^i), not underlyng GDP growth rate (θ_t)
- Estimates θ_t using the Kalman Filter (KF)

Cross-sectional disagreement:

- Cross-sectional disagreement is an important feature of survey data.
- Introduce heterogeneity among agents in *prior-mean* and *learning*.

1. Prior-mean heterogeneity
   Mean of each agent's prior belief regarding θ_t at the start of the quarter: drawn from a normal distribution.
2. Learning heterogeneity
   Each agent's parameter in the state and observation equations: drawn from a normal distribution centered at the true parameter value.
   - Parametrized by signal-to-noise ratio (s) that determines the parameter distribution

Implication: Forecasters differ in starting points and learning rules

Framework: Incorporating Bayesian Agents

SPF forecasters as Bayesians:

- Observes only returns (r_t^i), not underlyng GDP growth rate (θ_t)
- Estimates θ_t using the Kalman Filter (KF)

Cross-sectional disagreement:

- Cross-sectional disagreement is an important feature of survey data.
- Introduce heterogeneity among agents in *prior-mean* and *learning*.

1. Prior-mean heterogeneity
   Mean of each agent's prior belief regarding θ_t at the start of the quarter: drawn from a normal distribution.
2. Learning heterogeneity
   Each agent's parameter in the state and observation equations: drawn from a normal distribution centered at the true parameter value.
   - Parametrized by signal-to-noise ratio (s) that determines the parameter distribution

Implication: Forecasters differ in starting points and learning rules

Framework: Incorporating Bayesian Agents

SPF forecasters as Bayesians:

- Observes only returns (r_t^i), not underlyng GDP growth rate (θ_t)
- Estimates θ_t using the Kalman Filter (KF)

Cross-sectional disagreement:

- Cross-sectional disagreement is an important feature of survey data.
- Introduce heterogeneity among agents in *prior-mean* and *learning*.

1. Prior-mean heterogeneity
   Mean of each agent's prior belief regarding θ_t at the start of the quarter: drawn from a normal distribution.
2. Learning heterogeneity
   Each agent's parameter in the state and observation equations: drawn from a normal distribution centered at the true parameter value.
   - Parametrized by signal-to-noise ratio (s) that determines the parameter distribution

Implication: Forecasters differ in starting points and learning rules

Framework: Learning the Cross-sectional Moments

- Denote $\mu_{i,t} \equiv \mathbb{E}_t^i [\theta_{t+1}]$ as agent i's period t expectation of growth at period $t+1$
- KF implies this law of motion for agent i's expectation: Kalman Filter Setup

$$\mu_{i,t} = c_{0,t}^i + c_{1,t}^i \mu_{i,t-1} + \left(\mathbf{c}_{2,t}^i\right)' \mathbf{r}_t$$

where c_1 and $\mathbf{c}_2$ are functions of underlying structural parameters

- We estimate the moments directly rather than keep track of the entire cross-section
- Averaging across all agents:

$$\mu_t \equiv \frac{1}{N}\sum_{i=1}^{N} \mu_{i,t} = \frac{1}{N}\sum_{i=1}^{N} c_{0,t}^i + \frac{1}{N}\sum_{i=1}^{N} c_{1,t}^i \mu_{i,t-1} + \left(\frac{1}{N}\sum_{i=1}^{N} \mathbf{c}_{2,t}^i\right)' \mathbf{r}_t$$

- Motivated by this expression, we use the following approximating moment:

$$\mu_t \approx c_1 \mu_{t-1} + \mathbf{c}_2' \mathbf{r}_t$$

- **Our task:** Estimate c_1 and $\mathbf{c}_2$

Framework: Learning the Cross-sectional Moments

- Denote $\mu_{i,t} \equiv \mathbb{E}_t^i [\theta_{t+1}]$ as agent i's period t expectation of growth at period $t+1$
- KF implies this law of motion for agent i's expectation: Kalman Filter Setup

$$\mu_{i,t} = c_{0,t}^i + c_{1,t}^i \mu_{i,t-1} + \left(\mathbf{c}_{2,t}^i\right)' \mathbf{r}_t$$

where c_1 and $\mathbf{c}_2$ are functions of underlying structural parameters

- We estimate the moments directly rather than keep track of the entire cross-section
- Averaging across all agents:

$$\mu_t \equiv \frac{1}{N}\sum_{i=1}^{N} \mu_{i,t} = \frac{1}{N}\sum_{i=1}^{N} c_{0,t}^i + \frac{1}{N}\sum_{i=1}^{N} c_{1,t}^i \mu_{i,t-1} + \left(\frac{1}{N}\sum_{i=1}^{N} \mathbf{c}_{2,t}^i\right)' \mathbf{r}_t$$

- Motivated by this expression, we use the following approximating moment:

$$\mu_t \approx c_1 \mu_{t-1} + \mathbf{c}_2' \mathbf{r}_t$$

- **Our task:** Estimate c_1 and $\mathbf{c}_2$

Framework: Learning the Cross-sectional Moments

- Denote $\mu_{i,t} \equiv \mathbb{E}_t^i [\theta_{t+1}]$ as agent i's period t expectation of growth at period $t+1$
- KF implies this law of motion for agent i's expectation: ▸ Kalman Filter Setup

$$\mu_{i,t} = c_{0,t}^i + c_{1,t}^i \mu_{i,t-1} + \left(\mathbf{c}_{2,t}^i\right)' \mathbf{r}_t$$

where c_1 and $\mathbf{c}_2$ are functions of underlying structural parameters

- We estimate the moments directly rather than keep track of the entire cross-section
- Averaging across all agents:

$$\mu_t \equiv \frac{1}{N}\sum_{i=1}^{N} \mu_{i,t} = \frac{1}{N}\sum_{i=1}^{N} c_{0,t}^i + \frac{1}{N}\sum_{i=1}^{N} c_{1,t}^i \mu_{i,t-1} + \left(\frac{1}{N}\sum_{i=1}^{N} \mathbf{c}_{2,t}^i\right)' \mathbf{r}_t$$

- Motivated by this expression, we use the following approximating moment:

$$\mu_t \approx c_1 \mu_{t-1} + \mathbf{c}_2' \mathbf{r}_t$$

- **Our task:** Estimate c_1 and $\mathbf{c}_2$

Framework: Learning the Cross-sectional Moments

- Denote $\mu_{i,t} \equiv \mathbb{E}_t^i \left[\theta_{t+1}\right]$ as agent i's period t expectation of growth at period $t+1$
- KF implies this law of motion for agent i's expectation: Kalman Filter Setup

$$\mu_{i,t} = c_{0,t}^i + c_{1,t}^i \mu_{i,t-1} + \left(\mathbf{c}_{2,t}^i\right)' \mathbf{r}_t$$

  where c_1 and $\mathbf{c}_2$ are functions of underlying structural parameters
- We estimate the moments directly rather than keep track of the entire cross-section
- Averaging across all agents:

$$\mu_t \equiv \frac{1}{N}\sum_{i=1}^{N} \mu_{i,t} = \frac{1}{N}\sum_{i=1}^{N} c_{0,t}^i + \frac{1}{N}\sum_{i=1}^{N} c_{1,t}^i \mu_{i,t-1} + \left(\frac{1}{N}\sum_{i=1}^{N} \mathbf{c}_{2,t}^i\right)' \mathbf{r}_t$$

- Motivated by this expression, we use the following approximating moment:

$$\mu_t \approx c_1 \mu_{t-1} + \mathbf{c}_2' \mathbf{r}_t$$

- **Our task:** Estimate c_1 and $\mathbf{c}_2$

Framework: Learning the Cross-sectional Moments

- Denote $\mu_{i,t} \equiv \mathbb{E}_t^i[\theta_{t+1}]$ as agent i's period t expectation of growth at period $t+1$
- KF implies this law of motion for agent i's expectation: ▸ Kalman Filter Setup

$$\mu_{i,t} = c_{0,t}^i + c_{1,t}^i \mu_{i,t-1} + \left(\mathbf{c}_{2,t}^i\right)' \mathbf{r}_t$$

  where c_1 and $\mathbf{c}_2$ are functions of underlying structural parameters

- We estimate the moments directly rather than keep track of the entire cross-section
- Averaging across all agents:

$$\mu_t \equiv \frac{1}{N}\sum_{i=1}^{N} \mu_{i,t} = \frac{1}{N}\sum_{i=1}^{N} c_{0,t}^i + \frac{1}{N}\sum_{i=1}^{N} c_{1,t}^i \mu_{i,t-1} + \left(\frac{1}{N}\sum_{i=1}^{N} \mathbf{c}_{2,t}^i\right)' \mathbf{r}_t$$

- Motivated by this expression, we use the following approximating moment:

$$\mu_t \approx c_1 \mu_{t-1} + \mathbf{c}_2' \mathbf{r}_t$$

- **Our task:** Estimate c_1 and $\mathbf{c}_2$

The RL Approach

The Standard Model: Given utility function and state transition dynamics, derive optimal action

RL Approach: State transition dynamics unknown $\Rightarrow$ learn optimal action through experience

- **Agent:** Econometrician seeks to learn...
- **Action:** Best way to update previous expectation estimate based on new observed asset returns
- **Objective:** Minimize Euclidean distance between estimated and true expectations at quarter-end

Implication: We seek to learn optimal linear learning rule g for updating daily growth expectations

- Use variant of COPDAC-Q algorithm (Silver (2014)) to estimate g

The RL Approach

The Standard Model: Given utility function and state transition dynamics, derive optimal action

RL Approach: State transition dynamics unknown $\Rightarrow$ learn optimal action through experience

- **Agent:** Econometrician seeks to learn...
- **Action:** Best way to update previous expectation estimate based on new observed asset returns
- **Objective:** Minimize Euclidean distance between estimated and true expectations at quarter-end

Implication: We seek to learn optimal linear learning rule g for updating daily growth expectations

- Use variant of COPDAC-Q algorithm (Silver (2014)) to estimate g

The RL Approach

The Standard Model: Given utility function and state transition dynamics, derive optimal action

RL Approach: State transition dynamics unknown $\Rightarrow$ learn optimal action through experience

- **Agent:** Econometrician seeks to learn...
- **Action:** Best way to update previous expectation estimate based on new observed asset returns
- **Objective:** Minimize Euclidean distance between estimated and true expectations at quarter-end

Implication: We seek to learn optimal linear learning rule g for updating daily growth expectations

- Use variant of COPDAC-Q algorithm (Silver (2014)) to estimate g

The RL Approach

The Standard Model: Given utility function and state transition dynamics, derive optimal action

RL Approach: State transition dynamics unknown $\Rightarrow$ learn optimal action through experience

- **Agent:** Econometrician seeks to learn...
- **Action:** Best way to update previous expectation estimate based on new observed asset returns
- **Objective:** Minimize Euclidean distance between estimated and true expectations at quarter-end

Implication: We seek to learn optimal linear learning rule g for updating daily growth expectations

- Use variant of COPDAC-Q algorithm (Silver (2014)) to estimate g

The RL Approach

The Standard Model: Given utility function and state transition dynamics, derive optimal action

RL Approach: State transition dynamics unknown $\Rightarrow$ learn optimal action through experience

- **Agent:** Econometrician seeks to learn...
- **Action:** Best way to update previous expectation estimate based on new observed asset returns
- **Objective:** Minimize Euclidean distance between estimated and true expectations at quarter-end

Implication: We seek to learn optimal linear learning rule g for updating daily growth expectations

- Use variant of COPDAC-Q algorithm (Silver (2014)) to estimate g

The RL Approach

The Standard Model: Given utility function and state transition dynamics, derive optimal action

RL Approach: State transition dynamics unknown $\Rightarrow$ learn optimal action through experience

- **Agent:** Econometrician seeks to learn...
- **Action:** Best way to update previous expectation estimate based on new observed asset returns
- **Objective:** Minimize Euclidean distance between estimated and true expectations at quarter-end

Implication: We seek to learn optimal linear learning rule g for updating daily growth expectations

- Use variant of COPDAC-Q algorithm (Silver (2014)) to estimate g

The RL Approach

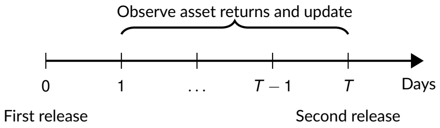
The Standard Model: Given utility function and state transition dynamics, derive optimal action

RL Approach: State transition dynamics unknown $\Rightarrow$ learn optimal action through experience

- **Agent:** Econometrician seeks to learn...
- **Action:** Best way to update previous expectation estimate based on new observed asset returns
- **Objective:** Minimize Euclidean distance between estimated and true expectations at quarter-end

Implication: We seek to learn optimal linear learning rule g for updating daily growth expectations

- Use variant of COPDAC-Q algorithm (Silver (2014)) to estimate g



Benchmarks

A. Naive Approach: Model μ_t as a random walk

B. Kalman Filter:

- Derive optimal Kalman gain expression given state (μ_t) and observation (r_t) equations ▸ Setup
- Estimation via maximum likelihood
- Number of parameters: $3m + 11$ where m is the number of assets

C. Mixed-Data Sampling (MIDAS) Regression:

- For each day within quarter, forecast end-of-quarter survey release using lagged asset returns ▸ Setup
- Use previous 90 days of returns following Ghysels & Wright (2009)
- Number of parameters: $m + 4$ per day for T days within quarter

RL's advantage: Far fewer parameters $\Rightarrow$ More efficient estimation of μ_t ($m + 1$ parameters)

Benchmarks

A. Naive Approach: Model μ_t as a random walk

B. Kalman Filter:

- Derive optimal Kalman gain expression given state (μ_t) and observation (r_t) equations ▸ Setup
- Estimation via maximum likelihood
- Number of parameters: $3m + 11$ where m is the number of assets

C. Mixed-Data Sampling (MIDAS) Regression:

- For each day within quarter, forecast end-of-quarter survey release using lagged asset returns ▸ Setup
- Use previous 90 days of returns following Ghysels & Wright (2009)
- Number of parameters: $m + 4$ per day for T days within quarter

RL's advantage: Far fewer parameters $\Rightarrow$ More efficient estimation of μ_t ($m + 1$ parameters)

Benchmarks

A. Naive Approach: Model μ_t as a random walk

B. Kalman Filter:

- Derive optimal Kalman gain expression given state (μ_t) and observation (r_t) equations ▸ Setup
- Estimation via maximum likelihood
- Number of parameters: $3m + 11$ where m is the number of assets

C. Mixed-Data Sampling (MIDAS) Regression:

- For each day within quarter, forecast end-of-quarter survey release using lagged asset returns ▸ Setup
- Use previous 90 days of returns following Ghysels & Wright (2009)
- Number of parameters: $m + 4$ per day for T days within quarter

RL's advantage: Far fewer parameters $\Rightarrow$ More efficient estimation of μ_t ($m + 1$ parameters)

Benchmarks

A. Naive Approach: Model μ_t as a random walk

B. Kalman Filter:

- Derive optimal Kalman gain expression given state (μ_t) and observation (r_t) equations Setup
- Estimation via maximum likelihood
- Number of parameters: $3m + 11$ where m is the number of assets

C. Mixed-Data Sampling (MIDAS) Regression:

- For each day within quarter, forecast end-of-quarter survey release using lagged asset returns Setup
- Use previous 90 days of returns following Ghysels & Wright (2009)
- Number of parameters: $m + 4$ per day for T days within quarter

RL's advantage: Far fewer parameters $\Rightarrow$ More efficient estimation of μ_t ($m + 1$ parameters)

Empirical Performance of RL

Results

Recursive estimation:

- Fit model on previous N quarters and apply to one quarter out-of-sample
- Baseline: Average models fit on previous $N = 40 - 60$ quarters ▸ Estimation Timeline

Evaluation:
We construct the **daily** series for each method and compute **quarterly** correlations with actual investor expectations from surveys.

	RL	Naive	MIDAS	KF
RMSE	0.449	0.588	0.916	39.103
R^2	0.823	0.647	0.392	0.0237

Results

Recursive estimation:

- Fit model on previous N quarters and apply to one quarter out-of-sample
- Baseline: Average models fit on previous $N = 40 - 60$ quarters Estimation Timeline

Evaluation:
We construct the **daily** series for each method and compute **quarterly** correlations with actual investor expectations from surveys.

	RL	Naive	MIDAS	KF
RMSE	0.449	0.588	0.916	39.103
R^2	0.823	0.647	0.392	0.0237

Validation with Macroeconomic Events

- Major daily changes in estimated growth expectations correspond to significant macroeconomic events

	Expectations Update (%)	Event
2011-08-08	-0.65	U.S. credit rating downgrade
2011-08-09	0.51	Fed promises to keep interest rates near zero for two years
2008-10-15	-0.51	Weak Fed economic forecasts, Bernanke comments
2008-10-28	0.50	Unclear
2011-08-04	-0.45	Weak jobs data, Japan weakens Yen, ECB re-enters bond market
2008-10-09	-0.44	Unclear
2009-03-23	0.44	Treasury announces TARP
2008-09-29	-0.43	House rejects bank bailout plan
2011-08-11	0.40	Jobless claims fall, strong earnings
2009-03-10	0.38	Citi earnings positive (were expected to be negative)

Testing the “Fed Information Effect”

Omitted Variable Bias in Low-Frequency Regressions

Usual Fed Information Effect regression: $\mathbb{E}_{t+15}[g_Q] - \mathbb{E}_{t-15}[g_Q] = \beta_0 + \boldsymbol{\beta_1} Shock_t + \epsilon_t$

- $\mathbb{E}_{t+15}[g_Q]$ and $\mathbb{E}_{t-15}[g_Q]$ are one-month apart surveyed expectations around monetary event at day t
- E.g. Nakamura & Steinsson (2018) use monthly Blue Chip forecasts and find $\boldsymbol{\beta_1} > 0$.

Omitted variable: Economic news released between day $t-15$ and day $t-1$ (in ϵ_t)

- The estimate $\hat{\beta}_1$ will be positively biased if:

$$Corr(\mathbb{E}_{t+15}[g_Q] - \mathbb{E}_{t-15}[g_Q], Econ\ News_{t-15:t-1}) > 0, \quad Corr(Shock_t, Econ\ News_{t-15:t-1}) > 0$$

Omitted Variable Bias in Low-Frequency Regressions

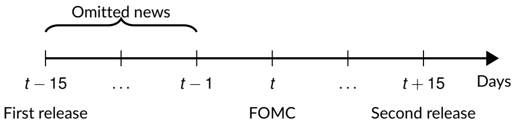
Usual Fed Information Effect regression: $\mathbb{E}_{t+15}[g_Q] - \mathbb{E}_{t-15}[g_Q] = \beta_0 + \boldsymbol{\beta_1} Shock_t + \epsilon_t$

- $\mathbb{E}_{t+15}[g_Q]$ and $\mathbb{E}_{t-15}[g_Q]$ are one-month apart surveyed expectations around monetary event at day t
- E.g. Nakamura & Steinsson (2018) use monthly Blue Chip forecasts and find $\boldsymbol{\beta_1} > 0$.

Omitted variable: Economic news released between day $t-15$ and day $t-1$ (in ϵ_t)

- The estimate $\hat{\beta}_1$ will be positively biased if:

$$Corr(\mathbb{E}_{t+15}[g_Q] - \mathbb{E}_{t-15}[g_Q], Econ\ News_{t-15:t-1}) > 0, \quad Corr(Shock_t, Econ\ News_{t-15:t-1}) > 0$$



Our Test of the Fed Information Effect

- Our daily growth expectations series on day $t-1$ already incorporates $\textit{Econ News}_{t-15:t-1}$.

- Therefore, we run:

$$\mathbb{E}_t[g_Q] - \mathbb{E}_{t-1}[g_Q] = \beta_0 + \beta_1 \textit{Shock}_t + +\epsilon_t$$

  where

$$\textit{Corr}(\mathbb{E}_t[g_Q] - \mathbb{E}_{t-1}[g_Q], \textit{Econ News}_{t-15:t-1}) = 0$$

- $\textit{Shock}_t$ is monetary policy news shock from Nakamura & Steinsson (2018)
  - First principal component of 30-minute changes in five interest rate futures around FOMC announcements

Daily Autocorrelation

Our Test of the Fed Information Effect

- Our daily growth expectations series on day $t-1$ already incorporates $\mathit{Econ\ News}_{t-15:t-1}$.

- Therefore, we run:

$$\mathbb{E}_t\left[g_Q\right]-\mathbb{E}_{t-1}\left[g_Q\right]=\beta_0+\beta_1 \mathit{Shock}_t++\epsilon_t$$

where

$$\mathit{Corr}\left(\mathbb{E}_t\left[g_Q\right]-\mathbb{E}_{t-1}\left[g_Q\right], \mathit{Econ\ News}_{t-15:t-1}\right)=0$$

- Shock_t is monetary policy news shock from Nakamura & Steinsson (2018)
  - First principal component of 30-minute changes in five interest rate futures around FOMC announcements

Daily Autocorrelation

Our Test of the Fed Information Effect

- Our daily growth expectations series on day $t-1$ already incorporates $\mathit{Econ\ News}_{t-15:t-1}$.

- Therefore, we run:
$$\mathbb{E}_t\left[g_Q\right] - \mathbb{E}_{t-1}\left[g_Q\right] = \beta_0 + \beta_1 \mathit{Shock}_t + +\epsilon_t$$
where
$$\mathit{Corr}\left(\mathbb{E}_t\left[g_Q\right] - \mathbb{E}_{t-1}\left[g_Q\right], \mathit{Econ\ News}_{t-15:t-1}\right) = 0$$

- Shock_t is monetary policy news shock from Nakamura & Steinsson (2018)
  - First principal component of 30-minute changes in five interest rate futures around FOMC announcements

▸ Daily Autocorrelation

Results

Comparison to Nakamura & Steinsson (2018)

	Full Sample 2005:03–2014:12	Full Sample Ex. 2008:06-2009:06	NS (2018) 2000:01–2014:12
Panel A. Response to Policy News Shock			
Policy news shock	-0.83	-0.82	1.04
	$(t: -3.632)$	$(t: -2.938)$	$(t: 2.971)$
Observations	71	63	90
Panel B. Response to Fed Funds Rate (FFR) Shock			
FFR Shock	-0.39	-0.38	N/A
	$(t: -1.914)$	$(t: -2.028)$	
Observations	71	63	

- **Negative coefficients:** Hawkish surprises are viewed as contractionary
- **Implication:** No evidence of a Fed Information Effect.

Conclusion

Conclusion

Main Findings:

- RL + asset prices $\Rightarrow$ High-frequency real GDP growth expectations
- Estimated daily series attains R^2 of 82.3% vs. original quarterly SPF series

Implications:

1. High-frequency series provides sharp tool for empirical work
   - Can help shed light on sources and mechanisms of expectations formation
2. Application to test Fed Information Effect – We find no evidence of this effect.

Conclusion

Main Findings:

- RL + asset prices $\Rightarrow$ High-frequency real GDP growth expectations
- Estimated daily series attains R^2 of 82.3% vs. original quarterly SPF series

Implications:

1. High-frequency series provides sharp tool for empirical work
   - Can help shed light on sources and mechanisms of expectations formation
2. Application to test Fed Information Effect – We find no evidence of this effect.

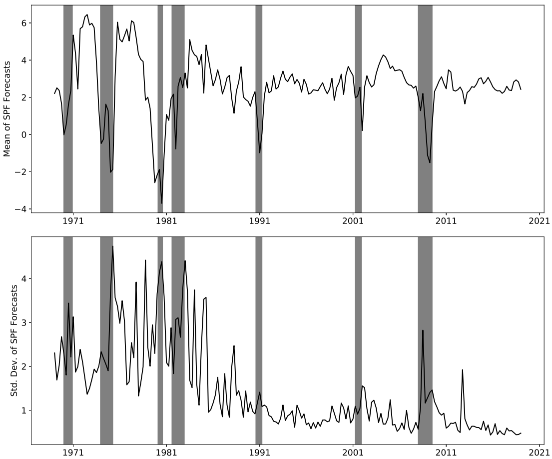
Appendix

Appendix: SPF Summary Statistics (Winsorized at 5%)

	# Forecasters	# Months	CX Mean	CX Median	CX Std	Real GDP Growth
Mean	35.956	114	2.531	2.510	0.683	2.520
Std Dev	6.014	114	0.929	0.946	0.228	2.307
Autocorr(1)	.	114	0.742	0.735	0.730	0.359

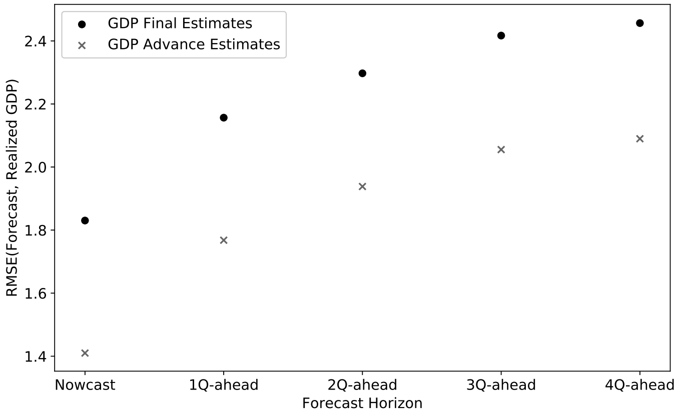
Back

Appendix: SPF Forecast Cyclicality



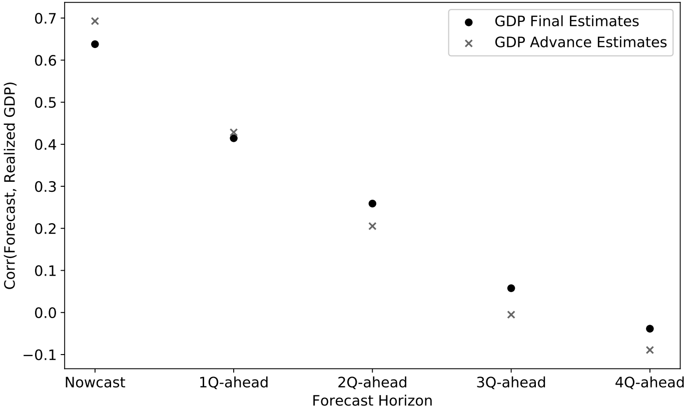
Back

Appendix: SPF Forecast Accuracy (RMSE)



Back

Appendix: SPF Forecast Accuracy (Correlation)



Back

Appendix: Do Asset Prices Matter?

Table 3: **Regressions of Forecast Innovations on Asset Returns**

Asset 1	Asset 2	Coeff. on Asset 1	Coeff. on Asset 2	R-squared
CRSP Value-weighted Return	5YR Fixed-term Index Return	0.036	-0.199	0.383
5YR Fixed-term Index Return	Change in BAA-10Y Spread	-0.178	-0.013	0.344
5YR Fixed-term Index Return	Change in Weighted Average of Forex Value	-0.242	-0.036	0.334
5YR Fixed-term Index Return	Change in VIX index	-0.225	-0.003	0.328
5YR Fixed-term Index Return	Slope of Yield Curve	-0.221	0.013	0.317
5YR Fixed-term Index Return	Change in AAA-10Y Spread	-0.213	-0.003	0.316
CRSP Value-weighted Return	Change in BAA-10Y Spread	0.030	-0.022	0.261
CRSP Value-weighted Return	Change in AAA-10Y Spread	0.038	-0.016	0.253
Slope of Yield Curve	Change in BAA-10Y Spread	0.029	-0.027	0.234
Change in BAA-10Y Spread	Change in VIX index	-0.029	-0.001	0.224
Change in AAA-10Y Spread	Change in BAA-10Y Spread	-0.002	-0.027	0.222

Appendix: Kalman Filter Setup

Rearranging the state-space equations and expression for returns yields:

- State equations:

$$\begin{aligned}\theta_{t+1} &= \mu + \delta\theta_t + \epsilon_{t+1} \\ \zeta_{t+1} &= \tau + \psi\zeta_t + \xi_{t+1}\end{aligned}$$

- Observation equations:

$$\forall i = 1, \dots, d, \quad r^i_{t+1} = \begin{bmatrix} \gamma & \left(\beta^i + \frac{\delta\beta^i}{1-\rho\delta}\right) & -\frac{\delta\beta^i}{1-\rho\delta} & -\frac{\phi^i}{1-\rho\psi} & \frac{\phi^i}{1-\rho\psi} \end{bmatrix} \begin{bmatrix} 1 \\ \theta_{t+1} \\ \theta_t \\ \zeta_{t+1} \\ \zeta_t \end{bmatrix} + \nu_{t+1}$$

- Kalman gain is thereby a linear combination of current asset returns $\left(r^i_{t+1}, \forall i\right)$ and lag 1-day expectation

Back

Appendix: Cross-Sectional Kalman Filter

Rearranging the state-space equations and expression for returns yields:

- State equations:

$$\begin{aligned}\theta_{t+1} &= \mu + \delta\theta_t + \epsilon_{t+1} \\ \zeta_{t+1} &= \tau + \psi\zeta_t + \xi_{t+1} \\ \mu_{t+1} &= \mathbf{c}_2'(\mathbf{1}\gamma + \mathbf{a}\mu + \mathbf{c}\tau) + \mathbf{c}_2'(\mathbf{a}\delta + \mathbf{b})\theta_t + \mathbf{c}_2'\mathbf{c}(\psi - 1)\zeta_t + c_1\mu_t\end{aligned}$$

- Observation equation:

$$\mathbf{c}_2'\mathbf{r}_t = \mu_t - c_1\mu_{t-1}$$

Back

Appendix: Cross-Sectional MIDAS

- Let d_t be a day on which we observe $y_t = \mu_t$, the quarterly-observed surveyed CX mean
- Let r^i_τ be the return of asset i on day τ
- We seek to forecast y_t on each day $d_{t-1} < \tau < d_t$(i.e. each day between survey releases)
- For each such day, fit the following model:

$$y_t = \alpha^\tau + \rho^\tau y_{t-1} + \sum_{i=1}^{m} \beta^\tau_i \gamma^\tau(L) r^i_\tau + \epsilon_t$$

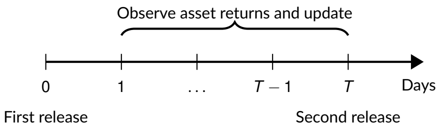
where $\gamma^\tau(L)$ is a lag-polynomial of order l. Thus,

$$\gamma^\tau(L) r^i_\tau = \sum_{s=\tau-l+1}^{\tau} \gamma^\tau_s r^i_s$$

- To limit the number of parametes, we use the beta lag specification from Ghysels & Wright (2009)
  - Parameterizes $\gamma^\tau(L)$ with only two parameters

Back

Appendix: Estimation Timeline



Back

Appendix: Daily Estimated Series Summary Statistics

	RL	Naive	MIDAS	KF
Panel A: Daily Series				
Mean	2.476	2.401	2.445	32.888
Std Dev	1.044	0.946	0.911	19.139
Autocorr(1)	0.997	0.997	0.673	0.960
Skewness	-2.431	-2.475	-2.912	2.652
Excess Kurtosis	7.006	7.208	27.095	12.608
Panel B: Change in Daily Series				
Mean	0.002	0.000	-.001	0.570
Mean of Absolute Values	0.036	0.000	0.385	0.597
Std. Dev.	0.060	0.000	0.736	1.923
Skewness	-0.670	0.000	0.920	6.998
Kurtosis	17.282	0.000	75.606	61.905

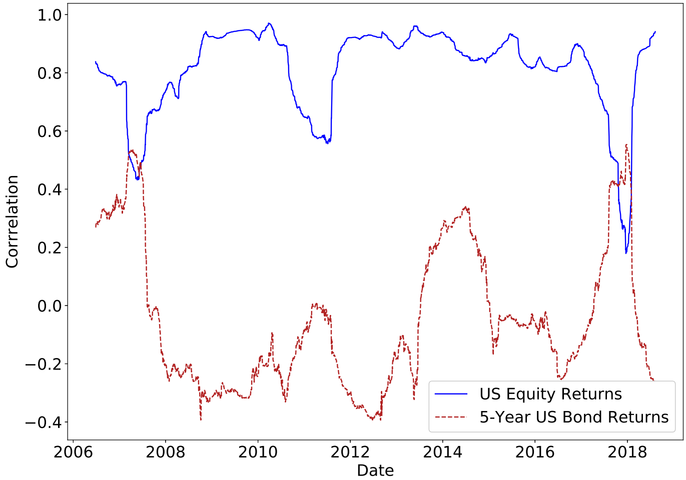
Back

Appendix: Unconditional Daily Correlations

	USA	ret_5yr	RL	MIDAS	KF	Naive
USA	1.00	-0.41	0.87	0.11	0.01	-
ret_5yr	-0.41	1.00	-0.16	-0.03	0.02	-
RL	0.87	-0.16	1.00	0.14	0.01	-
MIDAS	0.11	-0.03	0.14	1.00	-0.01	-
KF	0.01	0.02	0.01	-0.01	1.00	-
Naive	-	-	-	-	-	-

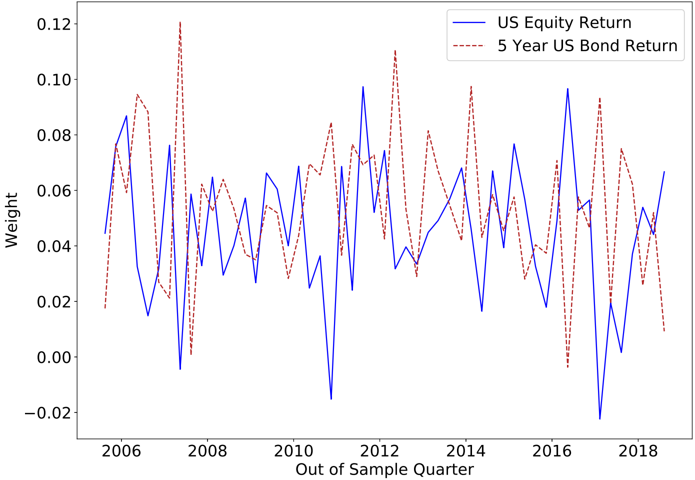
Back

Appendix: 1-Year Rolling Daily Correlations



Back

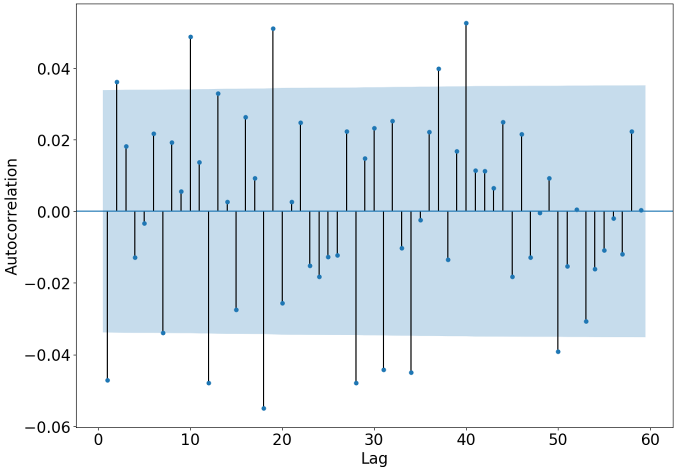
Appendix: CX Mean Policy Weights



Back

Appendix: CX Mean Autocorrelation

First difference of our daily growth expectations series displays no autocorrelation across days



Back