Discussion of

Dividend Momentum and Stock Return Predictability: A Bayesian Approach

By Juan Antolin-Diaz, Ivan Petrella, and Juan F. Rubio-Ramirez

@WFA 2022

Discussant: “J” Jiacui Li

Outline

- 1) How this paper fits into the literature

- 2) My understanding/assessment of the contributions
  - Caveat: I'm a low-tech person

1. Putting this paper into context

Background

- "Excess volatility" puzzle
  - Shiller (1981): stock prices are too volatile

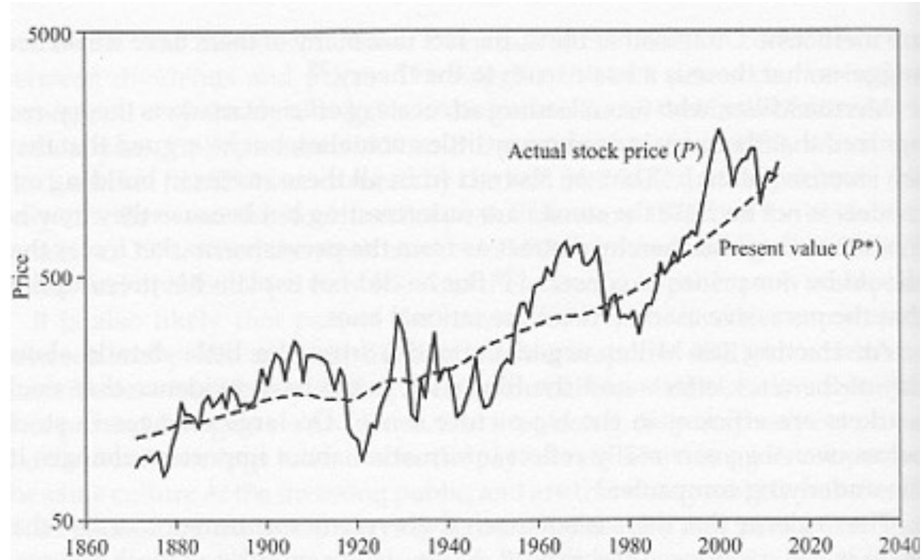

Figure 11.2
Stock Prices and Dividend Present Values, 1871–2013

- Campbell-Shiller (CS, 1988) provided a formal decomposition of returns into *cash flows* and *"discount rates"* (aka non-cash flows)

This paper proposes improvements to the standard VAR(1)-based Campbell-Shiller (CS) exercise

- Two main findings in CS exercise:
  - 1) "Discount rate" variation is a large fraction of aggregate stock price fluctuations
  - 2) Returns are *predictable* using price/dividend ratios

- This paper:
  - 1) Corrects a common mistake
    - In VAR(1), we do NOT have to drop the dividend equation
  - 2) Imposes a "skeptical prior" in a Bayesian approach
    - Even with skeptical priors, returns still appear predictable (weaker)
    - Technically very accomplished!
      - e.g. restrict the prior to satisfy CS accounting identity (importance sampling)

As a low-tech person, my assessments:

- Contribution 1 (correcting mistake): clearly correct
  - People should change their current practice!

- Contribution 2 (Bayesian estimation): this seems more like a general methodological contribution
  - This might be useful elsewhere (as the paper suggests)

- The two contributions appear orthogonal to each other

2. What is the mistake being corrected?

Existing work dropped a *non-redundant* equation

- Full VAR(1):

$$\underbrace{\begin{bmatrix} \Delta d_{t+1} \\ pd_{t+1} \\ r_{t+1} \end{bmatrix}}_{\mathbf{y}_{t+1}} = \underbrace{\begin{bmatrix} c^d \\ c^{pd} \\ c^r \end{bmatrix}}_{\mathbf{\Phi}_0} + \underbrace{\begin{bmatrix} \phi_{d,d} & \phi_{d,pd} & \phi_{d,r} \\ \phi_{pd,d} & \phi_{pd,pd} & \phi_{pd,r} \\ \phi_{r,d} & \phi_{r,pd} & \phi_{r,r} \end{bmatrix}}_{\mathbf{\Phi}_1} \underbrace{\begin{bmatrix} \Delta d_t \\ pd_t \\ r_t \end{bmatrix}}_{\mathbf{y}_t} + \underbrace{\begin{bmatrix} u^d_{t+1} \\ u^{pd}_{t+1} \\ u^r_{t+1} \end{bmatrix}}_{\mathbf{u}_{t+1}}$$

- Traditionally, people drop the first equation due to *perceived* collinearity from the CS identity:

$$r_{t+1} \approx \kappa + \rho pd_{t+1} - pd_t + \Delta d_{t+1}$$

- This paper: there is still the $\boldsymbol{pd_t}$ term! Thus not colinear

Paper shows this makes a difference

- There is "dividend momentum": $\Delta d_t \uparrow \rightarrow \Delta d_{t+1} \uparrow$ AND $r_{t+1} \uparrow$
  - … consistent with "prices not fully reflecting dividend persistence"

- How much does this change the CS headline results?

$$r_{t+1} - \mathbb{E}_t r_{t+1} = \underbrace{(\mathbb{E}_{t+1} - \mathbb{E}_t) \sum_{j=0}^{\infty} \rho^j \Delta d_{t+j+1}}_{NCF_{t+1}} - \underbrace{(\mathbb{E}_{t+1} - \mathbb{E}_t) \sum_{j=1}^{\infty} \rho^j r_{t+j+1}}_{NDR_{t+1}}$$

	2-variable VAR omitting dividend growth			3-variable VAR		
	Total	u^{pd}_{t+1}	u^{d}_{t+1}	Total	u^{pd}_{t+1}	u^{d}_{t+1}
$Var(NDR_{t+1})$	0.032 [0.022, 0.052]	99.7% [99.4%, 99.9%]	0.3% [0.1%, 0.6%]	0.030 [0.018, 0.059]	89.8% [78.6%, 95.3%]	10.2% [4.7%, 21.4%]
$Var(NCF_{t+1})$	0.006 [0.004, 0.013]	26.1% [4.3%, 60.7%]	73.9% [39.3%, 95.7%]	0.013 [0.008, 0.029]	7.9% [0.7%, 34.2%]	92.0% [65.7%, 99.2%]
$Corr(NDR_{t+1}, NCF_{t+1})$	55.4% [20.3%, 81.0%]	50.3% [14.9%, 77.5%]	4.4% [2.4%, 6.7%]	51.7% [17.4%, 79.9%]	16.9% [−16.2%, 51.6%]	29.3% [19.0%, 43.1%]

Let's be clear about what the mistake is

- Not adding Δd_t as a state variable, *per se*, is not a cardinal mistake

- The CS approach does not specify the state variables y_t
  - CS's approach quantifies Discount Rate $= Var(\sum_{j\geq 0} \rho^{j-1} \cdot E\ (r_{t+j}|y_t))$ and calls the residual "cash flow"
- Thus, it is always good to add return-predicting variables (e.g. Δd_t)!
  - Campbell himself added a few later: interest rate, volatility…

- Not adding Δd_t because we assume it is redundant is a mistake

3. Bayesian VAR

My personal view on this

- As researchers, I think many have accepted the main CS findings about (ex-post) return predictability and excess volatility
  - The Bayesian analysis tries to do a better job learning from the aggregate time series
- In addition to the aggregate data, we also learn from "soft" sources:
  - Behavior of market participants
  - Institutional constraints/frictions
  - How elasticity is the market to demand movements
- This is even more true for <u>investors</u> than <u>researchers</u>

Summary

- The paper improves the Campbell-Shiller exercise
  - 1) Corrects a long-standing mistake – which we should all take heed of
  - 2) Devises a Bayesian VAR approach

- I think the (impressive) Bayesian approach might be generally useful elsewhere, but I'm not equipped to judge

- I learned a lot from the paper and wish the authors best of luck