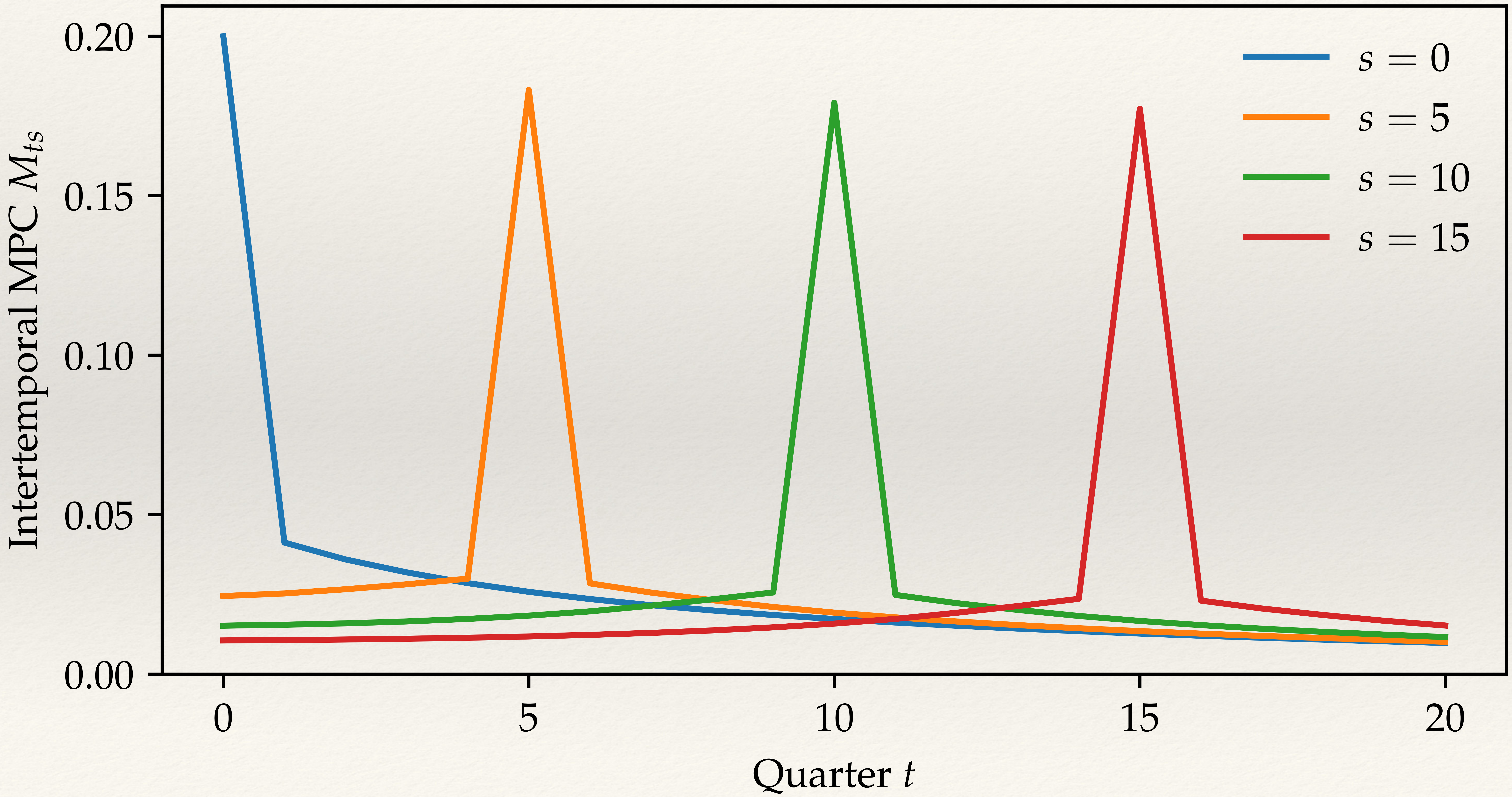
Introduction to the Sequence Space and Jacobians

Matthew Rognlie

NBER Heterogeneous-Agent Macro Workshop, 2025

Calculating sequence-space Jacobians

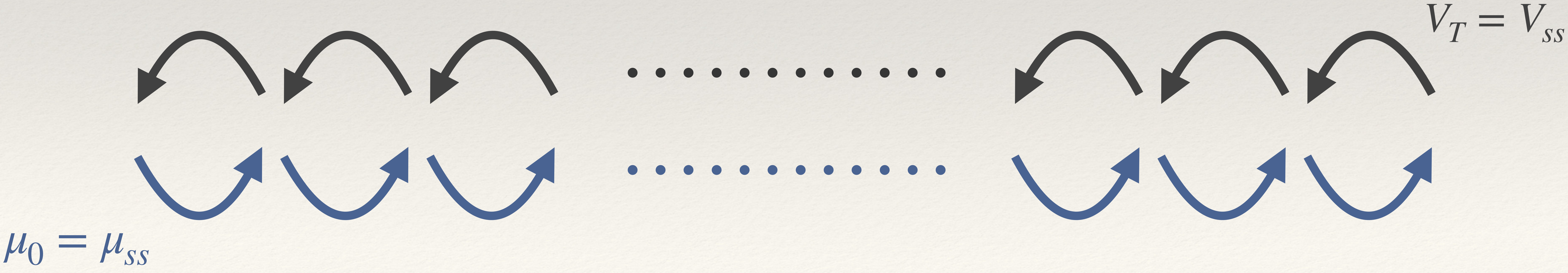
One sequence-space Jacobian: intertemporal MPCs



Here we're plotting a few **columns** of the sequence-space Jacobian **M**

How do we calculate sequence-space Jacobians?

- Sequence-space Jacobians are awesome if we have them
- But how do we get them the first place?
- Each column is an impulse response to perturbation only at s…
- Do we need to redo this process T times, once for each s, at cost $O(NT^2)$?



We can do better

- The "direct" or **"brute-force"** method is costly:
  - if $N \gg T$, then $O(NT^2)$ work to get Jacobians swamps $O(T^3)$ cost of matrix operations
  - (still not totally useless, especially if we can reuse them)

- Fortunately, there's a better way: the **"fake news algorithm"**
  - Need (roughly) *single* backward and forward pass, not one for each s
  - Reduces bottleneck steps to $O(NT)$
  - Reference: SSJ paper (Auclert, Bardóczy, Rognlie, Straub, Econometrica 2021)

General setup (similar to SSJ paper notation)

- Let superscript s denote infinitesimal shock $dZ_s = dx$ at date s
  - Income at all other dates remains in steady state
- Can iterate backward to get policy functions $\mathbf{c}_t^s$ and transition matrix over discretized states $\mathbf{\Lambda}_t^s$ at each date, which we represent as flattened vectors
- Distribution (over discretized states) and aggregate consumption given by

$$\mathbf{D}_{t+1}^s = (\mathbf{\Lambda}_t^s)'\mathbf{D}_t^s$$

$$C_t^s = (\mathbf{c}_t^s)'\mathbf{D}_t^s$$

- Then calculate intertemporal MPCs $M_{ts} = \partial C_t / \partial Z_s$ as dC_t^s / dx

Insight: only need to iterate backward once!

- Iterate backward separately to recalculate $\mathbf{c}_t^s$ and $\mathbf{\Lambda}_t^s$ for each s? **No!**
- Why? Because **only distance to the shock matters** for policy function:

$$\mathbf{c}_t^s = \mathbf{c}_{t+h}^{s+h} \text{ for any } h$$

- So, just consider one shock at maximal horizon $s = T - 1$, then write (same for $\mathbf{\Lambda}$)

$$\mathbf{c}_t^s = \begin{cases} \mathbf{c}_{ss} & s < t \\ \mathbf{c}_{T-1-(s-t)}^{T-1} & s \geq t \end{cases}$$

Very helpful, but still lots of work

- Backward iteration often costliest, so this is a big help!
- But still, for each *s*, need to iterate forward on distribution
- Economized on top steps but not bottom:

What's going on?

- We care about aggregate $C_t^s = (\mathbf{c}_t^s)'\mathbf{D}_t^s$ [or, more specifically, $M_{t,s} \equiv dC_t^s/dx$]
- We have $\mathbf{c}_t^s = \mathbf{c}_{t+h}^{s+h}$, but that's not true for $\mathbf{D}_t^s$: generally $\mathbf{D}_t^s \neq \mathbf{D}_{t+h}^{s+h}$
- Theorem: to first order,

$$d\mathbf{D}_t^s - d\mathbf{D}_{t-1}^{s-1} = (\mathbf{\Lambda}_{ss}')^{t-1} d\mathbf{D}_1^s$$

- Why? If shock happens at s instead of $s-1$, **one more period to anticipate it**
  - $\rightarrow$ affects date 0 policy $\rightarrow$ affects distribution date-1 distribution $d\mathbf{D}_1^s$
  - $\rightarrow$ carries over to date t distribution via $t-1$ applications of $(\mathbf{\Lambda}_{ss}')^{t-1}$

Effect on aggregates

- We have $d\mathbf{D}_t^s - d\mathbf{D}_{t-1}^{s-1} = (\mathbf{\Lambda}'_{ss})^{t-1} d\mathbf{D}_1^s$
- Effect on $dC_t^s - dC_{t-1}^{s-1} = \mathbf{c}'_{ss}(d\mathbf{D}_t^s - d\mathbf{D}_{t-1}^{s-1})$ is therefore:

$$\mathbf{c}'_{ss}(\Lambda'_{ss})^{t-1} d\mathbf{D}_1^s \qquad (\equiv F_{t,s} \cdot dx)$$

- The matrix $F_{t,s}$ is closely related to Jacobian $M_{t,s}$ via $F_{t,s} = M_{t,s} - M_{t-1,s-1}$
- Can reconstruct $M_{t,s}$ from diagonals $F_{t,s}$ (defining $F_{t,s} \equiv M_{t,s}$ for t or $s = 0$):

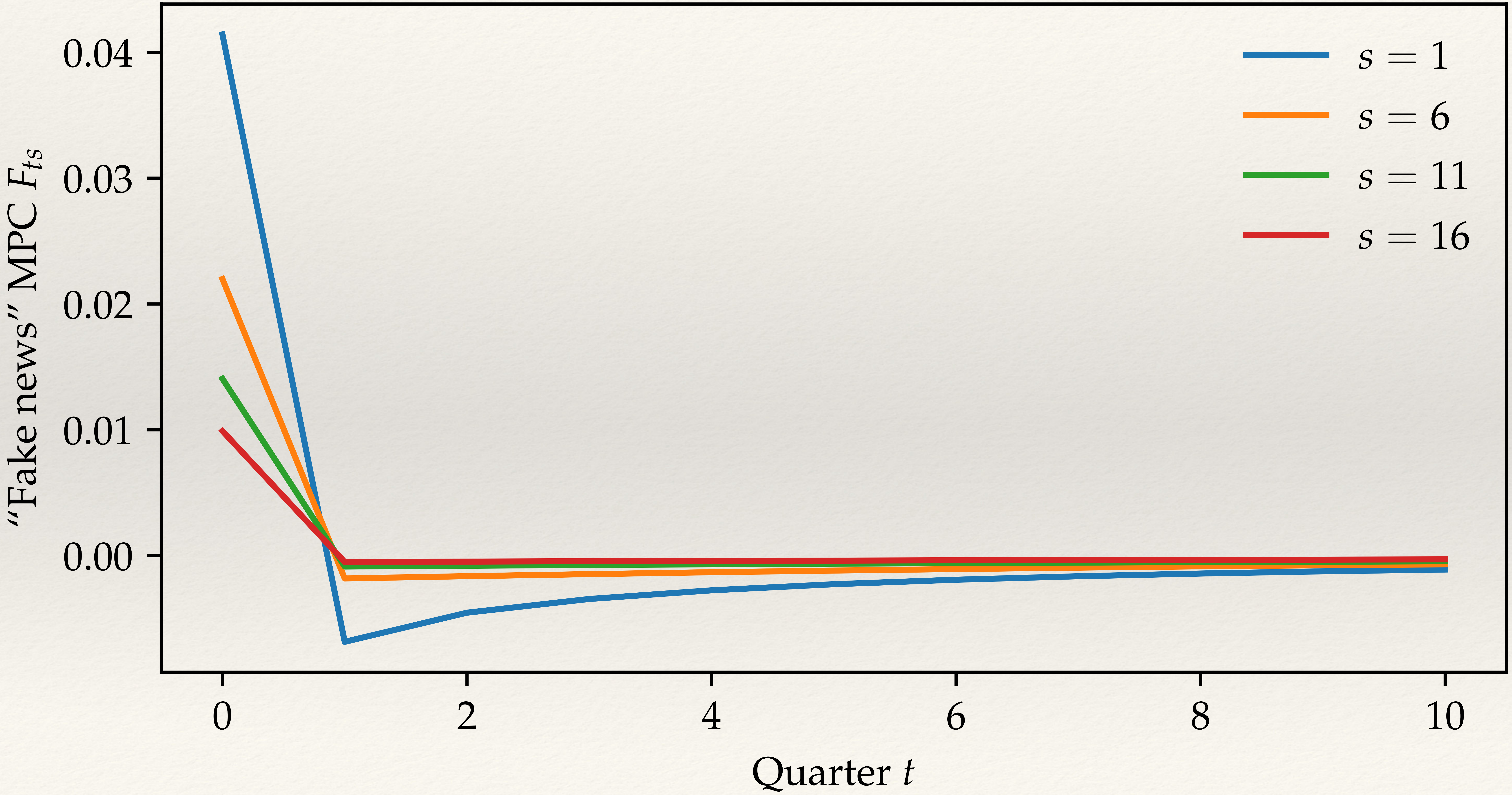
$$M_{3,4} = F_{3,4} + F_{2,3} + F_{1,2} + F_{0,1}$$

What is this F ("fake news matrix")?

- For $t, s > 0$, we have $F_{t,s} = M_{t,s} - M_{t-1,s-1}$
- Why are $M_{t,s}$ and $M_{t-1,s-1}$ different?
  - Because former has **one extra period of anticipation**
  - $F_{t,s}$ is the effect at t of having thought, at 0, that there would be shock at s

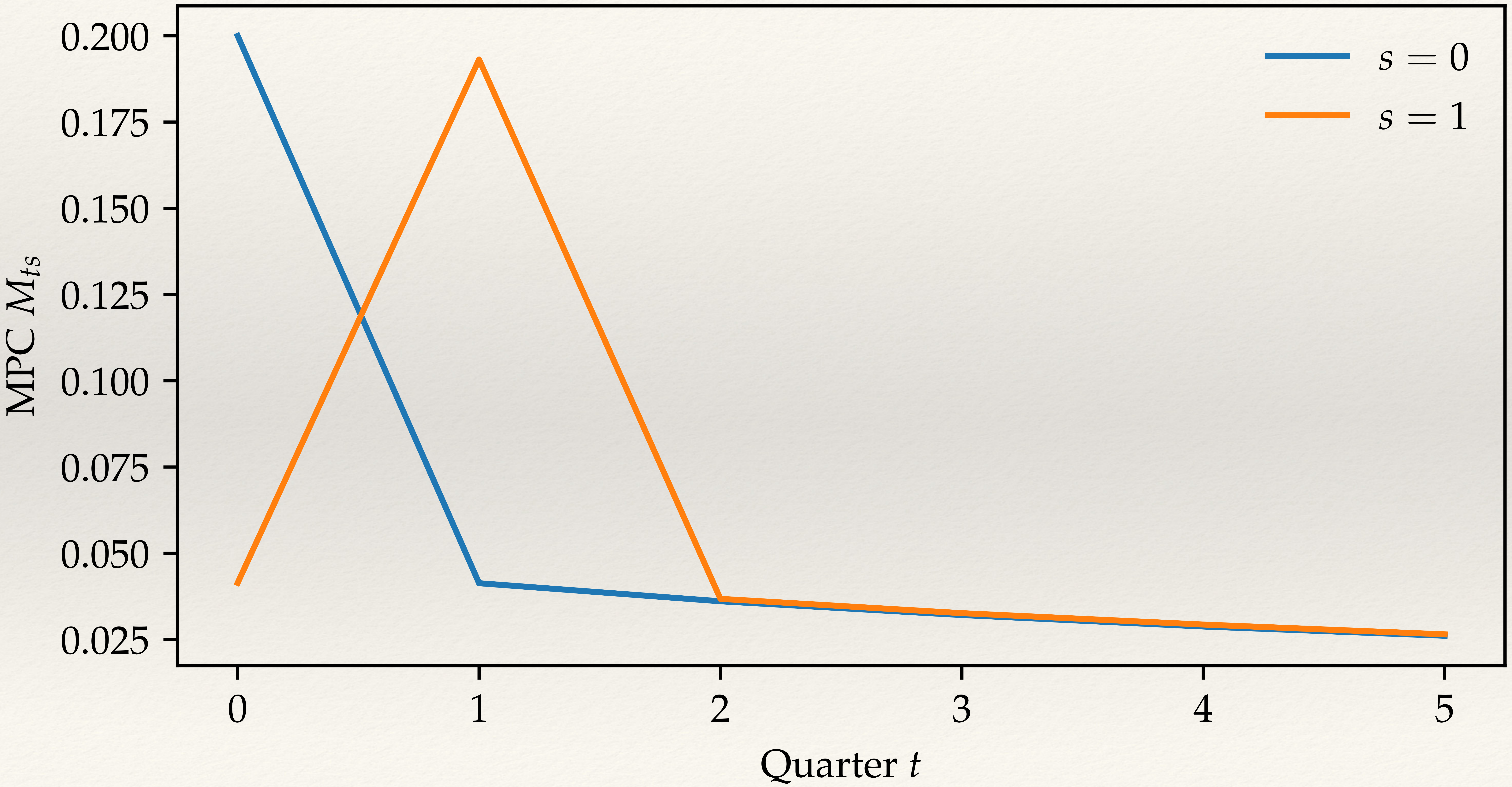
- One interpretation: "fake news shock"
  - $F_{\cdot,s}$ is impulse response to shock at s announced at 0, rescinded at 1

Visualizing columns of **F**

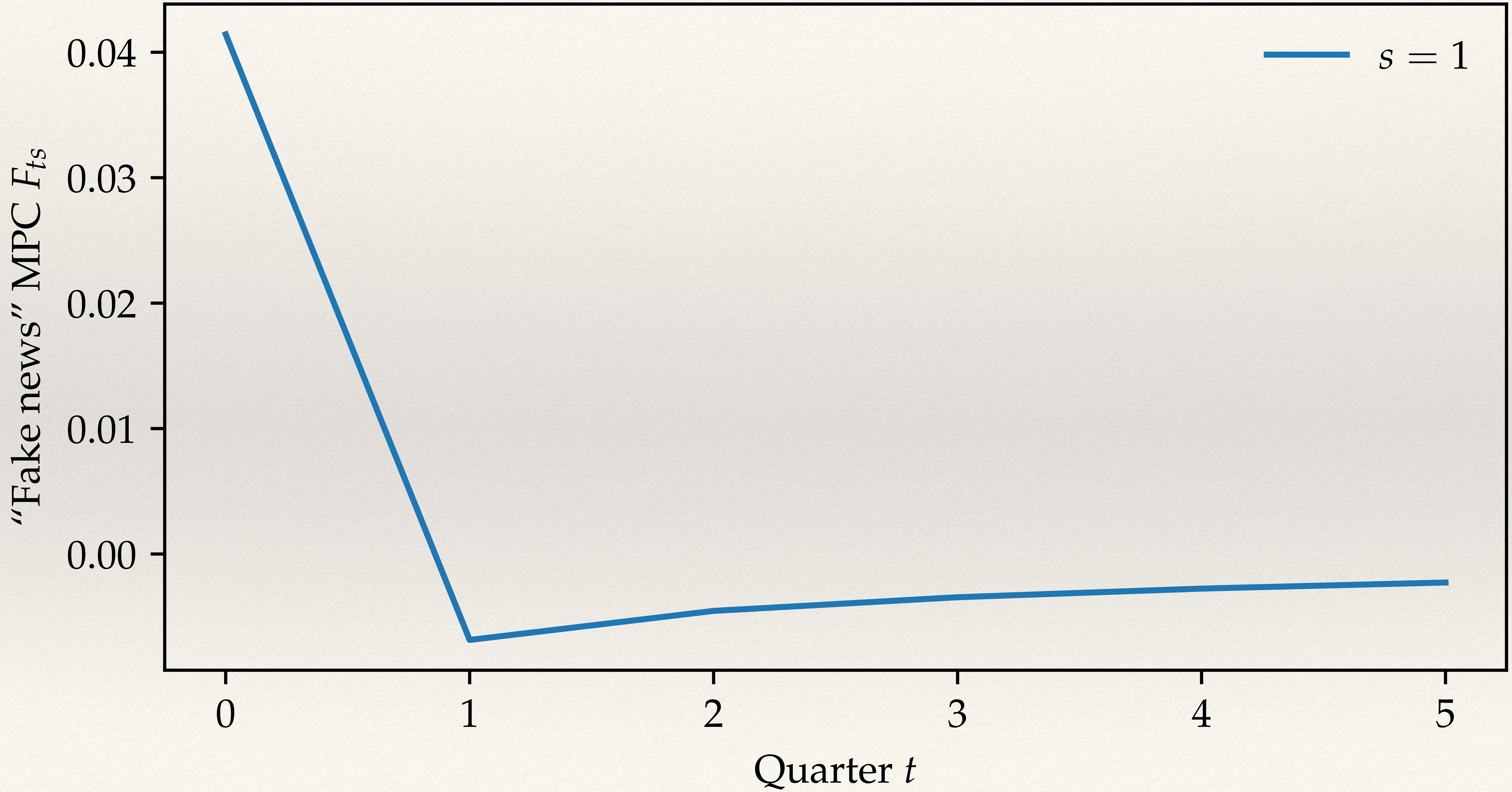


In response to "fake news shock" about future income, you consume more at date 0, then have fewer resources going forward

Difference between $M_{t,1}$ and $M_{t-1,0}$…



… is exactly $F_{t,1}$



Where we stand now

- Reduced finding Jacobian $\mathbf{M}$ to "fake news matrix" $\mathbf{F}$
- Simple formula for $F_{t,s}$ when $t > 0$:

$$F_{t,s}dx = \mathbf{c}_{ss}'(\mathbf{\Lambda}_{ss}')^{t-1}d\mathbf{D}_1^s$$

- Problem: still seems like a lot of work to apply $\mathbf{\Lambda}_{ss}'$ repeatedly to each $d\mathbf{D}_1^s$!
- Solution: evaluate formula from the **left, not the right**!
- Calculate **"expectation functions"** $\mathcal{E}_t \equiv (\mathbf{\Lambda}_{ss})^t\mathbf{c}_{ss}$ **only once**, then evaluate $\mathcal{E}_t' d\mathbf{D}_1^s$
  - $\mathcal{E}_t$ is expected c in t periods for a household who follows steady-state policy

Cracked it open, now have four-step algorithm

- **Step 1:** iterate backward once from shock $dZ_{T-1} = dx$ to obtain all $\mathbf{\Lambda}_t^s$, $\mathbf{y}_t^s$
  - define $\mathcal{Y}_s dx \equiv (d\mathbf{c}_0^s)'\mathbf{D}_{ss}$ and $\mathcal{D}_s dx \equiv (d\mathbf{\Lambda}_0^s)'\mathbf{D}_{ss}$
- **Step 2:** repeatedly apply Λ_{ss} to calculate expectation functions $\mathcal{E}_t \equiv (\Lambda_{ss})^t \mathbf{c}_{ss}$
- **Step 3:** form fake news matrix, which is $F_{0,s} = \mathcal{Y}_s$ and $F_{t,s} = \mathcal{E}'_{t-1}\mathcal{D}_s$ $(t > 0)$
- **Step 4**: calculate all $M_{t,s}$ by cumulatively summing diagonals of $F_{t,s}$

- First 2 steps are $O(NT)$, step 3 is $O(NT^2)$ but can be written as giant matrix multiplication (super efficient, never the bottleneck), step 4 is $O(T^2)$

Summary: the "fake news algorithm"

- Most complex of ideas so far, but now sequence-space Jacobians are practical!
  - Key step is only $O(NT)$, far better than the $O(N^3)$ of state-space methods
  - Example was iMPCs $\mathbf{M}$, but same method for any other Jacobian
  - Various implementation details (for multiple inputs / outputs, numerical vs. automatic differentiation, …): see SSJ paper and appendix
- Reducing Jacobians to "fake news matrices" an interesting step in own right
  - Isolate effects of information, useful for deviations from FIRE

What is a sequence-space solution?

Think about a stochastic economy

- So far we've done **"MIT shocks"**: one-time shocks starting from steady state, where new path becomes known at $t = 0$
- What if shocks **keep hitting** the economy?
- Deficit-financed tax cut example: suppose that

$$T_t = T_{ss} + \sum_{s=0}^{\infty} a_s \epsilon_{t-s}$$

  where $\epsilon_t \equiv \sigma \bar{\epsilon}_t$, with σ scaling size of shocks, and $\bar{\epsilon}_t$ symmetric around 0 and iid with variance 1, determined at date t

- What are implications for path of Y_t?

Sequence-space solution

- Realized output at date t depends on all past realized ϵ_t
- In a stationary world, can write nonlinear solution (won't formally derive):

$$Y_t \equiv Y(\sigma; \epsilon_t, \epsilon_{t-1}, \ldots)$$

  which depends on realized ϵ_t, and also σ because it scales future shocks

- Can then look to **first order** in σ around $\sigma = 0$:

$$\frac{dY_t}{d\sigma} = \frac{\partial Y}{\partial \sigma} + \frac{\partial Y}{\partial \epsilon}\bar{\epsilon}_t + \frac{\partial Y}{\partial \epsilon_{-1}}\bar{\epsilon}_{t-1} + \ldots$$

Simplifying insight

- To first order around $\sigma = 0$:

$$\frac{dY_t}{d\sigma} = \frac{\partial Y}{\partial \sigma} + \frac{\partial Y}{\partial \epsilon}\bar{\epsilon}_t + \frac{\partial Y}{\partial \epsilon_{-1}}\bar{\epsilon}_{t-1} + \dots$$

- Insight: must have $\frac{\partial Y}{\partial \sigma} = 0$!
    - Why? Symmetric shock distribution, **doesn't matter** if we scale by σ or $-\sigma$!
- So to first order, effect of shocks is an MA process:

$$\frac{dY_t}{d\sigma} = \frac{\partial Y}{\partial \epsilon}\bar{\epsilon}_t + \frac{\partial Y}{\partial \epsilon_{-1}}\bar{\epsilon}_{t-1} + \dots$$

Connection to MIT shocks

- An "MIT shock" is a one-time shock to steady state, with no uncertainty
- Corresponds to $\epsilon_0 \neq 0$, where $\sigma = 0$ and $\epsilon_t = 0$ for all $t \neq 0$
- To first order in ϵ_0, the impulse response to MIT shock is therefore

$$\frac{dY_t^{MIT}}{d\epsilon_0} = \frac{\partial Y}{\partial \epsilon_{-t}}$$

  where Y on right is our sequence-space solution

- So we get first-order coefficients in **general sequence-space solution** from impulse to an MIT shock: **MIT shock impulse = first-order MA coefficients**

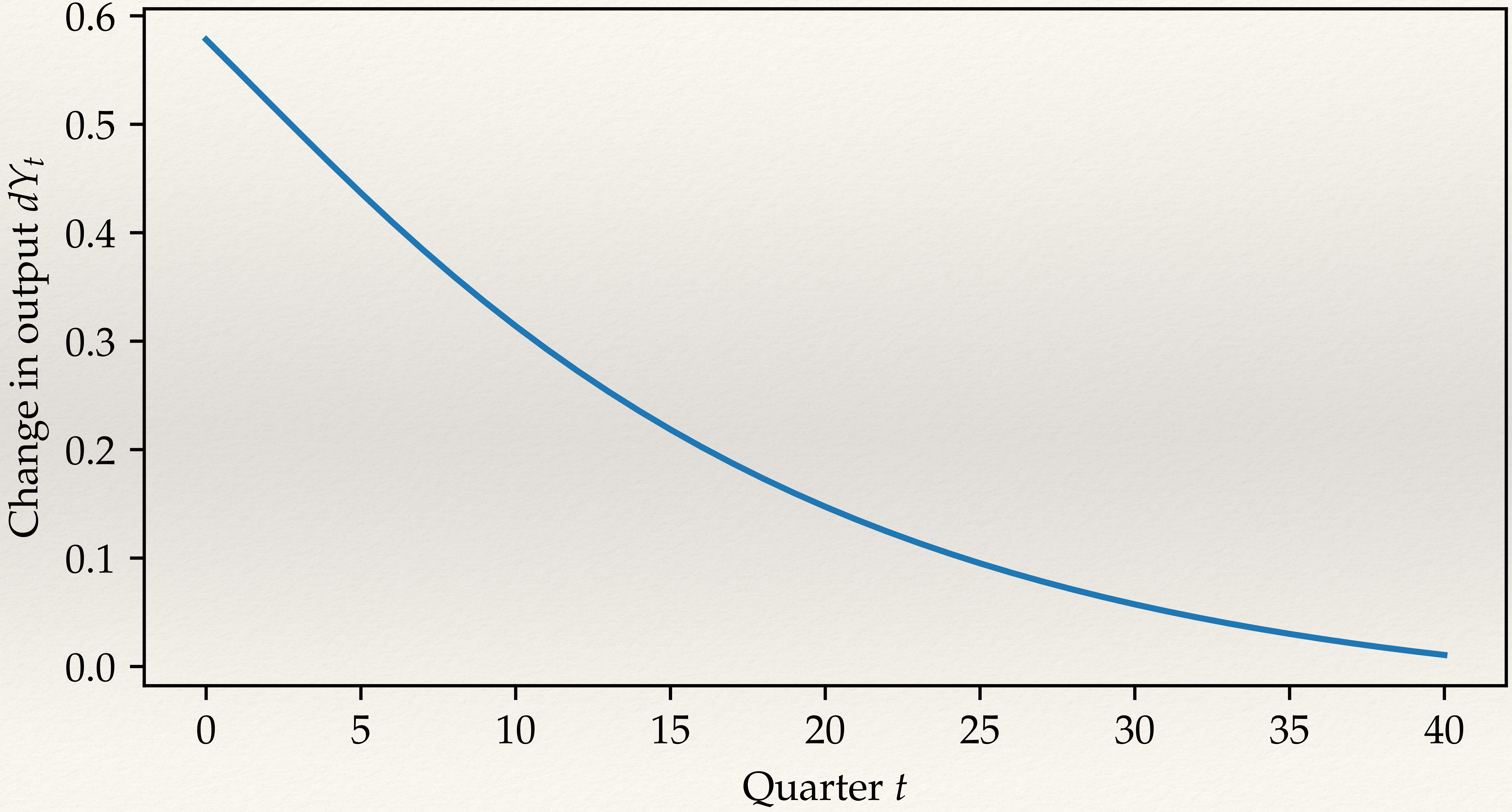
Simulation almost free

- Solve for impulse response to a small MIT shock
  - e.g. what we saw in last lecture for fiscal policy, can use SSJs to solve
- Then, can **simulate** time series to first order in σ, [writing $\epsilon_t = \sigma \bar{\epsilon}_t$]

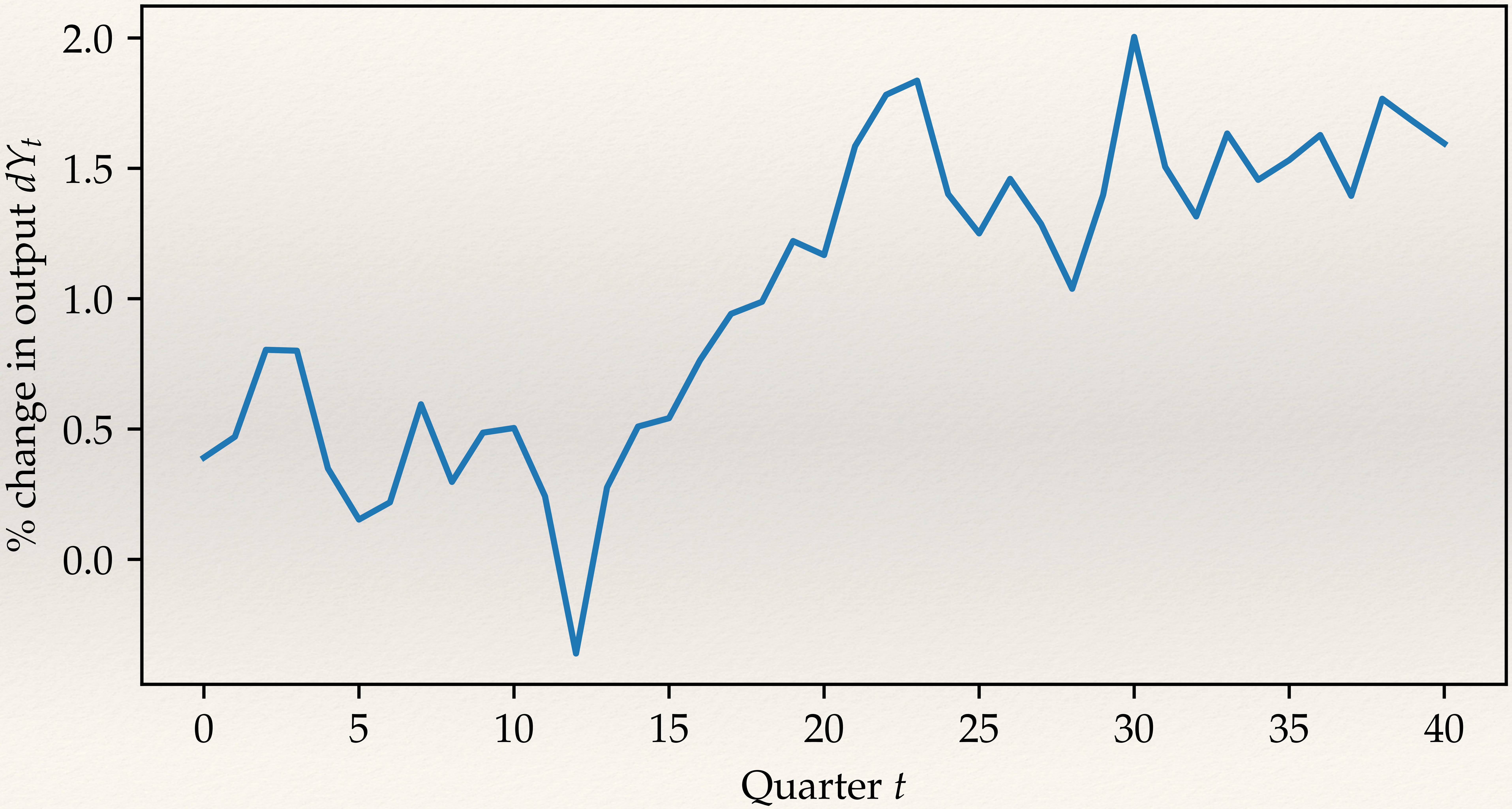
$$dY_t = \frac{\partial Y}{\partial \epsilon}\epsilon_t + \frac{\partial Y}{\partial \epsilon_{-1}}\epsilon_{t-1} + \dots$$

  for any path of $\{\epsilon_t\}$, taking off all $\partial Y / \partial \epsilon$ from MIT shock impulse response
- Only do work to solve MIT shock once, **then almost free**!
  - Insight of **Boppart, Krusell, Mitman (2018)**

Example: start with MIT shock impulse response



Then layer on top of itself to simulate



Analytical second moments

- Often we simulate to obtain moments of the simulated data, e.g. variances and autocorrelations
- But Monte Carlo slow and introduces sampling error, better to write solution

$$dY_t = \frac{\partial Y}{\partial \epsilon}\epsilon_t + \frac{\partial Y}{\partial \epsilon_{-1}}\epsilon_{t-1} + \dots$$

  and **analytically** find covariance

$$\text{Cov}(dY_t, dY_{t'}) = \sigma^2 \sum_{s=0}^{\infty} \frac{\partial Y}{\partial \epsilon_s}\frac{\partial Y}{\partial \epsilon_{s+t'-t}}$$

- **Much faster** in practice, can generalize to multiple series, speed up with FFT

Summing up

- To first order:

Impulse response to MIT shock = MA coefficients in stochastic economy

- Can use to efficiently simulate or get second moments
- Either is almost free once we have the MIT shock impulse

Advantages of the sequence space and Jacobians

Any shock, any heterogeneity

- In fiscal policy lecture, we could solve

$$d\mathbf{Y} = \mathbf{A}^{-1}d\mathbf{B} + d\mathbf{T}$$

- Once we've calculated inverse asset Jacobian $\mathbf{A}^{-1}$, we can solve for response to **any time path** of $d\mathbf{B}$ and $d\mathbf{T}$ almost instantly
- Similar "general equilibrium Jacobian" mapping in more complex cases

- Suppose we want 100 different types, over and above our heterogeneity
  - Just take weighted average of the $\mathbf{A}$s to get economy-wide $\mathbf{A}$

Some advantages of the sequence space: summary

1. Can get response to **any shock**
2. Can easily handle almost **any heterogeneity**
3. Can simulate, get **any second moments**, use to **estimate model** [to come!]
4. Can implement **non-rational expectations** [to come!]
5. Can get **informative decompositions** [e.g. $d\mathbf{Y} = d\mathbf{G} - \mathbf{M}d\mathbf{T} + \mathbf{M}d\mathbf{Y}$]

(These advantages carry over in part to any MIT shock / sequence-space method, but best by far when we're using Jacobians!)