Dynamic Mode Decomposition of a CEX Panel*

Thomas J. Sargent
New York University

Yatheesan J. Selvakumar
New York University

December 12, 2023

Abstract

We state sufficient conditions on matrices $\mathbf{A}, \mathbf{C}, \mathbf{G}, \mathbf{R}$ defining a linear state-space model that allow us to infer them from a reduced-rank first-order vector autoregression (VAR). We use a Dynamic Mode Decomposition (DMD) to compute that VAR and show how dynamic modes relate to hidden Markov states in the state-space representation. We compute a reduced-rank first order VAR for quantiles for income, post tax and transfer income, and consumption from the US Consumer Expenditure Survey. We relate our findings to a "neo-classical synthesis" that was widely practiced by leading twentieth-century macroeconomists and still is today.

Keywords: Singular Value Decomposition, Dynamic Mode Decomposition, Vector Autoregression, Innovations Representation, Pseudo Innovations Representation, Business Cycles, Inequality.

*We thank Patrick Kehoe and Elena Pastorini for helpful discussions.

1 Introduction

Leading practitioners of 20th century macroeconomics accepted a "neoclassical synthesis" that separated public policies intended to attenuate business cycles from policies intended to redistribute income. That synthesis justified models that were designed to explain aggregate quantities and prices and that are cast in terms of representative agents.[1] Thus, James Tobin and other members of President John F. Kennedy's economic team believed that "a rising tide lifts all boats" and practiced a macroeconomics that intentionally ignored distribution effects.[2] Empirical evidence assembled by the National Bureau of Economic Research that culminated in Burns and Mitchell (1946) had isolated a one-dimensional reference business cycle that theorists practicing the "neoclassical synthesis" sought to explain.

This paper is squarely in that Burns and Mitchell tradition. We use a Dynamic Mode Decomposition (DMD)[3] to extract aggregate US business cycle dynamics from Consumer Expenditure Survey (CEX) data between 1990 and 2021. The resulting descriptive statistical model of US business cycle dynamics is an example of what Koopmans (1947) called a "non-structural Kepler stage" model that compresses data and reveals patterns that "structural Newton stage" models would explain in terms of parameters that describe market structures and decision makers' preferences, constraints, and information flows.[4] Section 2 describes how DMD estimates a reduced-rank first-order vector autoregression in a setting in which least-squares regression coefficients are underdetermined. It provides pseudo code that section 3 then applies to CEX data.

[1] For an account of the neoclcassical synthesis and its relationship to old and new heterogeneous agent Keynesian models, please see Sargent (2023). For evidence that Robert E. Lucas, Jr., adhered to it, please see Sargent (2015).

[2] See Heller (1966) and Tobin (2015) for authoritative accounts of New Frontier economic policy advice and Blinder (2022) for a faithful account of President Kennedy's approach to macroeconomic policy.

[3] For example, see Brunton and Kutz (2022, sec. 7.2).

[4] Koopmans regarded Burns and Mitchell (1946) as such a "Kepler" stage model of business cycles, in contrast to the structural, simultaneous stochastic difference equation models of business cycles that could be constructed with tools developed by Koopmans (1950), Hood and Koopmans (1953), and Marschak (1953). The present analysis is the spirit of Sargent and Sims (1977), though as we shall see in section 4 , the statistical model of hidden factors is quite different.

Section 3 infers "dynamic modes" that underlie the evolution of CEX percentiles, then computes and plots them. Section 4 goes on to describe how dynamic modes are related to other objects that appear in what we call a "pseudo innovations representation". Section 4 relates a pseudo innovations representation to the authentic innovations representation that is associated with a particular linear Gaussian state space model.[5] That link allows us to interpret DMD modes as estimates of hidden factors in that state-space model. We state conditions on a linear state space representation that are sufficient to make an infinite-order VAR implied by an authentic innovations representation align with a reduced-rank first-order VAR inferred by the DMD algorithm. These conditions require that hidden factors in the state-space model follow univariate first-order autoregressions with shocks that are possibly correlated contemporaneously, and that each period a long column vector of observables provides enough information about those hidden factors. As the length of the observation vector grows, the pseudo innovations representation aligns better and better with the authentic innovations representation. This occurs because a big enough collection of contemporaneous noisy measurements of linear combinations of the hidden factors contains the same information about the hidden states as would be provided by an infinite history of those measurements.[6] When these conditions prevail, it is possible to recover matrices defining the state-space model from objects computed by DMD, a finding that justifies a fast algorithm for estimating parameters of a state-space system. Section 5 llustrates, checks, and stress-tests our section 4 theoretical findings by conducting two "lab experiments" with an example theoretical model. Section 6 concludes.

[5]See Ljungqvist and Sargent (2018, secs. 2.7-2.10) for a compact description of innovations representations and their link to vector autoregressions.

[6]That averaging observations over large cross-sections can accelerate learning about hidden dynamic factors is reminiscent of Chamberlain and Rothschild (1982). More recent examples include Forni et al. (2000) and Forni and Lippi (2001)

2 Reduced Rank First-Order VAR

2.1 Data and Model

Let $\mathbf{y}_t$ be an an $M \times 1$ vector of variables at $t = 1, \ldots, T+1$, and assume that $M > T$, so there are more variables than time periods. Horizontally stack $\mathbf{y}_t$ vectors to create two $M \times T$ "tall and skinny" matrices $\mathbf{Y}$ and $\mathbf{Y}'$:

$$\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_T] \tag{1}$$

$$\mathbf{Y}' = [\mathbf{y}_2, \mathbf{y}_3, \ldots, \mathbf{y}_{T+1}]. \tag{2}$$

We estimate a first-order vector autoregression

$$\mathbf{y}_t = \mathbf{B}\,\mathbf{y}_{t-1} + \mathbf{a}_t\,, \quad \mathbf{a}_t \perp \mathbf{y}_{t-1} \tag{3}$$

$$\mathbb{E}[\mathbf{a}_t\,\mathbf{a}_t^\top] = \mathbf{\Omega}. \tag{4}$$

Use $M(T+1)$ data points to estimate the M^2 parameters in $\mathbf{B}$. Since $M^2 > M(T+1)$, a least squares estimator $\widehat{\mathbf{B}}$ of $\mathbf{B}$ is underdetermined, so we choose[7]

$$\widehat{\mathbf{B}} = \underset{\text{rank}(B)=N}{\arg\min} \,||\,\mathbf{Y}' - B\,\mathbf{Y}\,||_F \tag{5}$$

To compute $\widehat{\mathbf{B}}$, we first represent $\mathbf{Y}$ with a reduced Singular Value Decomposition (SVD)[8]

$$\mathbf{Y} = \widetilde{\mathbf{U}}\widetilde{\mathbf{\Sigma}}\widetilde{\mathbf{V}}^\top$$

[7] $||\,\mathbf{D}\,||_F$ denotes the Frobenius norm $\sum_{i,j} \mathbf{D}_{i,j}^2$.

[8] Anderson (1951) and Anderson (1999) approach estimating reduced-rank regressions by first computing an unrestricted least-squares regression. This approach is infeasible in our $M > T$ setting.

where $\widetilde{\mathbf{U}}$ is $M \times T$, $\widetilde{\mathbf{\Sigma}}$ is $T \times T$ and $\widetilde{\mathbf{V}}^\top$ is $T \times T$. We compress $\mathbf{Y}$ and approximate it in terms of its N largest singular values:

$$\mathbf{Y} \approx \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top,$$

where from now on we set $\mathbf{Y}$ equal to this compressed version $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top$ of the original data set, $\mathbf{U} = \widetilde{\mathbf{U}}[:, : N]$, $\mathbf{\Sigma} = \widetilde{\mathbf{\Sigma}}[: N, : N]$ has N singular values as its only non-zero entries, and $\mathbf{V}^\top = \widetilde{\mathbf{V}}^\top[: N, :]$. Here $\mathbf{U}$ is $M \times T$, $\mathbf{V}$ is $T \times N$, $\mathbf{\Sigma}$ is $N \times N$, and $\mathbf{V}^\top$ is $N \times T$. We use this reduced-order SVD approximation of the original $\mathbf{Y}$ matrix to compute

$$\widehat{\mathbf{B}} = \mathbf{Y}' \, \mathbf{Y}^+,$$

where $\mathbf{Y}^+ = \mathbf{V} \, \mathbf{\Sigma}^{-1} \, \mathbf{U}^\top$ is a generalized inverse of $\mathbf{Y}$ that verifies $\mathbf{Y}^+ \, \mathbf{Y} = \mathbf{I}_{T \times T}$.

We use these inputs to represent our reduced-rank first-order VAR in terms of N underlying "dynamic modes" by implementing the following steps:[9]

1. Use $\mathbf{U}$ and $\widehat{\mathbf{B}}$ to construct $\widetilde{\mathbf{B}}_{N \times N} = \mathbf{U}^\top \, \widehat{\mathbf{B}} \, \mathbf{U}$.

2. Construct an eigen decomposition

$$\widetilde{\mathbf{B}} = \mathbf{W} \, \mathbf{\Lambda} \, \mathbf{W}^{-1}$$

   where columns of the $N \times N$ matrix $\mathbf{W}$ are eigenvectors of $\widetilde{\mathbf{B}}$ and eigenvalues of $\widetilde{\mathbf{B}}$ appear on the diagonal of the diagonal matrix $\mathbf{\Lambda}$.

3. Define

$$\mathbf{\Phi}_{M \times N} \equiv \mathbf{Y}' \, \mathbf{V} \, \mathbf{\Sigma}^{-1} \, \mathbf{W} \, .$$

4. Recognize a finding of Tu et al. (2014) that columns of $\mathbf{\Phi}$ are eigenvectors of $\widehat{\mathbf{B}}$

[9] Here we are following and extending steps laid out by Brunton and Kutz (2022, sec. 7.2).

that share eigenvalues with $\widetilde{\mathbf{B}}$, so that

$$\widehat{\mathbf{B}} = \mathbf{\Phi}\mathbf{\Lambda}\mathbf{\Phi}^{+} \tag{6}$$

where $\mathbf{\Phi}$ is $M \times N$, $\mathbf{\Lambda}$ is $N \times N$ and $\mathbf{\Phi}^{+}$ is the $N \times M$ (left) generalized inverse that verifies $\mathbf{\Phi}^{+}\mathbf{\Phi} = I_{N\times N}$.

5. Define an $N \times 1$ vector of "dynamic modes" by[10]

$$\widetilde{\mathbf{x}}_t \equiv \mathbf{\Phi}^{+}\,\mathbf{y}_t\,.$$

6. Use representation (6) of $\widehat{\mathbf{B}}$ to express the estimated reduced rank first-order VAR as

$$\mathbf{y}_t = \mathbf{\Phi}\mathbf{\Lambda}\mathbf{\Phi}^{+}\,\mathbf{y}_{t-1} + \widehat{\mathbf{a}}_t, \quad \widehat{\mathbf{a}}_t \perp \mathbf{y}_{t-1} \tag{7}$$

7. Multiply both sides of equation (7) by the $N \times M$ matrix $\mathbf{\Phi}^{+}$ to obtain $\widetilde{\mathbf{x}}_t = \mathbf{\Lambda}\widetilde{\mathbf{x}}_{t-1} + \mathbf{\Phi}^{+}\widehat{\mathbf{a}}_t$. Use it together with $\widetilde{\mathbf{x}}_{t-1} = \mathbf{\Phi}^{+}\,\mathbf{y}_{t-1}$ in equation (7) to form the system

$$\begin{aligned} \widetilde{\mathbf{x}}_t &= \mathbf{\Lambda}\widetilde{\mathbf{x}}_{t-1} + \mathbf{\Phi}^{+}\widehat{\mathbf{a}}_t \\ \mathbf{y}_t &= \mathbf{\Phi}\mathbf{\Lambda}\widetilde{\mathbf{x}}_{t-1} + \widehat{\mathbf{a}}_t \end{aligned} \tag{8}$$

8. Construct moving average representation for modes $\widetilde{\mathbf{x}}_t$:

$$\widetilde{\mathbf{x}}_{t+j} = \mathbf{\Lambda}^{j}\widetilde{\mathbf{x}}_t + \sum_{s=0}^{j-1} \mathbf{\Lambda}^{s}\mathbf{\Phi}^{+}\,\mathbf{a}_{t+j-s}$$

[10] In section 4, we'll link these modes to an $N \times 1$ hidden state vector $\mathbf{x}_t$ that is defined implicitly by $\widetilde{\mathbf{x}}_t \equiv \mathbb{E}[\mathbf{x}_t \mid \mathbf{y}_t]$. We'll also define another projection $\widehat{\mathbf{x}}_t = \mathbb{E}[\mathbf{x}_t \mid \mathbf{y}_{t-1}]$.

9. Construct j step ahead conditional covariances of modes:

$$\mathbb{E}(\widetilde{\mathbf{x}}_{t+j} - \mathbb{E}\,\widetilde{\mathbf{x}}_{t+j}|\widetilde{\mathbf{x}}_t)(\mathbf{x}_{t+j} - \mathbb{E}\,\widetilde{\mathbf{x}}_{t+j}|\widetilde{\mathbf{x}}_t)^\top = \sum_{s=0}^{j-1} \mathbf{\Lambda}^s \mathbf{\Phi}^+ \widehat{\mathbf{\Omega}} (\mathbf{\Phi}^+)^\top \mathbf{\Lambda}^s \tag{9}$$

In sections 4 and 5, we describe conditions under which outcomes of the DMD algorithm can be associated with objects that define a linear, Gaussian hidden Markov model. Those sections indicate indicate additional purposes a DMD analysis can serve. We set the stage for these by first proceeding to our application to quantiles from US CEX data.

3 US income and consumption dynamics

We construct representation (8) and use it to describe salient features of US business cycles from 1990 to 2021. We want to study connections between cross-section dynamics and the macroeconomic averages that macroeconomic models are designed to describe and understand. We characterize cross-section distributions by their quantiles.[11]

3.1 Data matrix Y

We gather data on private income, post-tax income and consumption from quarterly waves of the Interview section of the Consumer Expenditure Survey between 1990 and 2021.

For **Consumption** we sum total expenditure in the current month (TOTEXPCQ) and total expenditure in the previous month (TOTEXPPQ). We do this because, while households are interviewed once every three months, timings may not match calendar quarters. Summing both measures is a common approach that the BLS has suggested.

[11] Let $F_t[0, B] \to [0, 1]$ be a cumulative distribution function for a possibly bounded nonnegative random variable at t. Associated with c.d.f. F is a "quantile function" $Q : [0, 1] \to [0, B]$ that under some regularity conditions is the inverse of F. Here we'll work with "percentiles" defined as $Q(\frac{1}{i}), i = 1, \ldots, 100$.

Private income consists of a subset of categories that sum to before-tax income. Due to changes in definitions over time, the relevant codes are FINCBTAX (1990-2004) and FINCBTXM (2004-2021). FINCBTXM also changed its definition in 2013, but retained the same code name. We categorize the sub-categories of FINCBT* series into either private income or transfer income. Private income is defined as the sum of wages, business income, financial income, income from rental properties and pensions or annuities from any source. Transfer income is all other sub-categories of before-tax income that are not private income. Table 6 in Appendix A describes the categorization.

Post-tax income is private income plus transfers (i.e. before-tax income) minus taxes paid. Relevant codes are FINCATAX (1990-2004), FINCATXM (2004 - 2013) and FINCATXEM (2013-2021).

The Interview survey asks about income over the past 12 months, while it asks about quarterly amounts of consumption. As recommended by the BLS, we divide all income data by four. We remove all households that consume more than their reported annual incomes in a single quarter.

From these data, we constructed time series of cross-section distributions for consumption, private income, and post-tax income. We'll describe how we did this for consumption and note that we followed the same steps to construct panels for the other two variables. For wave t, we rank all $\tilde{I}$ households by their consumption levels $c_{1,t}, c_{2,t}, \ldots, c_{\tilde{I},t}$. Then, we split them into one hundred equally sized bins that we call percentiles. If $\tilde{I}$ is divisible by 100, then each percentile contains $I = \frac{\tilde{I}}{100}$ households. If not, the bottom 99 bins each contain I households and the 100th bin contains the remainder. Except for this detail, define $\tilde{q}_{p,t}$ as mean consumption of households in percentile p:

$$\tilde{q}_{p,t} = \frac{1}{I} \sum_{j=1}^{I} c_{j+I(p-1),t} \quad \text{for } p = 1, \ldots, 100$$

Quantity $\tilde{q}_{p,t}$ is the consumption level of the p-th percentile of the consumption

distribution. We repeated this step for every wave to obtain a time-series of consumption at each percentile. We noticed suspicious "jumps" in the constructed time-series for private and post-tax income in Q1-2013 and suspect that these come from the aforementioned changes in variable definitions. To respond to this situation, we spliced the pre-jump and post-jump data in Q1-2013 by setting Q1-2013 consumption equal to Q4-2012, and recursively adding subsequent changes in consumption in the post-jump data to the adjusted series. We deflated each time-series by the Personal Consumption Expenditures price index to obtain real objects and then seasonally adjusted. Call the cleaned, real, seasonally adjusted consumption percentiles $q_{p,t}$.

We removed 1st and 2nd percentile bins, because consumption for those are negative in some periods. Since earnings in the CEX are top coded, we also removed the 99th percentile bins. We formed quarterly growth rates so that our consumption growth variable is

$$y_{p,t}^{cons} = \log q_{p,t} - \log q_{p,t-1} \quad p = 3, \ldots, 98 \text{ and } t = 1, \ldots, T$$

Proceeding in this way for the two income concepts, our final data set is a quarterly, seasonally adjusted time-series of private income, post-tax income, and consumption growth distributions from 1990 to 2021.

We stack our three cross-sections to form matrix $\mathbf{Y}$. Consequently, $M = 291$ and

$$\mathbf{Y} = \begin{bmatrix} y_{3,1}^{private} & \cdots & y_{3,T}^{private} \\ \vdots & \cdots & \vdots \\ y_{98,1}^{private} & \cdots & y_{98,T}^{private} \\ y_{3,1}^{post-tax} & \cdots & y_{3,T}^{post-tax} \\ \vdots & \cdots & \vdots \\ y_{98,1}^{post-tax} & \cdots & y_{98,T}^{post-tax} \\ y_{3,1}^{cons} & \cdots & y_{3,T}^{cons} \\ \vdots & \cdots & \vdots \\ y_{98,1}^{cons} & \cdots & y_{98,T}^{cons} \end{bmatrix} \tag{10}$$

To interpret "dynamic modes" that we shall recover from our reduced-rank first-order VAR, we calculate seasonally adjusted growth rates of the first two moments of the private income, post-tax income and consumption distributions. We describe how we do this for the consumption series, again noting that we used the same procedure to compute cross section moments for the two income concept series.

To obtain one cross-section moment, we take the cross-section mean of consumption $\bar{q}_t := \frac{1}{97}\sum_{p=3}^{98} q_{p,t}$. Then we seasonally adjust and take log differences $\log \bar{q}_t - \log \bar{q}_{t-1} =: \mu_t^{cons}$. (Taking means across quantiles $q_{p,t}$ or households $c_{i,t}$ yields the same quantities.)

To obtain another moment, we compute the cross-section variance of consumption $s_t := \frac{1}{97}\sum_{p=3}^{98}(q_{p,t} - \bar{q}_t)^2$. Then we seasonally adjust and take log differences $\log s_t - \log s_{t-1} =: \sigma_t^{cons}$.[12] The final objects, μ_t^{cons} and σ_t^{cons}, represent growth rates of the first two moments of the consumption distribution.

[12] An alternative procedure would compute the variance across households $s_t := \frac{1}{I}\sum_{i=1}^{I}(c_{i,t} - \bar{q}_t)^2$. Note that in constructing σ_t^{cons}, we are ignoring the variances within each percentile bin. We also compute the variance from percentiles as described in Chang et al. (2021, Appendix A) with very similar empirical results.

3.2 A Reduced-Rank First-Order VAR

We set $N = 2$ and apply the steps in Section 2 to data matrix $\mathbf{Y}$ and extract what DMD researchers call "dynamic modes", i.e., the N components of $\widetilde{\mathbf{x}}_t = \mathbf{\Phi}^+ \mathbf{y}_t$. Thus, mode i is $\widetilde{\mathbf{x}}_{i,t} = \mathbf{\Phi}^+_{:,i} \mathbf{y}_t$, where $\mathbf{\Phi}^+_{:,i}$ is the ith column of the $M \times N$ matrix $\mathbf{\Phi}^+$. The modes evolve according to the first block of equations in the dynamic system (8) or of the corresponding dynamic system (8) with one-step ahead prediction errors in the dynamic modes having been orthogonalized via a Cholesky decomposition. i.e., as $\widetilde{\mathbf{x}}_{t+1} = \mathbf{\Lambda} \widetilde{\mathbf{x}}_t + \mathbf{H}\, \mathbf{e}_{t+1}$. Our algorithm infers

$$\mathbf{\Lambda} = \begin{bmatrix} 0.83 & 0 \\ 0 & 0.72 \end{bmatrix}, \quad \mathbf{H}\,\mathbf{H}^\top = \begin{bmatrix} 0.11 & 0.05 \\ 0.05 & 0.13 \end{bmatrix}, \tag{11}$$

so both modes are persistent and their innovations have a similar variances. Since $\mathbf{H}\,\mathbf{H}^\top$ is not diagonal, the modes are correlated. The one-step ahead conditional correlation between the two modes is 0.48. By setting $h = 100$ in formula (9) we approximate an unconditional covariance matrix to be

$$\begin{bmatrix} 0.33 & 0.14 \\ 0.14 & 0.26 \end{bmatrix}$$

with an associated unconditional correlation of 0.46 between the two modes. These objects will be useful below in rationalizing our findings against the existing results in the literature.

Figure 1 displays a mode $\widetilde{\mathbf{x}}_{1,t}$, standardized to zero mean and unit standard deviation. The dark line in Figure 1 plots mode $\widetilde{\mathbf{x}}_{1,t}$, i.e., the inner product of $\mathbf{y}_t$ with $\mathbf{\Phi}^+_{:,1}$; the three colored dashed lines show standardized growth of average private income ($\mu_t^{private}$), after-tax income ($\mu_t^{post-tax}$) and consumption (μ_t^{cons}) defined at the bottom of section 3.1.

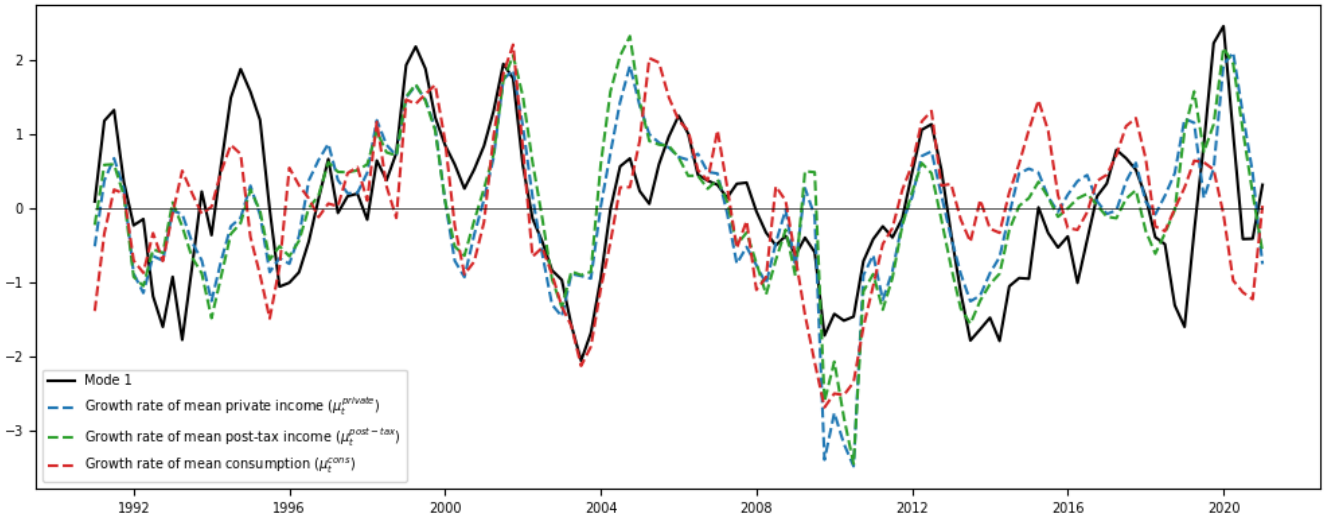

Figure 1: Mode 1, $\tilde{\mathbf{x}}_{1,t} = \mathbf{\Phi}^{+}_{:,1} y_t$ (standardized)

,

	$\mu_t^{private}$		$\mu_t^{post-tax}$		μ_t^{cons}	
Constant	0.32	0.31	0.36	0.36	0.19	0.18
	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Mode 1	0.66		0.64		0.48	
	(0.08)		(0.07)		(0.06)	
Mode 2		0.85		0.81		0.42
		(0.07)		(0.06)		(0.07)
R-squared	0.38	0.55	0.41	0.57	0.32	0.21

Table 1: Regression results of growth of first moment of distributions (μ_t) on modes

Table 1 shows the results of six bivariate regressions of the growth of average incomes and consumption (the dashed lines in Figure 1) on modes 1 and 2. The coefficients on both modes are significant and positive in all regressions, and the R^2 is also substantial. The R^2 associated with mode 1 is higher than mode 2 for consumption, and vice versa for income. Evidently, mode 1 is highly correlated with cross-section averages for all three cross-sections, so it seems to be Burns and Mitchell's unidimensional "reference cycle" index of business cycles.

To indicate how quantiles of our three cross-sections respond to mode 1, figure 2 plots the first column of $\mathbf{\Phi}$. All components (or "loadings") are positive: as mode 1

rises, so do all quantiles of incomes and consumption.

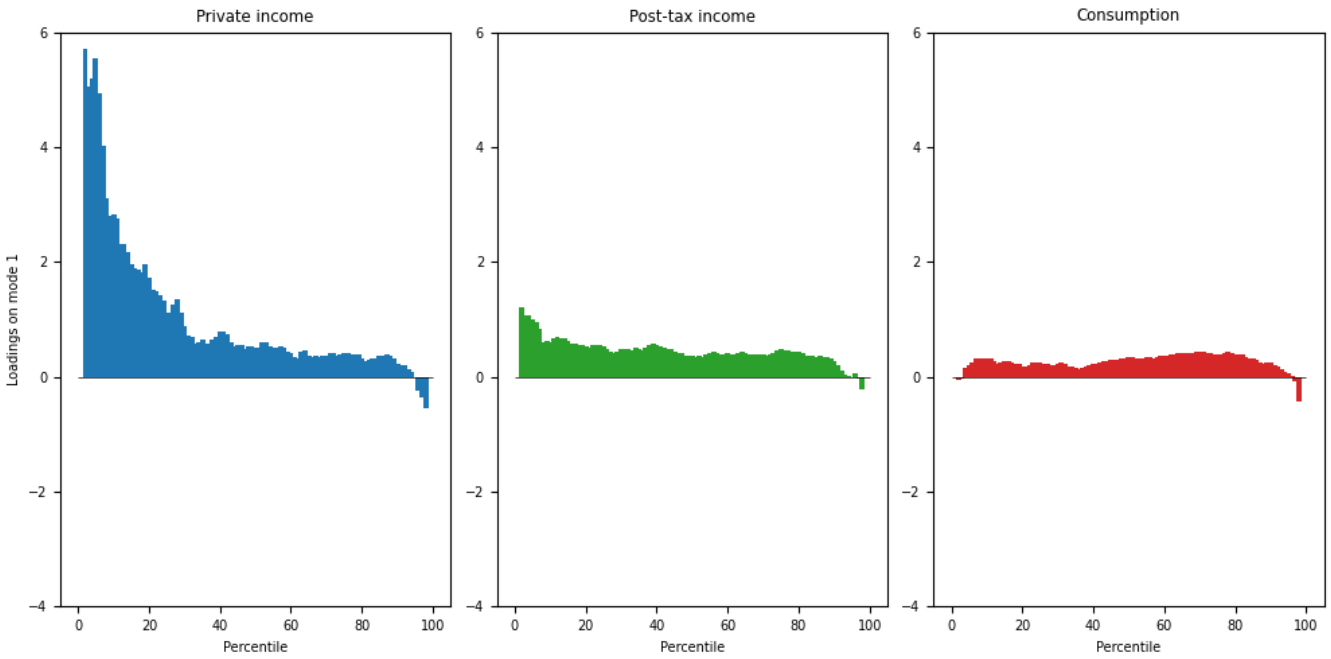

Figure 2: Loadings $\mathbf{\Phi}_{\cdot,1}$ of quantiles on mode 1

Low quantiles of private income growth are more sensitive to mode 1 than are high quantiles. All quantiles of total income growth are less sensitive to mode 1 than are corresponding quantiles of private income growth. Most consumption quantiles are even less sensitive.

Figure 3 shows mode $\tilde{\mathbf{x}}_{2,t}$. The dark line in Figure 3 plots standardized mode $\widetilde{\mathbf{x}}_{2,t}$ while the three colored dotted lines show standardized growth rates of the cross-section standard deviations of income and consumption variables. Table 2 shows counterparts to the Table 1 regressions that use cross-section standard deviations, instead of cross-section averages. Coefficients on both modes are positive and significant, and R^2 statistics of the mode 2 regressions are high for all three variables.

To indicate how quantiles of our three cross-sections respond to mode 2, figure 2 plots the second column of $\mathbf{\Phi}$. When mode 2 rises, higher quantiles of all three cross-sections increase more than lower quantiles. Private income growth falls for low

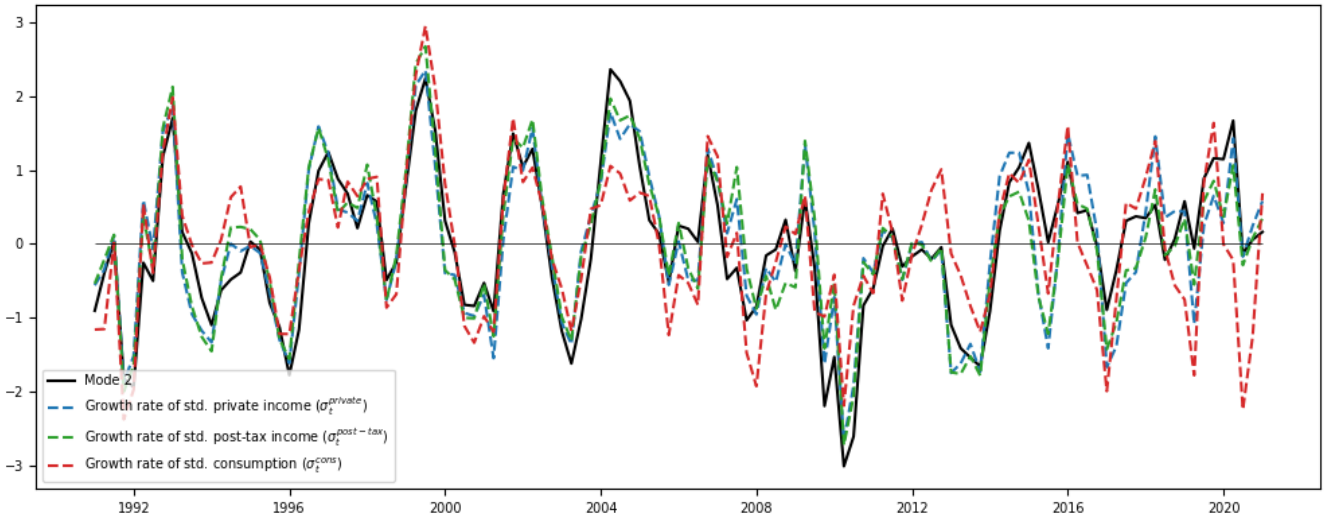

Figure 3: Mode 2, $\tilde{\mathbf{x}}_{2,t} = \mathbf{\Phi}^{+}_{:,2} y_t$ (standardized)

	$\sigma_t^{private}$		$\sigma_t^{post-tax}$		σ_t^{cons}	
Constant	0.39	0.39	0.43	0.42	0.11	0.10
	(0.11)	(0.06)	(0.12)	(0.06)	(0.13)	(0.09)
Mode 1	0.50		0.76		0.70	
	(0.20)		(0.20)		(0.21)	
Mode 2		2.1		2.19		1.98
		(0.10)		(0.11)		(0.15)
R-squared	0.05	0.77	0.11	0.79	0.08	0.59

Table 2: Regressions of growth of second moment of distributions (σ_t) on modes

quantiles. Mode 2 seems to be an "inequality mode".

3.3 Connections with other data summaries

Our DMD detects different responses of income and consumption quantiles to two dynamic modes. The reference cycle mode 1 affects low quantiles of private income the most, and provokes smaller responses of other quantiles. In contrast, responses of all consumption quantiles to model 1 are similar.

Various papers have inferred that higher quantiles consumption are more responsive to business cycles than lower ones. Parker and Vissing-Jorgensen (2009) regress the

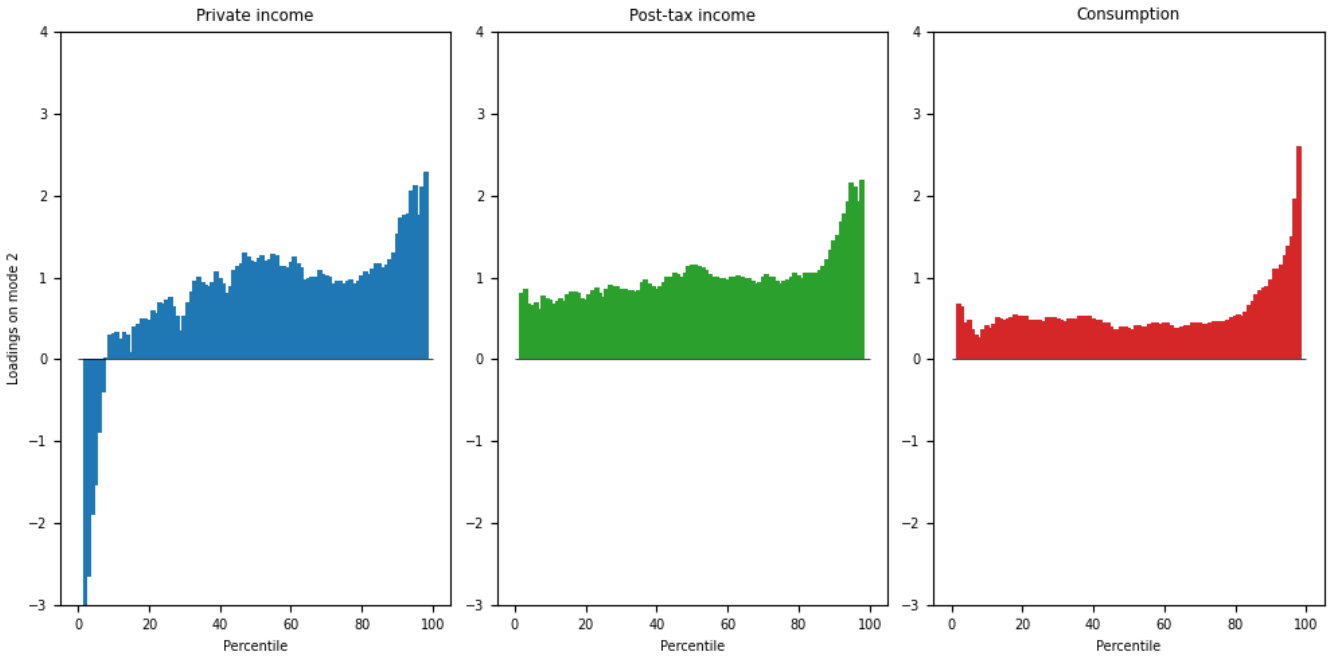

Figure 4: Loadings $\mathbf{\Phi}_{\cdot,2}$ of quantiles on mode 2

annual change in log consumption on the coincident change in log aggregate real per capital consumption using CEX data. They find that the sensitivity of the top 10% of the distribution is substantially higher than that of the lower 80%. They find a qualitatively similar relationship for income using data from Piketty and Saez (2003). Guvenan et al. (2014) and Guvenan et al. (2017) compute a similar regression on Social Security Administration data and find that exposures of earnings to aggregate variables are "U-shaped" with respect to the earnings level.

Are our results compatible with these? To find out, we substitute our DMD estimates of $\mathbf{\Lambda}, \mathbf{H},$ and $\mathbf{\Phi}$ in system (8) and simulate data of length $T = 1000$. For every period t and each variable, we calculate cross-sectional means, i.e.,

$$\bar{y}_t^{private} = \frac{1}{97}\sum_{i=3}^{98} y_{i,t}^{private}, \quad \bar{y}_t^{post-tax} = \frac{1}{97}\sum_{i=3}^{98} y_{i,t}^{post-tax}, \quad \bar{y}_t^{cons} = \frac{1}{97}\sum_{i=3}^{98} y_{i,t}^{cons}$$

We then compute time-series regressions of percentiles on the corresponding aggregate.

For private income, the regression is

$$y_{i,t}^{private} = \alpha_i^{private} + \beta_i^{private} \bar{y}_t^{private} + \epsilon_{i,t}^{private} \quad \text{for } i = 3, \ldots, 98 \tag{12}$$

Figure (5) plots β_i as a function of the percentile rank.[13] The left chart is for private income and shows that households in the low quantiles have high betas on aggregate private income growth; those in the middle of the distribution have lower betas, while betas are slightly higher for upper quantiles. This pattern is qualitatively similar to the "U-shaped" betas described by Guvenan et al. (2014).[14] The right panel shows corresponding betas for the consumption distribution. It shows that the highest quantiles of consumption are most responsive to aggregate consumption growth, in line with the findings of Parker and Vissing-Jorgensen (2009) and others. Thus, computing regressions like those in the literature on data simulated from our estimated DMD model yields similar results.

According to our DMD, the aggregate variables used in regressions (12) are actually combinations of two correlated dynamic modes. Consequently, estimated β_is are functions of DMD loadings $\mathbf{\Phi}_{[i,:]}$ on both modes. Increasing consumption loadings on inequality mode 2 drive the increasing β_i^{cons}, not loadings on the reference cycle mode 1. A similar effect, albeit smaller quantitatively appears in the left panel. Our DMD model indicates that those "U-shaped" private income betas found by Guvenan et al. (2014) are being driven by fluctuations in the inequality mode 2, not the reference cycle mode 1.

[13] Almost identical plots arise when we weight quintile growth rates by the corresponding income/consumption shares. We calculate simple means in favor of clarity.

[14] Since the Social Security Administration data used by Guvenan et al. (2014) is not top coded, the authors are able to include very high income individuals – the top 0.1% of households – in their analysis. This difference might explain why our income betas are not as distinctly "U-shaped".

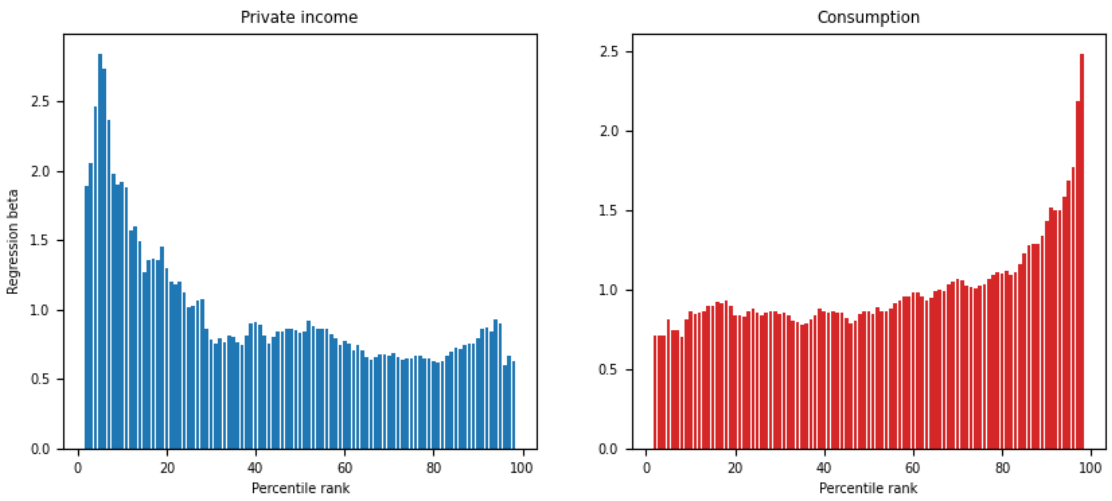

Figure 5: Aggregate regression betas

3.4 Impulse responses

To construct impulse response functions with orthogonal shocks, it is useful to calculate a lower triangular Cholesky decomposition $\mathbf{H}\,\mathbf{H}^{\top} = \mathbf{\Phi}^{+}\widehat{\mathbf{\Omega}}(\mathbf{\Phi}^{+})^{\top}$ and then transform the first equation of representation (8) to

$$\widetilde{\mathbf{x}}_t = \mathbf{\Lambda}\widetilde{\mathbf{x}}_{t-1} + \mathbf{H}\,\mathbf{e}_t \tag{13}$$

where $\mathbf{e}_t$ is a standardized random vector. Figure 6 plots impulse responses to a 1 standard deviation increase in the first orthogonalized shock. Since innovations $\widehat{\mathbf{a}}_t$ to the modes are correlated, the first shock affects both modes (top left panel). The first shock affects the lower private income quantiles much more than it does the higher quantiles. Differences in responses are much lower for after-tax income and consumption quantiles.

The impulse responses also imply differential effects to income and consumption inequality. Figure 7 plots the impulse response of the difference between the 95th

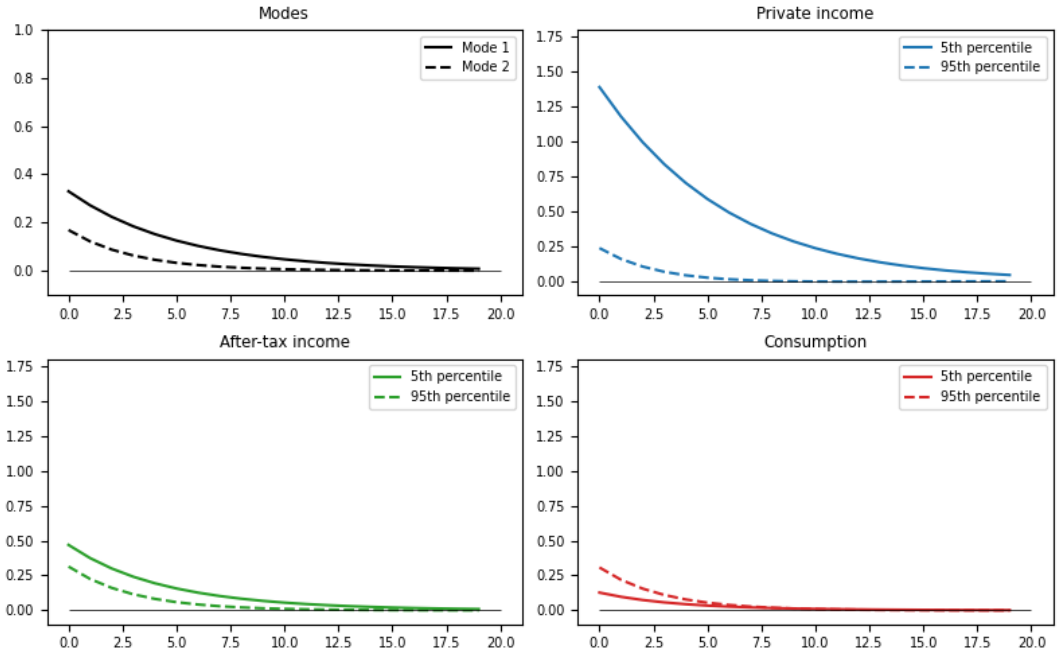

Figure 6: Impulse response to orthogonalized shock 1

percentile and the 5th percentile to a one-standard deviation **decrease** in shock 1. It shows that private income growth inequality increases substantially, and consumption inequality falls. These results are also consistent with existing literature on the business cycle dynamics of inequality. Using CPS data, Meyer and Sullivan (2013) find income inequality rose during the 2009 recession, while consumption inequality fell. In a more recent paper, Meyer and Sullivan (2023) confirm their original findings and attribute this dichotomy to asset ownership, the prices of which fell dramatically during the 2009 recession. In response to the negative shock, post-tax income inequality rises, albeit much smaller in magnitude than private income. Its reaction is consistent with government transfers being an important redistribution mechanism, as documented in Heathcote et al. (2023).[15]

Another implication of the impulse response to shock 1 is that consumption in-

[15] Heathcote et al. (2023) also infer the importance of household income pooling in the muting of consumption inequality. Since our private income variable is already at the household level, our current analysis is unable to speak to that insight.

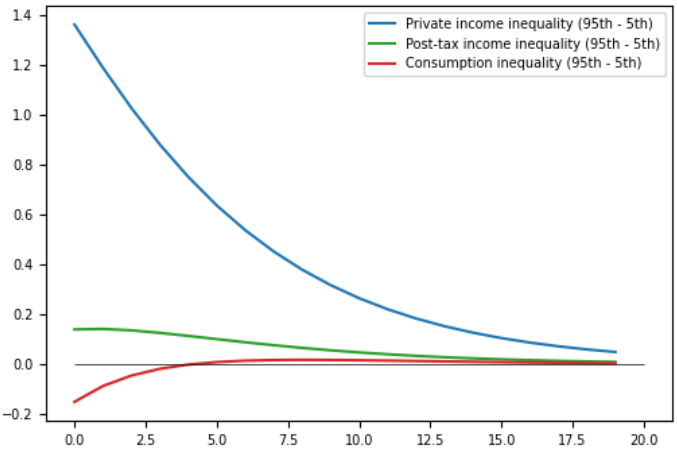

Figure 7: Impulse response of 95th - 5th percentile income and consumption to *decrease* in shock 1

equality – defined as the cross-sectional standard deviation – arising from shock 1 is "smoother" than income inequality. To quantify this, we simulate a long sample ($T = 1000$) from system (8), setting shock 2 always equal to zero. For each period, we calculated an inequality measure for each variable, defined as a cross-sectional standard deviation. The result is a time-series of our measure of inequality $\{\hat{\sigma}_t\}$ for each variable. Table 3 shows sample variances of inequality for all three variables for the simulated data; these are our estimates of the variances of our inequality measures attributable to shock 1. [16]

	Private income	Post-tax income	Consumption
Variance of inequality	0.11	0.002	0.001

Table 3: Estimated variance of inequality due to shock 1 from simulated data

[16] Explicitly, for consumption, the calculation is

$$\frac{1}{1000}\sum_{t=1}^{1000}\left(\hat{\sigma}_t^{cons} - \frac{1}{1000}\sum_{t=1}^{1000}\hat{\sigma}_t^{cons}\right)^2$$

The analagous computations are also done for private income and post-tax income.

The variance of private income inequality is higher than are the variances of post-tax income and consumption. This is consistent with findings of Heathcote et al. (2010) (see figure 13), who plot the different measures of inequality in disposable income and non-durable consumption between 1980 and 2005. These inequality measures include the variance of the log, the Gini coefficient, the P50-P10 ratio and the P90-P50 ratio. In all four cases, the time-series of consumption inequality appears, to the naked eye, less volatile than that of disposable income.

Figure 8 plots impulse responses to a 1 standard deviation increase of the second orthogonalized shock. By virtue of the lower triangular cholesky decomposition, the second shock affects only the second mode. Across our three variables, differences in responses of the 5th and 95th percentiles are largest for private income. For private income, the 5th quantile falls by 0.6ppts while the 95th quantile *increases* by a similar magnitude. It seems that shock 2 "redistributes" private income. Not so for post-tax income and consumption quantiles. Responses of all quantiles are positive, though they are larger in magnitude for higher quantiles.

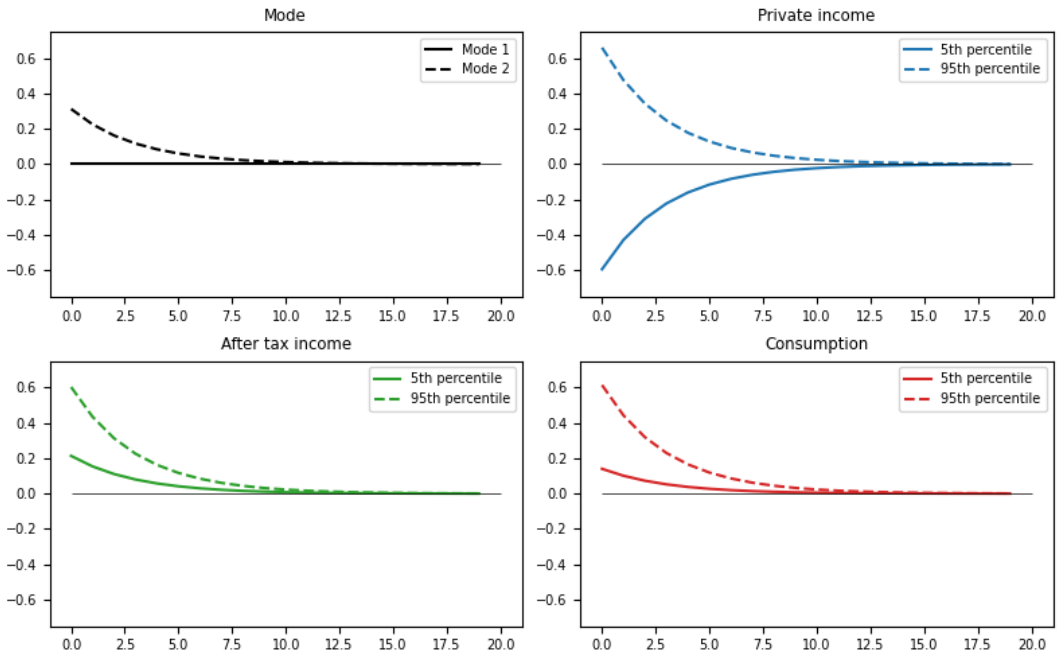

Figure 8: Impulse response to orthogonalized shock 2

4 Two Innovations Representations

In section 2, we described how to construct a reduced-rank first-order vector autoregression and then use it to cast a linear state space representation (8) in terms of dynamic modes $\widetilde{\mathbf{x}}_t$. In this section, we use this system to construct a related one that we call a **pseudo innovations representation** that is cast in terms of a different $N \times 1$ state vector $\widehat{\mathbf{x}}_t$. To accomplish this, we begin by recalling that in section 2, we implicitly defined an $N \times 1$ hidden state by a projection

$$\widetilde{\mathbf{x}}_t = \mathbb{E}\, \mathbf{x}_t \,|\, \mathbf{y}_t \,.$$

We now define a distinct projection

$$\widehat{\mathbf{x}}_t = \mathbb{E}\, \mathbf{x}_t \,|\, \mathbf{y}_{t-1} \,.$$

System (8) implies that $\widehat{\mathbf{x}}_t$ is related to $\widetilde{\mathbf{x}}_{t-1}$ by

$$\widehat{\mathbf{x}}_t = \mathbf{\Lambda}\widetilde{\mathbf{x}}_{t-1} = \mathbf{\Lambda}\mathbf{\Phi}^+ \mathbf{y}_{t-1} \,. \tag{14}$$

With equation (14) in mind, multiply both sides of the first equation of (8) by $\mathbf{\Lambda}$ and substitute $\widehat{\mathbf{x}}_t$ for $\mathbf{\Lambda}\widetilde{\mathbf{x}}_{t-1}$ in the second equation of (8) to obtain the following **pseudo innovations representation**:

$$\begin{aligned} \widehat{\mathbf{x}}_{t+1} &= \mathbf{\Lambda}\widehat{\mathbf{x}}_t + \mathbf{\Lambda}\mathbf{\Phi}^+\widehat{\mathbf{a}}_t \\ \mathbf{y}_t &= \mathbf{\Phi}\widehat{\mathbf{x}}_t + \widehat{\mathbf{a}}_t, \;\; \widehat{\mathbf{a}}_t \perp \mathbf{y}_{t-1} \,. \end{aligned} \tag{15}$$

It is enlightening to compare pseudo innovations representation (15) with an authentic time-invariant innovations representation

$$\begin{aligned} \widehat{\mathbf{x}}_{t+1} &= \mathbf{A} \;\widehat{\mathbf{x}}_t + \mathbf{K}\,\mathbf{a}_t \\ \mathbf{y}_t &= \mathbf{G} \;\widehat{\mathbf{x}}_t + \mathbf{a}_t, \;\; \mathbf{a}_t \perp \mathbf{y}^{t-1} \end{aligned} \tag{16}$$

that is associated with the linear state-space model

$$\begin{aligned} \mathbf{x}_{t+1} &= \mathbf{A}\,\mathbf{x}_t + \mathbf{C}\,\mathbf{w}_{t+1} \\ \mathbf{y}_t &= \mathbf{G}\,\mathbf{x}_t + \mathbf{v}_t, \end{aligned} \tag{17}$$

where shocks $\mathbf{w}_{t+1} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{N\times N})$, measurement errors $\mathbf{v}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_{M\times M})$ and $\mathbf{w}_s \perp \mathbf{v}_\tau$ for all s, τ; here $\mathbf{A}$ is $N \times N$, $\mathbf{C}$ is $N \times N$ and $\mathbf{G}$ is $M \times N$. Now $\widehat{\mathbf{x}}_t = \mathbb{E}[\mathbf{x}_t \,|\, \mathbf{y}^{t-1}]$, and $\mathbf{a}_t = \mathbf{y}_t - \mathbb{E}[\mathbf{y}_t \,|\, \mathbf{y}^{t-1}]$, $\mathbf{a}_t \perp \mathbf{a}_s \,\forall t \neq s$ for $\mathbf{y}^t = \{\mathbf{y}_s\}_{s<t}$ and the Hilbert space

$\mathcal{H}(\mathbf{a}^t) = \mathcal{H}(\mathbf{y}^t)$. Furthermore, $\mathbf{\Omega} \equiv \mathbb{E}[\mathbf{a}_t \, \mathbf{a}_t^\top] = \mathbf{G} \, \mathbf{\Sigma}_\infty \, \mathbf{G}^\top + \mathbf{R}$, where $\mathbf{\Sigma}_\infty$ and $\mathbf{K}$ satisfy

$$\begin{aligned} \mathbf{\Sigma}_\infty &= \mathbb{E}[\mathbf{x}_t - \widehat{\mathbf{x}}_t][\mathbf{x}_t - \widehat{\mathbf{x}}_t]^\top \\ &= \mathbf{C} \, \mathbf{C}^\top + \mathbf{K} \, \mathbf{R} \, \mathbf{K}^\top + (\mathbf{A} - \mathbf{K} \, \mathbf{G}) \mathbf{\Sigma}_\infty (\mathbf{A} - \mathbf{K} \, \mathbf{G})^\top \\ \mathbf{K} &= \mathbf{A} \, \mathbf{\Sigma}_\infty \, \mathbf{G}^\top (\mathbf{G} \, \mathbf{\Sigma}_\infty \, \mathbf{G}^\top + \mathbf{R})^{-1}. \end{aligned} \tag{18}$$

Comparing systems (15) and (16) tempts us to match objects according to

$$\mathbf{A} = \mathbf{\Lambda}, \quad \mathbf{K} = \mathbf{\Lambda} \mathbf{\Phi}^+, \quad \mathbf{G} = \mathbf{\Phi} \tag{19}$$

However, the distinct least-squares orthogonality conditions $\widehat{\mathbf{a}}_t \perp \mathbf{y}_{t-1}$ and $\mathbf{a}_t \perp \mathbf{y}^{t-1}$ are tell-tale signs that a pseudo-innnovations representation (15) is associated with a **first-order** VAR, while an authentic innovations representation (16) is associated with an **infinite-order** VAR. This means that in general it is not appropriate to expect the mapping in (19) to prevail.

However, in subsection 4.1, we describe restrictions on the linear state space model (17) that make a first-order VAR satisfy least-squares orthogonality conditions that raise its status to become an infinite-order VAR, thereby creating connections (19).[17] Under those restrictions we can infer parameters of an authentic linear state space model (17) from our reduced-dimension first-order VAR objects $\mathbf{\Lambda}, \mathbf{\Phi}$.

4.1 Restricted Linear state-space model

We impose the following restrictions on state-space system (17).

1. $M >> N$

2. $\mathbf{A}$ is a diagonal

[17] This requires that the innovations $\widehat{\mathbf{a}}_t$ be orthogonal to the linear space spanned by $\mathbf{y}^{t-1}$ and not just orthogonal to $\mathbf{y}_{t-1}$.

3. $\mathbf{G}$ has full column rank

Item 1 confines us to a "tall and skinny" $\mathbf{Y}$ settings. Item 2 asserts that each component of $\mathbf{x}_t$ follows an AR(1) process with shocks that can be correlated across components. Item 3 requires columns of $\mathbf{G}$ to be linearly independent.[18,19]

To facilitate connecting pseudo and authentic innovation representations, we recycle notation and temporarily define $\widetilde{\mathbf{x}}_t \equiv \mathbb{E}[\mathbf{x}_t \,|\, \mathbf{y}^t]$, so that now

$$\begin{aligned} \widetilde{\mathbf{x}}_t &= \widehat{\mathbf{x}}_t + \mathbf{L}\,\mathbf{a}_t \\ \mathbf{L} &= \boldsymbol{\Sigma}_\infty\,\mathbf{G}^\top(\mathbf{G}\,\boldsymbol{\Sigma}_\infty\,\mathbf{G}^\top + \mathbf{R})^{-1} \end{aligned} \tag{20}$$

and consequently

$$\begin{aligned} \widetilde{\mathbf{x}}_{t+1} &= \mathbf{A}\,\widetilde{\mathbf{x}}_t + \mathbf{L}^+\widehat{\mathbf{a}}_{t+1} \\ \mathbf{y}_{t+1} &= \boldsymbol{\Phi}\boldsymbol{\Lambda}\widetilde{\mathbf{x}}_t + \widehat{\mathbf{a}}_{t+1} \end{aligned} \tag{21}$$

4.2 Connections

Innovations representation (16) implies an infinite-order VAR for the $\{\mathbf{y}_t\}$ process:

$$\mathbf{y}_t = \sum_{j=1}^{\infty} \mathbf{B}_j^\infty\,\mathbf{y}_{t-j} + \mathbf{a}_t \tag{22}$$

$$\mathbb{E}[\mathbf{a}_t\,\mathbf{y}_{t-j}^\top] = \mathbf{0} \quad \text{for all } j \geq 1$$

$$\mathbb{E}[\mathbf{a}_t\,\mathbf{a}_t^T] = \boldsymbol{\Omega} = \mathbf{G}\,\boldsymbol{\Sigma}_\infty\,\mathbf{G}^\top + \mathbf{R} \tag{23}$$

$$\mathbf{B}_j^\infty = \mathbf{G}(\mathbf{A} - \mathbf{K}\,\mathbf{G})^{j-1}\,\mathbf{K} \quad \forall j \geq 1 \tag{24}$$

[18] Item 3 is an identifying assumption on $\mathbf{G}$ that precludes rewriting the state-space model with fewer factors. To see this, suppose $\mathbf{G}$ was rank $N-1$. Then it would be possible to rewrite the linear state-space model with $N-1$ factors, implying that only $N-1$ factors are identified.

[19] It is useful to compare our restrictions with those coming from a principle components analysis of Bai and Ng (2008) who describe a set of possible identifying restrictions on $\mathbf{G}$ and $\mathbf{x}_t$ that involve either (a) orthogonality of the hidden factors, or (b) orthogonality of the loadings $\mathbf{G}$, and/or (c) some zero restrictions on elements of $\mathbf{G}$. By contrast, item 3 does not restrict any particular elements of $\mathbf{G}$, a useful feature for our section 3 application.

where $\text{rank}(\mathbf{B}_j^\infty) = N \quad \forall j \geq 1$. Under the three restrictions we have imposed on state-space system (17), $(\mathbf{A} - \mathbf{K}\,\mathbf{G}) \approx \mathbf{0}$, so that (24) implies

$$\mathbf{B}_j^\infty \approx \begin{cases} \mathbf{G}\,\mathbf{K} & j = 1 \\ \mathbf{0} & j > 1 \end{cases} \tag{25}$$

Consequently, the infinite order VAR (22) becomes a first order VAR

$$\mathbf{y}_t = \mathbf{B}_1^\infty\,\mathbf{y}_{t-1} + \mathbf{a}_t, \tag{26}$$

which is equivalent to the reduced-rank first VAR (3) with $\mathbf{B} \approx \mathbf{B}_1^\infty$. This means that the pseudo-innovations representation *approximates* an innovations representation well. Consequently, our restrictions on state-space system (17) rationalize connections asserted in equations (19). We can proceed to infer $\mathbf{R}$ and $\mathbf{C}\,\mathbf{C}^\top$. From the above formulas for $\mathbf{K}$ and $\mathbf{\Sigma}_\infty$, it follows that

$$\mathbf{A}^\top - \mathbf{G}^\top\,\mathbf{K}^\top = \left[\mathbf{I} - \mathbf{G}^\top\,\mathbf{\Omega}^{-1}\,\mathbf{G}\,\mathbf{\Sigma}_\infty\right]\mathbf{A}^\top.$$

When $\mathbf{A} - \mathbf{K}\,\mathbf{G} = \mathbf{0}$, this formula implies that $\mathbf{I} = \mathbf{G}^\top\,\mathbf{\Omega}^{-1}\,\mathbf{G}\,\mathbf{\Sigma}_\infty$, which implies that $\mathbf{\Sigma}_\infty$ satisfies

$$\mathbf{\Sigma}_\infty = (\mathbf{G}^\top\,\mathbf{\Omega}^{-1}\,\mathbf{G})^{-1}.$$

Then from equation from (23), we infer

$$\mathbf{R} = \mathbf{\Omega} - \mathbf{G}\,\mathbf{\Sigma}_\infty\,\mathbf{G}^\top \tag{27}$$

When $\mathbf{A} - \mathbf{K}\,\mathbf{G} = \mathbf{0}$, equation (18) implies $\mathbf{\Sigma}_\infty = \mathbf{C}\,\mathbf{C}^\top + \mathbf{K}\,\mathbf{R}\,\mathbf{K}^\top$, which in turn implies

$$\mathbf{C}\,\mathbf{C}^\top = \mathbf{\Sigma}_\infty - \mathbf{K}\,\mathbf{R}\,\mathbf{K}^\top . \tag{28}$$

Next, we can use system (17) to solve the discrete Lyapunov equation

$$\mathbf{V}_\mathbf{x} = \mathbf{A}\,\mathbf{V}_\mathbf{x}\,\mathbf{A}^\top + \mathbf{C}\,\mathbf{C}^\top \tag{29}$$

for $\mathbf{V}$, the stationary covariance matrix of hidden state vector $\mathbf{x}$. Then we can compute the covariance matrix decomposition

$$\mathbf{V}_\mathbf{y} = \mathbf{G}\,\mathbf{V}_\mathbf{x}\,\mathbf{G}^\top + \mathbf{R}\,. \tag{30}$$

4.3 Algorithm

The preceding findings present us with pseudo-code for estimating $\mathbf{A}, \mathbf{C}\,\mathbf{C}^\top, \mathbf{G}, \mathbf{R}$ for our state-space model (17) that we describe in algorithm 1.

4.3.1 A computational detail

In step 9. of Algorithm 1, it is possible that $\widehat{\mathbf{\Omega}}$ is ill-conditioned, which poses a problem for matrix inversion in step 10. One option is to calculate the generalized inverse. We do this by computing the Singular Value Decomposition (SVD) of $\widehat{\mathbf{\Omega}}$. Recycling previous SVD notation,

$$\widehat{\mathbf{\Omega}} = \widetilde{\mathbf{U}}\widetilde{\mathbf{\Sigma}}\widetilde{\mathbf{V}}^\top$$

For a chosen k, we construct $\mathbf{U} = \widetilde{\mathbf{U}}[:, :\,k]$, $\mathbf{\Sigma} = \widetilde{\mathbf{\Sigma}}[:\,k, :\,k]$ and $\mathbf{V}^\top = \widetilde{\mathbf{V}}^\top[:\,k, :]$. Then, $\widehat{\mathbf{\Omega}}^{-1} = \mathbf{V}\,\mathbf{\Sigma}^{-1}\,\mathbf{U}^\top$ where $\mathbf{\Sigma}^{-1} = \text{diag}(\frac{1}{\mathbf{\Sigma}_{1,1}}, \ldots, \frac{1}{\mathbf{\Sigma}_{k,k}})$. In the CEX application, we

Algorithm 1 Pseudo-code for inferring $\mathbf{A}, \mathbf{G}, \mathbf{R}, \mathbf{C}\,\mathbf{C}^\top$

1. Set number of modes (factors) N
2. Compute $\mathbf{\Phi}$, $\mathbf{\Lambda}$ and $\widehat{\mathbf{B}}$ via DMD.
3. Approximate $\mathbf{A}$ with $\mathbf{\Lambda}$.
4. Approximate $\mathbf{G}$ with $\mathbf{\Phi}$.
5. Approximate $\mathbf{L}$ with $\mathbf{\Phi}^+$ and $\mathbf{K}$ with $\widehat{\mathbf{K}} = \mathbf{\Lambda}\mathbf{\Phi}^+$
6. Approximate $\widetilde{\mathbf{x}}_t$ with $\mathbf{\Phi}^+\,\mathbf{y}_t$.
7. Approximate $\widehat{\mathbf{x}}_t$ with $\mathbf{\Lambda}\widetilde{\mathbf{x}}_{t-1}$.
8. Approximate $\mathbb{E}[\mathbf{y}_{t+j} \,|\, \mathbf{y}_t] = \mathbf{G}\,\mathbf{A}^j\,\widetilde{\mathbf{x}}_t$ with $\mathbf{\Phi}\mathbf{\Lambda}^j\widetilde{\mathbf{x}}_t$.
9. Approximate $\mathbf{\Omega}$ with
$$\widehat{\mathbf{\Omega}} = \frac{1}{T-1}\sum_{t=1}^{T}\widehat{\mathbf{a}}_t\widehat{\mathbf{a}}_t^\top$$
where $\widehat{\mathbf{a}}_t = \mathbf{y}_t - \widehat{\mathbf{B}}\,\mathbf{y}_{t-1}$.
10. Approximate $\mathbf{\Sigma}_\infty$ with
$$\widehat{\mathbf{\Sigma}}_\infty = (\mathbf{\Phi}^\top\widehat{\mathbf{\Omega}}^{-1}\mathbf{\Phi})^{-1}$$
11. Approximate $\mathbf{R}$ with
$$\widehat{\mathbf{R}} = \widehat{\mathbf{\Omega}} - \mathbf{\Phi}\widehat{\mathbf{\Sigma}}_\infty\mathbf{\Phi}^\top$$
12. Approximate $\mathbf{C}\,\mathbf{C}^\top$ with
$$\widehat{\mathbf{C}\,\mathbf{C}^\top} = \widehat{\mathbf{\Sigma}}_\infty - \widehat{\mathbf{K}}\widehat{\mathbf{R}}\widehat{\mathbf{K}}^\top$$

choose k to satisfy

$$\frac{\sum_{i=1}^{k} \widetilde{\mathbf{\Sigma}}_{i,i}}{\sum_{i=1}^{T} \widetilde{\mathbf{\Sigma}}_{i,i}} \geq 0.975.$$

4.4 State-space objects recovered from our estimated first-order VAR

As presented in Section 4, estimates of $\mathbf{A}$ and $\mathbf{G}$ are immediate from the DMD algorithm, provided in (11) and Figures 2 and 4. We use steps 10 and 12 of Pseudo-code 1 to recover shock covariance matrix $\mathbf{C}\,\mathbf{C}^\top$ and conditional covariance $\mathbf{\Sigma}_\infty$. We use step 11 of Pseudo-code 1 to estimate $\mathbf{R}$. We obtain

$$\widehat{\mathbf{C}\,\mathbf{C}}^\top = \begin{bmatrix} 0.018 & 0.015 \\ 0.015 & 0.055 \end{bmatrix} \quad \widehat{\mathbf{\Sigma}}_\infty = \begin{bmatrix} 0.055 & 0.030 \\ 0.030 & 0.078 \end{bmatrix},$$

while Figure 9 plots the diagonal of $\widehat{\mathbf{R}}$, the variance of the measurement error. Measurement error variances are highest at the very low and very high quantiles, and relatively small for middle quantiles.

We use equations (29) and (30) to compute $\mathbf{V_y}$. We decompose unconditional variances $\text{diag}(\mathbf{V_y})$ into parts attributable to the factors ($\text{diag}(\mathbf{G}\,\mathbf{V_x}\,\mathbf{G}^\top)$) and parts attributable to the measurement error ($\text{diag}(\mathbf{R})$). We present these in Figure 10. In the middle quantiles for both income variables, the proportion of the variance explained by the factors is around 50%. This proportion falls to around 20% for both low and high quantiles. For consumption, for most quantiles only around 20% of the variance is explained by the factors.

To form a frequency-by-frequency counterpart of variance decomposition (30), we

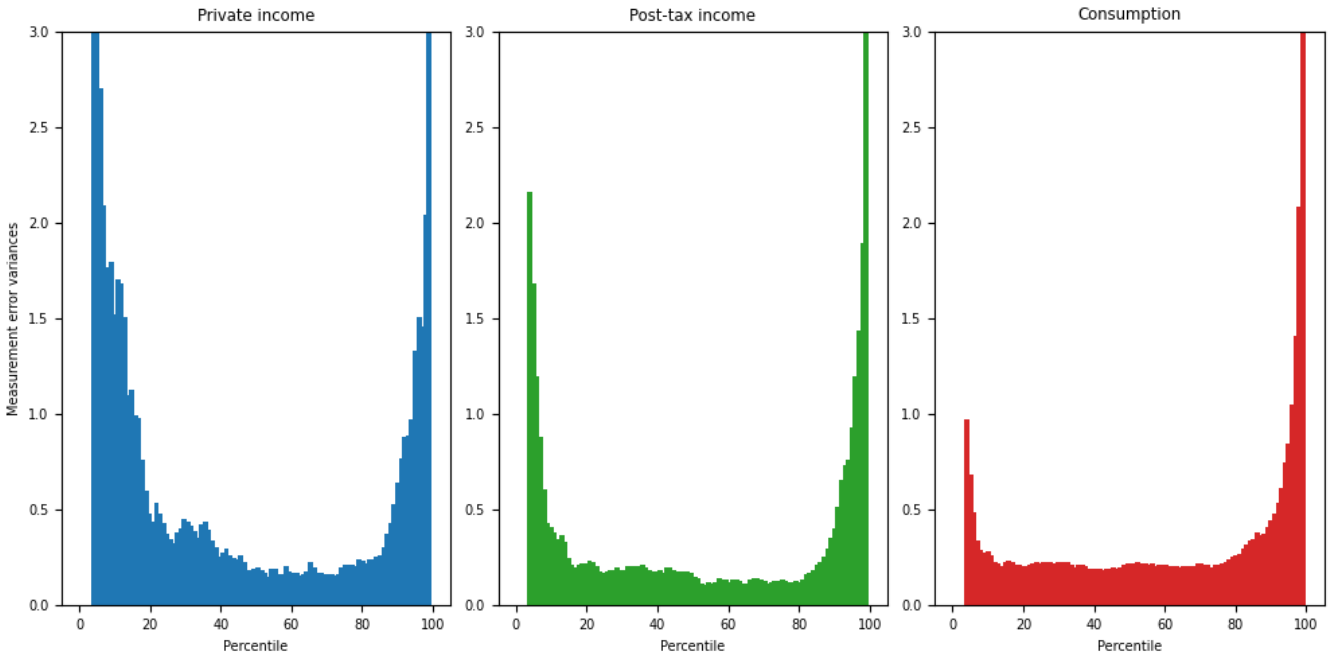

Figure 9: Measurement error variances

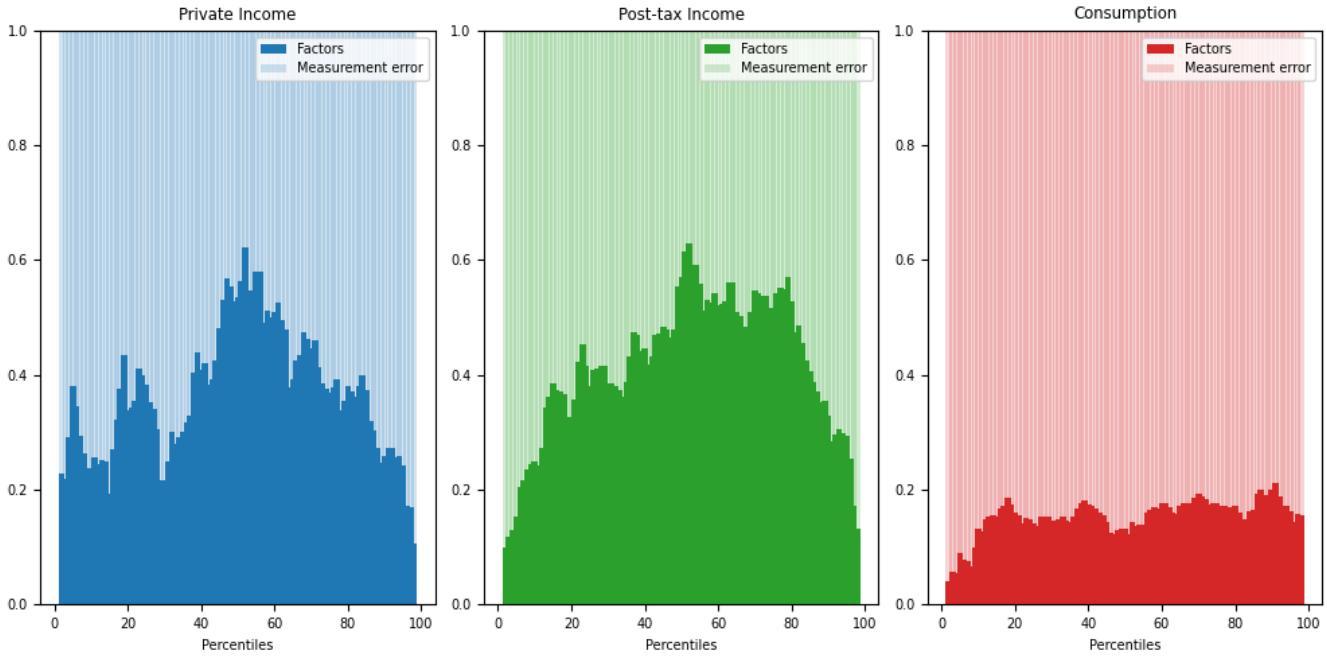

Figure 10: Decomposition of unconditional variance $\mathbf{V_y} = \mathbf{G}\,\mathbf{V_x}\,\mathbf{G}^\top + \mathbf{R}$ of $\mathbf{y}$.

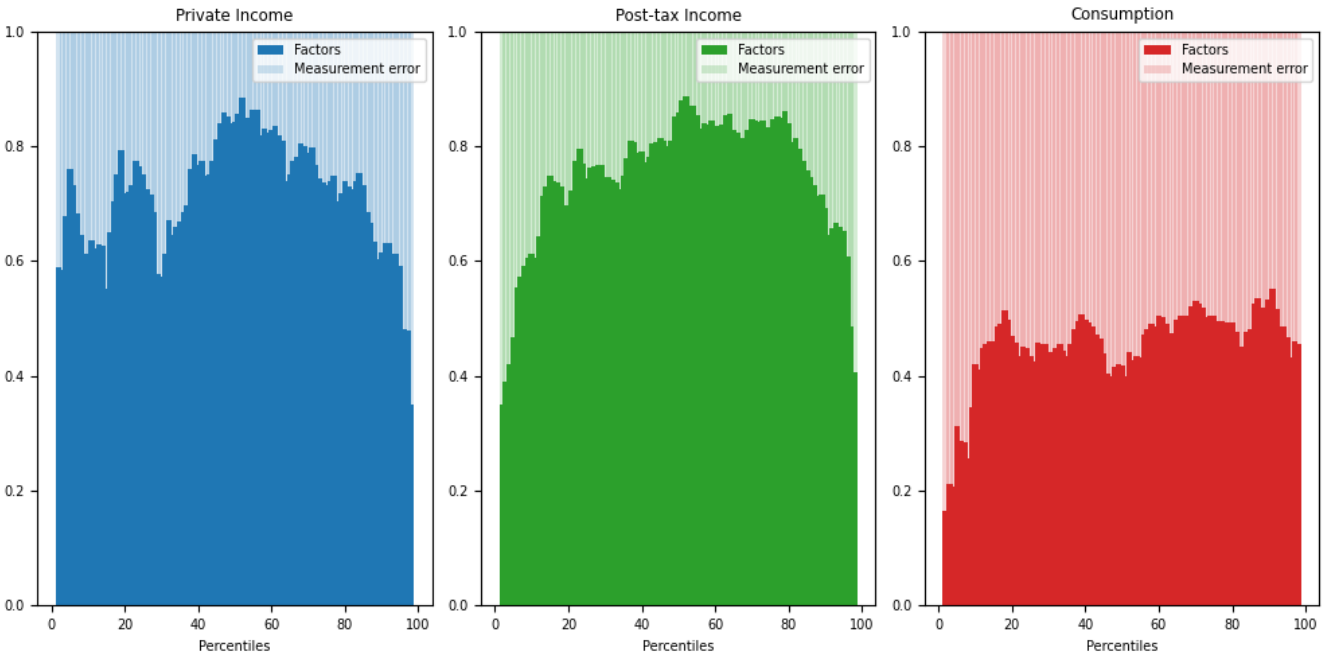

Figure 11: Spectral density decomposition of $\mathbf{y}$ at eight-year frequency ($\omega = \frac{2\pi}{32}$)

form the spectral densities of processes $\mathbf{x}$ and $\mathbf{y}$ at at frequencies $\omega \in [0, 2\pi]$:

$$S_{\mathbf{x}}(\omega) = [\mathbf{I} - \mathbf{A}\, e^{-\omega j}]^{-1}\, \mathbf{C}\, \mathbf{C}^{\top} [\mathbf{I} - \mathbf{A}^{\top}\, e^{\omega j}]^{-1}$$
$$S_{\mathbf{y}}(\omega) = \mathbf{G}\, S_{\mathbf{x}}(\omega)\, \mathbf{G}^{\top} + \mathbf{R} \tag{31}$$

Figures 11 and 12 decompose $\text{diag}(S_{\mathbf{y}}(\omega))$ into a part $\text{diag}(\mathbf{G}\, S_{\mathbf{x}}(\omega)\, \mathbf{G}^{\top})$ attributable to the factors and a part $\text{diag}(\mathbf{R})$ attributable to measurement errors at frequencies corresponding to periods of 8 years and 20 years, respectively.[20] Figures 11 and 12 indicate that the two hidden factors explain large percentages of the variances of quantiles at 20-year and 8-year frequencies.

[20]Since our data is quarterly, for a eight-year period $\omega = \frac{2\pi}{32}$ and for a twenty-year period $\omega = \frac{2\pi}{80}$

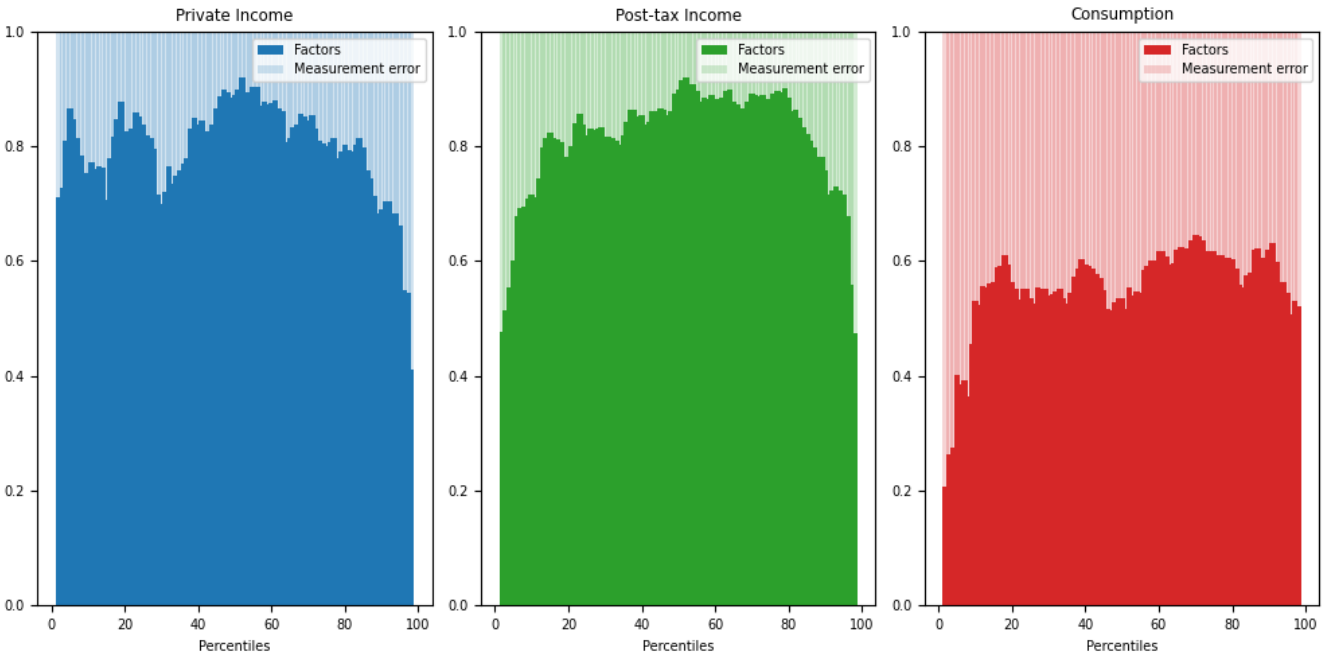

Figure 12: Spectral density decomposition of $\mathbf{y}$ at twenty-year frequency ($\omega = \frac{2\pi}{80}$)

5 Laboratory

We apply our algorithm to artificial outcomes generated by a state-space model in which $N = 2$, M is an even integer greater than 2, and

$$\mathbf{A} = \begin{bmatrix} 0.9 & 0 \\ 0 & 0.7 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 0.5 & 0.4 \\ 0 & 0.5 \end{bmatrix} \tag{32}$$

$$\widetilde{\mathbf{G}} = \begin{bmatrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ 0 & 1 \\ \vdots & \vdots \\ 0 & 1 \end{bmatrix}, \quad \mathbf{R} = 0.25\mathbf{I}_M$$

The first $\frac{M}{2}$ rows of $\widetilde{\mathbf{G}}$ are $[1, 0]$ and the second $\frac{M}{2}$ rows are $[0, 1]$, so $\widetilde{\mathbf{G}}$ has full column rank. To aid comparison for different M, we normalize columns of $\widetilde{\mathbf{G}}$ to have norm $\sqrt{M}$. Call the normalized matrix $\mathbf{G}$. Within this setting, we first study large-sample outcomes in which we know population moments. Then we report averages of statistics from repeated random samples.

5.1 Population Objects

From $\mathbf{A}, \mathbf{C}, \mathbf{G}, \mathbf{R}$ we compute population objects $\mathbf{B}, \mathbf{K}, \mathbf{\Sigma}_\infty, \mathbf{\Omega}, \mathbf{B}_1^\infty$ by first enlisting the quantecon class **LinearStateSpace** to compute population moments of the stationary distribution of the process $\{\mathbf{x}_t, \mathbf{y}_t\}$ and the associated population cross-covariagrams $\mathbb{E}[\mathbf{x}_t \mathbf{x}_t^\top] \equiv \mathbf{\Sigma}_\mathbf{x}$, $\mathbb{E}[\mathbf{y}_t \mathbf{y}_t^\top] \equiv \mathbf{\Sigma}_\mathbf{y}$, and $\mathbb{E}[\mathbf{x}_t \mathbf{y}_t^\top] \equiv \mathbf{\Sigma}_{\mathbf{x}\mathbf{y}}$. These allow us to compute population first-order autoregressive coefficients $\mathbf{B}$ via

$$\mathbf{B} = \mathbf{G}\,\mathbf{A}\,\mathbf{\Sigma}_\mathbf{x}\,\mathbf{G}^\top\,\mathbf{\Sigma}_\mathbf{y}^{-1}$$

After that, we use the quantecon class **Kalman** to compute the innovations representation and the associated $\mathbf{K}, \mathbf{\Sigma}_\infty$. This allows us to compute the population $\mathbf{B}_1^\infty$ and one-step-ahead conditional covariance matrix $\mathbf{\Omega}$ via

$$\begin{aligned}\mathbf{B}_1^\infty &= \mathbf{G}\,\mathbf{K} \\ \mathbf{\Omega} &= \mathbf{G}\,\mathbf{\Sigma}_\infty\,\mathbf{G}^\top + \mathbf{R}\,.\end{aligned}$$

Next, we use equation (27) to calculate a population analogue of $\widehat{\mathbf{R}}$ by setting $\mathbf{G} = \mathbf{\Phi}$. We then use $\widehat{\mathbf{R}}$ as an input to step **12** from Pseudo-code 1 to compute a version of $\widehat{\mathbf{C}\,\mathbf{C}}^\top$ to approximate $\mathbf{C}\,\mathbf{C}^\top$.

We performed these calculations for $M = 2, M = 300, M = 1000$ to obtain outcomes we report in Table 4. Frobenius norms reported in the first three rows describe how well a pseudo innovations representation approximates an authentic one. As anticipated,

as M gets larger, $||\mathbf{A} - \mathbf{K}\,\mathbf{G}||_F$ approaches zero, as do approximation error measures $M^{-1}||\mathbf{B} - \mathbf{B}_1^{\infty}||_F$ and $M^{-1}||\mathbf{K} - \mathbf{A}\,\mathbf{G}^{+}||_F$. In addition, $\widehat{\mathbf{C}\,\mathbf{C}^{\top}}$ approaches $\mathbf{C}\,\mathbf{C}^{\top}$ and $\widehat{\mathbf{R}}$ approaches $\mathbf{R}$.

Object	$M = 2$	$M = 300$	$M = 1000$
$\|\|\mathbf{A} - \mathbf{K}\,\mathbf{G}\|\|_F$	0.5	0.02	0.003
$M^{-1}\|\|\mathbf{B} - \mathbf{B}_1^{\infty}\|\|_F$	0.2	$3e^{-5}$	$3e^{-6}$
$M^{-1}\|\|\mathbf{K} - \mathbf{A}\,\mathbf{G}^{+}\|\|_F$	0.25	$3e^{-6}$	$2e^{-7}$
$M^{-1}\|\|\widehat{\mathbf{R}} - \mathbf{R}\|\|_F$	0.2	0.001	0.0004
$\|\|\widehat{\mathbf{C}\,\mathbf{C}^{\top}} - \mathbf{C}\,\mathbf{C}^{\top}\|\|_F$	0.5	0.004	0.001

Table 4: Population objects

These findings with population objects set the stage for an experiment in which we construct repeated samples of $M \times T$ data matrices $\mathbf{Y}$, to be described in the next subsection.

5.2 Sample Counterparts

We apply our algorithm to samples paths $\{\mathbf{y}_t\}_{t=1}^{T+1}$ generated by state-space system (32). For each sample $j = 1, \ldots J$, we create data matrices $\mathbf{Y}^{(j)}$ and $\mathbf{Y}'^{(j)}$. We apply Pseudo-code 1 to acquire the following objects for sample j: $\mathbf{\Lambda}^{(j)}, \mathbf{\Phi}^{(j)}, \widehat{\mathbf{K}}^{(j)}, \widehat{\mathbf{\Omega}}^{(j)}, \widehat{\mathbf{\Sigma}}_{\infty}^{(j)}, \widehat{\mathbf{R}}^{(j)}, \widehat{\mathbf{C}\,\mathbf{C}}^{\top(j)}$. Then we calculate element-wise means across all samples. For example, we compute

$$\widehat{E}[\mathbf{\Lambda}] \equiv \frac{1}{J}\sum_{j=1}^{J} \mathbf{\Lambda}^{(j)}$$

where $\widehat{E}[\mathbf{\Lambda}]$ is an $N \times N$ matrix of sample means. We then subtract the sample mean from its population counterpart to approximate mean errors of our sampled DMD estimators.

We set $T = 150$ and $J = 1000$. Table 5 reports Frobenius norms of mean errors.

Object	$M = 300$	$M = 1000$
$\|\|\widehat{E}[\mathbf{\Lambda}] - \mathbf{A}\|\|_F$	0.032	0.022
$M^{-1}\|\|\widehat{E}[\mathbf{\Phi}] - \mathbf{G}\|\|_F$	$4.0e^{-3}$	$2.5e^{-3}$
$M^{-1}\|\|\widehat{E}[\widehat{\mathbf{B}} - \mathbf{B}\|\|_F$	1.05^{-4}	$4.3e^{-5}$
$M^{-1}\|\|\widehat{E}[\widehat{\mathbf{K}}] - \mathbf{K}\|\|_F$	$2.5e^{-5}$	$4.0e^{-6}$
$M^{-1}\|\|\widehat{E}[\mathbf{\Phi}^+] - \mathbf{L}\|\|_F$	$4.0e^{-5}$	$5.9e^{-6}$
$M^{-1}\|\|\widehat{E}[\widehat{\mathbf{\Omega}}] - \mathbf{\Omega}\|\|_F$	5.0^{-3}	$4.0e^{-3}$
$\|\|\widehat{E}[\widehat{\mathbf{\Sigma}}_\infty] - \mathbf{\Sigma}_\infty\|\|_F$	0.25	0.20
$M^{-1}\|\|\widehat{E}[\widehat{\mathbf{R}}] - \mathbf{R}\|\|_F$	0.034	0.017
$\|\|E[\widehat{\mathbf{C}\,\mathbf{C}}^\top] - \mathbf{C}\,\mathbf{C}^\top\|\|_F$	0.20	0.19

Table 5: Sampled objects

6 Concluding remarks

We have used a Dynamic Mode Decomposition to represent salient statistical features of the cross-sectional dynamics of income and consumption in the Consumer Expenditure Survey. We detect two dominant, correlated dynamic modes: one seems to be a Burns and Mitchell (1946) reference cycle; the other seems to be an inequality factor. Loadings of cross-section quantiles on these modes reflects substantial government redistribution and provision of consumption insurance. Impulse responses of quantiles to innovations in the business cycle mode provoke different responses in earned income and consumption to a shock. It seems plausible that this "Kepler stage" finding can be reconciled with "Newton stage" structural models that government redistribution policies.

We link the DMD algorithm to a particular linear Gaussian state-space model. We describe restrictions on that state-space model that make noisy measurements of linear combinations of factors at one point in time equivalent to observing an infinite history of such noisy observations. This connection makes DMD a fast and inexpensive way to estimate state-space fundamentals for this special high-dimensional linear Gaussian factor model. Our application to CEX data illustrates that.

A Appendix

	Code Mneumonic		
	1990-2004	2004-2013	2013-2022
Private income			
Income from salary or wages	FSALARYX	FSALARYM	FSALARYM
Income from non-farm business	FNONFRMX	FNONFRMM	FSMPFRXM
Income from own farm	FFRMINCX	FFRMINCM	
Income from interest on savings accounts or bonds	INTEARNX	INTEARNM	INTRDVXM
Regular income earned from dividends, royalties, estates	FININCX	FININCXM	ROYESTXM
Income from pensions or annuities	PENSIONX	PENSIONM	RETSURVM
Net income or loss received from roomers or boarders	INCLOSSA	INCLOSAM	
Net income or loss received other rental properties	INCLOSSB	INCLOSBM	NETRENTM
Income from regular contributions from alimony and other	ALIOTHX	ALIOTHXM	
Income from care of foster children, cash scholarships	OTHRINCX	OTHRINCM	OTHRINCM
Transfer income			
Income from Social Security benefits and Railroad Benefit checks	FRRETIRX	FRRETIRM	FRRETIRM
Supplemental Security Income from all sources	FSSIX	FSSIXM	FSSIXM
Income from unemployment compensation	UNEMPLX	UNEMPLXM	
Income from workmen's compensation and veteran's payments	COMPENSX	COMPENSM	OTHREGXM
Income from public assistance including job training	WELFAREX	WELFAREM	WELFAREM
Income from other child support	CHDOTHX	CHDOTHXM	
Food stamps	JFDSTMPA		
Food stamps and electronic benefits	FOODSMPX	FOODSMPM	JFSAMTM

Table 6: Categorizing CEX income into private and transfers

References

ANDERSON, T. W. (1951): "Estimating Linear Restrictions on Regression Coefficients for Multivariate Normal Distributions," *The Annals of Mathematical Statistics*, 22, 327–351.

——— (1999): "Asymptotic distribution of the Reduced Rank Regression Estimator Under General Conditions," *The Annals of Statistics*, 27, 1141–1154.

BAI, J. AND S. NG (2008): "Large Dimensional Factor Analysis," *foundations and Trends in Econometrics*, 3, 89–163.

BLINDER, A. S. (2022): *A monetary and fiscal history of the United States, 1961–2021*, Princeton University Press.

BRUNTON, S. L. AND J. N. KUTZ (2022): *Data-Driven Science and Engineering: Machine Learnings, Dynamical Systems, and Control, second edition*, Cambridge University Press.

BURNS, A. F. AND W. C. MITCHELL (1946): *Measuring business cycles*, burn46-1, National bureau of economic research.

CHAMBERLAIN, G. AND M. ROTHSCHILD (1982): "Arbitrage, factor structure, and mean-variance analysis on large asset markets," National Bureau of Economic Research Cambridge, Mass., USA.

CHANG, W.-J., S. J. MONAHAN, A. OUAZAD, AND F. VASVARI (2021): "The higher moments of future earnings," *Accounting Review*, 96, 91–116.

FORNI, M., M. HALLIN, M. LIPPI, AND L. REICHLIN (2000): "The Generalized Dynamic-Factor Model: Identification and Estimation," *The Review of Economics and Statistics*, 82, 540–554.

FORNI, M. AND M. LIPPI (2001): "The Generalized Dynamic-Factor Model: Representation Theory," *Econometric Theory*, 17, 1113–1141.

GUVENAN, F., S. OZKAN, AND J. SONG (2014): "The Nature of Countercyclical Income Risk," *Journal of Political Economy*, 122, 621–660.

GUVENAN, F., S. SCHULHOFER-WOHL, J. SONG, AND M. YOGO (2017): "Worker Betas: Five Facts about Systematic Earnings Risk," *American Economic Review: Papers Proceedings*, 107, 398–403.

HEATHCOTE, J., F. PERRI, AND G. L. VIOLANTE (2010): "Unequal we stand: An empirical analysis of economic inequality in the United States, 1967–2006," *Review of Economic Dynamics*, 13, 15–51.

HEATHCOTE, J., F. PERRI, G. L. VIOLANTE, AND L. ZHANG (2023): "More Unequal We Stand? Inequality Dynamics in the United States, 1967–2021," *Review of Economic Dynamics*, 50, 235–266.

HELLER, W. W. (1966): *New dimensions of political economy*, Harvard University Press.

HOOD, W. C. AND T. C. KOOPMANS (1953): *Studies in econometric method*, Wiley.

KOOPMANS, T. C. (1947): "Measurement without theory," *The Review of Economics and Statistics*, 29, 161–172.

——— (1950): *Statistical inference in dynamic economic models*, New York: Wiley.

LJUNGQVIST, L. AND T. J. SARGENT (2018): *Recursive Macroeconomic Theory*, MIT Press, 4 ed.

MARSCHAK, J. (1953): "Economic Measurment for Policy and Prediction," in *Studies in econometric method*, ed. by W. C. Hood and T. C. Koopmans, Wiley, 1–26.

MEYER, B. D. AND J. X. SULLIVAN (2013): "Consumption and Income Inequality and the Great Recession," *American Economic Review: Papers Proceedings*, 103, 178–183.

——— (2023): "Consumption and Income Inequality in the United States since the 1960s," *Journal of Political Economy*, 131, 247–284.

PARKER, J. A. AND A. VISSING-JORGENSEN (2009): "Who Bears Aggregate Fluctuations and How?" *American Economic Review: Papers Proceedings*, 99, 399–405.

PIKETTY, T. AND E. SAEZ (2003): "Income Inequality in the United States, 1913-1998," *Quarterly Journal of Economics*, 118, 1–39.

SARGENT, T. J. (2015): "Robert E. Lucas Jr.'s Collected Papers on Monetary Theory," *Journal of Economic Literature*, 53, 43–64.

——— (2023): "HAOK and HANK Models," Keynote address for 2022 Bank of Chile conference.

SARGENT, T. J. AND C. A. SIMS (1977): "Business cycle modeling without pretending to have too much a priori economic theory," in *New methods in business cycle research*, Minneapolis, Minnesota: Federal Reserve Bank of Minneapolis, 145–168.

TOBIN, J. (2015): *The new economics one decade older*, Princeton University Press.

TU, J. H., C. W. ROWLEY, AND D. M. LUCHTENBURG (2014): "On Dynamic Mode Decomposition: Theory and Application," *Journal of Computational Dynamics*, 1, 391–421.