Spectral decomposition of the information about latent variables in dynamic macroeconomic models

Nikolay Iskrev

Abstract

In this paper, I show how to perform spectral decomposition of the information about latent variables in dynamic economic models. A model describes the joint probability distribution of a set of observed and latent variables. The amount of information transferred from the former to the latter is measured by the reduction of uncertainty in the posterior compared to the prior distribution of any given latent variable. Casting the analysis in the frequency domain allows decomposing the total amount of information in terms of frequency-specific contributions as well as in terms of information contributed by individual observed variables. I illustrate the usefulness of the proposed methodology with applications to two DSGE models taken from the literature.

Keywords: DSGE models, Frequency domain, Information content

JEL classification: C32, C51, C52, E32

Bank of Portugal, Av. Almirante Reis 71, 1150-012, Lisboa, Portugal, +351213130006
nikolay.iskrev@bportugal.pt

The views expressed in this paper are those of its author and do not necessarily reflect the views of Banco de Portugal or the Eurosystem.

1 Introduction

A pervasive challenge in empirical macroeconomic research is the estimation of latent variables by combining theoretical models and observational information. Examples abound and include endogenously determined variables, such as potential output and natural rates of interest or unemployment, as well as a plethora of exogenous random shocks driving business cycle fluctuations in modern macroeconomic models. These variables are typically not directly observable and to measure them requires estimating models that explicitly describe the joint dynamics of observed and latent variables. Having correctly specified and accurately measured latent endogenous variables and structural shocks is a key requirement for macroeconomic models to meet to be useful as tools for policy analysis and to be credible as story-telling devices.

The purpose of this paper is to show how to perform spectral decomposition of the information about latent variables in dynamic economic models. In particular, the proposed analysis reveals where in the frequency spectrum information about latent variables predominantly comes from, and how much of it is contributed by individual observed variables. The goal of this analysis is to enhance researchers' understanding of where in the data, according to a given model, information about unobservable quantities comes from. In doing so, the paper contributes to the recent literature aimed at improving the transparency of structural estimation in macroeconomic research.

The work that is most closely related to this paper is Iskrev (2019), where the question regarding the sources of information about latent variables is treated in the time domain. In that paper, the amount of information from observable variables about latent variables is quantified by comparing prior and posterior probability distributions and employing information-theoretic measures of uncertainty and information gain. Analysis in the time domain preserves information about the temporal order of the observable data in relation to the latent variables and allows to study the transfer of information between variables with arbitrary temporal patterns. The information pertaining to the temporal order of the variables is lost completely in the frequency domain. At the same time, it allows to decompose uncertainty and information into components of varying frequencies. Therefore, it reveals how much and where in the spectrum uncertainty is resolved for a given latent variable and what are the contributions from different observables. This is something that cannot be deduced in the time domain. The two approaches are therefore complementary.

The paper is also related to a growing literature on the feasibility of recovering structural shocks using reduced form models. Building upon the work of Hansen and Sargent (1980, 1991) and Lippi and Reichlin (1993, 1994), most of the research on this topic has focused on the issue of invertibility (or fundamentalness) in structural vector autoregressions, i.e. whether shocks from general equilibrium models can be recovered from the residuals of VARs (see Alessi et al. (2011) and Giacomini (2013) for useful overviews of this literature). Conditions for invertibility are discussed in Fernandez-Villaverde et al. (2007), Ravenna (2007), Franchi and Vidotto (2013), Franchi and Paruolo (2015)), while Giannone and Reichlin (2006) and Forni and Gambetti (2014) discuss how to test for lack of invertibility of structural VARs. Invertibility issues that are specific to DSGE models with news shocks are discussed in Leeper et al. (2013) and Blanchard et al. (2013). More recently, Soccorsi (2016) and Sala et al. (2016) proposed measures of the degree of non-invertibility, which quantify the discrepancies between true shocks and shocks obtained using non-fundamental VARs.[1] In another recent paper Chahrour and Jurado (2021) draw a distinction between invertibility on one hand, and what they call "recoverability" on the other, defining the latter as the feasibility of recovering structural shocks from all leads and lags of the observables variables. They argue that recoverability is often what matters in applied research and present a necessary and sufficient condition one can use to check if shocks in linear models are recoverable.

Similar to that literature, the analysis in the present paper can be used to determine whether the shocks in a given model are recoverable given a set of observed variables. Furthermore, as in Soccorsi (2016) and particularly Sala et al. (2016), a measure is provided of the degree to which any individual shock, or an endogenous latent variable, can be recovered. In particular, the proposed spectral measures of information gain are defined with respect to a particular unobserved variable and show how much of the prior uncertainty about it, within a given frequency band, is removed due to observing a given set of model variables.

While the existing research on invertibility is concerned with the usefulness of VAR–based tools for empirical validation of structural models, the purpose of the analysis presented here is to understand the properties of structural macroeconomic models in terms of how much and from where in the spectrum information transfers between observed and unobserved model variables. Therefore, identifying the principal sources of

[1]Simulation evidence that non-invertible VARs may in some cases produce good approximations of the true structural shocks are provided in Sims (2012) and Beaudry et al. (2015).

information is of primary interest rather than the total amount of information about a given shock or endogenous latent variable. To that end, I define and apply measures of frequency band-specific conditional information gains that quantify the amount of additional information contributed by a subset of variables, given the information contained in the remaining observed variables, at a given band of frequencies. As the analysis of the models considered in the application section shows, the conclusions one draws may be very different depending on what the conditional variables are.

The remainder of the paper is organized as follows. Section 2 reviews the relevant information-theoretic and frequency domain concepts and defines measures of information gains from observable with respect to latent variables. It also shows how the measures can be evaluated for linear Gaussian DSGE models. Section 3 illustrates the proposed analysis in two applications. One is a small-scale New-Keynesian model employed by Uribe (2021) to investigate the nature and empirical importance of monetary policy shocks that produce neo-Fisherian dynamics, i.e. move interest rates and inflation in the same direction over the short run. The second is a medium-scale New Keynesian model estimated by Justiniano et al. (2011) in order to investigate whether investment shocks are important drivers of business cycle fluctuations. For both models, a complete investigation of the sources of information about all structural shocks is presented. Section 4 concludes.

2 Methodology

The purpose of this section is three-fold. First, to introduce some basic information-theoretic concepts and use them to define a measure of information gain for variables with a multivariate complex Gaussian distribution. Second, to review relevant properties of the spectral representation of a stationary Gaussian vector process and present frequency domain measures of information gain. Third, to show how to apply the measures in the context of DSGE models to evaluate the information contributions with respect to latent variables across observed variables and frequencies.

2.1 Quantifying information gains

Consider a $(n_{\boldsymbol{y}} + 1)$-dimensional random vector $\boldsymbol{z} = [\boldsymbol{y}', x]'$ whose joint probability density function is $f(\boldsymbol{y}, x)$. How much information about x is gained when a realization of $\boldsymbol{y}$ is observed? Information theory provides the framework and tools to answer such questions. Specifically, entropy is a measure of the uncertainty associated with a random variable, and mutual information is a measure of the information shared by two random variables. Formally, if $f(x)$ is the marginal probability density function of x, and $\boldsymbol{S}_x$ is the support of x, the entropy $\mathrm{H}(x)$ of $f(x)$ is defined as

$$\mathrm{H}(x) = -\int_{\boldsymbol{S}_x} f(x) \ln\left(f(x)\right) dx = -\mathrm{E}\ln f(x). \tag{2.1}$$

The amount of information about x is measured as the reduction in uncertainty, i.e. the entropy $\mathrm{H}(x)$, relative to some base distribution. The mutual information of the random variables $\boldsymbol{y}$ and x is defined as

$$\mathrm{I}(\boldsymbol{y}, x) = \int_{\boldsymbol{S}_{\boldsymbol{y}}} \int_{\boldsymbol{S}_x} f(\boldsymbol{y}, x) \ln \frac{f(\boldsymbol{y}, x)}{f(\boldsymbol{y}) f(x)} d\boldsymbol{y} dx \tag{2.2}$$

where $\boldsymbol{S}_{\boldsymbol{y}}$ is the support of $\boldsymbol{y}$. The information interpretation of (2.2) follows from the fact that it can be expressed in terms of entropy as

$$\mathrm{I}(\boldsymbol{y}, x) = \mathrm{H}(x) - \mathrm{H}(x|\boldsymbol{y}). \tag{2.3}$$

where $\mathrm{H}(x|\boldsymbol{y}) = -\mathrm{E}\ln f(x|\boldsymbol{y})$ is the entropy of the conditional probability density function of x given $\boldsymbol{y}$. Therefore, $\mathrm{I}(\boldsymbol{y}, x)$ has an intuitive interpretation as the reduction of the uncertainty about x due to observing $\boldsymbol{y}$. It can be shown (see Granger and Lin (1994)) that $\mathrm{H}(x) \geq \mathrm{H}(x|\boldsymbol{y})$ with equality if and only if $f(\boldsymbol{y}, x) = f(\boldsymbol{y}) f(x)$. Hence, unless $\boldsymbol{y}$ and x are independent, observing $\boldsymbol{y}$ provides information about x. If we partition $\boldsymbol{y}$ into two sub-vectors $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$, we can express the conditional mutual information of x and $\boldsymbol{y}_1$ given $\boldsymbol{y}_2$ as

$$\mathrm{I}(\boldsymbol{y}_1, x|\boldsymbol{y}_2) = \mathrm{H}(x|\boldsymbol{y}_2) - \mathrm{H}(x|\boldsymbol{y}_1, \boldsymbol{y}_2) \tag{2.4}$$

The conditional mutual information tells us how much of the uncertainty about x that remains after $\boldsymbol{y}_2$ is observed is removed by observing also $\boldsymbol{y}_1$. Now, let the joint density function $f(\boldsymbol{y}, x)$ be the $(n_{\boldsymbol{y}} + 1)$-dimensional complex Gaussian distribution,

$$\mathcal{N}_{\mathcal{C}}\left(\begin{pmatrix} \mathbf{0} \\ 0 \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_{\boldsymbol{yy}} & \boldsymbol{\Sigma}_{\boldsymbol{y}x} \\ \boldsymbol{\Sigma}_{x\boldsymbol{y}} & \Sigma_{xx} \end{pmatrix}\right) \tag{2.5}$$

Then, both the marginal distribution of x and the conditional distribution of x given $\boldsymbol{y}$ are univariate complex Gaussian distributions, with covariances given by Σ_{xx}, and $\boldsymbol{\Sigma}_{x|\boldsymbol{y}} = \Sigma_{xx} - \boldsymbol{\Sigma}_{x\boldsymbol{y}}\boldsymbol{\Sigma}_{\boldsymbol{yy}}^{-1}\boldsymbol{\Sigma}_{\boldsymbol{y}x}$, respectively. Using this, it is straightforward to show that the mutual information of $\boldsymbol{y}$ and x is

$$\mathrm{I}(\boldsymbol{y}, x) = \mathrm{H}(x) - \mathrm{H}(x|\boldsymbol{y}) = .5 \ln\left(\frac{\Sigma_{xx}}{\boldsymbol{\Sigma}_{x|\boldsymbol{y}}}\right) \tag{2.6}$$

From $\mathrm{H}(x) \geq \mathrm{H}(x|\boldsymbol{y})$ it follows that mutual information is positive unless $\boldsymbol{y}$ and x are independent in which case it is zero. On the other hand, if the variables are perfectly dependent i.e. there exists a one-to-one function g such that $x = g(\boldsymbol{y})$, observing $\boldsymbol{y}$ is equivalent to observing x. In that case $\boldsymbol{\Sigma}_{x|\boldsymbol{y}} = 0$ and $\mathrm{I}(\boldsymbol{y}, x) = \infty$. It is common in practice to normalize the measure to be in the interval $[0,\ 1]$. This can be achieved using the following monotonous increasing transformation (see e.g. Joe (1989) or Granger and Lin (1994))

$$\mathrm{I}^*(\boldsymbol{y}, \boldsymbol{z}) = 1 - \exp\left(-2\mathrm{I}(\boldsymbol{y}, \boldsymbol{z})\right) \tag{2.7}$$

Applying this transformation to (2.6) results in the following measure of information gain:

$$\mathrm{IG}_{\boldsymbol{y}\to x} = \left(\frac{\Sigma_{xx} - \boldsymbol{\Sigma}_{x|\boldsymbol{y}}}{\Sigma_{xx}}\right) \times 100, \tag{2.8}$$

The interpretation of $\mathrm{IG}_{\boldsymbol{y}\to x}$ is the following: it measures the reduction in uncertainty about x due to observing vector $\boldsymbol{y}$, as a percent of the unconditional (prior) uncertainty about x. Similarly, when $\boldsymbol{y}$ is partitioned into $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$, we can define the conditional information gain of $\boldsymbol{y}_1$ with respect to x, given $\boldsymbol{y}_2$ as

$$\mathrm{IG}_{\boldsymbol{y}_1\to x|\boldsymbol{y}_2} = \left(\frac{\boldsymbol{\Sigma}_{x|\boldsymbol{y}_2} - \boldsymbol{\Sigma}_{x|\boldsymbol{y}}}{\Sigma_{xx}}\right) \times 100, \tag{2.9}$$

The interpretation of $\mathrm{IG}_{\boldsymbol{y}_1\to x|\boldsymbol{y}_2}$ is the following: it shows the amount of uncertainty about x left after observing $\boldsymbol{y}_2$ that is removed by observing also $\boldsymbol{y}_1$, as a percent of the unconditional uncertainty about x.

2.2 Information gains in the frequency domain

Let $\boldsymbol{z}_t \in \mathbb{R}^{n_{\boldsymbol{z}}}$ for $t \in \mathbb{Z}$ be $n_{\boldsymbol{z}}-$dimensional stationary Gaussian time series with

$$\mathrm{E}\,\boldsymbol{z}_t = \mathbf{0} \qquad t \in \mathbb{Z} \tag{2.10}$$

$$\operatorname{cov}(\boldsymbol{z}_t, \boldsymbol{z}_{t-h}) = \boldsymbol{\Gamma}(h) \qquad t, h \in \mathbb{Z} \tag{2.11}$$

If $\boldsymbol{Z} = [\boldsymbol{z}_1', \boldsymbol{z}_2', \ldots, \boldsymbol{z}_T']'$ is a $T \times n_{\boldsymbol{z}}-$dimensional realization the process, the joint distribution of $\boldsymbol{Z}$, as well as the marginal and conditional distributions of any subset of components of $\boldsymbol{Z}$, will be Gaussian. Therefore, in the time domain, the information gain measures from the previous section can be applied directly to quantify the information gained with respect to any realization of a component of $\boldsymbol{z}$ due to observing a sample of realizations of a subset of the remaining components of the process (see Iskrev (2019)).

In the frequency domain, the information gains analysis proceeds by applying the discrete Fourier transform to the values of $\boldsymbol{Z}$:

$$Z(\omega_j) = (2\pi T)^{-1/2} \sum_{t=1}^{T} \boldsymbol{z}_t e^{-it\omega_j} \tag{2.12}$$

for the Fourier frequencies $\omega_j = 2\pi j/T$, where $j \in \{j \in \mathbb{Z} : -\pi < 2\pi j/T \leq \pi\}$.

Due to the linearity of the discrete Fourier transform, the joint Gaussianity is preserved. Furthermore, it can be shown that $Z(\omega_j)$ behave asymptotically as independent complex Gaussian random variables with zero mean and a covariance matrix equal to $f(\omega_j)$, where $f_{\boldsymbol{z}\boldsymbol{z}}(\omega) \in \mathbb{C}^{n_{\boldsymbol{z}} \times n_{\boldsymbol{z}}}$ is the spectral density matrix of $\boldsymbol{z}(t)$ at frequency ω (see Brillinger (1981, Theorem 4.4.1)),

$$f_{\boldsymbol{z}\boldsymbol{z}}(\omega) = (2\pi)^{-1} \sum_{h=-\infty}^{\infty} \Gamma(h) e^{-ih\omega} \tag{2.13}$$

The asymptotic independence of the Fourier coefficients $Z(\omega_j)$ across frequencies implies that information gain analysis may be conducted on a frequency-by-frequency basis. In particular, (asymptotically) there is no information about a given component of the series at a frequency ω_j that comes from components at any other frequency $\omega_l,\ l \neq j$. Furthermore, the complex Gaussianity of the distribution implies that information analysis at a given frequency ω can be performed using the information gain measures from Section 2.1. To be more concrete, consider a partition of $\boldsymbol{z}_t$ into a $n_{\boldsymbol{y}}-$dimensional vector $\boldsymbol{y}_t$ and a scalar x_t, and let $\boldsymbol{y}(\omega)$ and $x(\omega)$ be their respective discrete Fourier transforms at a frequency $\omega \in (-\pi, \pi]$. The spectral density matrix of $[\boldsymbol{y}_t', x_t]'$ is given by

$$f_{\boldsymbol{zz}}(\omega) = \begin{bmatrix} f_{\boldsymbol{yy}}(\omega) & f_{\boldsymbol{y}x}(\omega) \\ f_{x\boldsymbol{y}}(\omega) & f_{xx}(\omega) \end{bmatrix} \tag{2.14}$$

and the frequency-specific information gain of $\boldsymbol{y}(\omega)$ with respect to $x(\omega)$ is

$$\mathrm{IG}_{\boldsymbol{y}\to x}(\omega) = \left(\frac{f_{xx}(\omega) - f_{x|\boldsymbol{y}}(\omega)}{f_{xx}(\omega)}\right) \times 100 \tag{2.15}$$

where $f_{x|\boldsymbol{y}}(\omega) = f_{xx}(\omega) - f_{x\boldsymbol{y}}(\omega) f_{\boldsymbol{yy}}^{-1}(\omega) f_{\boldsymbol{y}x}(\omega)$ is the partial spectrum of x given $\boldsymbol{y}$ (Priestley (1981)). Furthermore, if we split $\boldsymbol{y}_t$ into $\boldsymbol{y}_{1t}$ and $\boldsymbol{y}_{2t}$ and let $\boldsymbol{y}_1(\omega)$ and $\boldsymbol{y}_2(\omega)$ be their respective discrete Fourier transforms, the frequency-specific conditional information gain of $\boldsymbol{y}_1(\omega)$ with respect to $x(\omega)$ given $\boldsymbol{y}_2(\omega)$ is

$$\mathrm{IG}_{\boldsymbol{y}_1\to x|\boldsymbol{y}_2}(\omega) = \left(\frac{f_{x|\boldsymbol{y}_2}(\omega) - f_{x|\boldsymbol{y}}(\omega)}{f_{xx}(\omega)}\right) \times 100 \tag{2.16}$$

The interpretation of $\mathrm{IG}_{\boldsymbol{y}\to x}(\omega)$ and $\mathrm{IG}_{\boldsymbol{y}_1\to x|\boldsymbol{y}_2}$ is the same as before (see Section 2.1), except that now information is defined in terms of the reduction of uncertainty about x at a given frequency ω due to information from $\boldsymbol{y}$ (or conditionally, from $\boldsymbol{y}_1$), also at frequency ω. In practice, we are usually interested not in a single frequency but rather in a band of frequencies, such as low, business cycle, or high frequencies. Frequency band-specific measure of information gain may be obtained by replacing in (2.15) and (2.16) the frequency-specific spectrum and conditional spectrum of x with their integrated versions,

$$\mathrm{IG}_{\boldsymbol{y}\to x}(\boldsymbol{\omega}) = \left(\frac{f_{xx}(\boldsymbol{\omega}) - f_{x|\boldsymbol{y}}(\boldsymbol{\omega})}{f_{xx}(\boldsymbol{\omega})}\right) \times 100 \tag{2.17}$$

$$\mathrm{IG}_{\boldsymbol{y}_1\to x|\boldsymbol{y}_2}(\boldsymbol{\omega}) = \left(\frac{f_{x|\boldsymbol{y}_2}(\boldsymbol{\omega}) - f_{x|\boldsymbol{y}}(\boldsymbol{\omega})}{f_{xx}(\boldsymbol{\omega})}\right) \times 100 \tag{2.18}$$

where $\boldsymbol{\omega} = \{\omega : \omega \in [\underline{\omega}, \overline{\omega}] \cup [-\overline{\omega}, -\underline{\omega}]\}$ denotes the frequency band of interest, $f_{xx}(\boldsymbol{\omega}) = \int_{\omega\in\boldsymbol{\omega}} f_{xx}(\omega)\mathrm{d}\omega$, and $f_{x|\boldsymbol{y}}(\boldsymbol{\omega}) = \int_{\omega\in\boldsymbol{\omega}} f_{x|\boldsymbol{y}}(\omega)\mathrm{d}\omega$. The interpretation remains the same, except that now the uncertainty and information about x are with respect to the frequency band $\boldsymbol{\omega}$. Note that $\mathrm{IG}_{\boldsymbol{y}\to x}(\boldsymbol{\omega})$ can be written also as

$$\mathrm{IG}_{\boldsymbol{y}\to x}(\boldsymbol{\omega}) = \int_{\omega\in\boldsymbol{\omega}} \mathrm{IG}_{\boldsymbol{y}\to x}(\omega)\frac{f_{xx}(\omega)}{f_{xx}(\boldsymbol{\omega})}\mathrm{d}\omega \tag{2.19}$$

Therefore, the information gain for a selected band of frequencies $\boldsymbol{\omega}$ is given simply by the weighted sum of the frequency-specific information gains, with weights equal to the contribution of each frequency to the total variance of x in $\boldsymbol{\omega}$. Similarly, the conditional information gain (2.18) can be expressed as a weighted sum of the frequency-specific conditional information gains.

A special case of the band-specific information gain is when $\boldsymbol{\omega}$ covers the full spectrum, i.e. when $\underline{\omega} = 0$ and $\overline{\omega} = \pi$. Let $\overline{\boldsymbol{\omega}} = \{\omega : \omega \in [0, \pi]\cup(0, -\pi]\}$. Then, it is straightforward to show that information gain takes the form:

$$\mathrm{IG}_{\boldsymbol{y}\to x}(\overline{\boldsymbol{\omega}}) = \left(\frac{\mathrm{var}(x_t) - \mathrm{var}(x_t|\boldsymbol{y}_{t-\tau}, \tau \in \mathbb{Z})}{\mathrm{var}(x_t)}\right) \times 100 \tag{2.20}$$

Therefore, in addition to the obvious frequency domain interpretation, it has a time-domain interpretation, namely the per cent reduction of the unconditional variance of x_t as a result of observing the infinite sequence of past, present, and future values of $\boldsymbol{y}_t$. Similarly, the full spectrum version of the conditional information gain measure of $\boldsymbol{y}_1$ with respect to x given $\boldsymbol{y}_2$ is

$$\mathrm{IG}_{\boldsymbol{y}_1\to x|\boldsymbol{y}_2}(\overline{\boldsymbol{\omega}}) = \left(\frac{\mathrm{var}\,(x_t|\boldsymbol{y}_{2t-\tau}, \tau \in \mathbb{Z}) - \mathrm{var}(x_t|\boldsymbol{y}_{t-\tau}, \tau \in \mathbb{Z})}{\mathrm{var}(x_t)}\right) \times 100 \tag{2.21}$$

The interpretation of (2.21) is the following: it shows the amount of uncertainty about x_t left after observing the infinite sequence of past, present, and future values of $\boldsymbol{y}_2$ that is removed by observing also the infinite sequence of past, present, and future values of $\boldsymbol{y}_1$, as a percent of the unconditional uncertainty about x_t.

2.3 DSGE models

A linearized DSGE model can be expressed as a recursive equilibrium law of motion given by the following system of equations:

$$\boldsymbol{y}_t = \boldsymbol{C}(\boldsymbol{\theta})\boldsymbol{v}_{t-1} + \boldsymbol{D}(\boldsymbol{\theta})\boldsymbol{u}_t \tag{2.22}$$

$$\boldsymbol{v}_t = \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{v}_{t-1} + \boldsymbol{B}(\boldsymbol{\theta})\boldsymbol{u}_t \tag{2.23}$$

$$\boldsymbol{u}_t = \boldsymbol{G}(\boldsymbol{\theta})\boldsymbol{u}_{t-1} + \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim \mathcal{N}\left(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}(\boldsymbol{\theta})\right) \tag{2.24}$$

where $\boldsymbol{y}_t$ is a $n_{\boldsymbol{y}}$ vector of observed variables, $\boldsymbol{v}_t$ is a $n_{\boldsymbol{v}}$ vector of endogenous state variables, $\boldsymbol{u}_t$ is a $n_{\boldsymbol{u}}$ vector of exogenous state variables, and $\boldsymbol{\varepsilon}_t$ is a $n_{\boldsymbol{u}}$ vector of exogenous shocks. The matrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$, and $\boldsymbol{G}$ are functions of the structural parameters of the model, collected in the $n_{\boldsymbol{\theta}}$ vector $\boldsymbol{\theta}$.

In practice, latent variables researchers might be interested are endogenous variables, such as output gap, exogenous shocks, such as total factor productivity (TFP), or innovations to exogenous shocks, such as the innovation to the TFP shock. In other words, and using the notation from sections 2.1 and 2.2, the latent variable x_t will be an element of $\boldsymbol{v}_t$, $\boldsymbol{u}_t$, or $\boldsymbol{\varepsilon}_t$, while the vector of observed variables is $\boldsymbol{y}_t$. Evaluating the unconditional and conditional information gain measures requires knowing the spectral and cross-spectral densities of x_t, $\boldsymbol{y}_t$, and individual elements of $\boldsymbol{y}$. Those can be obtained from the joint spectral density matrix of $\boldsymbol{z}_t = [\boldsymbol{y}_t', \boldsymbol{v}_t', \boldsymbol{u}_t', \boldsymbol{\varepsilon}_t']'$, which is given by (see Uhlig (1999)):

$$\boldsymbol{f}_{\boldsymbol{zz}}(\omega) = \frac{1}{2\pi}\boldsymbol{W}(\omega,\boldsymbol{\theta})\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}(\boldsymbol{\theta})\boldsymbol{W}(\omega,\boldsymbol{\theta})^{*} \tag{2.25}$$

where

$$\boldsymbol{W}(\omega,\boldsymbol{\theta}) = \begin{bmatrix} \boldsymbol{C}(\boldsymbol{\theta})e^{-i\omega} & \boldsymbol{D}(\boldsymbol{\theta}) & \mathbf{O}_{n_{\boldsymbol{y}},n_{\boldsymbol{u}}} \\ \boldsymbol{I}_{n_{\boldsymbol{v}}} & \mathbf{O}_{n_{\boldsymbol{v}},n_{\boldsymbol{u}}} & \mathbf{O}_{n_{\boldsymbol{v}},n_{\boldsymbol{u}}} \\ \mathbf{O}_{n_{\boldsymbol{u}},n_{\boldsymbol{v}}} & \boldsymbol{I}_{n_{\boldsymbol{u}}} & \mathbf{O}_{n_{\boldsymbol{v}},n_{\boldsymbol{u}}} \\ \mathbf{O}_{n_{\boldsymbol{u}},n_{\boldsymbol{y}}} & \mathbf{O}_{n_{\boldsymbol{u}},n_{\boldsymbol{u}}} & \boldsymbol{I}_{n_{\boldsymbol{u}}} \end{bmatrix} \times \tag{2.26}$$
$$\begin{bmatrix} \left(\boldsymbol{I}_{n_{\boldsymbol{v}}} - \boldsymbol{A}(\boldsymbol{\theta})e^{-i\omega}\right)^{-1}\boldsymbol{B}(\boldsymbol{\theta})\left(\boldsymbol{I}_{n_{\boldsymbol{u}}} - \boldsymbol{G}(\boldsymbol{\theta})e^{-i\omega}\right)^{-1} \\ \left(\boldsymbol{I}_{n_{\boldsymbol{u}}} - \boldsymbol{G}(\boldsymbol{\theta})e^{-i\omega}\right)^{-1} \\ \boldsymbol{I}_{n_{\boldsymbol{u}}} \end{bmatrix}$$

and the asterisk denotes matrix transposition and complex conjugation.

In business cycle research, it is common to divide the spectrum into three nonoverlapping intervals, corresponding to business cycle frequencies with periodicity between 6 and 32 quarters (as is standard in the literature, for example Stock and Watson (1999)), and frequencies above and below that interval, labeled as low and high frequencies, respectively. Let$\boldsymbol{\omega}^{BC}$, $\boldsymbol{\omega}^{L}$, and $\boldsymbol{\omega}^{H}$ denote the respective frequency bands. Then, the total information gain from $\boldsymbol{y}_t$ with respect to x_t can be decomposed as follows:

$$\begin{aligned} \mathrm{IG}_{\boldsymbol{y}\to x}(\overline{\boldsymbol{\omega}}) &= \mathrm{IG}_{\boldsymbol{y}\to x}(\boldsymbol{\omega}^{L})\frac{f_{xx}(\boldsymbol{\omega}^{L})}{f_{xx}(\overline{\boldsymbol{\omega}})} + \mathrm{IG}_{\boldsymbol{y}\to x}(\boldsymbol{\omega}^{BC})\frac{f_{xx}(\boldsymbol{\omega}^{BC})}{f_{xx}(\overline{\boldsymbol{\omega}})} \\ &+ \mathrm{IG}_{\boldsymbol{y}\to x}(\boldsymbol{\omega}^{H})\frac{f_{xx}(\boldsymbol{\omega}^{H})}{f_{xx}(\overline{\boldsymbol{\omega}})} \end{aligned} \tag{2.27}$$

Therefore, the total information gain is given by the weighted sum of the band-specific information gains, with weights equal to the contribution of each frequency band to the total variance of x.

Decomposing information gains across frequency bands is possible because of the mutual independence of the respective frequency components. Since the variables in $\boldsymbol{y}$ are typically correlated, the overall information about x cannot be decomposed into contributions of individual observed variables. What we can measure instead are the marginal contribution of a given observed variable y_i, as well as its conditional contribution given the information in other observed variables $\boldsymbol{y}_j \subset \boldsymbol{y}_{-i} \equiv \{\boldsymbol{y} \setminus y_i\}$. In particular, the following decomposition holds for any given frequency band $\boldsymbol{\omega}$:

$$\mathrm{IG}_{\boldsymbol{y}\to x}(\boldsymbol{\omega}) = \mathrm{IG}_{y_i \to x|\boldsymbol{y}_{-i}}(\boldsymbol{\omega}) + \mathrm{IG}_{\boldsymbol{y}_{-i}\to x}(\boldsymbol{\omega}) \tag{2.28}$$

The first term on the right hand side represents information in y_i about x that is not in $\boldsymbol{y}_{-i}$. Note that this includes both information that is unique to y_i, i.e. is independent from $\boldsymbol{y}_{-i}$, as well as information about x that emerges from observing y_i together with $\boldsymbol{y}_{-i}$. At the same time, some of the information in y_i about x is also in $\boldsymbol{y}_{-i}$, and is therefore captured by the second term in (2.27).

3 Applications

In this section, I show how the proposed measures can be used to investigate the sources and spectral domain distribution of information about structural shocks in dynamic macroeconomic models. I apply the methodology to two models taken from the literature:

the small-scale model of Uribe (2021), and the medium-scale model of Justiniano et al. (2011). Considering two models allows me to illustrate different elements of the analysis in a complementary fashion. The model of Uribe (2021) is much smaller, with only three observed variables, which makes it possible to present fully results regarding information interactions among those variables. This is not practicable in the case of the Justiniano et al. (2011) model, where I present only selected results and leave the rest for the Appendix. Another important difference is that the Uribe (2021) has more shocks than observables, and finding out how well each shock can be recovered is a relevant dimension of the analysis, in addition to investigating the main sources of information. This is not an issue in the second model, which, with its richer structure, larger number of shocks and observables, is much more representative of the medium-scale New Keynesian framework in the DSGE literature

3.1 Uribe (2021)

Uribe (2021) investigates the nature and empirical importance of monetary policy shocks that produce neo-Fisherian dynamics, i.e. move interest rates and inflation in the same direction over the short run. To that end, the author estimates a standard small-scale New-Keynesian model with price stickiness and habit formation, augmented with seven structural shocks. Full details about the model can be found in the original publication. Here I only describe those of its features that are directly relevant for the analysis which follows.

Firstly, three of the shocks are to monetary policy, which is described by the following policy rule:

$$\frac{1+I_t}{\Gamma_t} = \left[A \left(\frac{1+\Pi_t}{\Gamma_t} \right)^{\alpha_t} \left(\frac{Y_t}{X_t} \right)^{\alpha_y} \right]^{1-\gamma_I} \left(\frac{1+I_{t-1}}{\Gamma_{t-1}} \right)^{\gamma_I} e^{z_t^m}, \tag{3.1}$$

where I_t the nominal interest rate, Y_t is aggregate output, Π_t is the inflation rate, Γ_t is the inflation-target, X_t is a nonstationary productivity shocks, and z_t^m is a stationary interest-rate shock. The inflation target is defined as

$$\Gamma_t = X_t^m e^{z_t^{m2}}, \tag{3.2}$$

where X_t^m and z_t^{m2} are permanent and transitory components of the inflation target. It is assumes that X_t^m and X_t grow at a rates g_t^m and g_t, respectively.

There are two preference shocks affecting the lifetime utility function of the representative household, given by

$$\mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t e^{\xi_t} \left\{ \frac{\left[\left(C_t - \delta \tilde{C}_{t-1}\right)\left(1 - e^{\theta_t} h_t\right)^{\chi}\right]^{1-\sigma} - 1}{1-\sigma} \right\}, \tag{3.3}$$

where C_t is consumption, $\tilde{C}_t$ is the cross sectional average of consumption, h_t is hours worked, ξ_t is an intertemporal preference shock, and θ_t is a shock to labor supply.

In addition to X_t, there is also a stationary productivity shock z_t, which affects the production technology according to

$$Y_t = e^{z_t} X_t h_t^{\alpha}, \tag{3.4}$$

The five stationary shocks (ξ_t, θ_t, z_t, z_t^m, and z_t^{m2}) and the growth rates of the two non-stationary shocks (g_t and g_t^m) are all assumed to follow first-order autoregressive processes.

Uribe (2021) estimates the model using quarterly US data on three variables: per capita output growth ($\triangle y_t$), the interest-rate-inflation differential ($r_t = i_t - \pi_t$), and the change in the nominal interest rate ($\triangle i_t$). All variables are assumed to be observed with measurement errors. Thus, there are ten independent sources of randomness in the data and only three observables. Clearly, not all, if any, of the latent variables can be recovered fully. The purpose of the remainder of this section is to determine how well each structural shock can be recovered and where in the spectrum most of the information comes from, as well as what are the information contributions of different observed variables overall and across different frequency bands.

3.1.1 Information decomposition across frequency bands

Uribe (2021) solves the model by log-linear approximation of the equilibrium conditions around steady state. The linearity of the solution together with the assumption that the structural innovations and the measurement errors are Gaussian, implies that the joint distribution of (any subset of) the innovations, shocks, state and observed variables is also Gaussian. Therefore, the analysis of information gains can be conducted using the measures introduced in Section 2. In the analysis which follows I fix the parameter values at the mean of posterior distribution reported in Uribe (2021, Table 5).

Table 1 presents the total information gains for the seven shocks and their decompositions into gains from three frequency bands - low, business cycle and high frequencies, with periodicities of more than 32 quarters, between 6 and 32 quarters and less than 6 quarters, respectively. The results show that none of the shocks can be fully recovered from the observed variables. The largest reduction of uncertainty is with respect to the intertemporal preference shock (ξ_t) – by about 93%, and the permanent productivity shock (g_t) – by about 85%. The gains with respect to the three monetary policy-related shocks are between 15% and 18%. The least information is gained with respect to the labor supply (θ_t) and the transitory productivity shocks (z_t), with information gains for both of 1.8%.

Table 1: Information gain decomposition across frequency bands

		total	low	BC	high
ξ_t	preference	93.2	70.4 = 96.4 × 0.73	19.5 = 88.4 × 0.22	3.2 = 66.0 × 0.05
θ_t	labor supply	1.8	0.2 = 0.5 × 0.33	1.1 = 2.3 × 0.48	0.5 = 2.9 × 0.18
z_t	transitory productivity	1.8	0.2 = 0.5 × 0.32	1.1 = 2.2 × 0.49	0.5 = 2.9 × 0.19
g_t	permanent productivity	83.5	9.3 = 94.9 × 0.10	32.3 = 87.1 × 0.37	42.0 = 78.9 × 0.53
z_t^m	transitory interest rate	15.5	0.1 = 0.9 × 0.12	3.2 = 7.9 × 0.41	12.2 = 25.7 × 0.47
z_t^{m2}	transitory trend inflation	16.5	5.8 = 12.7 × 0.46	9.7 = 23.1 × 0.42	1.0 = 8.3 × 0.12
g_t^m	permanent trend inflation	18.0	7.2 = 69.4 × 0.10	7.0 = 18.3 × 0.38	3.9 = 7.5 × 0.51

Note: Information gain (IG) measures the reduction of uncertainty (variance) about a shock due to observing all three observed variables, as a *percent* of the unconditional uncertainty of the shock. The contribution from each frequency band is shown as the product of the IG for that band and the variance in that band as a *fraction* of the total variance. Thus, the units in the last three columns are $\% = \% \times \frac{\text{variance band}}{\text{variance total}}$.

Columns 3 to 5 of the table show the information gain contributions from each frequency band. Following the earlier discussion (see equation (2.27)), the total contribution in each case is shown as the product of two terms: the band-specific information gain, which measures the reduction of uncertainty as a percent of the uncertainty in that band, and the fraction of total uncertainty that originates in the given frequency band.

For six of the seven shocks uncertainty is concentrated in either low and business cycle frequencies, or high and business cycle frequencies. Specifically, in the first groups are the transitory trend inflation, transitory productivity, and the labor supply shocks. And in the second are the permanent productivity, transitory interest rate, and permanent trend inflation shocks. The one exception is the intertemporal preference shock for which the low frequencies are by far the main source of uncertainty. As can be expected, the largest gains generally come from parts of the spectrum where prior uncertainty is larger.

There are some notable exceptions, however. In particular, note that even though the low frequency band accounts for only 10% of the uncertainty about the permanent trend inflation shock, the information gain contribution from that band is largest than the business cycle frequency band, which accounts for 38% of the uncertainty, and much larger than the contribution from the high frequency band, which accounts for more than half of the total uncertainty. This is due to the fact that a much larger fraction of the uncertainty in the low frequencies is resolved by information provided by the observed variables than is the case for the higher frequencies. Similarly, note that for the labor supply and transitory productivity shocks, because of the relatively larger information gains from the higher end of the spectrum, the information contributions from there is larger than from the low frequencies, even though the low frequencies account for a significantly larger fraction of the prior uncertainty.

3.1.2 Information contributions by variables

Table 2 shows the conditional information gains for each observed variable for the full spectrum and the three frequency bands. The largest contribution by far is from output growth ($\triangle y_t$) with respect to the permanent productivity shock. Note that the conditional information gain of 83.4% is almost equal to the total gain (all observables) of 83.5% for that shock (see Table 1). This implies that the other two variables - the interest rate-inflation differential (r_t) and the change in the nominal interest rate ($\triangle i_t$) alone reduce the uncertainty about the permanent productivity shock by only 0.1%. This result holds for the full spectrum and the individual frequency bands. Output growth contributes less information for the other shocks, compared to r_t or $\triangle i_t$. The contributions of these variables with respect to the two trend inflation shocks are similar, with r_t being relatively more informative for the transitory trend inflation shock, while $\triangle i_t$ is more informative for the permanent one. In addition, r_t contributes much more information than either $\triangle y_t$ or $\triangle i_t$ with respect to the preference shock, while $\triangle i_t$ is the most informative observable with respect to the transitory interest rate shock, and, marginally, for the labor supply and transitory productivity shocks.

The ranking of variables in terms of their total information contributions is determined by the relative size of the information gains in the part of the spectrum from where a given shock receives the most total information (see Table 1). In several cases, the ranking changes with the frequency band. For instance, $\triangle i_t$ contributes significantly more information than r_t with respect to the intertemporal preference shock in the BC

Table 2: Conditional information gains

	shock	total			low			BC			high		
		$\triangle y_t$	r_t	$\triangle i_t$	$\triangle y_t$	r_t	$\triangle i_t$	$\triangle y_t$	r_t	$\triangle i_t$	$\triangle y_t$	r_t	$\triangle i_t$
ξ_t	preference	0.3	26.8	7.2	0.0	26.4	0.8	0.1	0.5	4.1	0.1	0.0	2.3
θ_t	labor supply	0.1	0.1	1.1	0.0	0.0	0.0	0.1	0.0	0.6	0.0	0.0	0.5
z_t	transitory productivity	0.1	0.0	1.1	0.0	0.0	0.0	0.1	0.0	0.6	0.0	0.0	0.5
g_t	permanent productivity	83.4	0.8	5.7	9.3	0.0	0.1	32.2	0.6	2.8	41.9	0.2	2.8
z_t^m	transitory interest rate	2.2	1.5	9.0	0.0	0.1	0.0	0.4	0.9	0.4	1.8	0.5	8.5
z_t^{m2}	transitory trend inflation	1.7	13.0	8.2	0.1	5.5	4.3	1.1	7.3	3.7	0.5	0.2	0.2
g_t^m	permanent trend inflation	0.5	10.4	15.6	0.0	4.7	7.0	0.2	5.2	5.6	0.3	0.5	3.0

Note: Conditional information gain measures the additional reduction of uncertainty (variance) about a shock due to observing a variable given that the other two variables are also observed, as a percent of the unconditional uncertainty of the shock. The observed variables are: output growth ($\triangle y_t$), interest-rate-inflation differential (r_t), and the change in the nominal interest rate ($\triangle i_t$). Due to rounding in some cases the band-specific contributions do not add up to the total values.

and high frequencies. At the same time, r_t contributes the most information with respect to the transitory interest rate shock in the low and BC frequencies, in spite of being the least informative variable in the high frequencies and overall. Similarly, $\triangle y_t$ is the least informative variable overall with respect to the transitory trend inflation shock, but has the largest contribution in the high frequency band.

It is worth emphasizing that the information gains shown in Table 2 are from observing a given variable *conditional* on already having observed the other two variables. As the observed variables are obviously not mutually independent, it is conceivable that in some cases the contributions are small because different variables share common information

Table 3: Unconditional information gains

	shock	total			low			BC			high		
		$\triangle y_t$	r_t	$\triangle i_t$	$\triangle y_t$	r_t	$\triangle i_t$	$\triangle y_t$	r_t	$\triangle i_t$	$\triangle y_t$	r_t	$\triangle i_t$
ξ_t	preference	3.5	84.6	66.0	0.9	69.6	44.0	2.0	14.6	18.9	0.6	0.4	3.1
θ_t	labor supply	0.0	0.6	1.7	0.0	0.1	0.2	0.0	0.5	1.0	0.0	0.0	0.5
z_t	transitory productivity	0.0	0.6	1.6	0.0	0.1	0.1	0.0	0.5	1.0	0.0	0.0	0.5
g_t	permanent productivity	76.7	0.1	0.1	9.0	0.0	0.0	28.7	0.0	0.0	39.1	0.0	0.0
z_t^m	transitory interest rate	0.7	5.8	11.5	0.0	0.1	0.0	0.2	2.5	1.8	0.5	3.2	9.7
z_t^{m2}	transitory trend inflation	2.2	5.3	0.9	0.2	1.1	0.1	1.6	3.8	0.6	0.4	0.4	0.2
g_t^m	permanent trend inflation	1.8	0.4	6.8	0.1	0.0	2.5	0.9	0.3	1.3	0.8	0.1	3.0

Note: Unconditional information gain measures the reduction of uncertainty (variance) about a shock due to observing a given variable, as percent of the unconditional uncertainty of the shock.

with respect to those shocks. To help find out if and when that is the case, Table 3 shows the unconditional information gains, i.e. the percent reduction of uncertainty about a given shock due to observing only one variable at a time.

The results reveal some notable differences between conditional and unconditional information gains. Most striking is the reduction in the contributions of the three observables with respect to the intertemporal preference shock. In particular, the information gains from r_t and $\triangle i_t$ change from, respectively, 85% and 66% unconditionally, to 27% and 7% conditionally. Similarly, the contribution of $\triangle y_t$ decreases from 3.5% to only 0.3%. This suggests that, to a large extent, the information in either one of the observable variables is not unique to them but is also contained in the other two. In other words, there is a significant degree of redundancy of the information about the intertemporal preference shock. Another, less striking, example of redundancy is the transitory interest rate shock, where the conditional information gains from r_t and $\triangle i_t$ are smaller than the unconditional ones.

Information redundancy is not the only possible consequence of the existing interdependence among observables. In the case of the permanent productivity shock, the conditional information gains for all variables are significantly larger than the unconditional ones. The same is true for the contributions of r_t and $\triangle i_t$ with respect to the permanent and transitory trend inflation shocks, as well as for the contribution of $\triangle y_t$ with respect to the transitory interest rate shock. In all of these cases there is a positive information complementarity instead of information redundancy, that is, information increases when variables are observed together.

Following Iskrev (2019), the degree of information complementarity between variables can be measured by comparing the joint information gain with respect to a shock to the individual gains. Specifically, the information complementarity between variables y_1 and y_2 conditional on variables $\boldsymbol{y}_3 \subset \{\boldsymbol{y} \setminus \boldsymbol{y}_{12}\}$ at frequency band $\boldsymbol{\omega}$ is defined as:

$$\mathrm{IC}_{\boldsymbol{y}_{12}\to x|\boldsymbol{y}_3}(\boldsymbol{\omega}) = \frac{\mathrm{IG}_{\boldsymbol{y}_{12}\to x|\boldsymbol{y}_3}(\boldsymbol{\omega})}{\mathrm{IG}_{y_1\to x|\boldsymbol{y}_3}(\boldsymbol{\omega}) + \mathrm{IG}_{y_2\to x|\boldsymbol{y}_3}(\boldsymbol{\omega})} - 1. \tag{3.5}$$

Negative values indicate negative complementarity, or information redundancy, between y_1 and y_2, and positive values indicate positive complementarity between the two variables. Since the information gain is non-negative, we have $\mathrm{IC}_{\boldsymbol{y}_{12}\to x|\boldsymbol{y}_3}(\boldsymbol{\omega}) \geq -1/2$, with equality when y_1 and y_2 are (conditionally on $\boldsymbol{y}_3$) functionally dependent, in which case $\mathrm{IG}_{\boldsymbol{y}_{12}\to x|\boldsymbol{y}_3}(\boldsymbol{\omega}) = \mathrm{IG}_{y_1\to x|\boldsymbol{y}_3}(\boldsymbol{\omega}) = \mathrm{IG}_{y_2\to x|\boldsymbol{y}_3}(\boldsymbol{\omega})$. A lack of information complementarity,

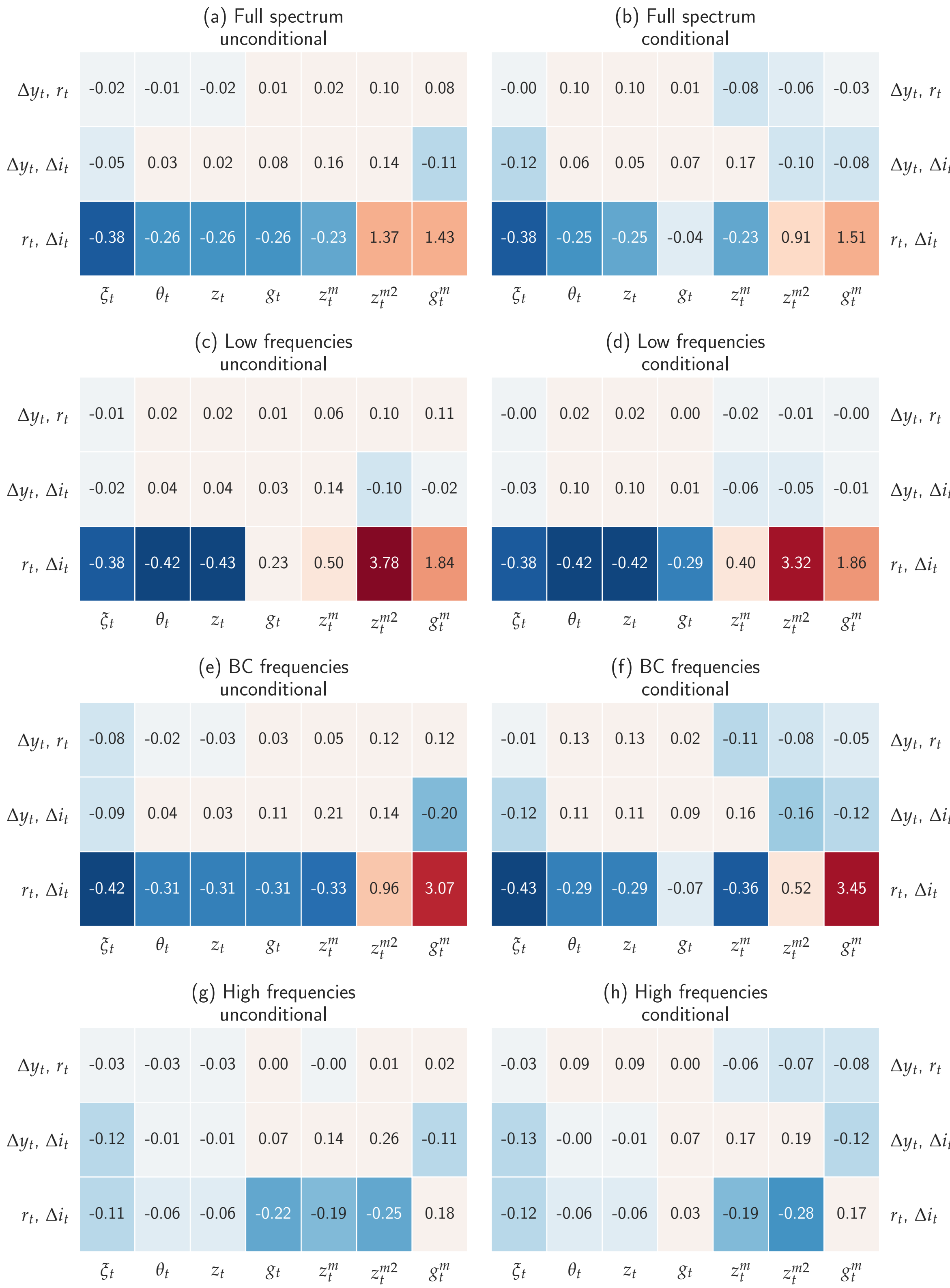

Figure 1: Pairwise information complementarity between observables with respect to shocks.

i.e. $\mathrm{IC}_{\boldsymbol{y}_{12}\to x|\boldsymbol{y}_3}(\boldsymbol{\omega}) = 0$ occurs when y_1 and y_2 are (conditionally on $\boldsymbol{y}_3$) independent, and hence $\mathrm{IG}_{\boldsymbol{y}_{12}\to x|\boldsymbol{y}_3}(\boldsymbol{\omega}) = \mathrm{IG}_{y_1\to x|\boldsymbol{y}_3}(\boldsymbol{\omega}) + \mathrm{IG}_{y_2\to x|\boldsymbol{y}_3}(\boldsymbol{\omega})$. Note that the conditioning could be with respect to any subset of observables, including the empty set, in which case we have unconditional complementarity between y_1 and y_2.

Figure 1 shows the unconditional and conditional information complementarities between all pairs of variables. The results are shown for the full spectrum as well as the three frequency bands. As already anticipated, the strongest complementarity overall is between r_t and $\triangle i_t$, and is negative for all shocks except the permanent and transitory trend-inflation shocks. Both unconditionally and conditionally the degree of complementarity tends to be significantly lower in the higher frequencies. Conditioning on the third observable in most cases preserves the sign of complementarity and reduces the magnitude. There are some notable exceptions to this pattern, however. For instance, the degree of complementarity between r_t and $\triangle i_t$ increases when conditioning on $\triangle y_t$, especially in the business cycle frequencies. Furthermore, the complementarity between the same variables with respect to the transitory productivity shock changes signs when conditioning on y_t, from positive to negative in the low frequencies, and from negative to positive in the high frequencies. At the same time, when evaluated over the full spectrum, the complementarity is strongly negative unconditionally and only weekly so, conditionally.

3.1.3 Information gains in the time domain

The time domain version of the full spectrum information gain measure (see equation (2.20)) is given by:

$$\mathrm{IG}_{\boldsymbol{Y}_T\to x_t} = \left(\frac{\mathrm{var}(x_t) - \mathrm{var}(x_t|\boldsymbol{Y}_T)}{\mathrm{var}(x_t)}\right) \times 100, \tag{3.6}$$

where $1 \leq t \leq T$ and $\boldsymbol{Y}_T = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_T\}$. The difference between the two measures is that, in the frequency domain, the information for any given x_t stems from the infinite past and future values of the observable variables. Therefore, for a given set of observed variables, the total amount of information is invariant to the temporal location of the latent variable. In contrast, in the time domain, it matters where the location of t is, relative to the beginning and the end of the sample. Thus, the value of time domain measure changes with t and is bounded from above by the value of the full spectrum frequency domain measure.

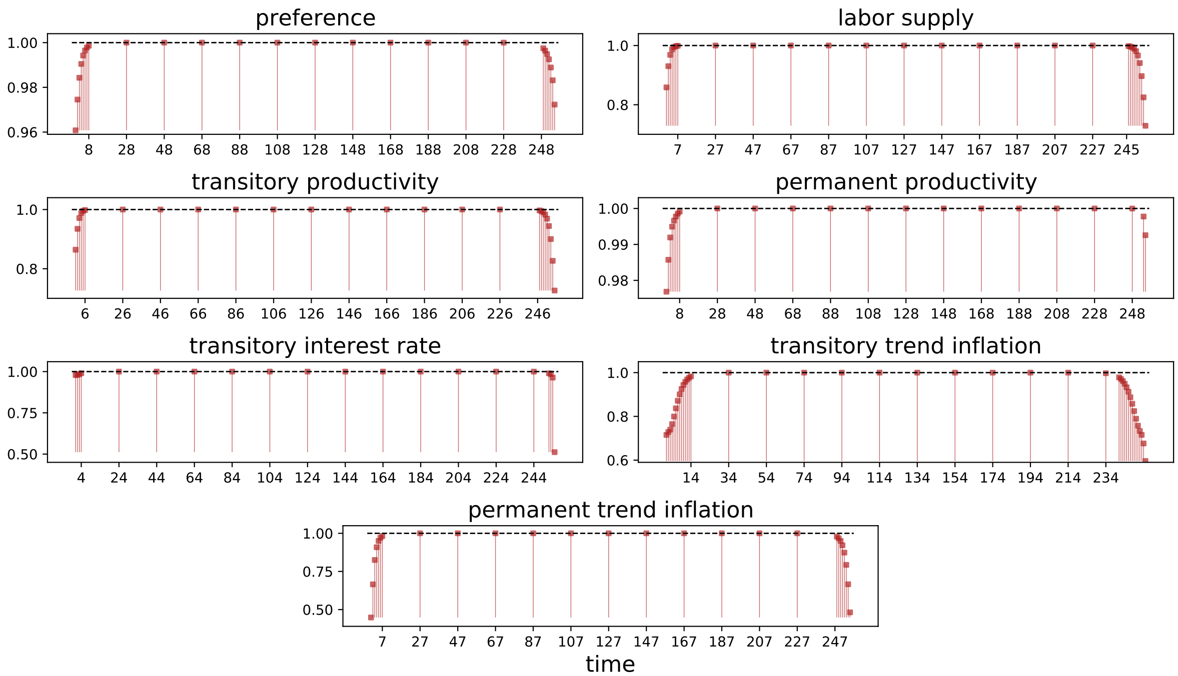

Figure 2: Total information gains in the time domain relative to the frequency domain (full spectrum).

Figure 2 compares the time and frequency domain information gains for the seven shocks in the model. Specifically, it shows the ratio of the time domain to the frequency domain measure for all values of t in a sample of $T = 255$ observations, which is the sample size in Uribe (2021). The results show that for most values of t the time and frequency domain information gains coincide. As anticipated, differences occur only at the beginning and end of the sample. For all shocks except the transitory trend inflation shock, for which convergence is somewhat slower, there are about ten observations or fewer on either end of the sample where the time domain information gains are smaller than the frequency domain ones.

3.1.4 Discussion of the results

As already noted, having more sources of uncertainty than the number of observed variables necessarily implies that the latent variables in the model cannot all be recovered fully. At the same time, as the results presented in Section 3.1.1 show, some shocks in the Uribe (2021) model are significantly better recoverable than others. The goal of this section is to develop a further understanding of these findings.

A natural question to ask is: why are the information gains with respect to the intertemporal preference and permanent productivity shocks so much larger than the gains for the remaining shocks, and in particular compared to those with respect to the labor supply and transitory productivity shocks? Intuitively, the amount of information one or more variables contain about another variable depends on the strength of their mutual dependence.[2] Furthermore, an insight gained from the frequency domain perspective is that the interactions need to be strong in the parts of the spectrum that are mainly responsible for the uncertainty of the latent variable. In addition, the extent to which information from multiple sources accumulates, in turn, depends on how interdependent they are among themselves. For instance, variables that are functionally dependent on other observed variables provide no useful information.[3]

Consider the intertemporal preference shock (ξ_t). According to the posterior mean estimates reported in Uribe (2021, Table 5), this shock is significantly more persistent and volatile than all other shocks. In particular, its volatility is an order of magnitude larger than the volatilities of all other shocks except the permanent productivity shock (g_t). The high degree of persistence explains why most of the uncertainty about ξ_t is concentrated in the lower end of the spectrum, as seen in Table 1. Furthermore, as seen from the same table, most of the uncertainty in the low frequencies is resolved by the information contained in the observed variables, which suggests that there are strong interactions between ξ_t and (some of) those variables. Since, in the present context, the variables have a clear causal direction, i.e. from shocks to endogenous variables, a natural way of describing their interactions is in terms of the shocks' impact on the observed variables. A convenient measure of the size of the total impact is each shock's contribution to the total variance of each variable. Figure 3 shows the individual contributions of the shocks as a percent of the total variances of the observables, as well as decompositions of the individual and total contributions in the low, BC, and high frequency bands. Note that the measurement errors also contribute to the variances, which is why the total contributions of the shocks sum up to less than 100%.

The results show that ξ_t drives most of the volatility in two of the observed variables – r_t and $\triangle i_t$, and, in the case of r_t, the contribution is mostly in the low frequencies. This

[2] In fact, the mutual information coefficient is commonly used to measure and test for statistical dependence between random variables (see e.g. Linfoot (1957), Joe (1989), and Granger and Lin (1994)).

[3] An example of this is output growth in the model estimated by Schmitt-Grohé and Uribe (2012), see Iskrev (2019) for details.

Shock	Band	$\triangle y$	r	$\triangle i$
all shocks	total	93.7	84.1	96.3
	low	8.7	64.8	8.8
	BC	37.2	17.3	40.7
	high	47.8	2.0	46.8
ξ	total	10.9	77.3	72.4
	low	0.2	62.7	7.3
	BC	4.3	13.8	34.8
	high	6.5	0.8	30.3
θ	total	0.1	0.4	2.4
	low	0.0	0.2	0.1
	BC	0.0	0.2	0.9
	high	0.0	0.0	1.4
z	total	0.1	0.4	2.3
	low	0.0	0.2	0.1
	BC	0.0	0.2	0.9
	high	0.0	0.0	1.4
g	total	76.6	0.0	0.1
	low	8.4	0.0	0.0
	BC	30.1	0.0	0.0
	high	38.2	0.0	0.0
z^m	total	0.7	2.0	11.9
	low	0.0	0.2	0.0
	BC	0.2	0.9	1.7
	high	0.5	0.8	10.1
z^{m2}	total	3.5	3.6	1.6
	low	0.1	1.3	0.0
	BC	1.7	1.9	0.8
	high	1.7	0.4	0.8
g^m	total	1.8	0.4	5.8
	low	0.1	0.2	1.4
	BC	0.9	0.2	1.5
	high	0.8	0.0	2.8
		$\triangle y$	r	$\triangle i$

Figure 3: Total and individual contributions of the shocks as a percent of the variances of the observables in the full spectrum and the low, business cycle, and high frequency bands. The difference to 100% is accounted for by the measurement error variances.

is consistent with the earlier findings that, of the three observed variables, r_t is the most informative and $\triangle y_t$ – the least informative one. Similarly, the second best recoverable shock – to permanent productivity, is responsible for the bulk of the volatility of the third variable – $\triangle y_t$, and particularly in the BC and high frequencies, which, as seen in Table 1, is also where most of the uncertainty of that shock stems from. The variance contributions of the remaining five shocks are significantly smaller, and account for only between 12%, in the case of transitory interest rate shock (z_t^m) with respect to $\triangle i_t$, and 2.3% – 2.4% in the case of both the labor supply (θ_t) and transitory productivity (z_t) shocks with respect again to $\triangle i_t$.

Equivalence between variance and information decompositions. Variance decompositions in dynamic structural models are typically obtained by shutting-off all shocks but one at a time and then computing the endogenous variables' variances or spectral densities (see for instance Fernández-Villaverde et al. (2016, Section 8)). This gives the contribution of each shock to the total variances or spectral densities of the endogenous variables. It is easy to see that the same quantities can be obtained using the information gain measures introduced in Section 2. Specifically, a shock's contribution to the variance of a variable is equal to the information gained, i.e. the reduction in variance, with respect to the variable due to knowing that shock. In other words, instead of information from observed variables to shocks, we measure the flow of information in the opposite direction – from shocks to observables. Of course, this only works when the shocks are mutually independent, which is also the assumption behind the standard variance decomposition approach. If shocks are mutually dependent one has to distinguish between conditional and unconditional variance contributions, as in the case of information from observed variables with respect to shocks.

To summarize, as expected, there is a clear link between, on the one hand, the shocks' contributions to the observed variables' volatilities and, on the other hand, the degree to which each shock can be recovered from information contained in those variables. At the same time, it is important to point out that the size of the contributions is not necessarily a good indicator of the variables' importance as sources of information about the shocks. For instance, the intertemporal preference shock contributes similar fractions of the variances of r_t and $\triangle i_t$. Yet, r_t is significantly more informative than $\triangle i_t$ about that shock. As noted earlier, this is due to the fact that the variance contributions are in different parts of the spectrum – the low frequencies in the case of r_t, and the BC and high frequencies, in the case of $\triangle i_t$. Since most of the variance of the preference shock comes from the low frequencies, r_t is significantly more informative than $\triangle i_t$. In other cases, it is the information interactions among the observed variables that affect their relative importance as sources of information. For instance, as can be seen in Table 2, the conditional contribution of information by $\triangle i_t$ with respect to the transitory trend inflation shock (z_t^{m2}) is much larger than that of $\triangle y_t$, in spite of the significantly larger fraction of the variance of $\triangle y_t$ attributed to that shock, compared to $\triangle i_t$. This is explained by the strong positive complementarity between r_t and $\triangle i_t$ in the BC and especially the low frequencies, which is where most of the uncertainty of that shocks is located. Lastly, small variance contributions of a shock does not necessarily imply that the shock cannot be recovered. In general, having the same number of non-redundant observables as the number of sources of uncertainty means that all shocks are fully recoverable. This is the case in the model I consider next.

3.2 Justiniano, Primiceri, and Tambalotti (2011)

Justiniano et al. (2011) (henceforth JPT) investigate whether investment shocks are important drivers of business cycle fluctuations. To that end, and expanding on their previous work in Justiniano et al. (2010), they estimate a New Keynesian model featuring imperfectly competitive goods and labor markets, as well as different nominal and real frictions such as sticky prices and wages, habit formation in consumption, variable capital utilization and investment adjustment costs. As in the previous section, here I outline only those features of the model that are relevant for the information decomposition analysis that follows.

The model has eight structural shocks in total, with three technology shocks, two of which are related to investment. Specifically, JPT distinguish between final and intermediate consumption, investment, and capital goods, each being produced in a different sector. They introduce a shock that affects the transformation of consumption into investment goods, and another shock that affects the transformation of investment goods into productive capital. The first, called investment-specific technology (IST) shock, is introduced via the production function in the investment good producing sector:

$$I_t = \Upsilon_t Y_t^I, \tag{3.7}$$

where I_t is the quantity of investment goods in efficiency units produced with Y_t^I units of the final good. Υ_t represents the IST and is assumed to be a non-stationary random process growing at a rate υ_t.

The second investment technology shock is introduced via the production technology in the capital good producing sector, which assumes that new capital, denoted with i_t, is produced from investment goods according to

$$i_t = \mu_t \left(1 - S\left(\frac{I_t}{I_{t-1}}\right)\right) \tag{3.8}$$

where S is an investment adjustment cost function, and μ_t is a stationary shock to the marginal efficiency of investment (MEI), assumed to be an AR(1) process.

The third technology shocks affects the production functions in the intermediate good producing sector according to:

$$Y_t(i) = \max\{A_t^{1-\alpha} K_t(i)^\alpha L_t(i)^{1-\alpha} - A_t \Upsilon_t^{\frac{\alpha}{1-\alpha}} F; 0\} \tag{3.9}$$

where $Y_t(i)$, $K_t(i)$, and $L_t(i)$ are the quantities of output produced, and effective capital and labor employed by intermediate good producer i. F represents fixed cost of production, and A_t is a common non-stationary neutral technology process, growing at rate z_t.

The final consumption good Y_t is produced by combining a continuum of intermediate goods, according to

$$Y_t = \left[\int_0^1 Y_t(i)^{\frac{1}{1+\lambda_{p,t}}}\right]^{1+\lambda_{p,t}} \tag{3.10}$$

where $\lambda_{p,t}$ is a stationary price markup shock following ARMA(1,1) process.

Similarly to the model in the previous section, there is a shock to the intertemporal preferences of the households populating the economy whose lifetime utility function is given by

$$\mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t b_t \left\{ \log\left(C_t - hC_{t-1}\right) - \varphi \frac{L_t(j)^{1+\nu}}{1+\nu} \right\}, \tag{3.11}$$

where C_t is consumption, b_t is the stationary intertemporal preference shock, assumed to follow an AR(1) process. JPT assume that there is a continuum of households $j \in [0,1]$, each one being a supplier of specialized labor denoted by $L_t(j)$. The specialized labor in turn is combined into homogenous labor input according to

$$L_t = \left[\int_0^1 L_t(i)^{\frac{1}{1+\lambda_{w,t}}}\right]^{1+\lambda_{w,t}} \tag{3.12}$$

where $\lambda_{w,t}$ is a stationary wage markup shock assumed to follow an ARMA(1,1) process.

The last two shocks are to government fiscal and monetary policy. Public spending G_t is a time-varying fraction of output,

$$G_t = \left(1 - \frac{1}{g_t}\right) Y_t \tag{3.13}$$

where the government spending shock g_t as a stationary AR(1) process.

Monetary policy consists of setting the nominal interest rate R_t according to the following policy rule:

$$\frac{R_t}{R} = \left(\frac{R_{t-1}}{R}\right)^{\rho_R} \left[\left(\frac{\pi_t}{\pi}\right)^{\phi_\pi} \left(\frac{X_t}{X_t^*}\right)^{\phi_X}\right]^{1-\rho_R} \left[\frac{X_t/X_{t-1}}{X_t^*/X_{t-1}^*}\right]^{\phi_{dX}} \varepsilon_{mp,t}, \tag{3.14}$$

where $e_{mp,t}$ is the monetary policy shock, R is the steady state of the nominal rate, π_t is the inflation rate, $X_t = C_t + I_t + G_t$ is actual real GDP and X_t^* is the level of GDP under flexible prices and wages and in the absence of markup shocks.

To summarize, there are eight shocks in the model, six stationary and two non-stationary. Two of the stationary shocks – to price and wage markups, follow ARMA(1,1) processes, and one – to monetary policy, is an i.i.d process. The remaining stationary shocks – to government spending, MEI, and intertemporal preferences, as well as the growth rates of the two non-stationary shocks – to IST and neutral technology, follow AR(1) processes. The disturbances to all shocks are assumed to be Gaussian, leading to a linear Gaussian state space representation of the solution of log-linear approximation of model.

JPT estimate the model using US data on hours worked ($h_t = \log L_t$), inflation (π_t), the nominal interest rate (R_t), and the growth rates of GDP ($x_t = \triangle \log X_t$), consumption ($c_t = \triangle \log C_t$), investment ($i_t = \triangle \log I_t$), real wages ($w_t = \triangle \log \frac{W_t}{P_t}$), and the relative price of investment ($\pi_t^i = \triangle \log \frac{P_{It}}{P_t}$). Unlike Uribe (2021), they do not allow for measurement errors in any of the series. As seen below, this implies that all eight shocks can be recovered fully with the information in the eight observed variables. In the remainder of this section I investigate the main sources of information for each shock in terms of observed variables and parts of the spectrum.

3.2.1 Information decomposition across frequency bands

Table 4 presents the total information gains for the eight shocks and their decompositions into gains from the low, BC, and high frequencies. As noted earlier, all shocks can be fully recovered from information in the observables, in the full spectrum as well as within each frequency band. The information contributions from the bands reflect the fraction of each shock's variance originating in that band.

Table 4: Information gain decomposition across frequency bands, JPT model

	shock	total	low	BC	high
z	neutral technology	100	$11.2 = 100 \times 0.11$	$40.0 = 100 \times 0.40$	$48.8 = 100 \times 0.49$
g	government	100	$96.1 = 100 \times 0.96$	$3.2 = 100 \times 0.03$	$0.7 = 100 \times 0.01$
v	IST	100	$8.4 = 100 \times 0.08$	$33.6 = 100 \times 0.34$	$58.0 = 100 \times 0.58$
λ_p	price mark-up	100	$51.7 = 100 \times 0.52$	$16.1 = 100 \times 0.16$	$32.2 = 100 \times 0.32$
λ_w	wage mark-up	100	$5.1 = 100 \times 0.05$	$27.3 = 100 \times 0.27$	$67.6 = 100 \times 0.68$
b	preference	100	$22.8 = 100 \times 0.23$	$49.9 = 100 \times 0.50$	$27.4 = 100 \times 0.27$
ε_{mp}	monetary policy	100	$6.3 = 100 \times 0.06$	$27.1 = 100 \times 0.27$	$66.7 = 100 \times 0.67$
μ	MEI	100	$47.4 = 100 \times 0.47$	$40.8 = 100 \times 0.41$	$11.7 = 100 \times 0.12$

Note: Information gain (IG) measures the reduction of uncertainty (variance) about a shock due to observing all three observed variables, as a *percent* of the unconditional uncertainty of the shock. The contribution from each frequency band is shown as the product of the IG for that band and the variance in that band as a *fraction* of the total variance. Thus, the units in the last three columns are $\% = \% \times \frac{\text{variance band}}{\text{variance total}}$

For six of the eight shocks uncertainty is distributed monotonically across the frequency bands, i.e. increases or decreases moving from low to high frequencies. Only for one of them – the government spending shock, uncertainty is concentrated in a single frequency band – the low frequencies, contributing 96% of the total variance. In the case of the MEI shock, most of the uncertainty is in the low and BC frequencies. For the other two technology shocks – neutral and IST, as well as the wage mark-up and the monetary policy shocks, uncertainty is mostly in the BC and high frequencies. In the case of the intertemporal preference shock, half of the uncertainty is in the BC frequencies, and the rest is divided almost evenly between the low and high frequencies. The other shock with a non-monotonic distribution of uncertainty is the price mark-up shocks, for which about half of the variance is due to the low frequencies, with a significant contribution by the high frequencies, and the least amount of uncertainty due to the BC frequencies.

3.2.2 Information contributions by variables

Table 5: Conditional information gains, JPT model

	shocks	total								low							
		x	c	i	h	w	π	R	π^i	x	c	i	h	w	π	R	π^i
z	neutral technology	15.6	0.0	0.2	46.4	0.0	0.0	0.0	0.1	0.9	0.0	0.1	1.5	0.0	0.0	0.0	0.0
g	government	45.5	52.8	18.3	0.0	0.0	0.0	0.0	0.0	42.6	49.9	14.8	0.0	0.0	0.0	0.0	0.0
υ	IST	0.0	0.0	0.0	0.0	0.0	0.0	0.0	97.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	8.1
λ_p	price mark-up	13.7	0.2	0.3	21.4	29.3	32.4	0.2	0.2	11.0	0.0	0.2	9.1	27.3	0.5	0.0	0.1
λ_w	wage mark-up	0.8	0.2	0.4	1.4	93.2	23.3	0.3	0.3	0.1	0.0	0.0	0.1	1.6	1.4	0.1	0.0
b	preference	1.4	28.5	7.3	11.2	2.5	0.7	5.6	0.0	1.0	4.2	6.1	6.6	2.3	0.6	3.4	0.0
ε_{mp}	monetary policy	0.3	3.1	0.2	10.2	0.1	12.1	92.6	0.0	0.1	0.1	0.0	1.5	0.0	4.4	4.8	0.0
μ	MEI	0.1	0.0	8.7	0.4	2.2	0.4	5.2	1.9	0.0	0.0	3.9	0.1	2.0	0.1	3.2	1.2
	shocks	BC								high							
		x	c	i	h	w	π	R	π^i	x	c	i	h	w	π	R	π^i
z	neutral technology	4.5	0.0	0.0	13.4	0.0	0.0	0.0	0.0	10.2	0.0	0.0	31.6	0.0	0.0	0.0	0.0
g	government	2.5	2.6	3.0	0.0	0.0	0.0	0.0	0.0	0.5	0.3	0.5	0.0	0.0	0.0	0.0	0.0
υ	IST	0.0	0.0	0.0	0.0	0.0	0.0	0.0	33.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	56.1
λ_p	price mark-up	1.5	0.0	0.0	6.7	1.9	7.9	0.1	0.0	1.2	0.1	0.1	5.7	0.1	24.0	0.1	0.0
λ_w	wage mark-up	0.3	0.1	0.1	0.4	24.6	8.7	0.2	0.2	0.4	0.1	0.2	1.0	67.1	13.2	0.1	0.1
b	preference	0.1	10.0	1.1	2.7	0.2	0.2	1.8	0.0	0.2	14.3	0.1	2.0	0.0	0.0	0.4	0.0
ε_{mp}	monetary policy	0.1	0.8	0.1	4.8	0.0	4.1	24.8	0.0	0.1	2.2	0.2	3.9	0.0	3.6	63.0	0.0
μ	MEI	0.0	0.0	1.7	0.2	0.1	0.2	1.6	0.3	0.0	0.0	3.2	0.2	0.0	0.1	0.4	0.3

Note: Conditional information gain measures the additional reduction of uncertainty (variance) about a shock due to observing a variable given that the other seven variables are also observed, as a percent of the unconditional uncertainty of the shock. The observed variables are: the growth rates of output (y), consumption (c), investment, and wages (w), the inflation rates for consumption (π) and investment (π^i), hours worked (h) and the nominal interest rate (r). Due to rounding in some cases the band-specific contributions do not add up to the total values.

Table 5 shows the conditional information gains for each observed variable in the full spectrum and the individual frequency bands. The three largest contributions, each exceeding 90%, are from the growth rate of the relative price of investment (π^i) with respect to the IST shock (υ), from the real wage growth (w) with respect to the wage mark-up shock (λ_w), and from the nominal interest rate (R) with respect to the monetary policy shock (ε_{mp}). As JPT show, the price of investment in terms of consumption goods coincides with the inverse of the IST process. Therefore, the IST growth rate process is fully recovered by observing π^i alone. A conditional information gain of 97.2% in the full spectrum implies that, in absence of π^i, information from the remaining variables reduces uncertainty about υ by 2.8%. In addition to IST shock, π^i also contributes information with respect to the MEI shock, although much less compared to other variables, and in particular the investment growth rate, which is the most informative variable for that shock. Other large conditional contributions are from the growth rates of output and consumption with respect to the government spending shock, and from hours worked with respect to the neutral technology shock. Consumption growth is also the most informative variable with respect to the intertemporal preference shock, while inflation is the most informative observable with respect to the price mark-up.

Table 6: Unconditional information gains, JPT model

	shocks	total								low							
		x	c	i	h	w	π	R	π^i	x	c	i	h	w	π	R	π^i
z	neutral technology	17.4	20.1	7.3	24.3	28.7	24.4	9.6	0.0	5.4	6.1	2.7	0.6	7.5	2.0	0.6	0.0
g	government	0.4	3.1	0.1	4.4	0.0	1.7	4.8	0.0	0.1	3.1	0.1	4.2	0.0	1.7	4.7	0.0
v	IST	1.3	0.1	2.0	0.6	0.1	0.1	0.2	100.0	0.0	0.1	0.1	0.0	0.0	0.1	0.1	8.4
λ_p	price mark-up	1.8	0.3	4.4	4.8	18.1	39.3	3.6	0.0	1.3	0.3	3.8	4.3	8.3	6.2	1.5	0.0
λ_w	wage mark-up	0.4	0.4	0.6	1.0	59.7	3.4	0.6	0.0	0.3	0.2	0.2	0.8	0.2	2.1	0.5	0.0
b	preference	7.4	61.9	0.9	6.9	0.0	1.7	9.7	0.0	0.3	3.6	0.2	0.6	0.0	0.6	1.7	0.0
ε_{mp}	monetary policy	3.5	1.2	2.8	3.1	0.0	1.7	57.6	0.0	0.2	0.1	0.1	0.3	0.0	0.4	0.2	0.0
μ	MEI	47.2	10.3	73.1	56.3	3.4	9.7	51.6	0.0	16.5	8.4	28.2	24.9	2.8	4.9	30.4	0.0
	shocks	BC								high							
		x	c	i	h	w	π	R	π^i	x	c	i	h	w	π	R	π^i
z	neutral technology	7.5	8.4	2.7	7.7	16.8	14.3	6.0	0.0	4.5	5.6	1.9	15.9	4.4	8.1	3.0	0.0
g	government	0.1	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.0	0.0	0.0	0.0
v	IST	0.2	0.0	0.3	0.1	0.0	0.0	0.1	33.6	1.0	0.0	1.7	0.5	0.0	0.0	0.0	58.0
λ_p	price mark-up	0.3	0.0	0.3	0.3	3.6	7.2	0.6	0.0	0.3	0.0	0.3	0.2	6.2	25.8	1.5	0.0
λ_w	wage mark-up	0.1	0.1	0.2	0.1	10.1	1.2	0.2	0.0	0.0	0.1	0.2	0.0	49.4	0.1	0.0	0.0
b	preference	4.0	34.9	0.5	3.9	0.0	0.9	5.6	0.0	3.2	23.5	0.2	2.3	0.0	0.2	2.4	0.0
ε_{mp}	monetary policy	1.1	0.4	0.8	1.2	0.0	0.7	8.1	0.0	2.2	0.7	1.9	1.6	0.0	0.6	49.3	0.0
μ	MEI	24.6	1.9	34.6	26.6	0.6	4.4	20.2	0.0	6.1	0.1	10.4	4.7	0.0	0.4	1.0	0.0

Note: Unconditional information gain measures the reduction of uncertainty (variance) about a shock due to observing a given variable, as percent of the unconditional uncertainty of the shock.

With a few exceptions, variables that contribute the most information overall are also the most informative ones within each frequency band. One of the exceptions is the contribution of wage growth with respect to the price markup shock, which is significantly larger than the contribution of inflation in the low frequency band, but much smaller in the BC and high frequencies, and thus overall. Another notable exception is the intertemporal preference shock where consumption growth is by far the most informative variable overall, even though in the low frequency band the conditional contributions of both hours worked and investment growth are much larger.

Table 6 shows results for the unconditional information gains. As discussed earlier, for a given variable and a shock, the difference between conditional and unconditional information gains indicates the existence of information complementarities with respect to that shock between the variable and other observables. The complementary may be positive or negative depending on whether the conditional gains are larger or smaller than the unconditional ones.

The most extreme case of positive complementarity is observed with respect to the government spending shock, where the largest unconditional gain – from R, is less than 5%, whereas there are two variables – c and x, with conditional gains exceeding 45%, and a third one – i, with conditional gain exceeding 18%. The obvious explanation for this result is the existing tight relationship among x, c, i, and g implied by the resource constraint of the economy. Since g is latent, joint information from pairs of the observed resource constraint variables is larger than the information contained in each of them individually. This intuition can be confirmed by applying the measure of information complementarity introduced earlier (see equation (3.5)). The top panel of Figure 4 shows the largest, in absolute value, unconditional and conditional information complementarities with respect to the government spending shock. In both cases, the largest positive complementarities are between pairs of resource constraint variables. For instance, the value of 3.2 in the case of x and c implies that, conditional on observing the remaining six variables, observing x and c together provides 2.2 times as much information about g as adding up the information from each of them individually.

The bottom panel of Figure 4 shows the most significant complementarities with respect to the MEI shock. As can be confirmed by comparing the values reported in Table 5 and Table 6, the MEI shock presents the most prominent case of negative information complementarities. In particular, the gain from i, which is the most informative variable for that shock, both conditionally and unconditionally, drops from more than 70% unconditionally, to less than 10%, conditionally. Similarly, the information gains from R, h, and x all drop from around 50% to about 5% or less. This implies that, to a large extent, information from these variables regarding the MEI shock is not unique to them but is also contained in other observed variables. As can be seen in Figure 4, different combinations of i, R, h, and x are among the variable pairs with the strongest negative information complementarity. In the case of i and x, the explanation again can be traced to their strong mutual dependence, together with c, implied by the resource constraint

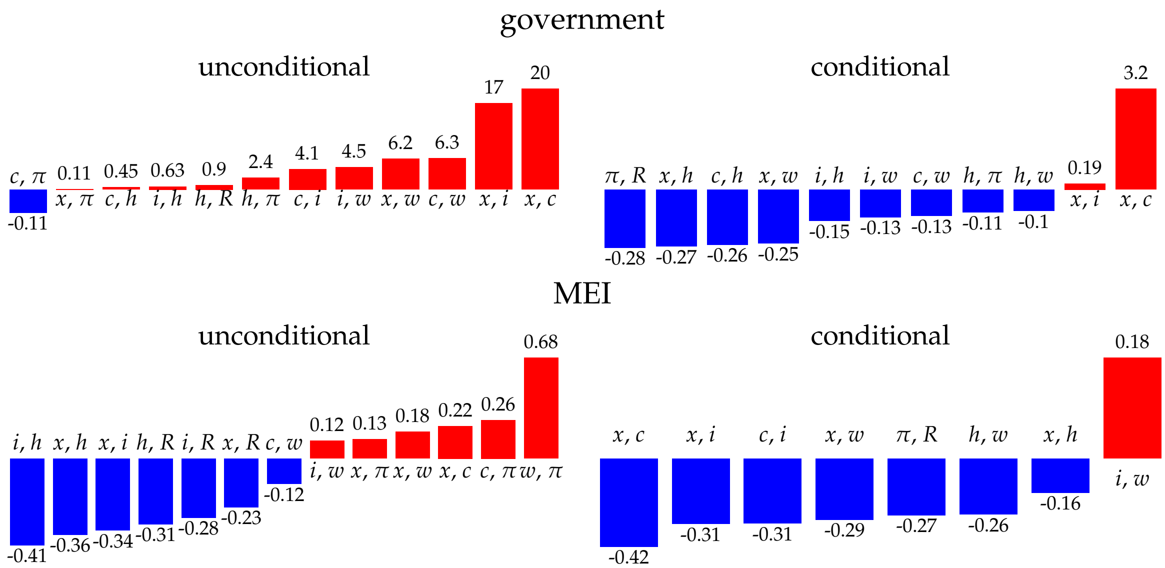

Figure 4: Largest pairwise information complementarities with respect to the government spending and MEI shocks, full spectrum.

of the economy. In fact, we should expect to find negative complementarity with respect to all shocks, other than g, between any two of the resource constraint variables when the third is among the conditional variables.[4] This is indeed the case as shown in more details in the Appendix. Also there, I report the pairwise complementarity coefficients for each frequency band. Examining those results can help explain, for example, the finding that the overall information complementarity between c and i with respect to the g shock is zero, which may be puzzling given the preceding argument for why all pairs of resource constraint variables should exhibit significant mutual complementarity. Indeed, the complementarities between c and i, c and x, and x and i are very strong in the BC and high frequencies. However, as seen earlier, almost all information about g is in the low frequencies, where the complementarity between c and i is zero.

The main conclusion of JPT is that the MEI shock is the key source of business cycle fluctuations whereas the IST shock plays no role. In particular, they show that the MEI shock is responsible for large fractions of the variances of GDP, investment, and hours *at business cycle frequencies*, and the contributions of the IST shock are nil. In

[4] At the risk of belaboring the obvious, consider the case where g is also observed, or, alternatively, where the g shock has zero variance. Then, because of the exact collinearity among them, the information in any variable entering the resource constraint with respect to all shocks is completely redundant given the information in the remaining resource constraint variables.

Figure 5, I report variance decompositions for the individual frequency bands as well as the overall contribution of each shock to the variances of the observed variables.[5] The results show that the MEI shock explains the bulk of the variances of x, i, and h in the full spectrum, not just the BC frequencies. In addition, the same shock contributes most of the volatility of R. This helps understand the earlier observation that R is the second most informative variable about μ (see Table 5). Note that the MEI shock contributes most of the variance in the low and BC frequencies of R – 65% and 55% of the total variance in those frequency bands, respectively, and those are the parts of the spectrum where most of the uncertainty about μ resides. Furthermore, unlike the resource constraint variables, a relatively smaller fraction of the information in R is redundant. The spectral decomposition of the contribution of the preference shock to the variance of c shows a relatively small impact in the low frequency band, which helps explain another observation made earlier – that in spite of being the most informative variable for that shock overall, c is not as informative about it in the low frequencies. Similarly, the fact that the contribution of the wage mark-up shock to the volatility of w is predominantly in the higher end of the spectrum is consistent with the dominant role of w as a source of information for that shock. In contrast, the low frequencies are a major source of uncertainty about the price mark-up shock, contributing more than half of its total variance. Given that most of that shock' contribution to the variance of h is also in the low frequency band, this helps understand why the importance of h as a source of information about λ_p is comparable to that of w and π, in spite of the much smaller fraction of total variance of h due to that shock.

As seen earlier, the importance of h is even greater in the case of the neutral technology shock, for which it is the observable with the largest conditional contribution of information. This might be hard to anticipate on the basis of the variance decomposition results, which show that more than half of the contribution of z is to the low frequency component of the variance, and only 0.3% of the total variance of h stems from the high frequency contribution of that shock. At the same time, the BC and high frequencies

[5]The JPT results are displayed in Table 3 of their article. There are several differences between their presentation and the one in Figure 5. One is that JPT show the variance contributions as fractions of the total variance in the business cycle frequencies. In my plot, the fractions are relative to the total variances in the full spectrum. To obtain comparable contributions, the values in the plot have to be multiplied by the fraction of the total variance of each shock due to the BC frequencies. Another difference is that for the trending variables (x, c, i, and w), JPT show decompositions for the levels, whereas I present results for the growth rates. Finally, the point estimates in JPT are the median values of the posterior distributions of the contributions. I present decompositions at the posterior median of the estimated parameters of the model.

account for almost 90% of the total information about z, and the bulk of the information contributed by h is within the high frequency band. As shown in more details in the Appendix, this result is due to, on the one hand, the strong positive complementarity between h and x, and, on the other hand, the also strong negative complementarities among x, c and i, as well as between π and w, and π and R. In other words, there is a substantial redundancy in the information about z in variables for which this shock is an important source of volatility. Furthermore, note that only 1% of the total variance of h originates in the high frequency band. Therefore, z is responsible for 30% of it, making it the second most important shock, after μ, for h in the high frequencies.

The last observation supports a point made earlier, with respect to the Uribe (2021) model, that the size of the variance contribution is not necessarily a good indicator of the variables' importance as sources of information about the shocks. As also pointed out earlier, it is possible that shocks are recoverable even if they play only a modest role as sources of volatility. In the JPT model, the monetary policy shock is responsible for at most 9.5% of the volatility of any observable, and the government spending shock contributes at most 7.3%. Yet both shocks are fully recoverable.

Shock	Band	x	c	i	h	w	π	R	π^i
all shocks	low	22.4	30.4	19.6	79.8	21.5	46.7	63.4	8.4
	BC	52.3	45.9	58.9	19.2	33.7	35.4	33.5	33.6
	high	25.3	23.7	21.5	1.0	44.7	17.9	3.1	58.0
z	total	23.2	30.4	8.4	6.6	33.1	22.2	8.1	0.0
	low	9.9	17.0	3.0	3.8	14.6	6.3	3.1	0.0
	BC	11.0	10.7	4.5	2.5	14.9	12.9	4.7	0.0
	high	2.4	2.7	0.9	0.3	3.6	3.0	0.2	0.0
g	total	7.3	2.2	0.1	2.1	0.0	0.5	1.1	0.0
	low	0.1	1.2	0.0	1.5	0.0	0.4	0.9	0.0
	BC	1.8	0.8	0.0	0.4	0.0	0.1	0.3	0.0
	high	5.4	0.2	0.0	0.1	0.0	0.0	0.0	0.0
v	total	0.7	0.3	1.1	0.4	0.1	0.4	0.9	100.0
	low	0.1	0.2	0.1	0.3	0.1	0.3	0.8	8.4
	BC	0.2	0.1	0.4	0.0	0.0	0.0	0.1	33.6
	high	0.4	0.0	0.6	0.0	0.0	0.0	0.0	58.0
λ_p	total	2.0	0.2	2.2	6.8	21.1	34.6	2.1	0.0
	low	0.8	0.1	0.9	6.2	4.9	6.4	0.9	0.0
	BC	1.0	0.0	1.1	0.6	7.5	14.2	1.0	0.0
	high	0.2	0.0	0.2	0.0	8.7	14.1	0.1	0.0
λ_w	total	1.5	1.8	1.0	26.2	44.0	28.2	10.4	0.0
	low	1.2	1.6	0.5	26.0	0.6	25.6	9.9	0.0
	BC	0.3	0.2	0.4	0.2	11.0	2.6	0.5	0.0
	high	0.0	0.0	0.1	0.0	32.4	0.0	0.0	0.0
b	total	7.3	56.5	1.0	3.3	0.0	2.0	8.4	0.0
	low	0.3	4.4	0.2	2.0	0.0	1.3	4.8	0.0
	BC	4.0	31.8	0.6	1.3	0.0	0.7	3.3	0.0
	high	2.9	20.3	0.2	0.1	0.0	0.1	0.3	0.0
ε_{mp}	total	3.7	1.3	2.8	4.9	0.0	3.9	9.5	0.0
	low	0.7	0.4	0.5	3.9	0.0	2.5	2.1	0.0
	BC	2.2	0.6	1.7	1.0	0.0	1.2	5.3	0.0
	high	0.8	0.3	0.6	0.0	0.0	0.2	2.1	0.0
μ	total	54.1	7.4	83.4	49.6	1.7	8.2	59.7	0.0
	low	9.3	5.6	14.3	36.2	1.3	3.9	40.9	0.0
	BC	31.7	1.6	50.1	13.0	0.3	3.7	18.4	0.0
	high	13.1	0.1	19.0	0.4	0.0	0.5	0.3	0.0
		x	c	i	h	w	π	R	π^i

Figure 5: Total and individual contributions of shocks as a percent of the variances of the observables in the full spectrum and the low, business cycle, and high frequency bands.

4 Conclusion

I have proposed a new framework for spectral decomposition of the information observables provide with respect to latent variables in dynamic macroeconomic models. Through this analysis, researchers can determine where in the frequency domain information about latent variables predominantly comes from, and evaluate the contributions of individual observed variables. In cases of information deficiency, the methodology can reveal what type of information is needed to better recover unobserved variables of interest. Having well-identified structural shocks and unobserved endogenous variables, such as potential output or natural rate of interest, is a key requirement for macroeconomic models to meet to be useful as tools for policy analysis and to be credible as story-telling devices. The methodology described in this paper will benefit both researchers who develop and estimate structural macroeconomic models, as well as the readers of such research, by improving their understanding and increasing the transparency of these models.

References

Alessi, L., M. Barigozzi, and M. Capasso (2011): "Non-Fundamentalness in Structural Econometric Models: A Review," *International Statistical Review*, 79, 16–47.

Beaudry, P., P. Fève, A. Guay, and F. Portier (2015): "When is Nonfundamentalness in VARs a Real Problem? An Application to News Shocks," Tech. rep., National Bureau of Economic Research.

Blanchard, O. J., J.-P. L'Huillier, and G. Lorenzoni (2013): "News, noise, and fluctuations: An empirical exploration," *The American Economic Review*, 103, 3045–3070.

Brillinger, D. R. (1981): *Time Series: Data Analysis and Theory*, vol. 36, SIAM.

Chahrour, R. and K. Jurado (2021): "Recoverability and Expectations-Driven Fluctuations," *Review of Economic Studies, forthcoming.*

Fernandez-Villaverde, J., J. Rubio-Ramirez, T. J. Sargent, and M. W. Watson (2007): "A, B, Cs (and Ds) of Understanding VARs," *American Economic Review*, 97.

Fernández-Villaverde, J., J. F. Rubio-Ramírez, and F. Schorfheide (2016): "Solution and estimation methods for DSGE models," *Handbook of Macroeconomics*, 2, 527–724.

Forni, M. and L. Gambetti (2014): "Sufficient information in structural VARs," *Journal of Monetary Economics*, 66, 124–136.

Franchi, M. and P. Paruolo (2015): "Minimality of state space solutions of DSGE models and existence conditions for their VAR representation," *Computational Economics*, 46, 613–626.

Franchi, M. and A. Vidotto (2013): "A check for finite order VAR representations of DSGE models," *Economics Letters*, 120, 100–103.

Giacomini, R. (2013): "The relationship between DSGE and VAR models," *Advances in Econometrics*, 31.

Giannone, D. and L. Reichlin (2006): "Does information help recovering structural shocks from past observations?" *Journal of the European Economic Association*, 4, 455–465.

Granger, C. and J.-L. Lin (1994): "Using the mutual information coefficient to identify lags in nonlinear models," *Journal of time series analysis*, 15, 371–384.

Hansen, H. and T. Sargent (1991): "Two Difficulties in Interpreting Vector Autoregressions," in *Rational Expectations Econometrics*, ed. by H. Hansen and T. Sargent, Westview Press.

Hansen, L. P. and T. J. Sargent (1980): "Formulating and estimating dynamic linear rational expectations models," *Journal of Economic Dynamics and control*, 2, 7–46.

Iskrev, N. (2019): "On the sources of information about latent variables in DSGE models," *European Economic Review*, 119, 318–332.

Joe, H. (1989): "Relative entropy measures of multivariate dependence," *Journal of the American Statistical Association*, 84, 157–164.

Justiniano, A., G. E. Primiceri, and A. Tambalotti (2010): "Investment shocks and business cycles," *Journal of Monetary Economics*, 57, 132–145.

——— (2011): "Investment shocks and the relative price of investment," *Review of Economic Dynamics*, 14, 102–121.

Leeper, E. M., T. B. Walker, and S.-C. S. Yang (2013): "Fiscal foresight and information flows," *Econometrica*, 81, 1115–1145.

Linfoot, E. H. (1957): "An informational measure of correlation," *Information and control*, 1, 85–89.

Lippi, M. and L. Reichlin (1993): "The dynamic effects of aggregate demand and supply disturbances: Comment," *The American Economic Review*, 83, 644–652.

——— (1994): "VAR analysis, nonfundamental representations, Blaschke matrices," *Journal of Econometrics*, 63, 307–325.

Priestley, M. B. (1981): *Spectral analysis and time series: probability and mathematical statistics*, 04; QA280, P7.

Ravenna, F. (2007): "Vector autoregressions and reduced form representations of DSGE models," *Journal of monetary economics*, 54, 2048–2064.

Sala, L., L. Gambetti, M. Forni, et al. (2016): "VAR Information and the Empirical Validation of DSGE Models," in *2016 Meeting Papers*, Society for Economic Dynamics, 260.

Schmitt-Grohé, S. and M. Uribe (2012): "What's News in Business Cycles," *Econometrica*, 80, 2733–2764.

Sims, E. R. (2012): "News, non-invertibility, and structural VARs," *Advances in Econometrics*, 28, 81.

Soccorsi, S. (2016): "Measuring nonfundamentalness for structural VARs," *Journal of Economic Dynamics and Control*, 71, 86–101.

Stock, J. H. and M. W. Watson (1999): "Business cycle fluctuations in US macroeconomic time series," *Handbook of macroeconomics*, 1, 3–64.

Uhlig, H. (1999): "A toolkit for analyzing nonlinear dynamic stochastic models easily," *Computational Methods for the Study of Dynamic Economies*, 30–61.

Uribe, M. (2021): "The neo-Fisher effect: Econometric evidence from empirical and optimizing models," *American Economic Journal: Macroeconomics, forthcoming.*

Appendix

A Uribe (2021) model

Table A1: Parameter values, Uribe (2021) model

	parameter	posterior mean
ϕ	price stickiness	146.000
α_π	coeff inflation in monetary policy rule	2.320
α_y	coeff output in monetary policy rule	0.188
γ_m	backward-looking component in inflation	0.606
γ_I	coeff lagged interest rate in monetary policy rule	0.242
δ	habit formation	0.258
ρ_ξ	AR preference	0.915
ρ_θ	AR labor supply	0.708
ρ_z	AR transitory productivity	0.700
ρ_g	AR permanent productivity	0.221
ρ_{gm}	AR permanent trend inflation	0.248
ρ_{zm}	AR transitory interest rate	0.306
ρ_{zm2}	AR transitory trend inflation	0.796
σ_ξ	std. preference	0.0287
σ_θ	std. labor supply	0.00164
σ_z	std. transitory productivity	0.00122
σ_g	std. permanent productivity	0.00758
σ_{gm}	std. permanent trend inflation	0.000848
σ_{zm}	std. transitory interest rate	0.000832
σ_{zm2}	std. transitory trend inflation	0.00131
σ_1^{me}	std. measurement error$\triangle y_t$	4.46e-06
σ_2^{me}	std. measurement errorr_t	4.55e-06
σ_3^{me}	std. measurement error $\triangle i_t$	1.74e-07

B Justiniano, Primiceri, and Tambalotti (2011)

Table B1: Parameter values, JPT (2011) model

	parameter	posterior median
α	capital share	0.169
ι_p	price indexation	0.113
ι_w	wage indexation	0.102
h	consumption habit	0.864
λ_p	SS mark-up goods prices	0.177
λ_w	SS mark-up wages	0.166
ν	inverse frisch elasticity	5.162
ξ_p	Calvo prices	0.783
ξ_w	Calvo wges	0.773
χ	Elasticity capital utilization cost	5.491
$S^{'}$	Investment adjustment costs	3.017
ϕ_π	Taylor rule inflation	1.735
ϕ_Y	Taylor rule output	0.059
ρ_R	Taylor rule smoothing	0.863
ρ_z	AR neutral technology growth	0.286
ρ_g	AR government spending	0.990
ρ_ν	AR IST growth	0.148
ρ_p	AR price mark-up	0.978
ρ_w	AR wage mark-up	0.968
ρ_b	intertemporal preference	0.583
θ_p	MA price mark-up	0.793
θ_w	MA wage mark-up	0.990
ϕ_{dy}	Taylor rule output growth	0.199
ρ_μ	AR MEI	0.807
σ_{mp}	std. monetary policy	0.216
σ_z	std. neutral technology growth	0.943
σ_g	std. government spending	0.362
σ_ν	std. IST growth	0.634
σ_p	std. price mark-up	0.222
σ_w	std. wage mark-up	0.310
σ_b	std. intertemporal preference	0.038
σ_μ	std. MEI	5.691

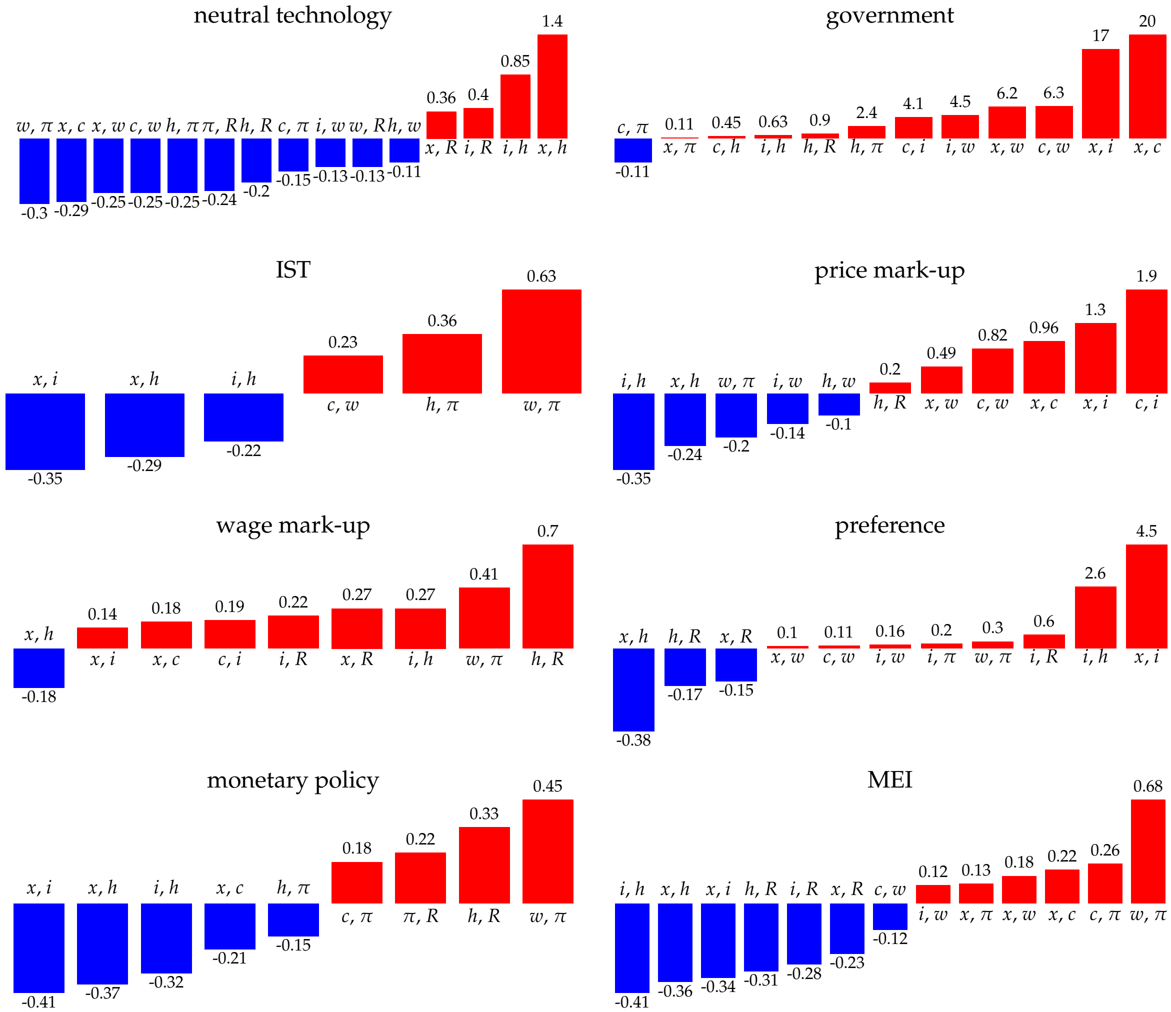

Figure B1: Largest unconditional pairwise information complementarities, all frequencies.

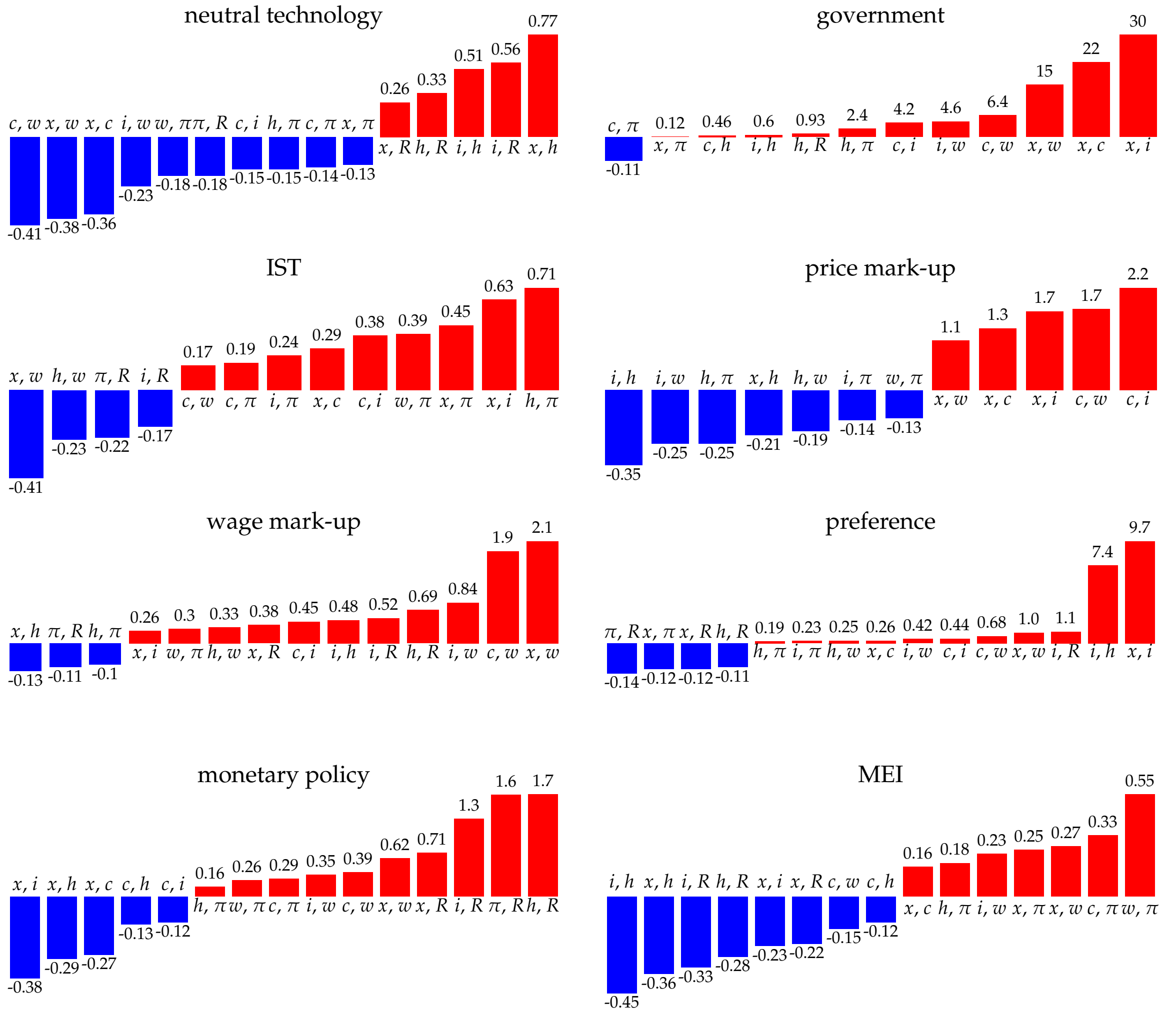

Figure B2: Largest unconditional pairwise information complementarities, low frequencies.

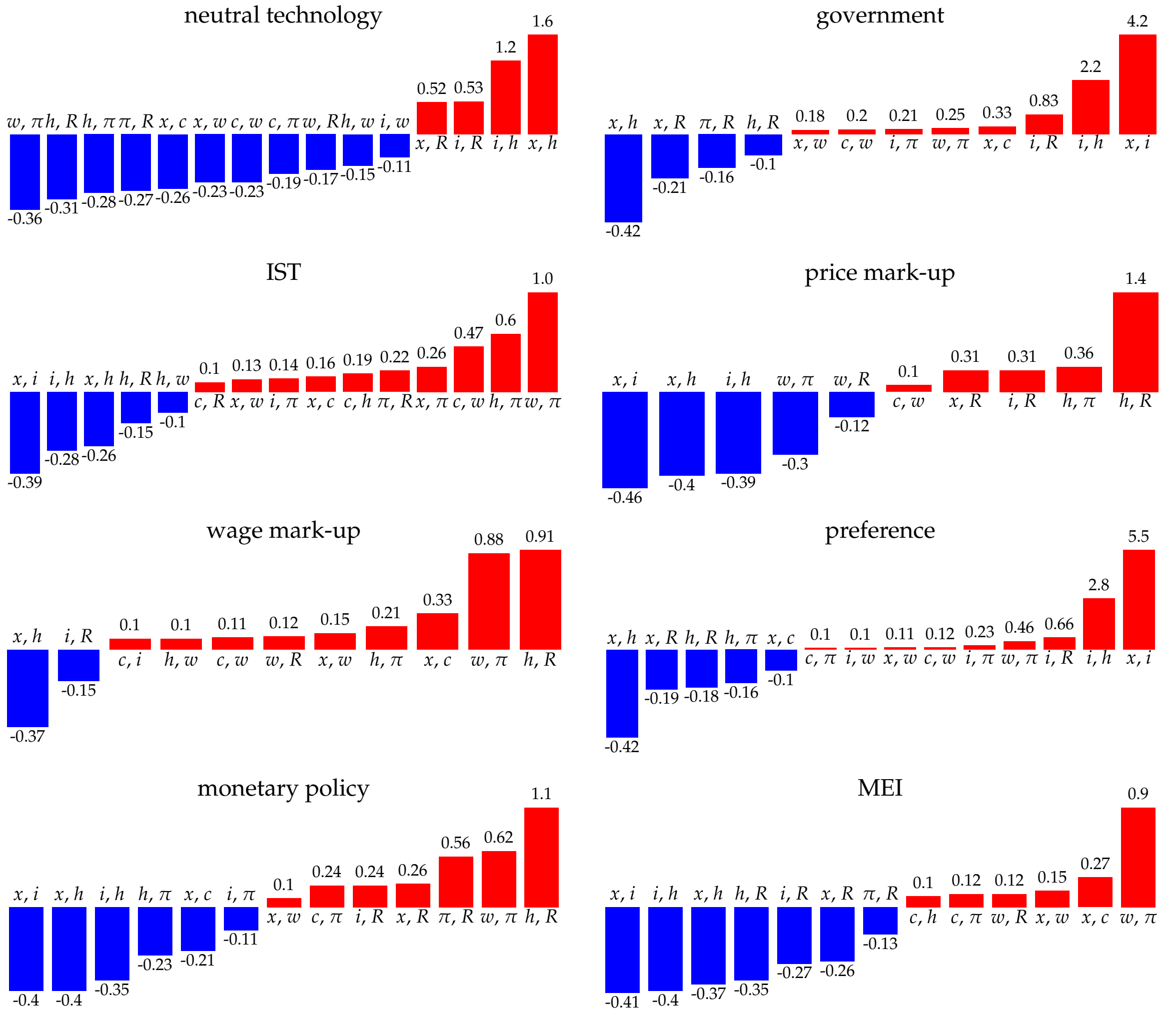

Figure B3: Largest unconditional pairwise information complementarities, BC frequencies.

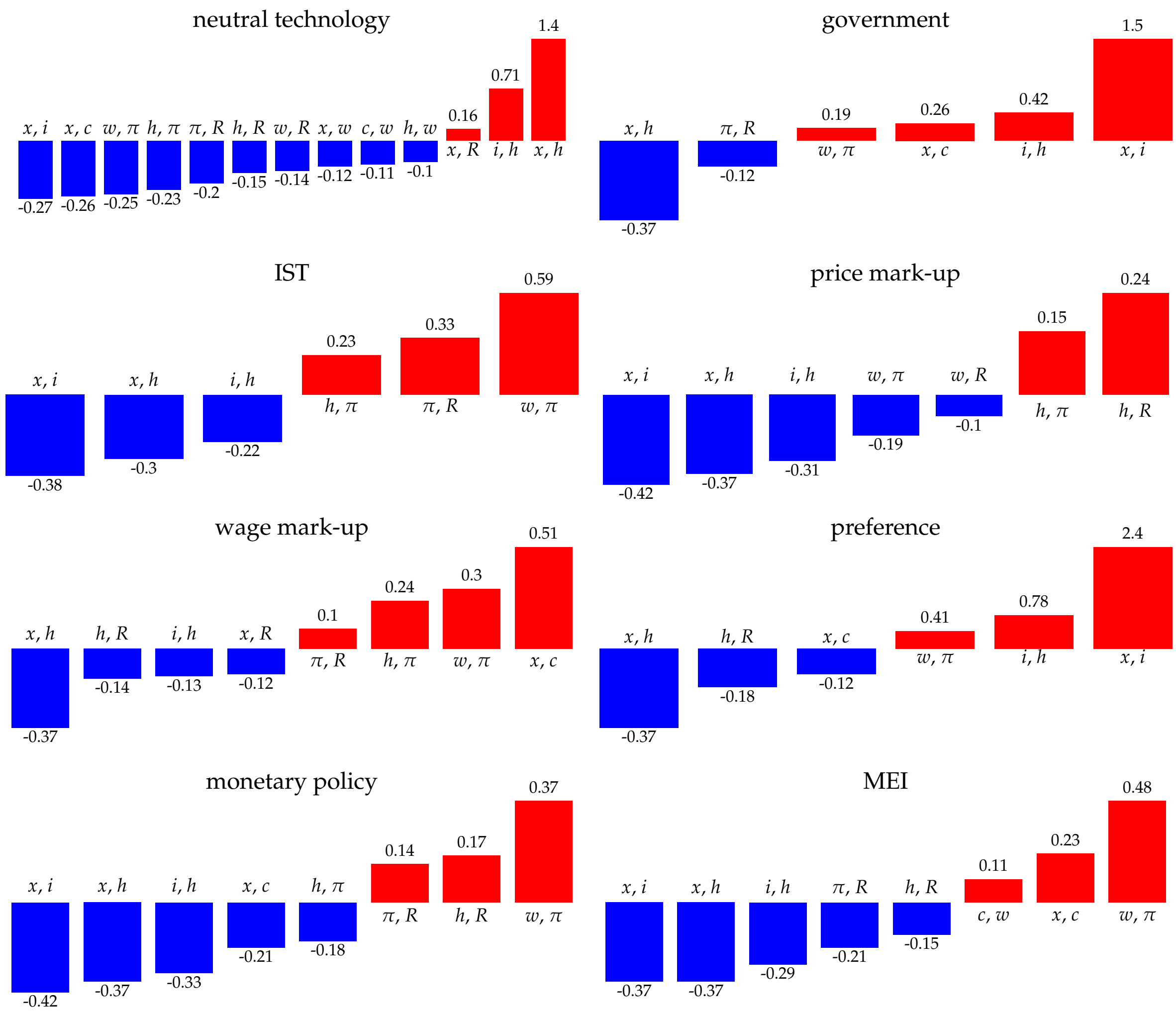

Figure B4: Largest unconditional pairwise information complementarities, high frequencies.

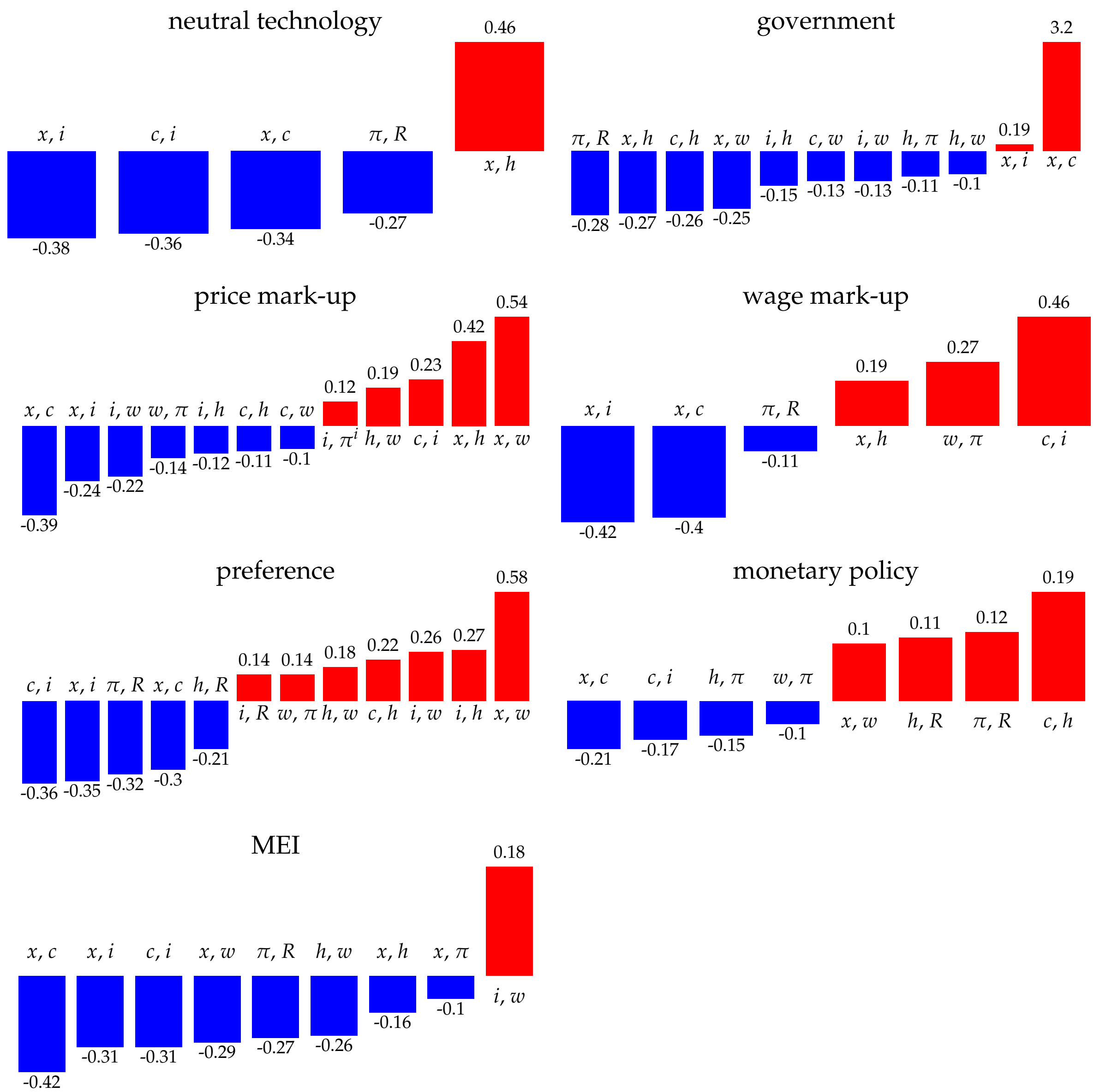

Figure B5: Largest conditional pairwise information complementarities, full spectrum.

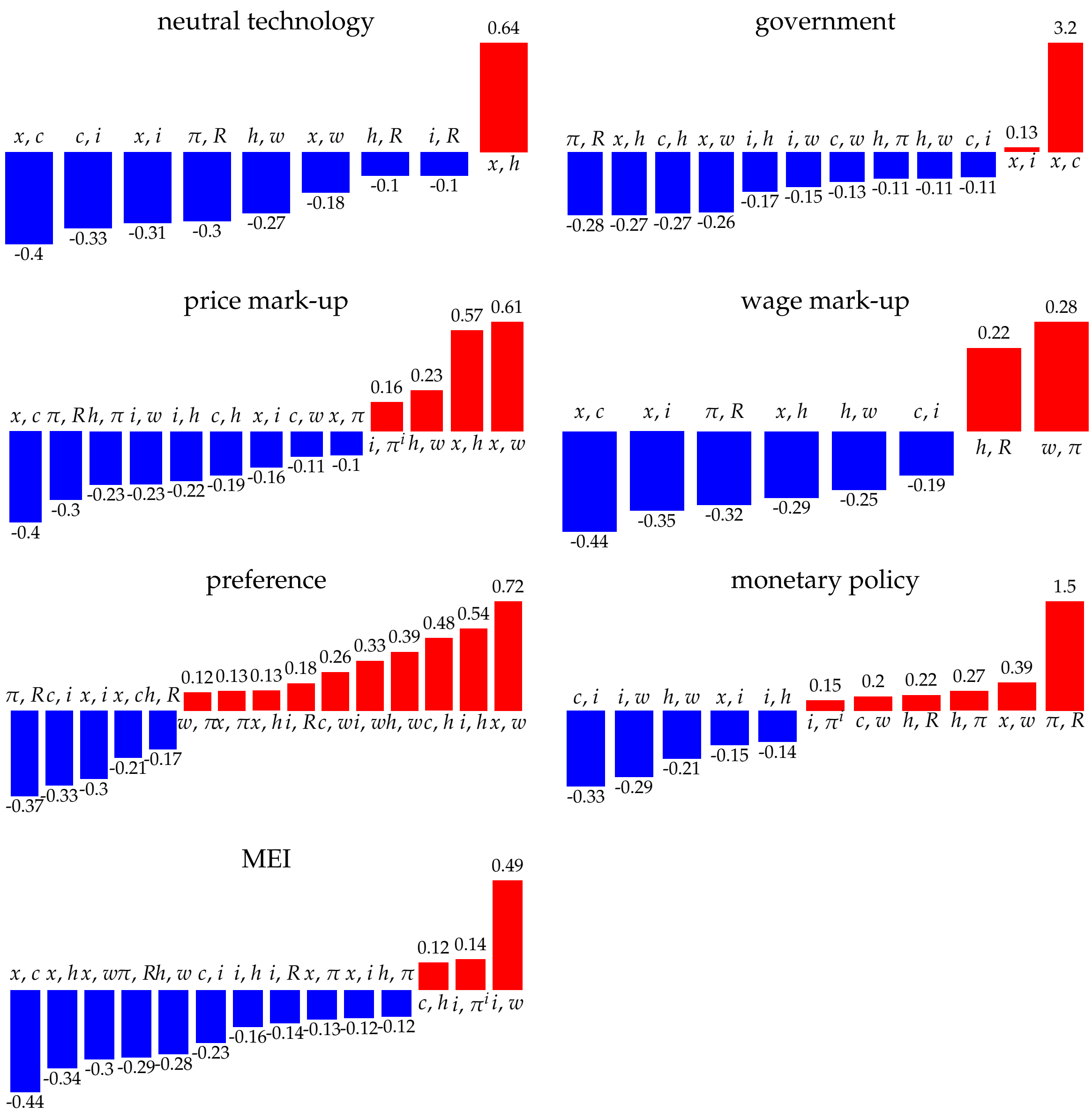

Figure B6: Largest conditional pairwise information complementarities, low spectrum.

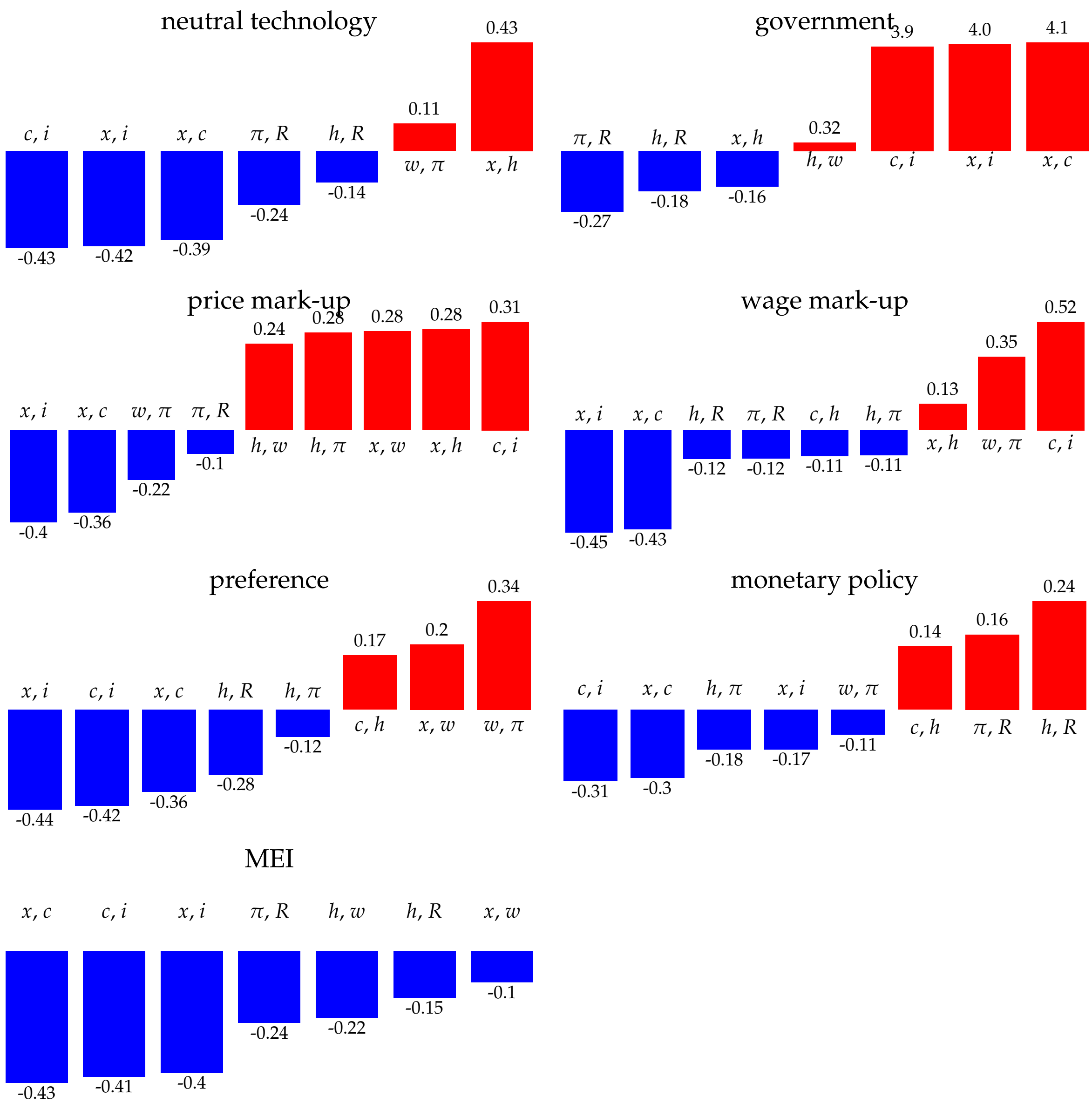

Figure B7: Largest conditional pairwise information complementarities, BC spectrum.

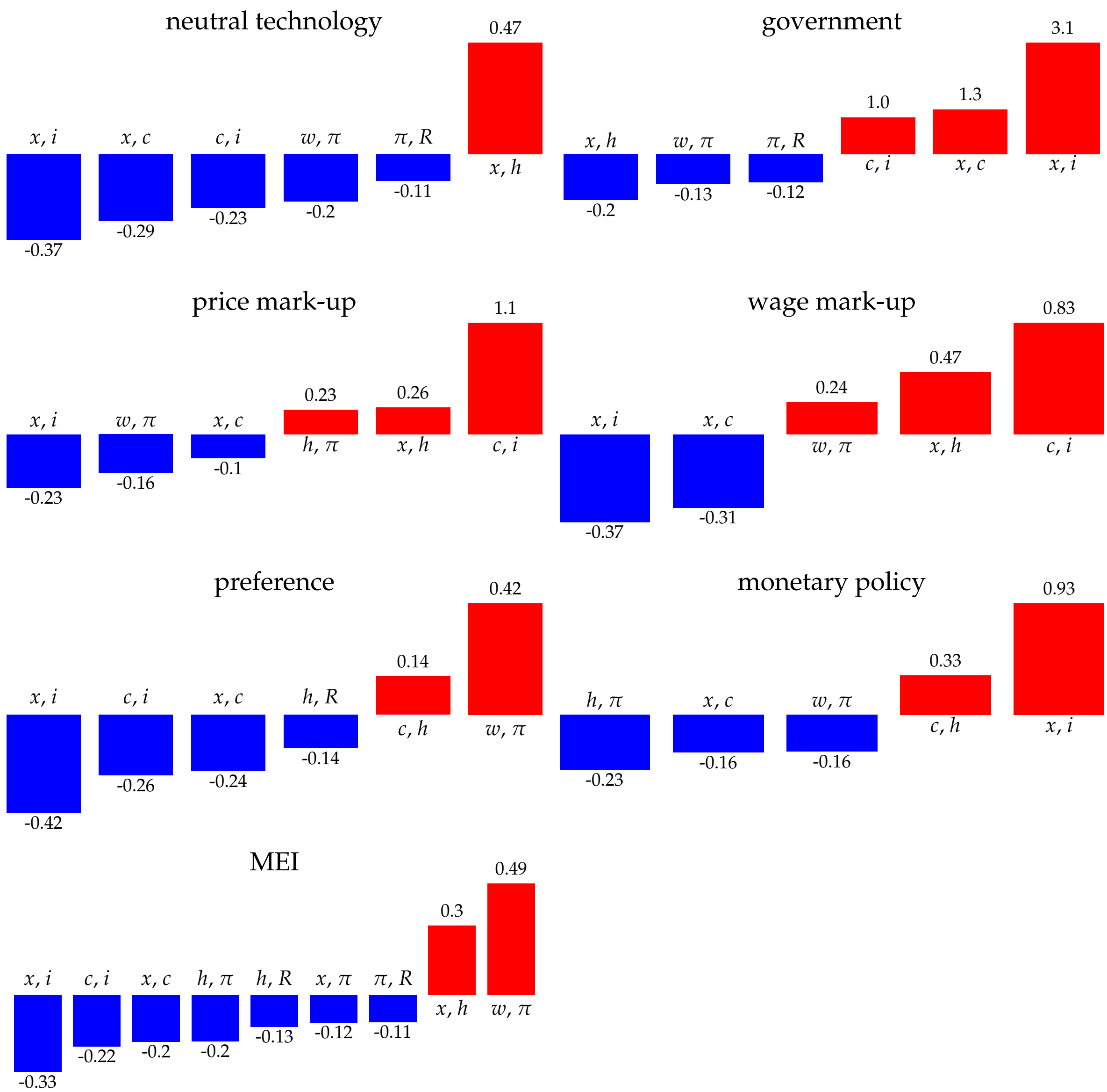

Figure B8: Largest conditional pairwise information complementarities, high spectrum.