Supplement to
Testing Dependence Among Serially Correlated Multi-category Variables*

M. Hashem Pesaran
Cambridge University
Allan Timmermann
University of California, San Diego
May 22, 2008

1 Overview

In this supplemental material, Section 2 provides a proof of the relationship between Pearson's Chi-squared test and the static trace test based on canonical correlations. Section 3 derives the test for independence with ordered alternatives introduced in Section 3.3 of the paper, while Section 4 introduces an iterated test method which is an alternative to the dynamically augmented reduced rank regression. Section 5 reports additional simulation results, and Section 6 reports the outcome of an empirical application to macroeconomic data on forecasts of economic recessions.

2 Relationship to Pearson's Chi-squared Test

In the special case where the realizations of X and Y are serially independent, the standard approach to testing independence of categorized variables is to arrange the outcomes in the form of a contingency table and then compute an appropriate test statistic from the individual cell frequencies. We show below that there is an exact relationship between the trace statistic and Pearson's contingency table χ^2-test of independence. To this end we first introduce some new notations.

When testing the independence of y_{it}, and x_{jt} for $i = 1, 2, ..., m_y$, $j = 1, ..., m_x$ the appropriate contingency table is given by

Y, X	1	2	$\cdots$	m_x	
1	n_{11}	n_{12}	$\cdots$	n_{1m_x}	$n_{1.}$
2	n_{21}	n_{22}	$\cdots$	n_{2m_x}	$n_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
m_y	n_{m_y1}	n_{m_y2}	$\cdots$	$n_{m_ym_x}$	$n_{m_y.}$
	$n_{.1}$	$n_{.2}$	$\cdots$	$n_{.m_x}$	n

Here n_{ij} is the frequency of the joint occurrence of y_{it} and x_{jt}, namely $n_{ij} = \sum_{t=1}^{T} y_{it}x_{jt}$, and

$$n_{i.} = \sum_{j=1}^{m_x} n_{ij} = \sum_{t=1}^{T} y_{it},\ n_{.j} = \sum_{i=1}^{m_y} n_{ij} = \sum_{t=1}^{T} x_{jt},\ n = \sum_{i=1}^{m_y} n_{i.} = \sum_{j=1}^{m_x} n_{.j} = T.$$

The familiar Pearson Chi-square test of independence is given by

$$\chi^2 = T\left(\sum_{i=1}^{m_y}\sum_{j=1}^{m_x}\frac{n_{ij}^2}{n_{i.}n_{.j}} - 1\right). \tag{1}$$

Proposition 1 *The Pearson Chi-square test for independence of data arranged in an $m_x \times m_y$ contingency table ($m_x \leq m_y$) is identical to a trace test based on the canonical correlations*

$$\sum_{i=1}^{m_y}\sum_{j=1}^{m_x}\frac{n_{ij}^2}{n_{i.}n_{.j}} - 1 = \sum_{i=1}^{m_x-1}\hat{\rho}_i^2,$$

where $\hat{\rho}_i$ is the sample estimate of the i^{th} canonical correlation between $\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_{m_y-1})$ and $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_{m_x-1})$, $\mathbf{y}_i = (y_{i1}, y_{i2}, ..., y_{iT})'$ and $\mathbf{x}_i = (x_{i1}, x_{i2}, ..., x_{iT})'$. Furthermore,

$$\sum_{i=1}^{m_x-1}\hat{\rho}_i^2 = Tr\left[\left(\mathbf{Y}'\mathbf{M}_\tau\mathbf{Y}\right)^{-1}\left(\mathbf{Y}'\mathbf{M}_\tau\mathbf{X}\right)\left(\mathbf{X}'\mathbf{M}_\tau\mathbf{X}\right)^{-1}\left(\mathbf{X}'\mathbf{M}_\tau\mathbf{Y}\right)\right],$$

where $\mathbf{M}_\tau = \mathbf{I}_T - T^{-1}\boldsymbol{\tau}_T\boldsymbol{\tau}_T'$, $\boldsymbol{\tau}_T = (1, 1, ..., 1)'$.

Proof. To establish the result we first write the various moment matrices in the trace expression in terms of n_{ij} notations. Since the events in the various categories are mutually exclusive, $\mathbf{Y}'\mathbf{Y}$ and $\mathbf{X}'\mathbf{X}$ will be diagonal matrices with their i^{th} diagonal elements given by $n_{i.} = \sum_{t=1}^{T} y_{it}$ and $n_{.j} = \sum_{t=1}^{T} x_{jt}$. Also the (i, j) element of $\mathbf{Y}'\mathbf{X}$ is n_{ij}, and, with $\boldsymbol{\tau}_{m_y-1}$ a $(m_y - 1) \times 1$ vector of ones

$$\boldsymbol{\tau}_T'\mathbf{Y} = \mathbf{h}_y' = (n_{1.}, n_{2.}, ..., n_{m_y-1.})', \ \ \boldsymbol{\tau}_T'\mathbf{X} = \mathbf{h}_x' = (n_{.1}, n_{.2}, ..., n_{.m_x-1})',$$

$$\left(\mathbf{Y}'\mathbf{Y}\right)^{-1}\mathbf{h}_y = \boldsymbol{\tau}_{m_y-1}, \ \ \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{h}_x = \boldsymbol{\tau}_{m_x-1},$$

hence $\mathbf{Y}'\mathbf{M}_\tau\mathbf{Y} = \mathbf{Y}'\mathbf{Y} - T^{-1}\mathbf{h}_y\mathbf{h}_y'$, $\mathbf{Y}'\mathbf{M}_\tau\mathbf{X} = \mathbf{Y}'\mathbf{X} - T^{-1}\mathbf{h}_y\mathbf{h}_x'$.

But $$\left(\mathbf{Y}'\mathbf{M}_\tau\mathbf{Y}\right)^{-1} = \left(\mathbf{Y}'\mathbf{Y}\right)^{-1} + \frac{T^{-1}\left(\mathbf{Y}'\mathbf{Y}\right)^{-1}\mathbf{h}_y\mathbf{h}_y'\left(\mathbf{Y}'\mathbf{Y}\right)^{-1}}{1 - T^{-1}\mathbf{h}_y'\left(\mathbf{Y}'\mathbf{Y}\right)^{-1}\mathbf{h}_y}.$$

Noting that $\mathbf{h}_y'\left(\mathbf{Y}'\mathbf{Y}\right)^{-1}\mathbf{h}_y = \sum_{i=1}^{m_y-1} n_{i.} = T - n_{m_y.}$ we have

$$\left(\mathbf{Y}'\mathbf{M}_\tau\mathbf{Y}\right)^{-1} = \left(\mathbf{Y}'\mathbf{Y}\right)^{-1} + \frac{\boldsymbol{\tau}_{m_y-1}\boldsymbol{\tau}_{m_y-1}'}{n_{m_y.}}, \left(\mathbf{X}'\mathbf{M}_\tau\mathbf{X}\right)^{-1} = \left(\mathbf{X}'\mathbf{X}\right)^{-1} + \frac{\boldsymbol{\tau}_{m_x-1}\boldsymbol{\tau}_{m_x-1}'}{n_{.m_x}}.$$

$$\begin{aligned}
\text{Therefore } \left(\mathbf{Y}'\mathbf{M}_\tau\mathbf{Y}\right)^{-1}\mathbf{Y}'\mathbf{M}_\tau\mathbf{X} &= \left[\left(\mathbf{Y}'\mathbf{Y}\right)^{-1} + \frac{\boldsymbol{\tau}_{m_y-1}\boldsymbol{\tau}_{m_y-1}'}{n_{m_y.}}\right]\left(\mathbf{Y}'\mathbf{X} - T^{-1}\mathbf{h}_y\mathbf{h}_x'\right) \\
&= \left(\mathbf{Y}'\mathbf{Y}\right)^{-1}\mathbf{Y}'\mathbf{X} - T^{-1}\left(\mathbf{Y}'\mathbf{Y}\right)^{-1}\mathbf{h}_y\mathbf{h}_x' \\
&\quad + \frac{\boldsymbol{\tau}_{m_y-1}\boldsymbol{\tau}_{m_y-1}'\mathbf{Y}'\mathbf{X}}{n_{m_y.}} - T^{-1}\frac{\boldsymbol{\tau}_{m_y-1}\boldsymbol{\tau}_{m_y-1}'\mathbf{h}_y\mathbf{h}_x'}{n_{m_y.}}.
\end{aligned}$$

Using these results, after some algebra we have

$$\begin{aligned}
\left(\mathbf{Y}'\mathbf{M}_\tau\mathbf{Y}\right)^{-1}\mathbf{Y}'\mathbf{M}_\tau\mathbf{X} &= \left(\mathbf{Y}'\mathbf{Y}\right)^{-\mathbf{1}}\mathbf{Y}'\mathbf{X}-\frac{\boldsymbol{\tau}_{m_x-1}\mathbf{q}_x'}{n_{m_y.}}, \\
\left(\mathbf{X}'\mathbf{M}_\tau\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{M}_\tau Y &= \left(\mathbf{X}'\mathbf{X}\right)^{-\mathbf{1}}\mathbf{X}'\mathbf{Y}-\frac{\boldsymbol{\tau}_{m_x-1}\mathbf{q}_y'}{n_{.m_x}},
\end{aligned}$$

where $\mathbf{q}_x = (n_{m_y1}, n_{m_y2}, ..., n_{m_y,m_x-1})'$ and $\mathbf{q}_y = (n_{1m_x}, n_{2m_x}, ..., n_{m_y-1,m_x})'$. Hence

$$\begin{aligned}
& Tr\left[\left(\mathbf{Y}'\mathbf{M}_\tau\mathbf{Y}\right)^{-1}\left(\mathbf{Y}'\mathbf{M}_\tau\mathbf{X}\right)\left(\mathbf{X}'\mathbf{M}_\tau\mathbf{X}\right)^{-1}\left(\mathbf{X}'\mathbf{M}_\tau\mathbf{Y}\right)\right] \\
= \; & Tr\left[\left(\mathbf{Y}'\mathbf{Y}\right)^{-1}\left(\mathbf{Y}'\mathbf{X}\right)\left(\mathbf{X}'\mathbf{X}\right)^{-1}\left(\mathbf{X}'\mathbf{Y}\right)\right] - \frac{\mathbf{q}_y'\left(\mathbf{Y}'\mathbf{Y}\right)^{-\mathbf{1}}\mathbf{Y}'\mathbf{X}\boldsymbol{\tau}_{m_x-1}}{n_{.m_x}} \\
& -\frac{\mathbf{q}_x'\left(\mathbf{X}'\mathbf{X}\right)^{-\mathbf{1}}\mathbf{X}'\mathbf{Y}\boldsymbol{\tau}_{m_y-1}}{n_{m_y.}} + \frac{\left(\mathbf{q}_x'\boldsymbol{\tau}_{m_x-1}\right)\left(\mathbf{q}_y'\boldsymbol{\tau}_{m_y-1}\right)}{n_{m_y.}n_{.m_x}}.
\end{aligned}$$

Consider now the various terms in this expression. First, since $\mathbf{Y}'\mathbf{Y}$ and $\mathbf{X}'\mathbf{X}$ are diagonal matrices and the typical element of $\mathbf{Y}'\mathbf{X}$ is n_{ij}, it readily follows that

$$Tr\left[\left(\mathbf{Y}'\mathbf{Y}\right)^{-1}\left(\mathbf{Y}'\mathbf{X}\right)\left(\mathbf{X}'\mathbf{X}\right)^{-1}\left(\mathbf{X}'\mathbf{Y}\right)\right] = \sum_{i=1}^{m_y-1}\sum_{j=1}^{m_x-1}\frac{n_{ij}^2}{n_{i.}n_{.j}}.$$

Also, $\mathbf{Y}'\mathbf{X}\boldsymbol{\tau}_{m_x-1} = \mathbf{h}_y - \mathbf{q}_y,\ \mathbf{X}'\mathbf{Y}\boldsymbol{\tau}_{m_y-1} = \mathbf{h}_x - \mathbf{q}_x,$

$$\begin{aligned}
\mathbf{q}_y'\left(\mathbf{Y}'\mathbf{Y}\right)^{-\mathbf{1}}\mathbf{Y}'\mathbf{X}\boldsymbol{\tau}_{m_x-1} &= \mathbf{q}_y'\boldsymbol{\tau}_{m_y-1} - \mathbf{q}_y'\left(\mathbf{Y}'\mathbf{Y}\right)^{-\mathbf{1}}\mathbf{q}_y \\
\mathbf{q}_x'\left(\mathbf{X}'\mathbf{X}\right)^{-\mathbf{1}}\mathbf{X}'\mathbf{Y}\boldsymbol{\tau}_{m_y-1} &= \mathbf{q}_x'\boldsymbol{\tau}_{m_x-1} - \mathbf{q}_x'\left(\mathbf{X}'\mathbf{X}\right)^{-\mathbf{1}}\mathbf{q}_x.
\end{aligned}$$

$$\begin{aligned}
\text{Finally } \mathbf{q}_x'\boldsymbol{\tau}_{m_x-1} &= \sum_{j=1}^{m_x-1} n_{m_yj} = n_{m_y.} - n_{m_ym_x}, \\
\mathbf{q}_y'\boldsymbol{\tau}_{m_y-1} &= \sum_{i=1}^{m_y-1} n_{im_x} = n_{.m_x} - n_{m_ym_x}, \\
\mathbf{q}_y'\left(\mathbf{Y}'\mathbf{Y}\right)^{-\mathbf{1}}\mathbf{q}_y &= \sum_{i=1}^{m_y-1}\frac{n_{im_x}^2}{n_{i.}},\ \ \mathbf{q}_x'\left(\mathbf{X}'\mathbf{X}\right)^{-\mathbf{1}}\mathbf{q}_x = \sum_{j=1}^{m_x-1}\frac{n_{m_yj}^2}{n_{.j}}.
\end{aligned}$$

$$\text{It follows that } \frac{\mathbf{q}_y'\left(\mathbf{Y}'\mathbf{Y}\right)^{-\mathbf{1}}\mathbf{Y}'\mathbf{X}\boldsymbol{\tau}_{m_x-1}}{n_{.m_x}} = 1 - \frac{n_{m_ym_x}}{n_{.m_x}} - \sum_{i=1}^{m_y-1}\frac{n_{im_x}^2}{n_{i.}n_{.m_x}},$$

$$\frac{\mathbf{q}_x'\left(\mathbf{X}'\mathbf{X}\right)^{-\mathbf{1}}\mathbf{X}'\mathbf{Y}\boldsymbol{\tau}_{m_y-1}}{n_{m_y.}} = 1 - \frac{n_{m_ym_x}}{n_{m_y.}} - \sum_{j=1}^{m_x-1}\frac{n_{m_yj}^2}{n_{m_y.}n_{.j}},$$

$$\frac{\left(\mathbf{q}_x'\boldsymbol{\tau}_{m_x-1}\right)\left(\mathbf{q}_y'\boldsymbol{\tau}_{m_y-1}\right)}{n_{m_y.}n_{.m_x}} = 1 - \frac{n_{m_y m_x}}{n_{m_y.}} - \frac{n_{m_y m_x}}{n_{.m_x}} + \frac{n_{m_y m_x}^2}{n_{m_y.}n_{.m_x}}.$$

Hence $$Tr\left[\left(\mathbf{Y}'\mathbf{M}_\tau\mathbf{Y}\right)^{-1}\left(\mathbf{Y}'\mathbf{M}_\tau\mathbf{X}\right)\left(\mathbf{X}'\mathbf{M}_\tau\mathbf{X}\right)^{-1}\left(\mathbf{X}'\mathbf{M}_\tau\mathbf{Y}\right)\right]$$

$$= \sum_{i=1}^{m_y-1}\sum_{j=1}^{m_x-1}\frac{n_{ij}^2}{n_{i.}n_{.j}} - 1 + \sum_{i=1}^{m_y-1}\frac{n_{im_x}^2}{n_{i.}n_{.m_x}} + \sum_{j=1}^{m_x-1}\frac{n_{m_y j}^2}{n_{m_y.}n_{.j}} + \frac{n_{m_y m_x}^2}{n_{m_y.}n_{.m_x}}$$

$$= \sum_{i=1}^{m_y}\sum_{j=1}^{m_x}\frac{n_{ij}^2}{n_{i.}n_{.j}} - 1.$$

as required. ■

3 Tests of Independence under Ordered Alternatives

This section derives the test of independence under ordered alternatives described in Section 3.3 of the paper.

Let π_{ij} be the joint probability that y and x fall in the i^{th} and j^{th} categories respectively, where $i = 1, 2, ..., m_y$ and $j = 1, 2, ..., m_x$. The y-categories are specified in terms of the thresholds, $a_0 < a_1 < ... < a_{m_y-1} < a_{m_y}$ and the x-categories in terms of $b_0 < b_1 < ... < b_{m_x-1} < b_{m_x}$. In both cases $a_0 = b_0 = -\infty$, and $a_{m_y} = b_{m_x} = +\infty$. The associated observed frequency for the joint occurrence of y and x in their i^{th} and j^{th} categories is denoted by n_{ij} and the relative frequencies by $\hat{\pi}_{ij} = n_{ij}/T$.

Under joint normality of the underlying latent variables and random draws (serially independent observations) the log-likelihood function of ordinal measure is given by

$$\ell(\mathbf{a}, \mathbf{b}, \rho) = \sum_{i=1}^{m_y}\sum_{j=1}^{m_x} n_{ij}\ln(\pi_{ij}), \tag{2}$$

where

$$\pi_{ij} = \Phi_2(a_i, b_j; \rho) - \Phi_2(a_{i-1}, b_j; \rho) - \Phi_2(a_i, b_{j-1}; \rho) + \Phi_2(a_{i-1}, b_{j-1}; \rho), \tag{3}$$

and $\Phi_2(a, b; \rho)$ is the distribution function of the bivariate standard normal distribution with ρ as its correlation coefficient.

As shown in Ronning and Kukuk (1996), under $\rho = 0$ the information matrix of the above (doubly) ordered probit model is block diagonal for ρ, $\mathbf{a}$ and $\mathbf{b}$, and hence the score (or Lagrangian Multiplier) statistic for testing the null of $\rho = 0$ is given by

$$S_\rho = \frac{\left[\frac{\partial\ell(\rho,\mathbf{a},\mathbf{b})}{\partial\rho}\mid_{\rho=0}\right]^2}{E\left[-\frac{\partial^2\ell(\rho,\mathbf{a},\mathbf{b})}{\partial^2\rho}\mid_{\rho=0}\right]}.$$

But it is easily seen that $\frac{\partial \ell(\rho,\mathbf{a},\mathbf{b})}{\partial \rho}\mid_{\rho=0}$, or $\frac{\partial \ell(0,\mathbf{a},\mathbf{b})}{\partial \rho}$ for short, is given by

$$\frac{\partial \ell(\mathbf{a},\mathbf{b},0)}{\partial \rho} = \sum_{i=1}^{m_y}\sum_{j=1}^{m_x} \frac{n_{ij}}{\pi_{ij}(0)} \frac{\partial \pi_{ij}(0)}{\partial \rho},$$

where $\pi_{ij}(0)$ stands for π_{ij} evaluated at $\rho = 0$. Using equations (14) and (15) in Ronning and Kukuk, and noting that $\Phi_2(a, b; 0) = \Phi(a)\Phi(b)$, where $\Phi(a)$ is the distribution function of a standard normal variate, we have

$$\pi_{ij}(0) = \pi_{i.}\pi_{.j},$$

$$\begin{aligned} \pi_{i.} &= \Phi(a_i) - \Phi(a_{i-1}), \\ \pi_{.j} &= \Phi(b_j) - \Phi(b_{j-1}), \end{aligned}$$

where $\pi_{i.}$ is the marginal probability that y falls in category i, and $\pi_{.j}$ is the marginal probability that x falls in category j. Also (see Olsson (1979))

$$\frac{\partial \Phi_2(a_i, b_j; \rho)}{\partial \rho} = \phi_2(u, v, \rho),$$

$$\phi_2(u, v, \rho) = (2\pi)^{-1}(1-\rho^2)^{-1/2} \exp\left[\frac{-1}{2(1-\rho^2)}\left(u^2 + v^2 - 2\rho uv\right)\right]$$

then

$$\frac{\partial \pi_{ij}(\rho)}{\partial \rho} = \phi_2(a_i, b_j; \rho) - \phi_2(a_{i-1}, b_j; \rho) - \phi_2(a_i, b_{j-1}; \rho) + \phi_2(a_{i-1}, b_{j-1}; \rho),$$

and

$$\frac{\partial \pi_{ij}(0)}{\partial \rho} = \left[\phi(a_i) - \phi(a_{i-1})\right]\left[\phi(b_i) - \phi(b_{i-1})\right].$$

Hence

$$\frac{\partial \ell(\mathbf{a},\mathbf{b},0)}{\partial \rho} = \sum_{i=1}^{m_y}\sum_{j=1}^{m_x} \frac{n_{ij}\left[\phi(a_i) - \phi(a_{i-1})\right]\left[\phi(b_j) - \phi(b_{j-1})\right]}{\left[\Phi(a_i) - \Phi(a_{i-1})\right]\left[\Phi(b_j) - \Phi(b_{j-1})\right]}.$$

For $\frac{\partial^2 \ell(\rho,\mathbf{a},\mathbf{b})}{\partial^2 \rho}$ we note that

$$\frac{\partial^2 \ln(\pi_{ij})}{\partial^2 \rho} = \frac{-1}{\pi_{ij}^2}\left(\frac{\partial \pi_{ij}}{\partial \rho}\right)^2 + \frac{1}{\pi_{ij}}\frac{\partial^2 \pi_{ij}}{\partial \rho^2},$$

and

$$\frac{\partial \phi_2(a, b; \rho)}{\partial \rho} = \frac{\phi_2(a, b; \rho)}{1-\rho^2}\left[ab + \rho\left(1 + a^2 + b^2 - 2\rho ab\right)\right].$$

Hence

$$\frac{\partial^2 \ln(\pi_{ij}(0))}{\partial^2 \rho} = \frac{-1}{\pi_{ij}^2(0)}\left(\frac{\partial \pi_{ij}(0)}{\partial \rho}\right)^2 + \frac{1}{\pi_{ij}(0)}\frac{\partial^2 \pi_{ij}(0)}{\partial \rho^2},$$

where

$$\frac{1}{\pi_{ij}(0)}\frac{\partial^2\pi_{ij}(0)}{\partial\rho^2}=\frac{\left[a_i\phi(a_i)-a_{i-1}\phi(a_{i-1})\right]\left[b_j\phi(b_j)-b_{j-1}\phi(b_{j-1})\right]}{\left[\Phi(a_i)-\Phi(a_{i-1})\right]\left[\Phi(b_j)-\Phi(b_{j-1})\right]}.$$

Putting the various terms together

$$\begin{aligned} E\left[-\frac{\partial^2\ell(\mathbf{a},\mathbf{b},\rho)}{\partial^2\rho}\mid_{\rho=0}\right] &= \sum_{i=1}^{m_y}\sum_{j=1}^{m_x}n_{ij}\frac{\left[\phi(a_i)-\phi(a_{i-1})\right]^2\left[\phi(b_j)-\phi(b_{j-1})\right]^2}{\left[\Phi(a_i)-\Phi(a_{i-1})\right]^2\left[\Phi(b_j)-\Phi(b_{j-1})\right]^2}- \\ & \sum_{i=1}^{m_y}\sum_{j=1}^{m_x}n_{ij}\frac{\left[a_i\phi(a_i)-a_{i-1}\phi(a_{i-1})\right]\left[b_j\phi(b_j)-b_{j-1}\phi(b_{j-1})\right]}{\left[\Phi(a_i)-\Phi(a_{i-1})\right]\left[\Phi(b_j)-\Phi(b_{j-1})\right]}. \end{aligned}$$

More compactly

$$\begin{aligned} E\left[-\frac{\partial^2\ell(\mathbf{a},\mathbf{b},\rho)}{\partial^2\rho}\mid_{\rho=0}\right] &= \sum_{i=1}^{m_y}\sum_{j=1}^{m_x}n_{ij}\frac{\left[\phi(a_i)-\phi(a_{i-1})\right]^2\left[\phi(b_j)-\phi(b_{j-1})\right]^2}{\pi_{i.}^2\pi_{.j}^2}- \\ & \sum_{i=1}^{m_y}\sum_{j=1}^{m_x}n_{ij}\frac{\left[a_i\phi(a_i)-a_{i-1}\phi(a_{i-1})\right]\left[b_j\phi(b_j)-b_{j-1}\phi(b_{j-1})\right]}{\pi_{i.}\pi_{.j}}. \end{aligned}$$

Finally, under $\rho=0$, the maximum likelihood estimates of the thresholds can be obtained from equation (16) and (17) of Ronning and Kukuk (1996) (or equations (13) and (14) of Olsson (1979)). Under $\rho=0$, equation (16) of Ronning and Kukuk gives (recall that when $\rho=0$, $\pi_{kj}=\pi_{k.}\pi_{.j}$)

$$\begin{aligned} \phi(a_k)\sum_{j=1}^{m_x}\left(\frac{n_{kj}}{\pi_{kj}}-\frac{n_{k+1,j}}{\pi_{k+1,j}}\right)\pi_{.j} &= \phi(a_k)\sum_{j=1}^{m_x}\left(\frac{n_{kj}}{\pi_{k.}\pi_{.j}}-\frac{n_{k+1,j}}{\pi_{k+1,.}\pi_{.j}}\right)\pi_{.j}=0, \text{ for } k=1,2,...,m_y-1 \\ \text{or } \sum_{j=1}^{m_x}\left(\frac{n_{kj}}{\pi_{k.}}-\frac{n_{k+1,j}}{\pi_{k+1,.}}\right) &= 0 \text{ which yields } \frac{n_{k.}}{\pi_{k.}}=\frac{n_{k+1,.}}{\pi_{k+1,.}}, \text{ for } k=1,2,...,m_y-1. \end{aligned}$$

Similarly, under $\rho=0$, equation (17) of Ronning and Kukuk (1996) yields

$$\frac{n_{.h}}{\pi_{.h}}=\frac{n_{.,h+1}}{\pi_{.,h+1}},\quad for\ h=1,2...,m_x-1.$$

Dividing by T, we have

$$\begin{aligned} \frac{\hat{\pi}_{k.}}{\pi_{k.}} &= \frac{\hat{\pi}_{k+1,.}}{\pi_{k+1,.}} \\ \frac{\hat{\pi}_{.h}}{\pi_{.h}} &= \frac{\hat{\pi}_{.,h+1}}{\pi_{.,h+1}}, \end{aligned}$$

and then it is easily seen that under $\rho=0$ these relations yield equations (18) and (19) of Ronning and Kukuk (1996) - namely

$$\begin{aligned} \Phi(\hat{a}_i)-\Phi(\hat{a}_{i-1}) &= \hat{\pi}_{i.}=\sum_{j=1}^{m_x}n_{ij}/T \\ \Phi(\hat{b}_j)-\Phi(\hat{b}_{j-1}) &= \hat{\pi}_{.j}=\sum_{i=1}^{m_y}n_{ij}/T, \end{aligned}$$

with $\Phi(\hat{a}_0) = \Phi(\hat{b}_0) = 0$ and $\Phi(\hat{a}_{m_y}) = \Phi(\hat{b}_{m_x}) = 1$. Thus

$$\begin{aligned}\hat{a}_i &= \Phi^{-1}(\hat{\pi}_{1.} + \hat{\pi}_{2.} + ... + \hat{\pi}_{i.}),\ i = 1, 2, ..., m_y \\ \hat{b}_j &= \Phi^{-1}(\hat{\pi}_{.1} + \hat{\pi}_{.2} + ... + \hat{\pi}_{.j}),\ j = 1, 2, ..., m_x\end{aligned}$$

Thus the score test can be computed as

$$S_\rho = \frac{T\left[\sum_{i=1}^{m_y}\sum_{j=1}^{m_x}\left(\frac{\hat{\pi}_{ij}}{\hat{\pi}_{i.}\hat{\pi}_{.j}}\right)[\phi(\hat{a}_i) - \phi(\hat{a}_{i-1})]\left[\phi(\hat{b}_j) - \phi(\hat{b}_{j-1})\right]\right]^2}{D}, \tag{4}$$

where

$$\begin{aligned}D &= \sum_{i=1}^{m_y}\sum_{j=1}^{m_x}\hat{\pi}_{ij}\frac{[\phi(\hat{a}_i) - \phi(\hat{a}_{i-1})]^2\left[\phi(\hat{b}_j) - \phi(\hat{b}_{j-1})\right]^2}{\hat{\pi}_{i.}^2\hat{\pi}_{.j}^2} - \\ &\quad \sum_{i=1}^{m_y}\sum_{j=1}^{m_x}\left(\frac{\hat{\pi}_{ij}}{\hat{\pi}_{i.}\hat{\pi}_{.j}}\right)[\hat{a}_i\phi(\hat{a}_i) - \hat{a}_{i-1}\phi(\hat{a}_{i-1})]\left[\hat{b}_j\phi(\hat{b}_j) - \hat{b}_{j-1}\phi(\hat{b}_{j-1})\right].\end{aligned} \tag{5}$$

4 Iterated Method

An alternative to dynamically augmenting the reduced rank regression is to adjust the moment matrices used in calculating the variance matrix of $\hat{\boldsymbol{\gamma}}$ to account for heteroskedasticity and autocorrelation in the errors. The F-statistic corresponding to equation (2) in the paper is then given by

$$F(\boldsymbol{\theta}) = \left(\frac{T - m_x}{m_x - 1}\right)\frac{\boldsymbol{\theta}'\mathbf{S}_{yx}\mathbf{H}_{xx}^{-1}(\boldsymbol{\theta})\mathbf{S}_{xy}\boldsymbol{\theta}}{\boldsymbol{\theta}'\left(\mathbf{S}_{yy} - \mathbf{S}_{yx}\mathbf{H}_{xx}^{-1}(\boldsymbol{\theta})\mathbf{S}_{xy}\right)\boldsymbol{\theta}},$$

where

$$\mathbf{H}_{xx}(\boldsymbol{\theta}) = \lim_{T\to\infty} E\left[\frac{1}{T}\sum_{s=1}^{T}\sum_{t=1}^{T}(\mathbf{x}_t - \bar{\mathbf{x}})(\mathbf{x}_s - \bar{\mathbf{x}})' u_t(\boldsymbol{\theta})u_s(\boldsymbol{\theta})\right],$$

$\bar{\mathbf{x}} = (\bar{x}_1, \bar{x}_2, ..., \bar{x}_{m_x-1})'$, $\bar{\mathbf{y}} = (\bar{y}_1, \bar{y}_2, ..., \bar{y}_{m_y-1})'$, and under $\boldsymbol{\gamma} = \mathbf{0}$, $u_t(\boldsymbol{\theta}) = \boldsymbol{\theta}'(\mathbf{y}_t - \bar{\mathbf{y}})$. Hence

$$\mathbf{H}_{xx}(\boldsymbol{\theta}) = \lim_{T\to\infty} E\left[\frac{1}{T}\sum_{s=1}^{T}\sum_{t=1}^{T}\boldsymbol{\theta}'(\mathbf{y}_t - \bar{\mathbf{y}})(\mathbf{x}_t - \bar{\mathbf{x}})(\mathbf{x}_s - \bar{\mathbf{x}})'(\mathbf{y}_s - \bar{\mathbf{y}})'\boldsymbol{\theta}\right],$$

can be viewed as the long run variance of $T^{-1/2}\sum_{t=1}^{T}\mathbf{d}_t(\boldsymbol{\theta})$, where $\mathbf{d}_t(\boldsymbol{\theta}) = \boldsymbol{\theta}'(\mathbf{y}_t - \bar{\mathbf{y}})(\mathbf{x}_t - \bar{\mathbf{x}})$. Since elements of $\mathbf{x}_t$ and $\mathbf{y}_t$ are bounded, $\mathbf{H}_{xx}(\boldsymbol{\theta})$ exists under general assumptions concerning the serial dependence and heteroskedasticity of the error terms, as set out in Newey and West (1987).

Unlike the serially independent case, the first order conditions for maximization of $LM(\boldsymbol{\theta})$ cannot get reduced to solving an eigenvalue problem. An asymptotically equivalent alternative (under $\boldsymbol{\gamma} = \mathbf{0}$) is to use a first-stage consistent estimate of

$\mathbf{H}_{xx}(\boldsymbol{\theta})$ that abstracts from the serial dependence of the errors. Such an estimator of $\boldsymbol{\theta}$ is given by equation (7) in the paper, and the first-stage estimate of $\mathbf{H}_{xx}(\boldsymbol{\theta})$ can be obtained by (using a Bartlett window)

$$\hat{\mathbf{H}}_{xx,h}(\hat{\boldsymbol{\theta}}_1) = \hat{\boldsymbol{\Gamma}}_0 + \sum_{j=1}^{h} \left(1 - \frac{j}{h+1}\right) (\hat{\boldsymbol{\Gamma}}_j + \hat{\boldsymbol{\Gamma}}_j'),$$

$$\hat{\boldsymbol{\Gamma}}_j = T^{-1} \sum_{t=j+1}^{T} \mathbf{d}_t(\hat{\boldsymbol{\theta}}_1)\mathbf{d}_{t-j}'(\hat{\boldsymbol{\theta}}_1), \text{ and } \mathbf{d}_t(\hat{\boldsymbol{\theta}}_1) = \hat{\boldsymbol{\theta}}_1'(\mathbf{y}_t - \bar{\mathbf{y}})(\mathbf{x}_t - \bar{\mathbf{x}})\,.$$

Using this estimator, one can solve the eigenvalue problem

$$\left(\mathbf{S}_{yx}\hat{\mathbf{H}}_{xx}^{-1}(\hat{\boldsymbol{\theta}}_1)\mathbf{S}_{xy} - \tilde{\rho}_1^2\mathbf{S}_{yy}\right)\tilde{\boldsymbol{\theta}}_1 = \mathbf{0},$$

where $\tilde{\rho}_1^2$ is the largest value of $\tilde{\rho}^2$ that solves $\left|\mathbf{S}_{yx}\hat{\mathbf{H}}_{xx}^{-1}(\hat{\boldsymbol{\theta}}_1)\mathbf{S}_{xy} - \tilde{\rho}^2\mathbf{S}_{yy}\right| = 0$. Under the null that $\boldsymbol{\gamma} = \mathbf{0}$, and conditional on the initial estimator of $\boldsymbol{\theta}$, $\hat{\boldsymbol{\theta}}_1$, the Trace test is now given by

$$T \times Trace\left[\tilde{\mathbf{S}}(\hat{\boldsymbol{\theta}}_1)\right] \overset{a}{\sim} \chi^2_{(m_x-1)^2}, \tag{6}$$

where $\tilde{\mathbf{S}}(\hat{\boldsymbol{\theta}}_1) = \mathbf{S}_{yy}^{-1}\mathbf{S}_{yx}\hat{\mathbf{H}}_{xx}^{-1}(\hat{\boldsymbol{\theta}}_1)\mathbf{S}_{xy}$. The estimate of $\boldsymbol{\theta}$ used for the estimation of $\mathbf{H}_{xx}(\boldsymbol{\theta})$ can be iterated upon as required until convergence is achieved, subject to the normalization restriction, $\boldsymbol{\theta}'\mathbf{S}_{yy}\boldsymbol{\theta} = 1$. It is often found, however, that the estimate of $H_{xx}(\theta)$ can be sensitive to the choice of kernel and estimation window, while dynamic augmentation methods are generally more robust, see, e.g. Andrews and Monahan (1992).

5 Additional Simulation Results

In this section we first provide critical values for the canonical correlation tests. These are required for the maximum canonical correlation test and can also be used in finite samples for the trace canonical correlation test. We next report simulation results for the Tavare (1983) test referred to in the paper. Finally, we report simulation results under higher order dynamics in the data generating process (assuming an AR(2) model) or under heteroskedasticity in the innovations to the model. We also present results for the iterated method introduced in Section 4 above.

5.1 Simulation of Critical Values

Critical values for the maximum canonical correlation test must be simulated whenever the number of categories exceeds two since the ordering of squared canonical correlations induces a non-standard distribution so the test statistic will not follow a chi-squared distribution even in large samples. Furthermore, even for the trace canonical correlation test $T \times Trace(\mathbf{S}_{yy}^{-1}\mathbf{S}_{yx}\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy})$ where ranking of the canonical correlations is not an issue, the chi-squared distribution is only achieved asymptotically, so the critical values will differ in small samples.

To compute critical values for the two-way case, we undertook the following simulation experiment. Letting m_x and m_y be selected from the set $\{2, 3, 4, 5\}$, we generated random draws from the multinomial distribution with $(m_x - 1)(m_y - 1)$ equally likely categories. We carried out $100,000$ replications and considered sample sizes of $T = 20, 50, 100, 500$ and $1,000$.

For the two-way case Table S1 reports 95% critical values for the maximum canonical correlation test and the trace canonical correlation test based on the eigenvalues of $S = \mathbf{S}_{yy}^{-1}\mathbf{S}_{yx}\mathbf{S}_{xx}^{-1}\mathbf{S}_{xy}$ multiplied by the sample size, T. Due to the symmetry of the setup, we only report results for $m_x \leq m_y$. There is no particular pattern in the critical values although, as the sample size rises, they clearly asymptote to the associated $\chi^2_{(m_x-1)(m_y-1)}$ distribution.

Turning to the three-way case, finite sample critical values for the test of conditional independence of Y and X given Z are the same as those in Table S1. For the joint independence test statistic in Theorem 2, new critical values are needed and these are provided in Table S2 which reports 95% critical values when $m_x = m_y = m_z \equiv m$. In most cases the critical values increase towards their asymptotic values from the $\chi^2_{2(m-1)^2}$ distribution as the sample size grows.

5.2 Size and Power of Tavare Test

Table S3 compares the performance of the Tavare test to that of the static and dynamically augmented canonical correlation tests. The simulation setup is described in the main paper, although we here let the serial correlation, as measured by ϕ, vary from 0.50 to 0.95. In the absence of serial correlation, the Tavare test has the correct size. However as the serial correlation gets stronger, the Tavare test tends to get oversized.

5.3 Higher order Dynamics and Heteroskedasticity

To supplement the simulation results based on the first-order autoregressive process in the main text, we also provide results for a stationary second-order autoregressive process of the following form

$$y_t = 1.3y_{t-1} - 0.4y_{t-2} + \varepsilon_t, \quad \varepsilon_t \sim N(0, 1).$$

Table S4 reports the size of the static, dynamically augmented and iterated tests under this data generating process. The results show that little is changed by simulating from an AR(2) process as opposed to an AR(1) process, and that the size of the dynamically augmented test is again well controlled across different samples, although the test is mildly oversized when T is very small and m is large (e.g. $T = 20$ and $m \geq 3$).

Next, we address the effect of heteroskedasticity on the simulation results. We do so by simulating from a first-order autoregressive process with heteroskedastic innovations generated by an autoregressive conditionally heteroskedastic (ARCH) process of the form

$$\begin{aligned} y_t &= \phi y_{t-1} + e_t, \quad e_t \sim N(0, \sigma_t^2) \\ \sigma_t^2 &= \alpha_0 + \alpha_1 e_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \end{aligned}$$

where $\alpha_0 = 1 - \alpha_1 - \beta_1$, $\alpha_1 = 0.1$ and $\beta_1 = 0.8$. A common measure of persistence for the conditional variance of this process is $\alpha_1 + \beta_1 = 0.9$, so we have chosen the parameters to represent quite persistent dynamics in the second moment of the underlying data.

Table S5 presents the size of the independence tests under this process. The results show that heteroskedasticity of this form also leaves our conclusions unchanged with regard to the behavior of the static and dynamically augmented tests.

5.4 Iterated Test

Finally, we investigate the performance of the iterated test described in Section 4 of the supplement. This method uses a HAC procedure with the number of lags chosen to be proportional to $T^{1/3}$, a procedure shown to be optimal for the Bartlett kernel by Andrews (1991), whereas for other kernels such as the Parzen or Tukey-Hanning kernels, the optimal rate is $T^{1/5}$.

The Monte Carlo results in Tables S4 and S5 as well as further results not shown here suggest that the size of the iterated test can be quite sensitive to the underlying process−in some cases being undersized while in others being oversized. This is linked to the sensitivity of estimates of the long-run variance matrix $\mathbf{H}_{xx}(\boldsymbol{\theta})$ noted in Section 4. Finite sample size distortions disappear, however, as the sample grows large.

6 Macroeconomic Application

As an additional empirical application, we make use of the binary recession indicator for the US economy published by the National Bureau of Economic Research (NBER). This comes close to being an official recession indicator for the US economy. We compare this indicator to the median probability that the economy is in a recession as reported by the Survey of Professional Forecasters for the current quarter and one-, two- and three-quarter-ahead horizons. All series are highly serially correlated. We use data over the sample from 1968q3 to 2007q3 to investigate whether economists can forecast recessions given the serial persistence in the recession indicator. Table S6 shows the outcome of our tests. Since $m = 2$, the trace and maximum canonical correlation tests are identical in this case.

Unsurprisingly both the static and dynamically augmented canonical correlation statistics decline as the forecast horizon is expanded. However, while the static canonical correlation test rejects at all forecast horizons, the dynamically augmented test only rejects at the current- and one-quarter-ahead horizons and so fails to indicate predictability at horizons beyond two quarters. The Tavare (1983) test and the test under ordered alternatives fall somewhere in between and indicate predictability of recessions up to two quarters ahead in time. This again shows that it is important to account for serial persistence when testing for correlation between time series of discrete random variables and suggests that conclusions can be sensitive to whether or not serial dependence is accounted for.

References

[1] Andrews, D.W.K., 1991, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator. Econometrica 60, 953-966.

[2] Andrews, D.W.K., and J. C. Monahan, 1992, An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator. Econometrica 60, 953-966.

[3] Newey, W.K. and K.D. West, 1987, A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, Econometrica, 55, 703-708.

[4] Olsson, U., 1979, Maximum Likelihood Estimation of the Polychloric Correlation Coefficient. Psychometrika 44, 443-460.

[5] Ronning, G. and M. Kukuk, 1996, Efficient Estimation of Ordered Probit Models. Journal of American Statistical Association 91, 1120-1129.

[6] Tavare, S., 1983, Serial Dependence in Contingency Tables. Journal of Royal Statistical Society B, 45, 1, 100-106.

Table S1. 95% Finite Sample Critical Values for Maximum Canonical Correlation and Trace Canonical Correlation Tests for Two-Way Tables

	Maximum Canonical Corr.				Trace Canonical Corr.			
T=20								
m_x/m_y	2	3	4	5	2	3	4	5
2	4.11	5.94	7.55	8.89	4.11	5.94	7.55	8.89
3		8.21	9.95	11.35		9.18	12.07	14.73
4			11.71	13.22			16.22	20.00
5				14.66				25.19
T=50								
m_x/m_y	2	3	4	5	2	3	4	5
2	3.92	6.06	7.81	9.37	3.92	6.06	7.81	9.37
3		8.56	10.51	12.19		9.51	12.48	15.28
4			12.61	14.51			16.68	20.69
5				16.49				25.89
T=100								
m_x/m_y	2	3	4	5	2	3	4	5
2	3.96	6.02	7.84	9.46	3.96	6.02	7.84	9.46
3		8.52	10.63	12.54		9.41	12.59	15.50
4			12.88	14.91			16.76	20.90
5				17.03				26.11
T=500								
m_x/m_y	2	3	4	5	2	3	4	5
2	3.87	6.05	7.79	9.49	3.87	6.05	7.79	9.49
3		8.54	10.71	12.62		9.44	12.60	15.51
4			13.14	15.17			16.96	21.02
5				17.38				26.21
T=1000								
m_x/m_y	2	3	4	5	2	3	4	5
2	3.84	6.03	7.86	9.45	3.84	6.03	7.86	9.45
3		8.61	10.70	12.71		9.55	12.57	15.58
4			13.14	15.19			16.95	20.97
5				17.48				26.34

Note: The table is based on 100,000 simulations under the null of no serial correlation in the data, m_x, m_y are the number of categories for x and y.

Table S2. 95% Finite Sample Critical Values for Maximum Canonical Correlation and Trace Canonical Correlation Tests of Joint Independence in Three-Way Tables

	Maximum Canonical Corr.			Trace Canonical Corr.		
T	$m = 2$	$m = 3$	$m = 4$	$m = 2$	$m = 3$	$m = 4$
20	5.98	11.43	15.32	5.98	14.80	27.10
50	6.03	12.28	17.76	6.03	15.34	28.28
100	6.04	12.45	18.53	6.04	15.42	28.70
500	5.97	12.58	19.02	5.97	15.46	28.75
1000	5.97	12.67	18.95	5.97	15.48	28.77

Note: The table is based on 100,000 simulations under the null of no serial correlation in the data. $m = m_x = m_y = m_z$ is the number of categories.

Table S3. Size and Power Comparisons of Tavare and the Dynamically Augmented Reduced Rank Regression Tests, Two-Way, Two Category Example

		Size ($r_{yx} = 0$)			Power ($r_{yx} = 0.80$)		
	T	Static	Tavare	Dyn. Aug.	Static	Tavare	Dyn. Aug.
$\varphi = 0.50$							
	20	0.072	0.048	0.059	0.714	0.633	0.621
	50	0.059	0.050	0.050	0.985	0.982	0.977
	100	0.088	0.043	0.048	1.000	1.000	1.000
	500	0.087	0.050	0.042	1.000	1.000	1.000
	1000	0.082	0.056	0.047	1.000	1.000	1.000
$\varphi = 0.80$							
	20	0.132	0.043	0.066	0.665	0.461	0.486
	50	0.197	0.088	0.068	0.955	0.907	0.862
	100	0.250	0.081	0.058	1.000	0.994	0.991
	500	0.253	0.091	0.059	1.000	1.000	1.000
	1000	0.231	0.085	0.061	1.000	1.000	1.000
$\varphi = 0.95$							
	20	0.189	0.049	0.070	0.615	0.356	0.394
	50	0.366	0.135	0.061	0.884	0.734	0.655
	100	0.507	0.184	0.064	0.970	0.887	0.794
	500	0.551	0.209	0.043	1.000	1.000	1.000
	1000	0.539	0.200	0.045	1.000	1.000	1.000

Table S4. Size of the Tests for Independence Under Second Order Process

		Trace Canonical Corr.			Maximum Canonical Corr.			Ordered
m	T	Static	Dyn. Augm.	Iterated	Static	Dyn. Augm.	Iterated	Alternatives
2	20	0.194	0.053	0.025	0.194	0.053	0.025	0.222
2	50	0.302	0.070	0.108	0.302	0.070	0.108	0.271
2	100	0.359	0.057	0.101	0.359	0.057	0.101	0.281
2	500	0.334	0.051	0.090	0.334	0.051	0.090	0.300
2	1000	0.335	0.043	0.072	0.335	0.043	0.072	0.340
3	20	0.196	0.109	0.043	0.189	0.096	0.039	0.312
3	50	0.404	0.068	0.023	0.400	0.066	0.011	0.356
3	100	0.450	0.061	0.033	0.459	0.061	0.028	0.372
3	500	0.500	0.052	0.057	0.501	0.052	0.062	0.377
3	1000	0.507	0.051	0.058	0.502	0.052	0.063	0.387
4	20	0.056	0.122	0.207	0.017	0.063	0.232	0.278
4	50	0.411	0.086	0.102	0.381	0.070	0.097	0.368
4	100	0.506	0.064	0.049	0.489	0.065	0.036	0.395
4	500	0.572	0.058	0.043	0.549	0.055	0.032	0.408
4	1000	0.601	0.046	0.048	0.575	0.052	0.045	0.409

Table S5. Size of the Independence Tests with Heteroskedastic Innovations

A. No Serial Correlation ($\varphi = 0$)

		Trace Canonical Corr.			Maximum Canonical Corr.			Ordered
m	T	Static	Dyn. Augm.	Iterated	Static	Dyn. Augm.	Iterated	Alternatives
2	20	0.041	0.055	0.005	0.041	0.055	0.005	0.033
2	50	0.033	0.047	0.028	0.033	0.047	0.028	0.029
2	100	0.070	0.060	0.051	0.070	0.060	0.051	0.042
2	500	0.059	0.055	0.053	0.059	0.055	0.053	0.043
2	1000	0.040	0.041	0.038	0.040	0.041	0.038	0.044
3	20	0.018	0.062	0.016	0.020	0.053	0.013	0.040
3	50	0.039	0.055	0.007	0.042	0.053	0.003	0.049
3	100	0.043	0.050	0.013	0.048	0.055	0.010	0.046
3	500	0.052	0.053	0.036	0.053	0.053	0.039	0.056
3	1000	0.057	0.056	0.047	0.059	0.061	0.048	0.059
4	20	0.007	0.094	0.109	0.002	0.040	0.116	0.041
4	50	0.024	0.060	0.029	0.023	0.046	0.025	0.049
4	100	0.042	0.063	0.022	0.042	0.056	0.010	0.051
4	500	0.050	0.051	0.033	0.052	0.053	0.021	0.050
4	1000	0.050	0.051	0.033	0.058	0.057	0.034	0.044

B. Serial Correlation ($\varphi = 0.8$)

		Trace Canonical Corr.			Maximum Canonical Corr.			Ordered
m	T	Static	Dyn. Augm.	Iterated	Static	Dyn. Augm.	Iterated	Alternatives
2	20	0.144	0.059	0.011	0.144	0.059	0.011	0.161
2	50	0.210	0.059	0.090	0.210	0.059	0.090	0.181
2	100	0.247	0.056	0.089	0.247	0.056	0.089	0.188
2	500	0.239	0.053	0.076	0.239	0.053	0.076	0.209
2	1000	0.217	0.055	0.063	0.217	0.055	0.063	0.218
3	20	0.092	0.082	0.028	0.083	0.075	0.021	0.215
3	50	0.213	0.059	0.016	0.219	0.060	0.0079	0.246
3	100	0.274	0.056	0.030	0.274	0.055	0.020	0.254
3	500	0.317	0.052	0.045	0.324	0.054	0.048	0.292
3	1000	0.340	0.063	0.056	0.348	0.059	0.057	0.311
4	20	0.028	0.122	0.163	0.006	0.059	0.179	0.204
4	50	0.192	0.067	0.073	0.181	0.059	0.067	0.272
4	100	0.252	0.070	0.037	0.254	0.060	0.028	0.271
4	500	0.344	0.047	0.037	0.340	0.053	0.032	0.316
4	1000	0.331	0.050	0.051	0.324	0.051	0.044	0.313

Table S6. Independence Tests Applied to Survey Forecasts of Economic Recessions

Forecast Horizon	Tavare	Static	Dyn. Augm.	Ordered Alternatives
Current quarter	21.74	61.76	20.73	22.46
	(0.000)	(0.000)	(0.000)	(0.000)
1 quarter ahead	17.61	49.03	9.26	19.01
	(0.000)	(0.000)	(0.002)	(0.000)
2 quarters ahead	9.67	20.91	0.47	8.85
	(0.002)	(0.000)	(0.493)	(0.003)
3 quarters ahead	2.27	5.43	0.17	2.70
	(0.132)	(0.020)	(0.680)	(0.100)

Notes: p-values are provided in brackets underneath the test statistics.