Regression-Based Mixed Frequency Granger Causality Tests: Short Version*

Eric Ghysels[†] Jonathan B. Hill[‡] Kaiji Motegi[§]

June 23, 2015

Abstract

Testing for Granger causality with mixed frequency data often involves many parametric restrictions relative to sample size, especially when there is a large wedge between sampling frequencies (e.g. weekly and quarterly data). In such a case the trilogy test statistics may not be well approximated by their asymptotic distribution. A bootstrap method can be employed to improve empirical size, but this generally results in a loss of power. This paper presents simple and remarkably powerful Granger causality tests applicable to any mixed frequency sampling data setting. Our tests are based on a simple dimension reduction technique for regression models. The procedure involves multiple parsimonious regression models where each model regresses a low frequency variable onto only one individual lag or lead of a high frequency variable. The lag or lead slope parameter is necessarily zero under the null hypothesis of non-causality. Our test is then based on a max test statistic that selects the largest squared estimator among all parsimonious regression models. Parsimony ensures sharper estimates and hence improved power in small samples. Inference requires simple simulation procedures because the test statistic has a non-standard limit distribution. We show via Monte Carlo simulations that the max test is more powerful than existing mixed frequency Granger causality tests in small samples. An empirical application examines Granger causality between weekly interest rate spread and quarterly economic growth in the U.S.

Keywords: Granger causality test, Local asymptotic power, Max test, Mixed data sampling (MIDAS), Sims test, Temporal aggregation.

---

*Full version and the supplemental appendix are posted at `http://www.aoni.waseda.jp/motegi/`

[†]Department of Economics and Department of Finance, Kenan-Flagler Business School, University of North Carolina at Chapel Hill. E-mail: `eghysels@unc.edu`

[‡]Department of Economics, University of North Carolina at Chapel Hill. E-mail: `jbhill@email.unc.edu`

[§]Faculty of Political Science and Economics, Waseda University. E-mail: `motegi@aoni.waseda.jp`

1 Introduction

Time series are often sampled at different frequencies, and it is well known that temporal aggregation adversely affects Granger causality. One of the most popular Granger causality tests is a Wald statistic based on multi-step ahead vector autoregression (VAR). Since standard VAR models are designed for single-frequency data, these tests often suffer from the adverse effect of temporal aggregation. In order to alleviate this problem, Ghysels, Hill, and Motegi (2013) develop a set of Granger causality tests that explicitly take advantage of data sampled at mixed frequencies. They accomplish this by extending Dufour, Pelletier, and Renault's (2006) VAR-based causality test using Ghysels' (2014) mixed frequency vector autoregression (MF-VAR). Although these tests avoid the undesirable effects of temporal aggregation, their applicability is limited because parameter proliferation in MF-VAR adversely affects the power of the tests. Indeed, if we let m be the ratio of sampling frequencies (e.g. $m = 3$ for monthly versus quarterly data), then for bivariate cases the MF-VAR is of dimension $m + 1$. Parameter proliferation occurs when m is large, and becomes worse as the VAR lag length increases. In these cases, Ghysels, Hill, and Motegi's (2013) Wald test exhibits size distortions, while a bootstrapped Wald test results in correct size but low power.

The present paper proposes new mixed frequency Granger causality tests that have several advantages: (i) they are regression-based and simple to implement, and (ii) they apply to a large m for example a week vs. quarter mixture. We postulate multiple parsimonious regression models where the j^{th} model regresses a low frequency variable x_L onto lags of x_L and only the j^{th} lag or lead of a high frequency variable x_H. Our test statistic is the *maximum* among squared estimators scaled and weighted properly. Although the max test statistic follows a non-standard asymptotic distribution under the null hypothesis of non-causality, a simulated p-value is readily available through an arbitrary number of draws from the null distribution. The max test is therefore straightforward to implement in practice. We prove that the max test is consistent for any form of causality from x_H to x_L. (Consistency for causality from x_L to x_H remains as an open question.)

In our Monte Carlo simulations, we compare finite sample properties of the max test based on mixed frequency [MF] data, a Wald test based on MF data, a max test based on low frequency [LF] data, and a Wald test based on LF data. We show that MF tests are more robust against complex (but realistic) causal patterns than LF tests. Further, the MF max test is more powerful than the MF Wald test in most cases due to the greater parsimony of the former.

The full version of the paper has a substantial amount of results omitted here. First, the short version discusses causality from x_H to x_L only, while the full version discusses the opposite direction as well. The latter exploits Sims' (1972) technique that regresses x_L onto leads of x_H. Second, the full paper conducts local power analysis. Third, here we report only a small part of the entire simulation results. Fourth, the full version presents complete proofs of theorems. Finally, the full paper discusses a wide range of potential applications of the max test.

This paper is organized as follows. Section 2 presents the max test statistic and derives its asymptotic properties. Section 3 runs Monte Carlo simulations, Section 4 presents an empirical application, and Section 5 concludes the paper. Tables and figures are collected at the end.

2 Max Test

This paper focuses on a bivariate case with a high frequency variable x_H and a low frequency variable x_L. We denote by m the number of high frequency time periods for each low frequency time period τ_L. We assume throughout that m is fixed (e.g. $m = 3$ months per quarter). In low frequency time period τ_L, we observe m high frequency realizations $\{x_H(\tau_L, 1),\ x_H(\tau_L, 2),\ \ldots,\ x_H(\tau_L, m)\}$ and only one low frequency realization $x_L(\tau_L)$. $x_H(\tau_L, 1)$ is the first high frequency observation of x_H in time τ_L, $x_H(\tau_L, 2)$ is the second, and $x_H(\tau_L, m)$ is the m-th.

The true data generating process (DGP) is assumed to be:

$$x_L(\tau_L) = \sum_{k=1}^{p} a_k x_L(\tau_L - k) + \sum_{j=1}^{pm} b_j x_H(\tau_L - 1, m + 1 - j) + \epsilon_L(\tau_L), \tag{1}$$

where $\{\epsilon_L(\tau_L)\}$ is a martingale difference sequence with variance σ_L^2. The index $j \in \{1, \ldots, pm\}$ is in high frequency terms, and the second argument $m + 1 - j$ of $x_H(\tau_L - 1, m + 1 - j)$ can be less than 1 since $j > m$ occurs when $p > 1$. It can be understood, without any confusion, that $x_H(\tau_L - i, j) = x_H(\tau_L, j - im)$ for $j = 1, \ldots, m$ and $i \geq 0$ (e.g. $x_H(\tau_L, 0) = x_H(\tau_L - 1, m)$, $x_H(\tau_L, -1) = x_H(\tau_L - 1, m - 1)$, $x_H(\tau_L, m + 1) = x_H(\tau_L + 1, 1)$, etc.).

Based on the classic theory of Dufour and Renault (1998) and the mixed frequency extension made by Ghysels, Hill, and Motegi (2013), we know that x_H does not Granger cause x_L given the mixed frequency information set *if and only if* $\boldsymbol{b} \equiv [b_1, \ldots, b_{pm}]' = \mathbf{0}_{pm\times 1}$. In order to test for non-causality $H_0 : \boldsymbol{b} = \mathbf{0}_{pm\times 1}$, we want a test statistic that obtains asymptotic power of one against any deviation from non-causality, achieves high power in finite samples, and does not produce size distortions in small samples when pm is large.

Before presenting the new test, it is helpful to review the existing mixed frequency Granger causality test proposed by Ghysels, Hill, and Motegi (2013). They work with what we call a *naïve regression model* that regresses x_L onto q low frequency lags and h high frequency lags of x_H:

$$x_L(\tau_L) = \sum_{k=1}^{q} \alpha_k x_L(\tau_L - k) + \sum_{j=1}^{h} \beta_j x_H(\tau_L - 1, m + 1 - j) + u_L(\tau_L). \tag{2}$$

Ghysels, Hill, and Motegi (2013) estimate the parameters in (2) by least squares and then test $H_0 : \beta_1 = \cdots = \beta_h = 0$ via a Wald test. Model (2) contains DGP (1) as a special case when $q \geq p$ and $h \geq pm$, hence the Wald test is trivially consistent if $q \geq p$ and $h \geq pm$.

A potential problem here is that pm, the true lag order of x_H, may be quite large when m takes a large value (e.g. weekly versus quarterly data). Including sufficiently many high frequency lags $h \geq pm$ generally results in size distortions for an asymptotic Wald test when the sample size T_L is small and pm is large. A bootstrap can be employed to eliminate size distortions, but this generally results in a loss of power. Of course, we may use a small number of lags $h < pm$ to ensure the Wald statistic is well characterized by its χ^2 limit distribution, but this results in an inconsistent test when there exists causality involving lags beyond h.

A main contribution of this paper is to resolve this trade-off by combining the following multiple *parsimonious regression models*:

$$x_L(\tau_L) = \sum_{k=1}^{q} \alpha_{k,j} x_L(\tau_L - k) + \beta_j x_H(\tau_L - 1, m + 1 - j) + u_{L,j}(\tau_L), \quad j = 1, \ldots, h. \tag{3}$$

Assuming that $q \geq p$, each model (3) is correctly specified under the null hypothesis of non-causality $\boldsymbol{b} = \mathbf{0}_{pm\times 1}$. Hence, if there is indeed non-causality, the least squares estimators $\hat{\beta}_j \xrightarrow{p} 0$, hence $\max_{1\leq j\leq h} \{\hat{\beta}_j^2\} \xrightarrow{p} 0$. Using this property, we propose a max test statistic:

$$\hat{\mathcal{T}} \equiv \max_{1\leq j\leq h} \left(\sqrt{T_L} w_{T_L,j} \hat{\beta}_j \right)^2,$$

where $\{w_{T_L,j} : j = 1, \ldots, h\}$ is a sequence of weighting schemes with limits $\{w_j : j = 1, \ldots, h\}$. As a standardization we assume $\sum_{j=1}^{h} w_{T_L,j} = 1$ without loss of generality. When we have no prior information about the weighting scheme, a simple choice of $w_{T_L,j}$ is a flat weight $1/h$.

The limit distribution of the max test statistic $\hat{\mathcal{T}}$ under non-causality can be derived as follows.

Theorem 2.1. Under $H_0 : \boldsymbol{b} = \mathbf{0}_{pm\times 1}$, we have that $\hat{\mathcal{T}} \xrightarrow{d} \max_{1\leq j\leq h} \mathcal{N}_j^2$ as $T_L \to \infty$, where $\boldsymbol{\mathcal{N}} \equiv [\mathcal{N}_1, \ldots, \mathcal{N}_h]'$ is distributed $N(\mathbf{0}_{h\times 1}, \boldsymbol{V})$ with positive definite covariance matrix $\boldsymbol{V}$.

See the full paper for a precise expression of $\boldsymbol{V}$. The mixed frequency max test statistic $\hat{\mathcal{T}}$ has a non-standard limit distribution under H_0 that can be easily simulated in order to compute an approximate p-value. We can construct a consistent estimator $\hat{\boldsymbol{V}}_{T_L} \xrightarrow{p} \boldsymbol{V}$ from sample (see the full paper). Draw R samples $\boldsymbol{\mathcal{N}}^{(1)}, \ldots, \boldsymbol{\mathcal{N}}^{(R)}$ independently from $N(\mathbf{0}_{h\times 1}, \hat{\boldsymbol{V}}_{T_L})$. Compute artificial test statistics $\hat{\mathcal{T}}^{(r)} \equiv \max_{1\leq j\leq h} (\boldsymbol{\mathcal{N}}_j^{(r)})^2$. An asymptotic p-value approximation for $\hat{\mathcal{T}}$ is $\hat{p} = (1/R) \sum_{r=1}^{R} I(\hat{\mathcal{T}}^{(r)} > \hat{\mathcal{T}})$, where $I(A)$ is the indicator function that equals one if event A occurs and zero otherwise.

We show that the max test is consistent (i.e. $\hat{\mathcal{T}} \xrightarrow{p} \infty$ whenever $\boldsymbol{b} \neq \mathbf{0}_{pm\times 1}$). The next theorem presents an analytical expression of the pseudo-true values β_j^*, which is identically the probability limits of $\hat{\beta}_j$.

Theorem 2.2. Consider model (3). The pseudo-true value of β_j, denoted by β_j^*, is the last element of

$$\left[E\left[\boldsymbol{x}_j(\tau_L - 1) \boldsymbol{x}_j(\tau_L - 1)' \right] \right]^{-1} \times E\left[\boldsymbol{x}_j(\tau_L - 1) \boldsymbol{X}_H(\tau_L - 1)' \right] \times \boldsymbol{b},$$

where $\boldsymbol{x}_j(\tau_L - 1) = [x_L(\tau_L - 1), \ldots, x_L(\tau_L - q), x_H(\tau_L - 1, m + 1 - j)]'$ and $\boldsymbol{X}_H(\tau_L - 1) = [x_H(\tau_L - 1, m + 1 - 1), \ldots, x_H(\tau_L - 1, m + 1 - pm)]'$.

Theorem 2.2 provides useful insights on the relationship between the underlying coefficient $\boldsymbol{b}$ and the pseudo-true value $\boldsymbol{\beta}^* \equiv [\beta_1^*, \ldots, \beta_h^*]'$. First, $\boldsymbol{\beta}^* = \mathbf{0}_{h\times 1}$ whenever there is non-causality (i.e $\boldsymbol{b} = \mathbf{0}_{pm\times 1}$), regardless of the relative magnitude of h and pm. Second, as the next theorem demonstrates, $\boldsymbol{b} = \mathbf{0}_{pm\times 1}$ whenever $\boldsymbol{\beta}^* = \mathbf{0}_{h\times 1}$, provided $h \geq pm$.

Theorem 2.3. Assume $h \geq pm$, then $\boldsymbol{\beta}^* = \mathbf{0}_{h\times 1}$ implies $\boldsymbol{b} = \mathbf{0}_{pm\times 1}$.

Theorem 2.3, which is one of the main results of this paper, implies the consistency of the max test. Assume the weight limits $w_j > 0$ for all $j = 1, \ldots, h$ so that we have a non-trivial result under the alternative. Suppose that there exists an arbitrary form of causality ($\boldsymbol{b} \neq \mathbf{0}_{pm\times 1}$), then Theorem 2.3 implies that at least one of $\{\beta_1^*, \ldots, \beta_h^*\}$ must be nonzero. Hence, the max test statistic $\hat{\mathcal{T}} = \max_{1\leq j\leq h}(\sqrt{T_L} w_{T_L,j} \hat{\beta}_j)^2$ must diverge to ∞.

Theorem 2.4 (Consistency)**.** Assume $h \geq pm$ and $w_j > 0$ for all $j = 1, \ldots, h$. Then $\hat{\mathcal{T}} \xrightarrow{p} \infty$ if $H_1 : \boldsymbol{b} \neq \mathbf{0}_{pm\times 1}$ is true.

3 Monte Carlo Simulations

We now conduct Monte Carlo simulations in order to compare the finite sample performance of the MF max test and MF Wald test. We also implement the low frequency [LF] max test and LF Wald test, which work with aggregated x_H, in order to see how temporal aggregation affects empirical power. The ratio of sampling frequencies is $m = 12$, approximately a week versus quarter mixture. The sample size is $T_L = 80$ quarters.

The true DGP is Ghysels' (2014) MF-VAR(1): $\boldsymbol{X}(\tau_L) = \boldsymbol{A}_1 \boldsymbol{X}(\tau_L - 1) + \boldsymbol{\epsilon}(\tau_L)$, where $\boldsymbol{X}(\tau_L) = [x_H(\tau_L, 1), \ldots, x_H(\tau_L, 12), x_L(\tau_L)]'$. Note that the last equation of this system corresponds to (1). We assume that the autoregressive coefficient of high frequency x_H is either $d = 0.2$ (transitory) or $d = 0.8$ (persistent). See the full paper for a complete specification.

The lower-left block of $\boldsymbol{A}_1$ corresponds to the key parameter $\boldsymbol{b}$ because they govern Granger causality from x_H to x_L. We consider four patterns for $\boldsymbol{b}$: *non-causality* $\boldsymbol{b} = \mathbf{0}_{12\times 1}$; *decaying causality* with alternating signs $b_j = (-1)^{j-1} \times 0.3/j$ for $j = 1, \ldots, 12$; *lagged causality* $b_j = 0.3 \times I(j = 12)$ for $j = 1, \ldots, 12$; and *sporadic causality* $(b_3, b_7, b_{10}) = (0.2, 0.05, -0.3)$ and all other $b_j = 0$.

We fit regression models that in all cases include two low frequency lags of x_L (i.e. $q = 2$). The number of high frequency lags of x_H used in the MF tests is $h_{MF} \in \{4, 8, 12, 24\}$. The number of low frequency lags of aggregated x_H used in the LF tests is $h_{LF} \in \{1, 2, 3, 4\}$. Low frequency tests use flow sampling ($x_H(\tau_L) = (1/12)\sum_{j=1}^{12} x_H(\tau_L, j)$) and stock sampling ($x_H(\tau_L) = x_H(\tau_L, 12)$). The max test weighting scheme is flat $\boldsymbol{W}_h = (1/h) \times \boldsymbol{I}_h$, and the number of draws from the limit distributions under H_0 is 5,000 for p-value computation. We use Gonçalves and Killian's (2004) parametric bootstrap with 499 bootstrap samples in order to better approximate the small sample Wald statistic distribution. The number of Monte Carlo samples drawn is 5,000 for max tests and 1,000 for bootstrapped Wald tests (due to the added computation time), and nominal size α is 0.05.

See Table 1 for the simulation results. Empirical size in both tests is fairly sharp, ranging across cases between 0.037 and 0.065 (Panel A). The max tests has sharp size due to its parsimonious specification, while the Wald test has sharp size due to the bootstrap.

Based on Panels B-D, we discuss in what sense MF tests are preferred to LF tests. It is not the case that MF tests always achieve higher power than LF tests. If a causal pattern is simple enough for LF tests to capture, then LF tests may have higher power than MF tests due to greater parsimony. See decaying causality with $(d, h_{MF}, h_{LF}) = (0.2, 4, 1)$ for example (Panel B.1). The LF max test with stock sampling has power 0.657, while the MF max test has power 0.482. Similarly, the LF Wald test with stock sampling has power 0.597, while the MF Wald test has power 0.527. (The LF test with flow sampling has virtually no power because flow aggregation offsets the positive and negative impacts of lagged x_H on x_L.)

A clear advantage of MF tests against LF tests emerges when there exist complicated causal patterns like sporadic causality. See Panel D.1 ($d = 0.2$) for example. The MF max and MF Wald tests have power $[0.167, 0.442]$ depending on h_{MF}, while the LF tests have virtually no power regardless of h_{LF}. This is because the low frequency lags of x_H are too coarse to capture the complex causality from disaggregated x_H to x_L. Since we do not know what kind of causality exists in practice, it is a safer strategy to use MF tests instead of LF tests.

We now consider the relative performance of the MF max test and MF Wald test. In a strong majority of cases across causal patterns $\boldsymbol{b}$, lag length h_{MF}, and persistence d, the max test has higher power than the Wald test (20 cases out of 24). In the four exceptions the differences are negligible, where the greatest spread being $0.482 - 0.527 = -0.045$ when there is decaying causality, $d = 0.2$, and $h_{MF} = 4$ (Panel B.1). In the 20 cases where the max test performs better, the difference in power is often substantial. Under lagged causality with $(d, h_{MF}) = (0.8, 24)$, for example, max test power is 0.907 while Wald test power is 0.498 (Panel C.2). Overall, the MF max test is more powerful than the MF Wald test in finite sample because of the parsimonious specification of the former.

4 Empirical Application

As an empirical illustration, we study Granger causality from a weekly interest rate spread to quarterly real GDP growth in the U.S. A decline in the interest rate spread has historically been regarded as a strong predictor of an immediate recession, but recent events place doubt on its use for such prediction. We use the year-to-year growth rate of seasonally-adjusted quarterly real GDP as a business cycle measure. In order to remove potential seasonal effects remaining after seasonal adjustment, we use annual growth (i.e. 4 quarter log-difference $\ln(y_t) - \ln(y_{t-4})$). The short and long term interests rates used for the term spread are respectively the federal funds (FF) rate and 10-year Treasury constant maturity rate. We aggregate each daily series into weekly series by picking the last observation in each week (recall that interest rates are stock variables). The sample period is January 5, 1962 to December 31, 2013, covering 2,736 weeks or 208 quarters.

Figure 1 shows the weekly 10-year rate, FF rate, their spread (10Y - FF), and quarterly GDP growth. The shaded areas represent recession periods defined by the National Bureau of Economic Research. In the first half of the sample period, a sharp decline of the spread seems to be immediately followed by a recession. In the second half there appears to be a weaker association, and a larger time lag between a spread drop and a recession.

The number of weeks contained in each quarter τ_L is not constant. We simplify the analysis by forcing a constant $m = 12$ by taking a sample average at the end of each τ_L, resulting in the modified spread $\{x_H^*(\tau_L, j)\}_{j=1}^{12}$ (see the full paper for details). This modification gives us a dataset with $T_L = 208$, $m = 12$, and thus $T = mT_L = 2,496$ high frequency observations.

In view of our 52-year sample period, we implement a rolling window analysis with a window width of 80 quarters. The first subsample covers the first quarter of 1962 through the fourth quarter of 1981 (written as 1962:I-1981:IV), the second one is 1962:II-1982:I, and the last one is 1994:I-2013:IV, equaling 129 subsamples.

The MF max test operates on parsimonious regression models:

$$x_L(\tau_L) = \alpha_{0,j} + \sum_{k=1}^{2} \alpha_{k,j} x_L(\tau_L - k) + \beta_j x_H^*(\tau_L - 1, 12 + 1 - j) + u_{L,j}(\tau_L). \quad j = 1, \ldots, 24,$$

which includes two quarters of lagged GDP growth (x_L) and 24 weeks of lagged interest rate spread (x_H^*). The MF Wald test operates on: $x_L(\tau_L) = \alpha_0 + \sum_{k=1}^{2} \alpha_k x_L(\tau_L - k) + \sum_{j=1}^{24} \beta_j x_H^*(\tau_L - 1, 12 + 1 - j) + u_L(\tau_L)$.

The LF max test is based on parsimonious models: $x_L(\tau_L) = \alpha_{0,j} + \sum_{k=1}^{2} \alpha_{k,j} x_L(\tau_L - k) + \beta_j x_H^*(\tau_L - j) + u_{L,j}(\tau_L)$ for $j = 1, 2, 3$. They have two quarters of lagged x_L and three quarters of lagged x_H^*. Since the interest rate spread is a stock variable, we let the aggregated high frequency variable be $x_H^*(\tau_L) = x_H^*(\tau_L, 12)$. Finally, the LF Wald test is performed on: $x_L(\tau_L) = \alpha_0 + \sum_{k=1}^{2} \alpha_k x_L(\tau_L - k) + \sum_{j=1}^{3} \beta_j x_H^*(\tau_L - j) + u_L(\tau_L)$.

Wald statistic p-values are computed based on Gonçalves and Killian's (2004) bootstrap, with $N = 999$ replications. Max statistic p-values are computed based on 100,000 draws from the limit distributions under non-causality.

Figure 2 plots p-values for tests of non-causality over the 129 subsamples. All tests except for the MF Wald test find significant causality in early periods. The MF max test detects significant causality prior to 1981:IV-2001:III, the LF max test detects significant causality prior to 1980:III-2000:II, and the LF Wald test detects significant causality prior to 1974:III-1994:II. The MF max test has the longest period of significant causality, arguably due to its high power, as shown in Section 3. These three tests all agree that there is non-causality in recent periods, possibly reflecting some structural change in the middle of the entire sample.

The MF Wald test, in contrast, suggests that there is significant causality only *after* subsample 1990:III-2010:II, which is somewhat counter-intuitive. This result may stem from parameter proliferation. The MF naïve regression model has many more parameters than any other model. In view of the intuitive test results, the MF max test seems to be preferred to the MF Wald test when the ratio of sampling frequencies m is large.

5 Conclusions

This paper proposes a new mixed frequency Granger causality test that achieves high power even when the ratio of sampling frequencies is large. This is accomplished by exploiting multiple parsimonious regression models where the j^{th} model regresses a low frequency variable x_L onto the j^{th} lag or lead of a high frequency variable x_H. Our resulting max test statistic then operates on the largest j^{th} lag or lead estimated parameter. Although the max test statistic follows a non-standard asymptotic distribution under the null hypothesis of non-causality, a p-value can be easily computed via a simulation method.

We prove the mixed frequency max test is consistent for Granger causality from x_H to x_L. We also show via Monte Carlo simulations that the max test is more powerful than existing mixed frequency Wald tests in small samples. An empirical application examines Granger causality between weekly interest rate spread and quarterly economic growth in the U.S. The mixed frequency max test yields an intuitive result that the interest rate spread causes GDP growth until about the year 2000, after which causality vanishes, while Wald and low frequency tests yield mixed results.

References

Dufour, J., D. Pelletier, and E. Renault (2006): "Short Run and Long Run Causality in Time Series: Inference," *Journal of Econometrics*, 132, 337–362.

Dufour, J., and E. Renault (1998): "Short Run and Long Run Causality in Time Series: Theory," *Econometrica*, 66, 1099–1125.

Ghysels, E. (2014): "Macroeconomics and the Reality of Mixed Frequency Data," *Journal of Econometrics*, (forthcoming).

Ghysels, E., J. B. Hill, and K. Motegi (2013): "Testing for Granger Causality with Mixed Frequency Data," Working paper, Dept. of Economics, University of North Carolina.

Gonçalves, S., and L. Killian (2004): "Bootstrapping Autoregressions with Conditional Heteroskedasticity of Unknown Form," *Journal of Econometrics*, 123, 89–120.

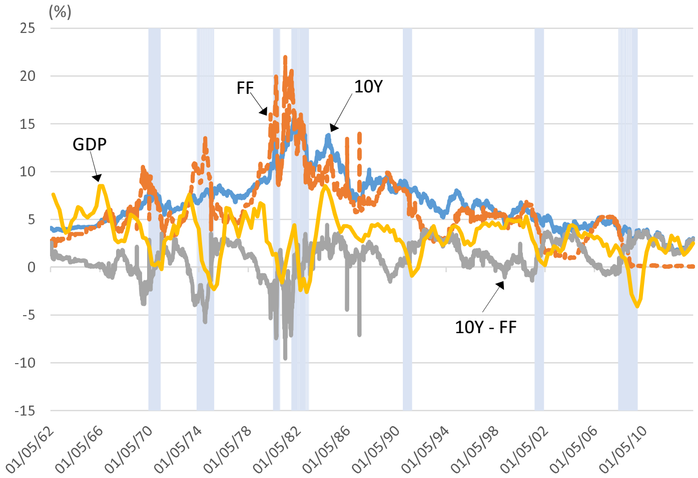
Sims, C. A. (1972): "Money, Income, and Causality," *American Economic Review*, 62, 540–552.

Table 1: Rejection Frequencies of High-to-Low Causality Tests Based on MF-VAR(1)

A. Non-Causality: $\boldsymbol{b} = \mathbf{0}_{12\times 1}$								B. Decaying Causality: $b_j = (-1)^{j-1}0.3/j$							
A.1 $d = 0.2$ (low persistence in x_H)								B.1 $d = 0.2$ (low persistence in x_H)							
	MF			LF (flow)		LF (stock)			MF			LF (flow)		LF (stock)	
h_{MF}	Max	Wald	h_{LF}	Max	Wald	Max	Wald	h_{MF}	Max	Wald	h_{LF}	Max	Wald	Max	Wald
4	.055	.054	1	.058	.048	.056	.045	4	.482	.527	1	.088	.062	.657	.597
8	.057	.037	2	.055	.040	.058	.039	8	.374	.412	2	.071	.058	.554	.469
12	.052	.048	3	.058	.065	.058	.042	12	.332	.335	3	.069	.052	.503	.429
24	.056	.038	4	.050	.054	.059	.039	24	.257	.229	4	.069	.053	.451	.342
A.2 $d = 0.8$ (high persistence in x_H)								B.2 $d = 0.8$ (high persistence in x_H)							
	MF			LF (flow)		LF (stock)			MF			LF (flow)		LF (stock)	
h_{MF}	Max	Wald	h_{LF}	Max	Wald	Max	Wald	h_{MF}	Max	Wald	h_{LF}	Max	Wald	Max	Wald
4	.060	.035	1	.058	.038	.057	.045	4	.794	.700	1	.309	.265	.883	.834
8	.058	.037	2	.056	.054	.053	.055	8	.697	.564	2	.235	.205	.798	.764
12	.055	.043	3	.054	.039	.058	.038	12	.642	.480	3	.198	.159	.748	.697
24	.052	.058	4	.059	.041	.059	.048	24	.534	.272	4	.189	.147	.719	.619
C. Lagged Causality: $b_j = 0.3 \times I(j = 12)$								D. Sporadic Causality: $(b_3, b_7, b_{10}) = (0.2, 0.05, -0.3)$							
C.1 $d = 0.2$ (low persistence in x_H)								D.1 $d = 0.2$ (low persistence in x_H)							
	MF			LF (flow)		LF (stock)			MF			LF (flow)		LF (stock)	
h_{MF}	Max	Wald	h_{LF}	Max	Wald	Max	Wald	h_{MF}	Max	Wald	h_{LF}	Max	Wald	Max	Wald
4	.053	.040	1	.125	.102	.061	.056	4	.248	.207	1	.058	.057	.062	.048
8	.052	.043	2	.104	.088	.082	.056	8	.184	.167	2	.052	.049	.051	.049
12	.405	.264	3	.088	.074	.068	.049	12	.442	.416	3	.050	.055	.054	.051
24	.325	.155	4	.083	.059	.061	.060	24	.352	.250	4	.059	.039	.058	.061
C.2 $d = 0.8$ (high persistence in x_H)								D.2 $d = 0.8$ (high persistence in x_H)							
	MF			LF (flow)		LF (stock)			MF			LF (flow)		LF (stock)	
h_{MF}	Max	Wald	h_{LF}	Max	Wald	Max	Wald	h_{MF}	Max	Wald	h_{LF}	Max	Wald	Max	Wald
4	.084	.060	1	.592	.518	.073	.052	4	.459	.305	1	.076	.052	.273	.228
8	.227	.141	2	.606	.562	.845	.779	8	.404	.404	2	.128	.101	.391	.402
12	.931	.731	3	.539	.527	.804	.749	12	.803	.740	3	.107	.067	.332	.325
24	.907	.498	4	.493	.438	.769	.662	24	.677	.524	4	.101	.071	.292	.335

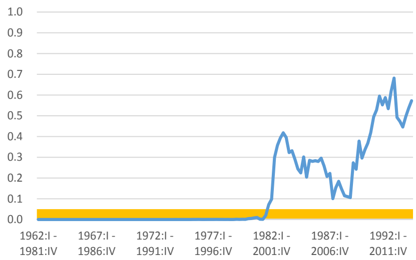
The DGP is MF-VAR(1) with a ratio of sampling frequencies $m = 12$, where b_j signifies the impact of $x_H(\tau_L - 1, m + 1 - j)$ on $x_L(\tau_L)$. Panels A, B, C, and D concern non-causality, decaying causality, lagged causality, and sporadic causality, respectively. The models estimated have two low frequency lags of x_L (i.e. $q = 2$). The max test statistic is computed using a flat weight $\boldsymbol{W}_h = (1/h) \times \boldsymbol{I}_h$, and the p-value is computed using 1,000 draws from the null limit distribution. The Wald test p-value is computed using the parametric bootstrap based on Gonçalves and Killian (2004), with 499 bootstrap replications. Nominal size is $\alpha = 0.05$. The number of Monte Carlo samples drawn is 5,000 for max tests and 1,000 for Wald tests.

Figure 1: Time Series Plot of U.S. Interest Rates and Real GDP Growth

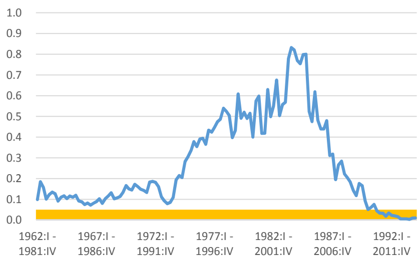


This figure plots weekly 10-year Treasury constant maturity rate, weekly effective federal funds rate, their spread 10Y - FF, and the quarterly real GDP growth from previous year. The sample period covers January 5, 1962 through December 31, 2013. The shaded areas represent recession periods defined by the National Bureau of Economic Research (NBER).

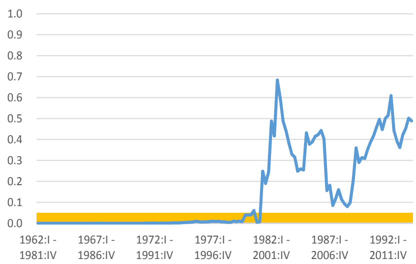
Figure 2: P-values for Tests of Non-Causality from Interest Rate Spread to GDP



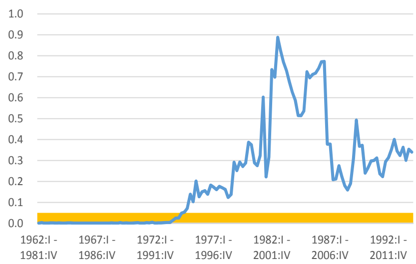
(a) MF Max Test



(b) MF Wald Test



(c) LF Max Test



(d) LF Wald Test

This figure plots rolling window p-values of Granger causality tests. MF tests concern weekly interest rate spread and quarterly GDP growth, while LF tests concern quarterly interest rate spread and GDP growth. The sample period is January 5, 1962 through December 31, 2013. The window size is 80-quarters. Any p-value in the shaded area indicates rejection of non-causality from the interest rate spread to GDP growth at the 5% level for that window.