Econometrics Working Paper EWP0008
ISSN 1485-6441

Testing for Two-step Granger Noncausality in Trivariate VAR Models

Judith A Giles

University of Victoria, Victoria, British Columbia, Canada

I. INTRODUCTION

Granger's (1969) popular concept of causality, based on work by Weiner (1956), is typically defined in terms of predictability for one period ahead. Recently, Dufour and Renault (1998) generalized the concept to causality at a given horizon h, and causality up to horizon h, where h is a positive integer that can be infinite ($1 \leq h < \infty$); see also Sims (1980), Hsiao (1982) and Lütkepohl (1993a) for related work. They show that the horizon h is important when auxiliary variables are available in the information set that are not directly involved in the noncausality test, as causality may arise more than one period ahead indirectly via these auxiliary variables, even when there is one period ahead noncausality in the traditional sense. For instance, suppose we wish to test for Granger noncausality (GNC) from Y to X with an information set consisting of three variables – X, Y and Z, and suppose that Y does not Granger cause X, in the traditional one-step sense. This does not preclude two-step Granger causality, which will arise when Y Granger causes Z and Z Granger causes X; the auxiliary variable Z enables predictability to result two periods ahead. Consequently, it is important to examine for causality at horizons beyond one period when the information set contains variables that are not directly involved in the GNC test.

Dufour and Renault (1998) do not provide information on testing for GNC when $h>1$; our aim is to contribute in this direction. In this chapter we provide an initial investigation of testing for two-

step GNC by suggesting two sequential testing strategies to examine this issue. We also provide information on the sampling properties of these testing procedures through simulation experiments. We limit our study to the case when the information set contains only three variables for reasons that we explain in Section II.

The layout of this chapter is as follows. In the next section we present our modeling framework and we discuss the relevant noncausality results. Section III introduces our proposed sequential two-step noncausality tests and provides details of our experimental design for the Monte Carlo study. Section IV reports the simulation results. The proposed sequential testing strategies are then applied to a trivariate data set concerned with money-income causality in the presence of an interest rate variable in Section V. Some concluding remarks are given in Section VI.

II. DISCUSSION OF NONCAUSALITY TESTS

A. Model Framework

We consider an n-dimensional vector time series $\{y_t: t=1,2,..., T\}$, which we assume is generated from a vector autoregression (VAR) of finite order p[1]:

$$y_t = \sum_{i=1}^{p} \Pi_i y_{t-i} + \varepsilon_t \qquad (1)$$

where Π_i is an $(n\times n)$ matrix of parameters, ε_t is an $(n\times 1)$ vector distributed as $IN(0, \Sigma)$, and (1) is initialized at t=-p+1,...,0; the initial values can be any random vectors including constants. Let y_t be partitioned as $y_t = (X_t^T, Y_t^T, Z_t^T)^T$, where, for Q=X,Y,Z, Q_t is an $(n_Q\times 1)$ vector, and $n_X+n_Y+n_Z=n$. Also, let Π_i be conformably partitioned as

[1] We limit our attention to testing GNC within a VAR framework to remain in line with the vast applied literature; the definitions proposed by Dufour and Renault (1998) are more widely applicable.

$$\Pi_i = \begin{bmatrix} \pi_{i,XX} & \pi_{i,XY} & \pi_{i,XZ} \\ \pi_{i,YX} & \pi_{i,YY} & \pi_{i,YZ} \\ \pi_{i,ZX} & \pi_{i,ZY} & \pi_{i,ZZ} \end{bmatrix}.$$

where, for Q,R=X,Y,Z, $\pi_{i,QR}$ is an $(n_Q \times n_R)$ matrix of coefficients.

B. h-step Noncausality

Suppose we wish to determine whether or not Y Granger noncauses X one period ahead, denoted as $Y \underset{1}{\not\rightarrow} X$, in the presence of the auxiliary variables contained in Z. Traditionally, within the framework we consider, this is examined via a test of the null hypothesis H_{01}: $P_{XY}=0$ where $P_{XY} = [\pi_{1,XY}, \pi_{2,XY},..., \pi_{p,XY}]$ using a Wald or Likelihood Ratio (LR) statistic. What does the result of this hypothesis test imply for GNC from Y to X at horizon h (>1), which we denote as $Y \underset{h}{\not\rightarrow} X$, and for GNC from Y to X up to horizon h that includes one-period ahead, denoted by $Y \underset{(h)}{\not\rightarrow} X$?

There are three cases of interest.

1. $n_Z=0$; i.e., there are no auxiliary variables in Z so that all variables in the information set are involved in the GNC test under study. Then, from Dufour and Renault (1998) Proposition 2.2, the four following properties are equivalent:

$$\text{(i) } Y \underset{1}{\not\rightarrow} X; \qquad \text{(ii) } Y \underset{h}{\not\rightarrow} X; \qquad \text{(iii) } Y \underset{(h)}{\not\rightarrow} X; \qquad \text{(iv) } Y \underset{(\infty)}{\not\rightarrow} X.$$

That is, when all variables are involved in the GNC test, support for the null hypothesis H_0: $P_{XY}=0$ is also support for GNC at, or up to, any horizon h. This is intuitive as there are no auxiliary variables available through which indirect causality can occur. For example, the bivariate model satisfies these conditions.

2. $n_Z=1$; i.e., there is one auxiliary variable in the information set that is not directly involved in the GNC test of interest. Then, from Dufour and Renault (1998) Proposition 2.4 and Corollary

3.6, we have that $Y \underset{(2)}{\not\rightarrow} X$ (which implies $Y \underset{(\infty)}{\not\rightarrow} X$) if and only if at least one of the two following conditions is satisfied:

C1. $Y \underset{1}{\not\rightarrow} (X^T, Z)^T$;

C2. $(Y^T, Z)^T \underset{1}{\not\rightarrow} X$.

The null hypothesis corresponding to condition C1 is H_{02}: $P_{XY}=0$ & $P_{ZY}=0$ where P_{XY} is as defined previously and $P_{ZY} = [\pi_{1,ZY}, \pi_{2,ZY},..., \pi_{p,ZY}]$, while that corresponding to condition C2 is H_{03}: $P_{XY}=0$ & $P_{XZ}=0$, where $P_{XZ} = [\pi_{1,XZ}, \pi_{2,XZ},..., \pi_{p,XZ}]$; note that the restrictions under test are linear. When one or both of these conditions holds there cannot be indirect causality from Y to X via Z, while failure of both conditions implies that we have either one-step Granger causality, denoted as $Y \underset{1}{\rightarrow} X$, or two-step Granger causality, denoted by $Y \underset{2}{\rightarrow} X$ via Z; that is, $Y \underset{(2)}{\rightarrow} X$. As Z contains only one variable it directly follows that two-step GNC implies GNC for all horizons, as there are no additional causal patterns, internal to Z, which can result in indirect causality from Y to X. For example, a trivariate model falls into this case.

3. $n_Z > 1$. This case is more complicated as causality patterns internal to Z must also be taken into account. If Z can be appropriately partitioned as $Z = (Z_1^T, Z_2^T)^T$ such that $(Y^T, Z_2^T)^T \underset{1}{\not\rightarrow} (X^T, Z_1^T)^T$, then this is sufficient for $Y \underset{(\infty)}{\not\rightarrow} X$. Intuitively, this result follows because the components of Z that can be caused by Y (those in Z_2) do not cause X so that indirect causality cannot arise at longer horizons. Typically, the zero restrictions necessary to test this sufficient condition in the VAR representation are nonlinear functions of the coefficients. Dufour and Renault (1998: 1117) note that such "restrictions can lead to Jacobian matrices of the restrictions having less than full rank under the null hypothesis", which may lead to nonstandard asymptotic null distributions for test statistics.

For these reasons, and the preliminary nature of our study, we limit our attention to univariate Z and, as the applied literature is dominated by cases with $n_X=n_Y=1$, consequently to a trivariate VAR model; recent empirical examples of the use of such models to examine for GNC include Friedman and Kuttner (1992), Kholdy (1995), Henriques and Sadorsky (1996), Lee et al. (1996), Riezman et al. (1996), Thornton (1997), Cheng (1999), Black et al. (2000) and Krishna et al. (2000), among many others.

C. Null Hypotheses, Test Statistics and Limiting Distributions

The testing strategies we propose to examine for two-step GNC within a trivariate VAR framework involve the hypotheses H_{01}, H_{02} and H_{03} detailed in the previous sub-section and the following conditional null hypotheses:

H_{04}: $P_{ZY}=0|P_{XY}=0$;

H_{05}: $P_{XZ}=0|P_{XY}=0$.

The null hypotheses involve linear restrictions on the coefficients of the VAR model and their validity can be examined using various methods, including Wald statistics, LR statistics and model selection criteria. We limit attention to the use of Wald statistics, though we recognize that other approaches may be preferable; this remains for future exploration.

Each of the Wald statistics that we consider in Section III is obtained from an appropriate model, for which the lag length p must be determined prior to testing; the selection of p is considered in Section III below. In general, consider a model where θ is an $(m\times1)$ vector of parameters and let R be a known nonstochastic $(q\times m)$ matrix with rank q. To test H_0: $R\theta=0$, a Wald statistic is

$$W = T\,\hat{\theta}^{T}R^{T}\{R\,\hat{V}[\hat{\theta}]R^{T}\}^{-1}R\,\hat{\theta} \qquad (2)$$

where $\hat{\theta}$ is a consistent estimator of θ and $\hat{V}[\hat{\theta}]$ is a consistent estimator of the asymptotic variance-covariance matrix of $\sqrt{T}(\hat{\theta}-\theta)$. Given appropriate conditions, W is asymptotically distributed as a $\chi^2(q)$ variate under H_0.

The conditions needed for W's null limiting distribution are not assured here, as y_t may be nonstationary with possible cointegration. Sims et al. (1990) and Toda and Phillips (1993, 1994) show that W is asymptotically distributed as a χ^2 variate under H_0 when y_t is stationary or it is nonstationary with "sufficient" cointegration; otherwise W has a nonstandard limiting null distribution that may involve nuisance parameters. The basic problem with nonstationarity is that a singularity may arise in the asymptotic distribution of the least squares (LS) estimators, as some of the coefficients or linear combinations of them may be estimated more efficiently with a faster convergence rate than $\sqrt{T}$. Unfortunately, there seems no basis for testing for "sufficient" cointegration within the VAR model, so that Toda and Phillips (1993, 1994) recommend that, in general, GNC tests should not be undertaken using a VAR model when y_t is nonstationary.

One possible solution is to map the VAR model to its equivalent vector error correction model (VECM) form and to undertake the GNC tests within this framework; see Toda and Phillips (1993, 1994). The problem then is accurate determination of the cointegrating rank, which is known to be difficult with the currently available cointegration tests due to their low power properties and sensitivity to the specification of other terms in the model, including lag order and deterministic trends. Giles and Mirza (1999) use simulation experiments to illustrate the impact of this inaccuracy on the finite sample performance of one-step GNC tests; they find that the practice of pretesting for cointegration can often result in severe over-rejections of the noncausal null.

An alternative approach is the "augmented lags" method suggested by Toda and Yamamoto (1995) and Dolado and Lütkepohl (1996), which results in asymptotic χ^2 null distributions for Wald statistics of the form we are examining, irrespective of the system's integration or cointegration properties. They show that overfitting the VAR model by the highest order of integration in the system eliminates the covariance matrix singularity problem. Consider the augmented VAR model

$$y_t = \sum_{i=1}^{p} \Pi_i y_{t-i} + \sum_{i=1}^{d} \Pi_{p+i} y_{t-p-i} + \varepsilon_t \qquad (3)$$

where we assume that y_t is at most integrated of order d (I(d)). Then, Wald test statistics based on testing restrictions involving the coefficients contained in $\Pi_1,...\Pi_p$ have standard asymptotic χ^2 null distributions; see Theorem 1 of Dolado and Lütkepohl (1996) and Theorem 1 of Toda and Yamamoto (1995). This approach will result in a loss in power, as the augmented model contains superfluous lags and we are ignoring that some of the VAR coefficients, or at least linear combinations of them, can be estimated more efficiently with a higher than usual rate of convergence. However, the simulation experiments of Dolado and Lütkepohl (1996), Zapata and Rambaldi (1997) and Giles and Mirza (1999) suggest that this loss is often minimal with it often resulting in more accurate GNC outcomes than the VECM method, which conditions on the outcome of preliminary cointegration tests. Accordingly, we limit our attention to undertaking Wald tests using the augmented model (3).

Specifically, assuming a trivariate augmented model with $n_X=n_Y=n_Z=1$, let θ be the 9(p+d) vector given by $\theta=\text{vec}[\Pi_1, \Pi_2, ..., \Pi_{p+d}]$, where vec denotes the vectorization operator that stacks the columns of the argument matrix. The LS estimator of θ is $\hat{\theta}$. Then, the null hypothesis H_{01}: $P_{XY}=0$ can be examined using the Wald statistic given by (2) with R being a selector matrix such

that $R\theta=vec[P_{XY}]$. We denote the resulting Wald statistic as W_1. Under our assumptions, W_1 is asymptotically distributed as a $\chi^2(p)$ variate under H_{01}.

The Wald statistic for examining H_{02}: $P_{XY}=0$ & $P_{ZY}=0$, denoted W_2, is given by (2) where $\hat{\theta}$ is the estimator of $\theta=vec[\Pi_1, \Pi_2, ..., \Pi_{p+d}]$ and R is a selector matrix such that $R\theta=vec[P_{XY}, P_{ZY}]$. Under our assumptions, the statistic W_2 has a limiting $\chi^2(2p)$ null distribution. Similarly, let W_3 be the Wald statistic for examining H_{03}: $P_{XY}=0$ & $P_{XZ}=0$. The statistic W_3 is then given by (2) with the selector matrix chosen to ensure that $R\theta=vec[P_{XY}, P_{XZ}]$; under our assumptions W_3 is an asymptotic $\chi^2(2p)$ variate under H_{03}.

We denote the Wald statistics for testing the null hypotheses H_{04}: $P_{ZY}=0|P_{XY}=0$ and H_{05}: $P_{XZ}=0|P_{XY}=0$ as W_4 and W_5 respectively. These statistics are obtained from the restricted model that imposes $P_{XY}=0$; then let θ^* be the vector of remaining unconstrained parameters and $\hat{\theta}^*$ be the LS estimator of θ^*, so that to test the conditional null hypothesis H_0: $R^*\theta^*=0$, a Wald statistic is $W^* = T\,\hat{\theta}^{*T}R^{*T}\{R^*\hat{V}[\hat{\theta}^*]R^{*T}\}^{-1}R^*\hat{\theta}^*$, where $\hat{V}[\hat{\theta}^*]$ is a consistent estimator of the asymptotic variance-covariance matrix of $\sqrt{T}(\hat{\theta}^*-\theta^*)$. Under our assumptions, W^* is asymptotically distributed as a χ^2 variate under H_0 with the rank of R^* determining the degrees of freedom; Theorem 1 of Dolado and Lütkepohl (1996) continues to hold in this restricted case as the elements of $\Pi_{p+1},...,\Pi_{p+d}$ are not constrained under the conditioning or by the restrictions under test. Our statistics W_4 and W_5 are then given by W^* with, in turn, the vector $R^*\theta^*$ equal to $vec(P_{ZY})$ and $vec(P_{XZ})$; in each case p is the degrees of freedom. We now turn, in the next section, to detail our proposed testing strategies.

III. PROPOSED SEQUENTIAL TESTING STRATEGIES AND MONTE CARLO DESIGN

A. Sequential Testing Strategies

Within the augmented model framework, the sequential testing procedures we consider are:

(M1)

$$\text{Test } H_{02} \text{ \& test } H_{03} \begin{cases} \text{If } H_{02} \text{ \& } H_{03} \text{ are rejected, reject the null hypothesis of } Y \underset{(2)}{\not\rightarrow} X. \\ \text{Otherwise, support the null hypothesis of } Y \underset{(2)}{\not\rightarrow} X. \end{cases}$$

(M2)

$$\text{Test } H_{01} \begin{cases} \text{If } H_{01} \text{ is rejected, reject the null hypothesis of } Y \underset{1}{\not\rightarrow} X. \\ \text{Otherwise, test } H_{04} \text{ \& test } H_{05} \begin{cases} \text{If } H_{04} \text{ \& } H_{05} \text{ are rejected, reject } Y \underset{2}{\not\rightarrow} X. \\ \text{Otherwise, support the null hypothesis of } Y \underset{2}{\not\rightarrow} X. \end{cases} \end{cases}$$

The strategy (M1) provides information on the null hypothesis H_{0A}: $Y \underset{(2)}{\not\rightarrow} X$ as it directly tests whether the conditions C1 and C2, outlined in the previous section, are satisfied. Each hypothesis test tests 2p exact linear restrictions. The approach (M1) does not distinguish between one-step and two-step GNC; that is, rejection here does not inform the researcher as to whether the causality arises directly from Y to X one-step ahead or whether there is GNC one-step ahead with the causality then arising indirectly via the auxiliary variable at horizon two. This is not a relevant concern when interest is in only answering the question of the presence of Granger causality at any horizon.

The strategy (M2), on the other hand, provides information on the horizon at which the causality, if any, arises. The first hypothesis test, H_{01}, undertakes the usual one-step test for direct GNC

from Y to X, which when rejected implies that no further testing is required as causality is detected. However, the possibility for indirect causality at horizon two via the auxiliary variable Z is still feasible when H_{01} is supported, and hence the second layer of tests that examine the null hypothesis H_{0B}: $Y \underset{2}{\not\rightarrow} X \mid Y \underset{1}{\not\rightarrow} X$. We require both of H_{04} and H_{05} to be rejected for a conclusion of causality at horizon two, while we accept H_{0B} when we support one or both of the hypotheses H_{04} and H_{05}. Each test requires examining the validity of p exact linear restrictions.

That the sub-tests in the two strategies have different degrees of freedom may result in power differences that may lead us to prefer (M2) over (M1). In our simulation experiments described below, we consider various choices of nominal significance level for each sub-test, though we limit each to be identical at say 100α%. We can say the following about the asymptotic level of each of the strategies. In the case of (M1) we are examining nonnested hypotheses using statistics that are not statistically independent. When both H_{02} and H_{03} are true we know from the laws of probability that the level is smaller than 200α%, though we expect to see asymptotic levels less than this upper bound dependent on the magnitude of the probability of the union of the events. When one of the hypotheses is false and the other is true, so that $Y \underset{(2)}{\not\rightarrow} X$ still holds, the asymptotic level for strategy (M1) is smaller than or equal to 100α%.

An asymptotic level of 100α% applies for strategy (M2) for testing for $Y \underset{1}{\not\rightarrow} X$, while it is at most 200α% for testing for H_{0B} when H_{04} and H_{05} are both true, and the level is at most 100α% when only one is true. As noted with strategy (M1), when both hypotheses are true we would expect levels that are much smaller than 200α%. Note also that W_4 and W_5 are not (typically) asymptotically independent under their respective null hypotheses, though the statistics W_4 and W_1 are asymptotically independent under H_{04} and H_{01}, given the nesting structure.

Finally, we note that the testing strategies are consistent, as each component test is consistent, which follows directly from the results presented by e.g., Dolado and Lütkepohl (1996: 375).

B. Monte Carlo Experiment

In this sub-section, we provide information on our small-scale simulation design that we use to examine the finite sample performance of our proposed sequential test procedures. We consider two basic data generating processes (DGPs), which we denote as DGPA and DGPB; for each we examine three cases: DGPA1, DGPA2, DGPA3, DGPB1, DGPB2 and DGPB3. To avoid potential confusion, we now write y_t as $y_t=[y_{1t}, y_{2t}, y_{3t}]^T$ to describe the VAR system and the GNC hypotheses we examine; we provide the mappings to the variables X, Y and Z from the last sub-section in a table. For all cases the series are I(1). The first DGP, denoted as DGPA, is

$$\begin{bmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0.5 & 0 \\ a & -0.5 & 1 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \\ y_{3t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix}$$

We consider a=1 for DGPA1, which implies one cointegrating vector. The parameter a is set to zero for DGPA2 and DGPA3, which results in two cointegrating vectors among the I(1) variables. Toda and Phillips (1994) and Giles and Mirza (1999) also use this basic DGP. The null hypotheses H_{01} to H_{05} are true for the GNC effects we examine for DGPA1 and DGPA3. Using Corollary 1 of Toda and Phillips (1993, 1994) it is clear that there is insufficient cointegration with respect to the variables whose causal effects are being studied for all hypotheses except H_{05}. That is, Wald statistics for H_{01} to H_{04} applied in the non-augmented model (1) do not have their usual χ^2 asymptotic null distributions, though the Wald statistics we use in the augmented model (3) have standard limiting null distributions for all the null hypotheses.

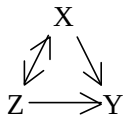
Our second basic DGP, denoted as DGPB, is

$$\begin{bmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{bmatrix} = \begin{bmatrix} 0.32 & b & 0.44 \\ 0 & 1.2325 & 0.31 \\ 0 & -0.285 & 0.62 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \\ y_{3t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix}$$

We set the parameter b=-0.265 for DGPB1 and b=0 for DGPB2 and DGPB3; for each DGP there is two cointegrating vectors. Zapata and Rambaldi (1997) and Giles and Mirza (1999) also use a variant of this DGP in their experiments. The null hypotheses H_{01}, H_{02} and H_{04} are each true for DGPB1 and DGPB2, while H_{03} and H_{05} are false. The cointegration for this DGP is sufficient with respect to the variables whose causal effects are being examined so that standard Wald statistics in the non-augmented model for H_{01}, H_{02} and H_{04} are asymptotic χ^2 variates under their appropriate null hypotheses.

We provide summary information on the DGPs in Table 1. We include the mappings to the variables X, Y and Z used in our discussion of the sequential testing strategies in the last sub-section, the validity of the null hypotheses H_{01} to H_{05}, and the GNC outcomes of interest. Though our range of DGPs is limited, they enable us to study the impact of various causal patterns on the finite sample performance of our strategies (M1) and (M2).

The last row of the table provides the '1-step GC map' for each DGP, which details the pair-wise 1-step Granger causal (GC) patterns. For example, the 1-step GC map for DGPB1 in terms of X, Y and Z is



This causal map is an example of a 'directed graph', because the arrows lead from one variable to another; they indicate the presence of 1-step GC, e.g., $X \underset{1}{\rightarrow} Y$. In addition to being a useful way to summarize the 1-step GC relationships, the 1-step GC map allows us to visualize the possibilities for 2-step GC. To illustrate, we consider the GC relations from Y to X. The map indicates that $Y \underset{1}{\not\rightarrow} X$, so that for $Y \underset{2}{\rightarrow} X$ we require $Y \underset{1}{\rightarrow} Z$ and $Z \underset{1}{\rightarrow} X$; directly from the map we see that the latter causal relationship holds but not the former – Z is not operating as an auxiliary variable through which 2-step GC from Y to X can occur.

As our aim is to examine the finite-sample performance of (M1) and (M2) at detecting 2-step GNC, and at distinguishing between GC at the different horizons, each of our DGPs imposes that $Y \underset{1}{\not\rightarrow} X$, but for two DGPs – DGPA2 and DGPB3 – $Y \underset{2}{\rightarrow} X$, while for the other DGPs we have $Y \underset{2}{\not\rightarrow} X$. This allows us to present rejection frequencies when the null hypothesis $Y \underset{2}{\not\rightarrow} X$ (or $Y \underset{(2)}{\not\rightarrow} X$) is true and false.

When the null is true we denote the rejection frequencies as FI(α), because they estimate the probability that the testing strategy makes a Type I error when the nominal significance level is α, that is, the probability that the testing strategy rejects a correct null hypothesis. The strategy (M1) provides information on the null hypothesis H_{0A}: $Y \underset{(2)}{\not\rightarrow} X$, which is true when we support both H_{02} and H_{03} or when only one is accepted. Let PIA(α) be the probability of a Type I error associated with testing H_{0A}, at nominal significance level α, using (M1), so PIA(α) = $\Pr_{H_{0A}}$(reject H_{03} & reject H_{02}|$Y \underset{(2)}{\not\rightarrow} X$). We estimate PIA($\alpha$) by FIA($\alpha$) = $N^{-1} \sum_{i=1}^{N} I(P_{3i} \leq \alpha \;\&\; P_{2i} \leq \alpha)$, where N denotes the number of Monte Carlo samples, P_{ji} is the i^{th} Monte Carlo sample's P

value associated with testing H_{0j} that has been calculated from a $\chi^2(2p)$ distribution using the statistic W_j for j=2,3, and $I(\cdot)$ is the indicator function.

We desire FIA(α) to be 'close' to the asymptotic nominal level for the strategy, which is bounded by 100α% when only one of H_{02} or H_{03} is true and by 200α% when both are true, though FIA(α) will differ from the upper bound because of our use of a finite number of simulation experiments and the asymptotic null distribution to calculate the P values; the latter is used because our statistics have unknown finite sample distributions. In particular, our statistics, though asymptotically pivotal, are not pivotal in finite samples, so FIA(α) (and also PIA(α)) depends on where the true DGP is in the set specified by H_{0A}. One way of solving the latter problem is to report the size of the testing strategy, which is the supremum of the PIA(α) values over all DGPs contained in H_{0A}. In principle, we could estimate the size as the supremum of the FIA(α) values, though this task is in reality infeasible here because of the multitude of DGPs that can satisfy H_{0A}. Accordingly, we report FIA(α) values for a range of DGPs.

The testing strategy (M2) provides information on two null hypotheses: H_{01}: $Y \underset{1}{\not\to} X$ and H_{0B}: $Y \underset{2}{\not\to} X | Y \underset{1}{\not\to} X$. The statistic W_1, used to test H_{01}, has asymptotic level 100α%. We estimate the associated probability of a Type I error, denoted as $PI1(\alpha) = Pr_{H_{01}}(\text{reject } H_{01})$, by $FI1(\alpha) = N^{-1} \times \sum_{i=1}^{N} I(P_{1i} \le \alpha)$, where P_{1i} is the P value associated with testing H_{01} for the i^{th} Monte Carlo sample and generated from a $\chi^2(p)$ distribution, α is the assigned nominal level, and $I(\cdot)$ is the indicator function. Further, let PIB(α) denote the probability of a Type I error for using (M2) to test H_{0B}, and let $FIB(\alpha) = N_{1*}^{-1} \sum_{i=1}^{N_{1*}} I(P_{5i} \le \alpha \ \& \ P_{4i} \le \alpha)$ be an estimator of PIB(α), where N_{1*} is the number

of Monte Carlo samples that accepted H_{01}, and P_{ji} is the P value for testing H_{0j} for the i^{th} Monte Carlo sample calculated from a $\chi^2(p)$ distribution, j=4,5. We report FI1(α) and FIB(α) values for our DGPs that satisfy H_{01} and H_{0B}. Recall that the upper bound on PIB(α) is 100a% when only one of H_{04} or H_{05} is true and it is 200α% when both are true.

We also report rejection frequencies for the strategies when $Y \underset{2}{\rightarrow} X$; these numbers aim to indicate 'powers' of our testing strategies. In econometrics, there is much debate on how powers should be estimated with many studies advocating that only so-called 'size-corrected' power estimates be provided. Given the breadth of the set under the null hypotheses of interest here, it is computationally infeasible to contemplate obtaining 'size-corrected' power estimates for our study. Some researchers approach this problem by obtaining a critical value for their particular DGP that ensures that the nominal and true probability of a Type I error are equal; they then claim, incorrectly, that the reported powers are 'size-corrected'. A further complication in attempting to provide estimates of size-corrected power is that the size-corrected critical value may often be infinite, which implies zero power for the test (e.g, Dufour, 1997).

Given these points, when dealing with a composite null hypothesis, Horowitz and Savin (2000) suggest that it may be preferable to form power estimates from an estimate of the Type I critical value that would be obtained if the exact finite-sample distribution of the test statistic under the true DGP were known. One such estimator is the asymptotic critical value, though this estimator may not be accurate in finite-samples. Horowitz and Savin (2000) advocate bootstrap procedures to estimate the pseudo-true value of the parameters from which an estimate of the finite sample Type I critical value can be obtained; this approach may result in higher accuracy in finite samples. In our study, given its preliminary nature, we use the asymptotic critical value to estimate the powers of our strategies; we leave the potential application of bootstrap procedures

for future research. We denote the estimated power for strategy (M1) that is associated with testing H_{0A} by $FIIA(\alpha) = N^{-1} \sum_{i=1}^{N} I(P_{3i} \leq \alpha \ \& \ P_{2i} \leq \alpha)$, and the estimated powers for strategy (M2) for testing H_{0B} by $FIIB(\alpha) = N_{1*}^{-1} \sum_{i=1}^{N_{1*}} I(P_{5i} \leq \alpha \ \& \ P_{4i} \leq \alpha)$ respectively.

For each of the cases outlined in Table 1 we examine four net sample sizes: T=50, 100, 200, 400. The number of Monte Carlo simulation repetitions is fixed to be 5000, and for each experiment we generate (T+100+6) observations from which we discard the first 100 to remove the effect of the zero starting values; the other 6 observations are needed for lagging. We limit attention to an identity innovation covariance matrix, though we recognize that this choice of covariance matrix is potentially restrictive and requires further attention in future research[2]. We generate FI and FII values for twelve values of the nominal significance level: α=0.0001, 0.0005, 0.001, 0.005, 0.01, 0.025, 0.05, 0.1, 0.15, 0.2, 0.25, and 0.30.

The conventional way to report the results is by tables, though this approach has two main drawbacks. First, there would be many tables and secondly, it is often difficult to see how changes in the sample size, DGPs and α values affect the rejection frequencies. In this paper, as recommended by Davidson and MacKinnon (1998), we use graphs to provide the information on the performance of our testing strategies. We use P value plots to report the results on FIA, FI1 and FIB; these plot the FI values against a nominal level. In the ideal case, each of our P values would be distributed as uniform (0,1), so that the resulting graph should be close to the 45^{o} line. Consequently, we can easily see when a strategy is over-rejecting, or under-rejecting or rejecting

[2] Note that Yamada and Toda (1998) show that the finite sample distribution of Wald statistics, such as those considered here, are invariant to the form of the innovation covariance matrix when the lag length order of the VAR is known. However, this result no longer holds once we allow for estimation of the lag order.

the right proportion reasonably often. The asymptotic nominal level for DGPB1 and DGPB2 is α for each of H_{01}, H_{0A} and H_{0B}, it is α for H_{01} for DGPA1 and DGPA3, and the upper bound is 2α for H_{0A} and H_{0B}. However, we anticipate levels less than 2α, as this upper bound ignores the probability of the union event, so we use α as the level to provide the 45^o line for the plots.

We provide so-called 'size-power' curves to report the information on the power of our testing strategies; Wilk and Gnanadesikan (1968) and Davidson and MacKinnon (1998). In our case we use FI values rather than size estimates and our power estimates are based on the FII values; accordingly, we call these graphs FI-FII curves. The horizontal axis gives the FI values, computed when the DGP satisfies the null hypothesis, and the vertical axis gives the FII values, generated when the DGP does not satisfy $Y \underset{2}{\not\rightarrow} X$ in a particular way. The lower left hand corner of the curve arises when the strategy always supports the null hypothesis, while the upper right hand corner results when the test always rejects the null hypothesis. When the power of a testing strategy exceeds its associated probability of a Type I error, the FI-FII curve lies above the 45^o line, which represents the points of equal probability.

To undertake the hypothesis tests of interest, we need to choose the lag order for the VAR, which is well known to impact on the performance of tests; see, Lütkepohl (1993b), Dolado and

l (1996), Giles and Mirza (1999), among others. We examine four approaches to specifying the lag order: p is correctly specified at 1, which we denote as the 'True' case; p is always over estimated by 1, that is, p=2, which we denote as 'Over'; and p is selected by two common goodness-of-fit criteria – Schwarz's (1978) Bayesian criterion (SC) and Akaike's (1973) information criterion (AIC). The AIC does not consistently estimate the lag order (e.g., Nishi, 1988; Lütkepohl, 1993b), as there is a positive probability of over estimation of p, which does not, nevertheless, affect consistent estimation of the coefficients, though over estimation may

result in efficiency and power losses. The SC is a consistent estimator of p, though evidence suggests that this estimator may be overly parsimonious in finite-samples, which may have a detrimental impact on the performance of subsequent hypothesis tests. In our experiments that use the AIC and SC we allow p to be at most 6. Though the 'True' case is somewhat artificial, we include it as it can be regarded as a best-case scenario. The over-specified case illustrates results on using a pre-specified, though incorrect, lag order.

IV. SIMULATION RESULTS

A. P Value Plots

Figure 1 shows P value plots for the testing strategy (M1) for DGPA1 when T=50, 100, 200 and 400 and for the four lag selection methods we outlined in the previous section. It is clear that the strategy systematically results in levels that are well below α irrespective of the sample size. For instance, the estimated probability of a Type I error is typically close to $\alpha/2$ when T=200, irrespective of the specified nominal level. In contrast, the strategy seems to work better for the smaller sample of 50 observations.

As expected from our discussion these observed features for DGPA1 differ with DGPB1 and DGPB2 when only one of the hypotheses is true. In Figure 2 we provide the P value plots for DGPB1, from which we see that the strategy (M1) systematically over-rejects, with the degree of over-rejection becoming more pronounced as T falls. Qualitatively, the P value plots for DGPA3 and DGPB2 match those for DGPA1 and DGPB1 respectively, so we do not include them here.

Over-specifying or estimating the lag order does little to alter the P value plots from those for the correctly specified model when T≥200, though there are some observable differences when T is smaller, in particular for T=50. The performance of the AIC and Over cases is always somewhat worse than for the SC, although we expect that this is a feature of the low-order VAR models that we examine for our DGPs. Figure 1 and Figure 2 show that the AIC rejects more often than does Over for small α, but this is reversed for higher levels, say greater than 10%.

We now turn to the P value plots for the procedure (M2). We provide the P value plots for testing H_{01}: $Y \underset{1}{\not\rightarrow} X$ for DGPA3 in Figure 3. This hypothesis forms the first part of strategy (M2) and is undertaken using the Wald statistic W_1. The plots for the other DGPs are qualitatively similar. It is clear that the test systematically over-rejects H_{01}, especially when T≤100 and irrespective of the lag selection approach adopted.

Figure 4 shows P value plots for the second part of strategy (M2) that examines $Y \underset{2}{\not\rightarrow} X | Y \underset{1}{\not\rightarrow} X$ for DGPA1. It is clear that there is a pattern of the FI values being less than α, with the difference increasing with T, irrespective of the lag selection approach. This systematic pattern is similar to that observed for strategy (M1) with this DGP, and also with DGPA3. Like then we do not observe this feature with DGPB1 and DGPB2, as we see from Figure 5 that shows P value plots for DGPB1, which are representative for both DGPB1 and DGPB2. Here the testing procedure works well, though there is a small tendency to over-reject for T≥100, while only the Over approach leads to over-rejection when T=50 with the other lag selection cases slightly under-rejecting. These over- and under-rejections become more pronounced as the nominal level rises.

B. FI-FII Plots

Figure 6 presents FIA-FIIA curves associated with strategy (M1) generated from DGPA1 (for which the null $Y \underset{(2)}{\not\rightarrow} X$ is true) and DGPA3 (for which the null $Y \underset{2}{\not\rightarrow} X$ is false in a particular way) when T=50 and T=100; the consistency property of the testing procedure is well established for this DGP and degree of falseness of the null hypothesis by T=100, so we omit the graphs for T=200 and T=400. Several results are evident from this figure. We see good power properties irrespective of the sample size, though this is a feature of the chosen value for the parameter "a" for DGPA3. Over-specifying the lag order by a fixed amount does not result in a loss in power, while there is a loss associated with use of the AIC.

The FIA-FIIA curves for strategy (M1) generated from DGPB2 (for which the null $Y \underset{(2)}{\not\rightarrow} X$ is true) and DGPB3 (for which the null $Y \underset{2}{\not\rightarrow} X$ is false in a particular way) when T=50 and T=100 are given in Figure 7; we again omit the curves for T=200 and T=400 as they merely illustrate the consistency feature of the strategy. In this case, compared to that presented in Figure 6, the FII levels are lower for a given FI level, which reflects the degree to which the null hypothesis is false. Nevertheless, the results suggest that the strategy does well at rejecting the false null hypothesis, especially when T≥100.

The FIB-FIIB curves for strategy (M2) for testing H_{0B} display qualitatively similar features to those just discussed for strategy (M1). To illustrate, Figure 8 provides the curves generated from DGPB2 and DGPB3; figures for the other cases are available on request. It is of interest to compare Figure 7 and Figure 8. As expected and irrespective of sample size, there are gains in

power from using strategy (M2) over (M1) when $Y \underset{1}{\not\rightarrow} X$ but $Y \underset{2}{\rightarrow} X$, as the degrees of freedom for the former strategy are half those of the latter.

Overall, we conclude that the strategies perform well at detecting a false null hypothesis, even in relatively small samples. Our results illustrate the potential gains in power with strategy (M2) over strategy (M1) when $Y \underset{1}{\not\rightarrow} X$ and $Y \underset{2}{\rightarrow} X$. In practice, this is likely useful as we anticipate that most researchers will first test for $Y \underset{1}{\not\rightarrow} X$, and only proceed to a second stage test for $Y \underset{2}{\not\rightarrow} X$ when the first test is not rejected. There is some loss in power in using the AIC to select the lag order compared with the other approaches we examine, though a study of DGPs with longer lag orders may alter this outcome. Our results support the finding from other studies that lag order selection is important.

V. EMPIRICAL EXAMPLE

To illustrate the application of the testing strategies on the outcome of GNC tests in trivariate systems, we have re-examined the data set used by Hoffman and Rasche (1996), which enables us to consider the well-studied issue of the causal relationships between money and income. We downloaded the quarterly, seasonally adjusted US time series data from the *Journal of Applied Econometrics* Data Archive; the data are originally obtained from Citibase. Let X be real money balances, which is calculated by deflating the nominal series by the GDP deflator; let Y be real income, represented by real GDP; and let Z be nominal interest rates, the auxiliary variable, which is represented by the Treasury bill rate. Both real balances and real GDP are expressed in natural logarithms. Allowing for lagging, we use observations from 1950:3 to 1995:2 (164 observations). Our Monte Carlo study ignored the realistic possibility that there may be deterministic trends, which we incorporate here by extending model (1) as:

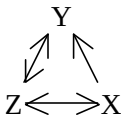
$$(y_t-\mu-\delta t) = \sum_{i=1}^{p} \Pi_i \, (y_{t-i}-\mu-\delta(t-i)) + \varepsilon_t \qquad (4)$$

where μ and δ are vectors of unknown coefficients. We can write (4) equivalently as

$$y_t = \mu^* + \delta^* t + \sum_{i=1}^{p} \Pi_i \, y_{t-i} + \varepsilon_t \qquad (5)$$

where $\mu^* = (-\Pi\mu+\Pi^*\delta)$, $\delta^* = -\Pi\delta$, $\Pi^* = \sum_{i=1}^{p} i\Pi_i$, and $\Pi = -\left(I_3 - \sum_{i=1}^{p} \Pi_i \right)$. The matrix Π is the usual potentially reduced rank matrix that indicates the number of cointegrating relationships. Often applied researchers impose $\delta=0$ a priori, which implies that μ^* is forced to zero when there is no cointegration; this seems a limiting restriction, so we use (5) as stated.

We assume that each of our time series is integrated of order one, as does also Hoffman and Rasche (1996); this implies that we augment our model with one extra lag of y. We use the AIC to choose the lag order p allowing for up to ten possible lags; the results support six lags. Table 2 reports asymptotic P values for examining for the six possible 1-step GNC relationships. Using a (nominal) 10% significance level the outcomes imply that the 1-step GC map is



from which it is clear that there is support for the common finding that real money balances Granger causes real income without feedback. However, the map illustrates the potential for 2-step GC from real income to real money balances to arise indirectly via the interest rate variable. To explore this possibility we apply our testing strategies (M1) and (M2); the asymptotic P values are reported in Table 2. Both strategies support 2-step Granger causality from real income to real money balances via the interest rate variable. The example serves to illustrate the changes in causality conclusions that may occur once an allowance is made for indirect causality via an auxiliary variable.

V. CONCLUDING REMARKS

This paper has provided two testing strategies for examining for 2-step Granger noncausality in a trivariate VAR system that may arise via the auxiliary variable that is included for explanatory power, but is not directly involved in the noncausality test under examination. The testing strategy (M1) can be applied without a prior test for the traditional 1-step noncausality as it tests for noncausality up to horizon two; however, this strategy does not provide information on whether detected causality arises directly one-step ahead or indirectly two-steps ahead. The strategy (M2) provides information on the latter, and is perhaps the more practically useful of the two procedures that we propose.

We investigated the finite sample properties of our sequential strategies through Monte Carlo simulations. Though the data generating processes that we employed were relatively simple and should be expanded on in a more elaborative study, our findings may be summarised as follows:

(i) The testing procedures perform reasonably well, irrespective of sample size and lag selection method.

(ii) The form of the underlying DGP can impact substantially on the probability of the Type I error associated with that DGP. The actual level is closer to a notion of an assigned level when the auxiliary variable is sufficiently causal with one of the variables under test. That is, when testing for $Y \underset{2}{\not\rightarrow} X$ via Z there are two causal relationships of interest: $Y \underset{1}{\not\rightarrow} Z$ and $Z \underset{1}{\not\rightarrow} X$. If only one of these is true, so that either $Y \underset{1}{\rightarrow} Z$ or $Z \underset{1}{\rightarrow} Y$, then our results suggest that we can have some confidence about the probability of the Type I that we are likely to observe. However, when both are true so that Z is not sufficiently

involved with X or Y, then the probability of the Type I error can be small and quite different from any nominal notion of a level that we may have.

(iii) The testing strategies do well at detecting when $Y \not\xrightarrow[2]{} X$ is false irrespective of the form of the DGP. Our results suggest that a sample size of at least 100 is preferable, though this depends, naturally, on the degree to which the null is false.

(iv) The choice of the lag length in the performance of the test is important. This issue requires further attention with DGPs of higher lag order.

An obvious extension of this work is to the development of tests that examine for multi-step noncausality in higher dimensional systems. This will typically involve testing for zero restrictions on multilinear functions of the VAR coefficients, which may result in Jacobian matrices of the restrictions having less that full rank under the null hypothesis and so lead to nonstandard asymptotic null distributions.

TABLE 1 DGP Descriptions

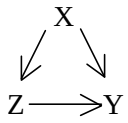
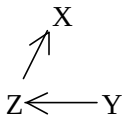
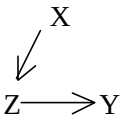
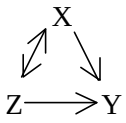
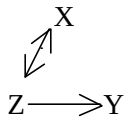
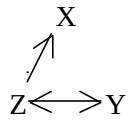
	DGPA1	DGPA2	DGPA3	DGPB1	DGPB2	DGPB3
X	y_1	y_3	y_1	y_2	y_2	y_1
Y	y_3	y_1	y_3	y_1	y_1	y_2
Z	y_2	y_2	y_2	y_3	y_3	y_3
H_{01}: $P_{XY}=0$	True	True	True	True	True	True
H_{02}: $P_{XY}=0 \& P_{ZY}=0$	True	False	True	True	True	False
H_{03}: $P_{XY}=0 \& P_{XZ}=0$	True	False	True	False	False	False
H_{04}: $P_{ZY}=0 \mid P_{XY}=0$	True	False	True	True	True	False
H_{05}: $P_{XZ}=0 \mid P_{XY}=0$	True	False	True	False	False	False
$Y \overset{1}{\not\to} X$	Yes	Yes	Yes	Yes	Yes	Yes
$Y \overset{2}{\not\to} X$	Yes	No	Yes	Yes	Yes	No
1-step GC map						

TABLE 2 Money-Income Example P Values

NULL HYPOTHESIS	P VALUE
$Y \underset{1}{\not\rightarrow} X$	0.170
$Z \underset{1}{\not\rightarrow} X$	<0.001
$X \underset{1}{\not\rightarrow} Y$	0.001
$Z \underset{1}{\not\rightarrow} Y$	<0.001
$X \underset{1}{\not\rightarrow} Z$	0.090
$Y \underset{1}{\not\rightarrow} Z$	0.086
<u>Strategy (M1)</u>	
H_{02}	0.070
H_{03}	<0.001
<u>Strategy (M2)</u>	
H_{04}	<0.001
H_{05}	0.095

Figure 1. P value plots for strategy (M1), FIA, DGPA1

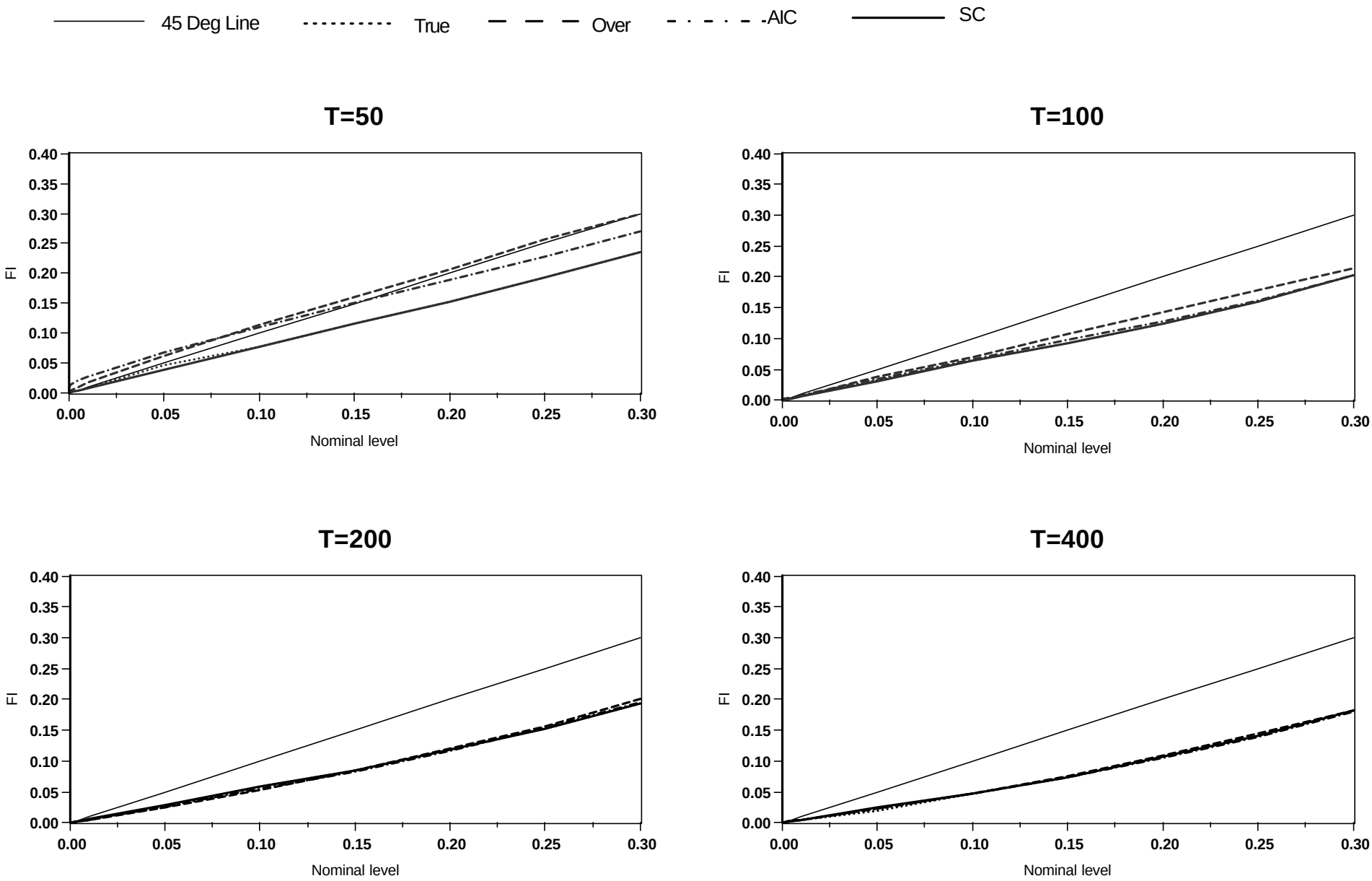


Figure 2. P value plots for strategy (M1), FIA, DGPB1

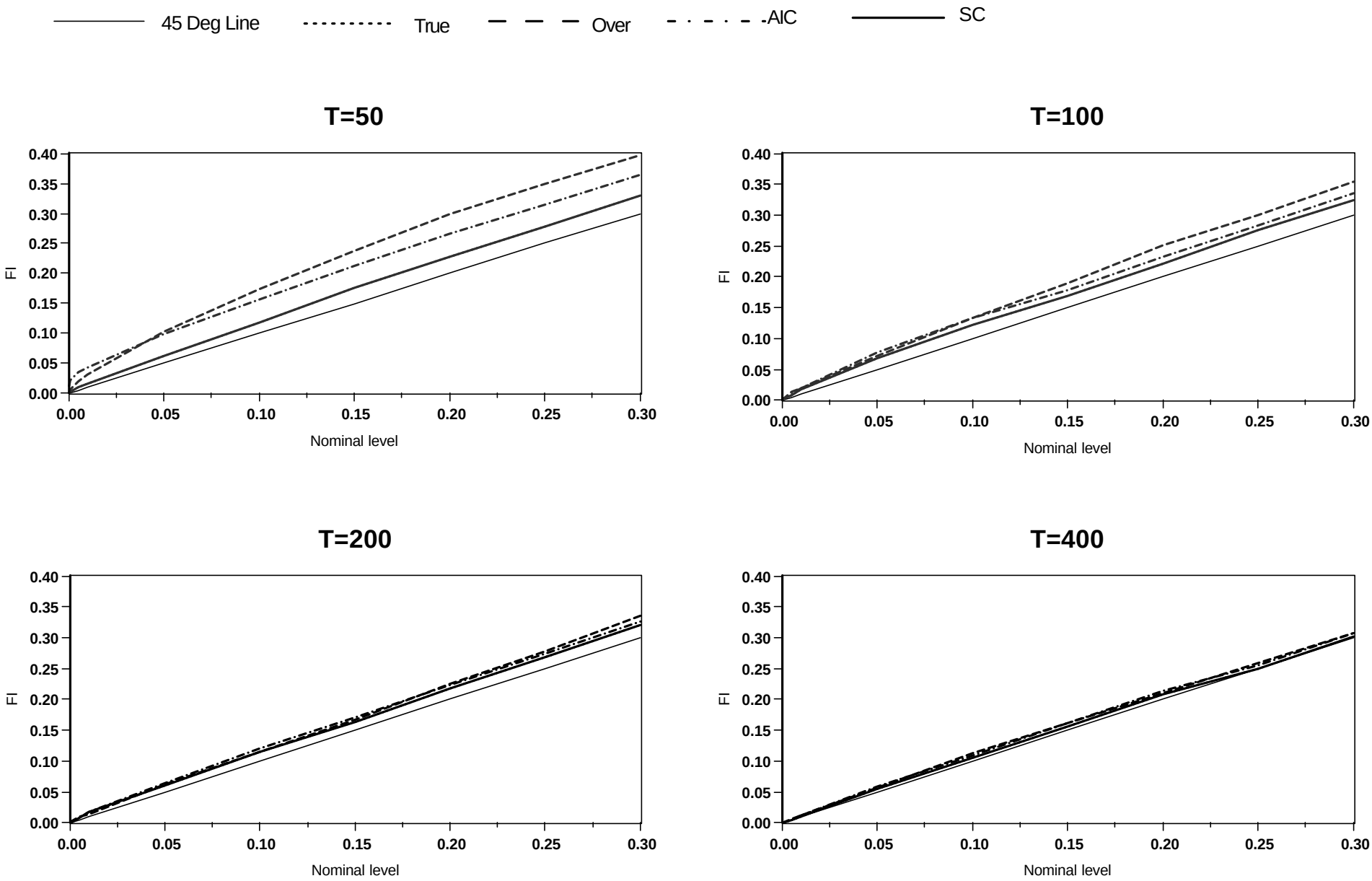


Figure 3. P value plots for strategy (M2), FI1, DGPA3

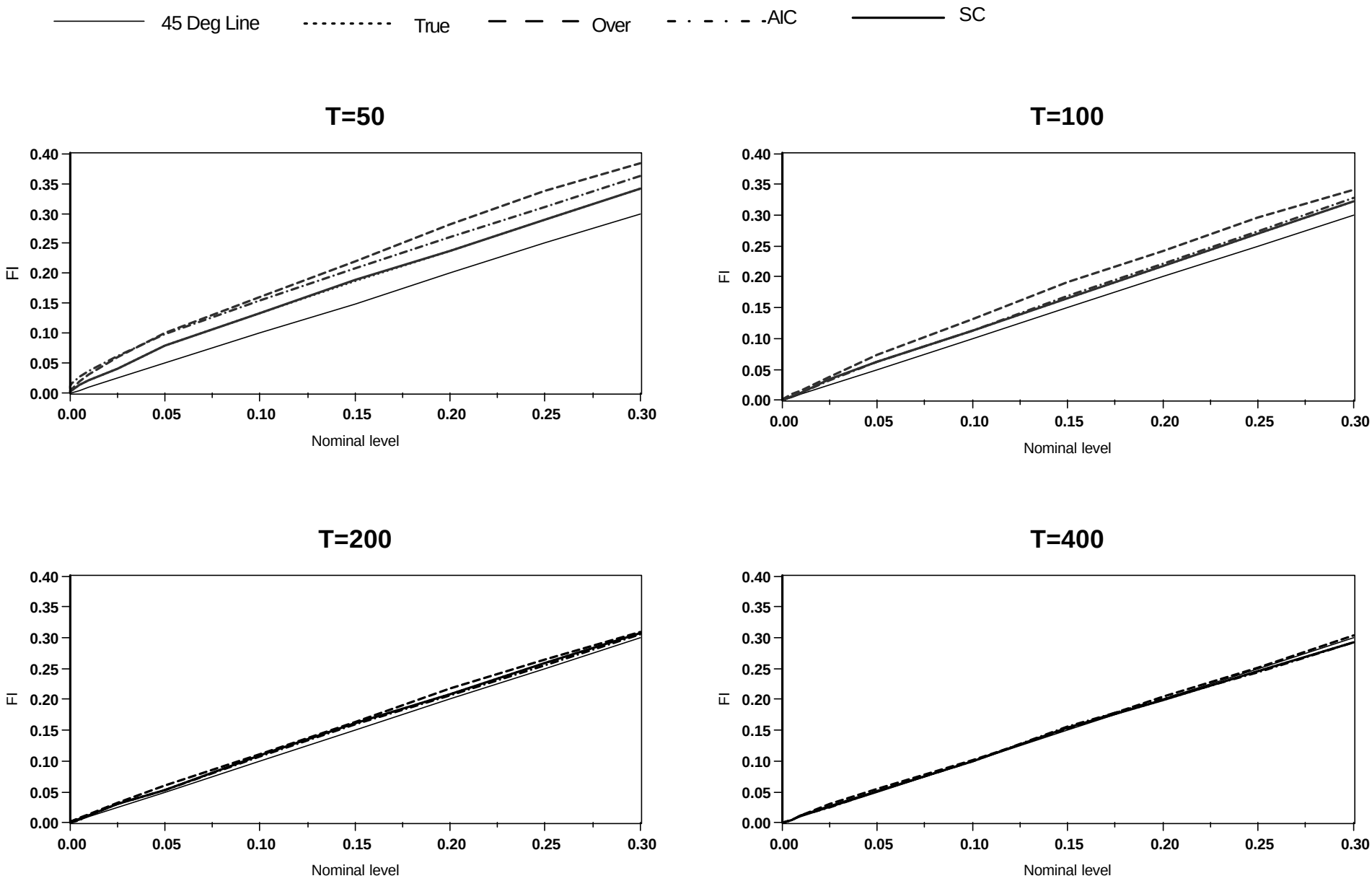


Figure 4. P value plots for strategy (M2), FIB, DGPA1

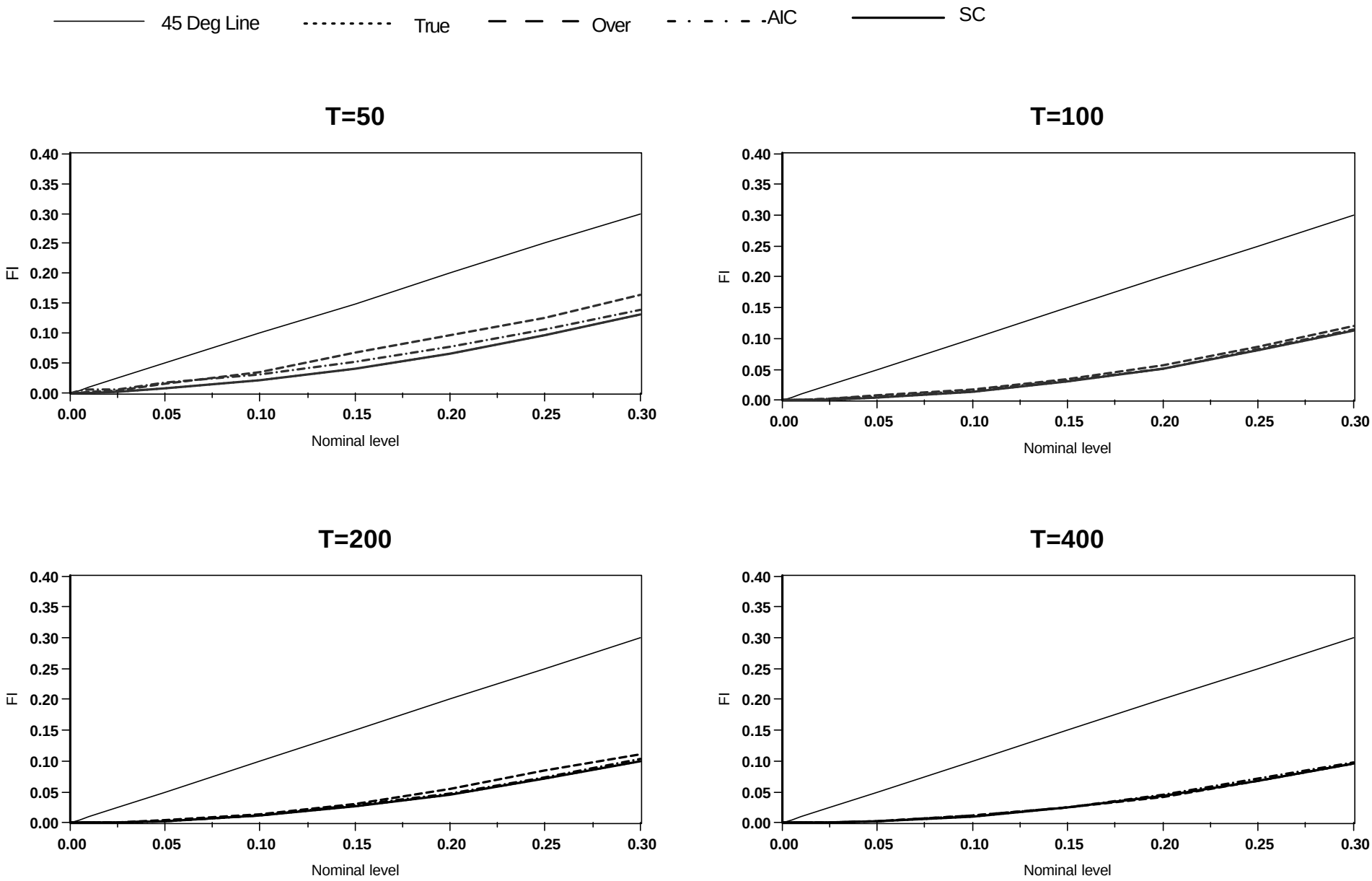


Figure 5. P value plots for strategy (M2), FIB, DGPB1

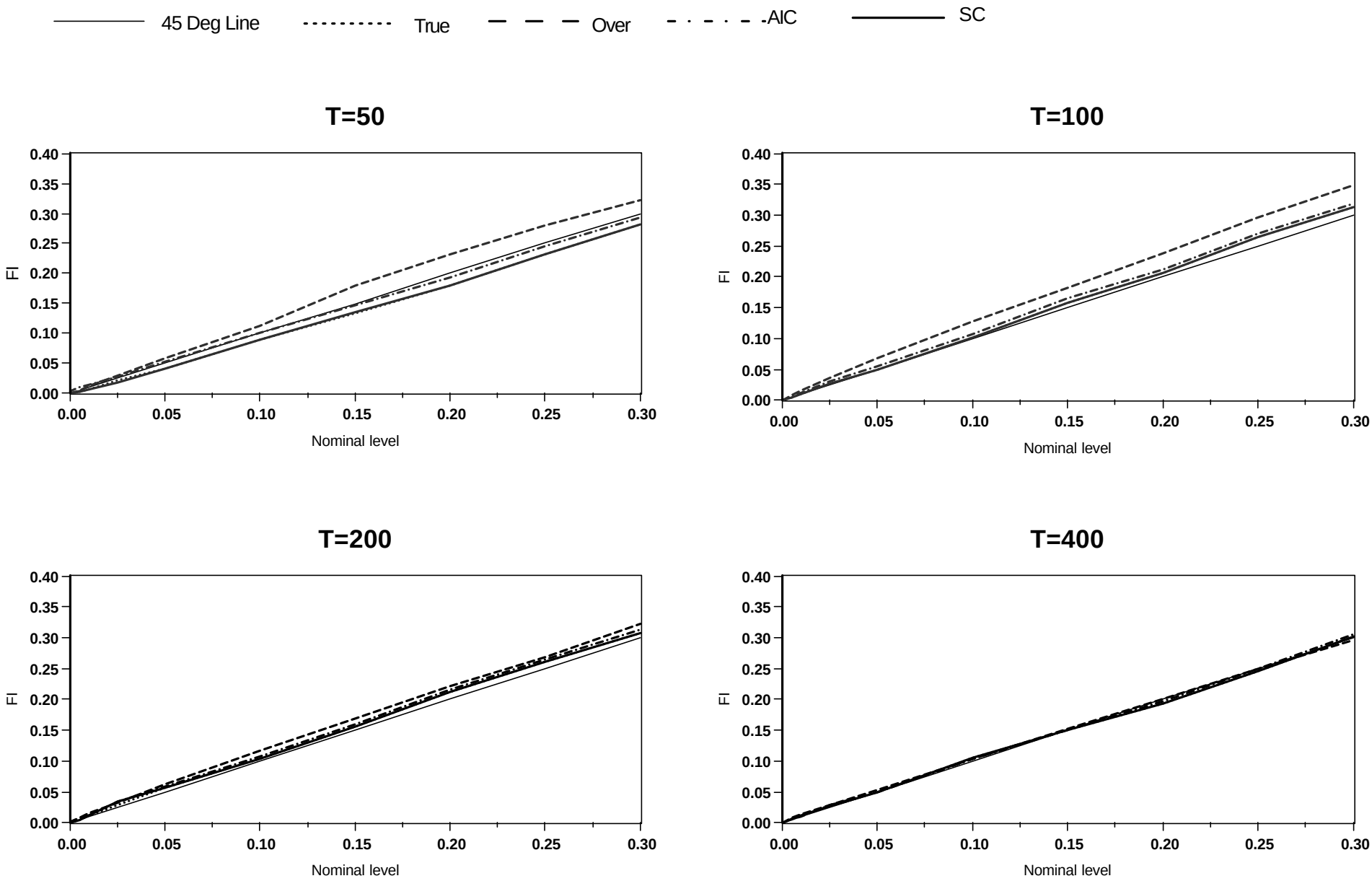


Figure 6. FIA-FIIA plots for strategy (M1), DGPA1 & DGPA2

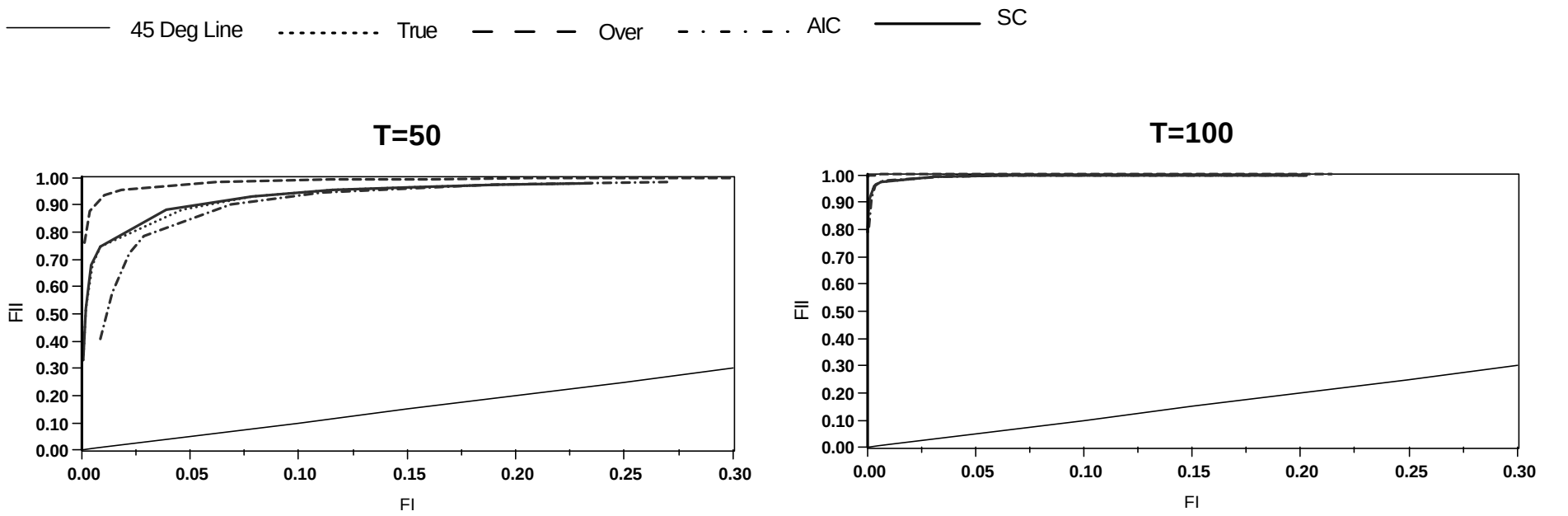


Figure 7. FIA-FIIA plots for strategy (M1), DGPB2 & DGPB3

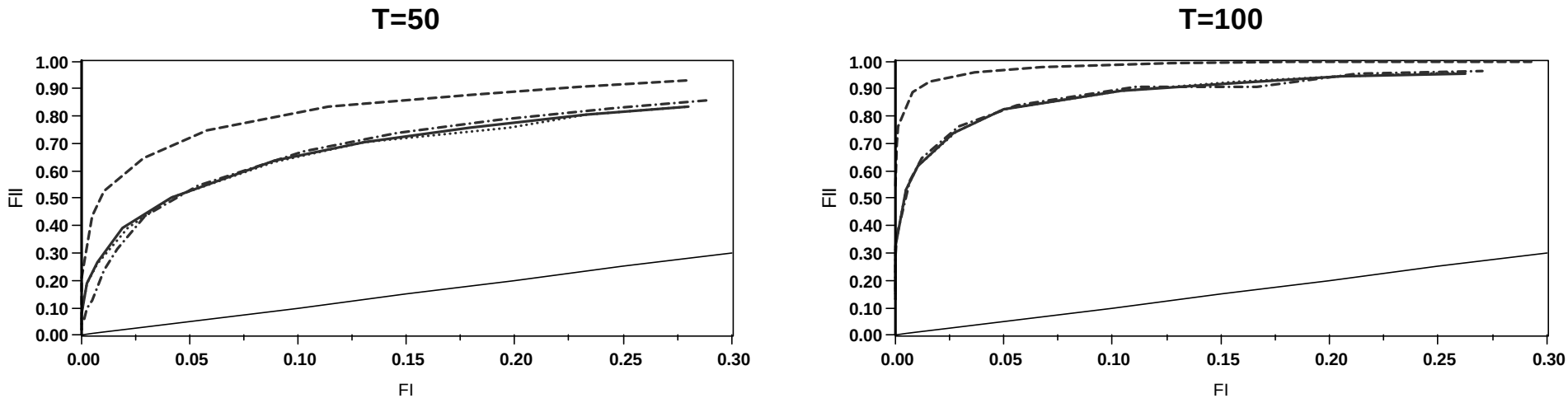
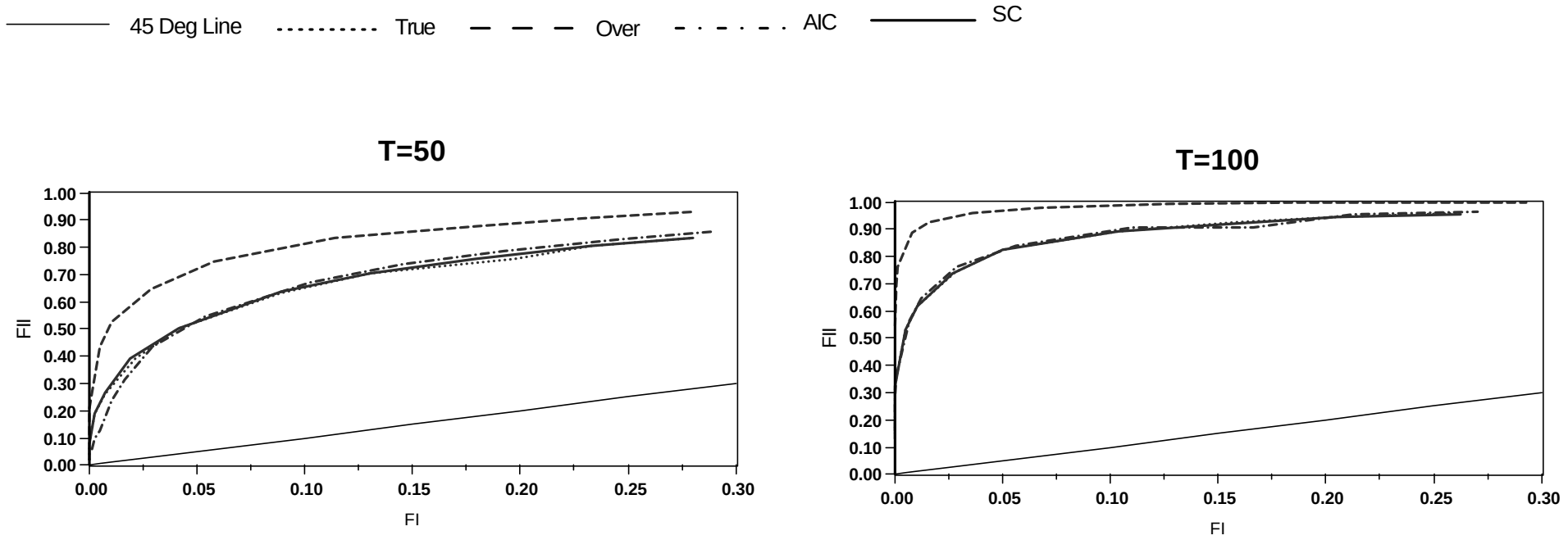


Figure 8. FIB-FIIB plots for strategy (M2), DGPB2 & DGPB3



ACKNOWLEDGMENTS

I am grateful to David Giles and Nilanjana Roy for helpful discussions and to other seminar participants at the University of Victoria for comments and suggestions. This research was supported by a grant from the Social Sciences and Humanities Council of Canada for the first author.

REFERENCES

H Akaike. Information theory and an extension of the maximum likelihood principle. In: BN Petrov and F Csáki, eds. 2nd International Symposium on Information Theory. Budapest:

-281, 1973.

DC Black, PR Corrigan and MR Dowd. New dogs and old tricks: do money and interest rates still provide information content for forecasts of output and prices? International Journal of Forecasting 16: 191-205, 2000.

BS Cheng. Beyond the purchasing power parity: testing for cointegration and causality between exchange rates, prices and interest rates. Journal of International Money and Finance 18: 911-924.

JJ Dolado and H Lütkepohl. Making Wald tests work for cointegrated VAR systems. Econometric Reviews 15: 369-385, 1996.

R Davidson and JG MacKinnon. Graphical methods for investigating the size and power of hypothesis tests. The Manchester School 66: 1-26, 1998.

J-M Dufour. Some impossibility theorems in econometrics with applications to structural and dynamic models. Econometrica 65: 103-132, 1997.

J-M Dufour and E Renault. Short run and long run causality in time series: theory. Econometrica 66: 1099-1125, 1998.

BM Friedman and KN Kuttner. Money, income, prices and interest rates. American Economic Review 82: 472-492, 1992.

JA Giles and S Mirza. Some pretesting issues on testing for Granger noncausality. Econometrics Working Paper EWP9914, Department of Economics, University of Victoria, 1999.

CWJ Granger. Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37: 424-459, 1969.

I Henriques and P Sadorsky. Export-led growth or growth-driven exports? The Canadian case. Canadian Journal of Economics 96: 540-555, 1996.

DL Hoffman and RH Rasche. Assessing forecast performance in a cointegrated system. Journal of Applied Econometrics 11, 495-517.

JL Horowtitz and NE Savin. Empirically relevant critical values for hypothesis tests: a bootstrap approach. Journal of Econometrics 95: 375-389, 2000.

C Hsiao. Autoregressive modeling and causal ordering of economic variables. Journal of Economic Dynamics and Control 4: 243-259, 1982.

S Kholdy. Causality between foreign investment and spillover efficiency. Applied Economics 27: 745-749, 1995.

K Krishna, A Ozyildirim and NR Swanson. Trade, investment and growth: nexus, analysis and prognosis. Mimeograph, Department of Economics, Texas A&M University, 2000.

J Lee, BS Shin and I Chung. Causality between advertising and sales: new evidence from cointegration. Applied Economics Letters 3: 299-301, 1996.

H Lütkepohl. Testing for causation between two variables in higher dimensional VAR models. In: H Schneeweiss and K Zimmerman, eds. Studies in Applied Econometrics. Heidelberg: Springer-Verlag, 1993a.

H Lütkepohl. Introduction to Multiple Time Series Analysis, 2nd Edition. Berlin: Springer-Verlag, 1993b.

R Nishi. Maximum likelihood principle and model selection when the true model is unspecified. Journal of Multivariate Analysis 27: 392-403, 1988.

RG Riezman, PM Summers and CH Whiteman. The engine of growth or its handmaiden? A time series assessment of export-led growth. Empirical Economics 21: 77-113, 1996.

G Schwarz. Estimating the dimension of a model. Annals of Statistics 6: 461-464, 1978.

CA Sims. Macroeconomics and reality. Econometrica 48: 10-48, 1980.

CA Sims, JH Stock and MJ Watson. Inference in linear time series models with some unit roots. Econometrica 58: 113-144, 1990.

J Thornton. Exports and economic growth: evidence from nineteenth century Europe. Economics Letters 55: 235-240, 1997.

HY Toda and PCB Phillips. Vector autoregression and causality. Econometrica 61: 1367-1393, 1993.

HY Toda and PCB Phillips. Vector autoregression and causality: a theoretical overview and simulation study. Econometric Reviews 13: 259-285, 1994.

HY Toda and T Yamamoto. Statistical inference in vector autoregressions with possibly integrated processes. Journal of Econometrics 66: 225-250, 1995.

N Weiner. The theory of prediction. In: EF Beckenback, ed.The Theory of Prediction. New York: McGraw-Hill, Chapter 8, 1956.

MB Wilk and R Gnanadesikan. Probability plotting methods for the analysis of data. Biometrika: 33, 1-17, 1968.

H Yamada and HY Toda. Inference in possibly integrated vector autoregressive models: some finite sample evidence. Journal of Econometrics 86: 55-95, 1998.

HO Zapata and AN Rambaldi. Monte Carlo evidence on cointegration and causation. Oxford Bulletin of Economics and Statistics 59: 285-298, 1997.