Lagged Variables as Instruments[*]

Yu Wang[†] and Marc Bellemare[‡]

July 2020

Abstract

Lagged explanatory variables are commonly used as instrumental variables (IVs) to address endogeneity concerns in empirical studies with observational data. Few theoretical studies, however, address whether those "lagged IVs" mitigate endogeneity. We develop a generic model in which dynamics among the endogenous explanatory variable and the unobserved confounders cannot be ruled out, and look analytically at the endogeneity of lagged IV estimates. We then use Monte Carlo simulations to illustrate our analytical findings. We show that when lagged IVs violate *only* the independence assumption, the lagged IV method mitigates endogeneity. When lagged IVs violate both the independence assumption and the exclusion restriction, the lagged IV method cannot mitigate endogeneity—and may even aggravate the problem. Both scenarios result in the likelihood of Type I close to one.

Keywords: Endogeneity, Instrumental Variables, Lagged variables, Treatment Effects, Causal Inference

---

[*] We thank John Freeman, Paul Glewwe, and Steve Miller for valuable comments and suggestions. All remaining errors are ours.
[†] Wang: Corresponding Author, Department of Applied Economics, University of Minnesota, email: wang5979@umn.edu.
[‡] Bellemare: Department of Applied Economics, University of Minnesota, email: mbellema@umn.edu.

I. INTRODUCTION

To address endogeneity concerns in empirical studies with observational data, it is not uncommon for researchers use the lag of an endogenous variable as an instrumental variable (IV) for endogenous variable. This strategy, which we will refer to as "lagged IV" throughout this paper, is popular among applied researchers because it requires no other variable as IV, which are notoriously difficult to find.

Though many researchers would readily admit that lagged variables may not be proper IVs because they are not strictly exogenous, it is often argued that lagged variables might alleviate endogeneity, at least to some extent (Anderson and Hsiao, 1981; Todd and Wolpin, 2003). Few formal theoretical analyses have been conducted to discuss whether the lagged IV method reduces the threat of endogeneity, however, and so applied researchers have little guidance regarding the conditions under which the lagged IV method could alleviate endogeneity concerns.

We analytically study the validity of the lagged IV method in response to endogeneity concern and present simulation results to illustrate our analytical results. We find that when the lagged IV neither has a direct causal impact on the outcome variable nor on the unobserved confounder, it only violates the independence assumption, but not the exclusion restriction, which are both necessary for the local average treatment effect (LATE) theorem to hold (Imbens and Angrist 1994; Angrist et al., 1996). In this case, the lagged IV estimate only consists of the restricted local average treatment on the treated (ATT). Comparatively, the naïve OLS estimate—one which ignores endogeneity concerns—consists of both the ATT plus some bias due to selection. As a result, when the lagged IV *only* violates the independence assumption, its estimate could be less biased than that of a naïve OLS estimate. In other words, the lagged IV method could serve to mitigate the endogeneity problem.

When the lagged IV has direct causal impact either on the outcome variable, on the unobserved confounder, or both, it violates not only the independence assumption, but also the exclusion restriction. In these three cases, the lagged IV estimate consists of both the relaxed local ATT and the local selection bias. Comparatively, the naïve OLS estimate consists once again of both the ATE plus some bias due to selection. As a result, when the lagged IV violates both the independence assumption *and* the exclusion restriction, its estimate may aggravate the endogeneity problem.

We set up a general model to compare the naïve OLS estimate with the lagged IV estimate. In our model, the outcome variable is determined both by an endogenous explanatory variable, an unobserved confounder, and (perhaps) the lagged explanatory variable. The explanatory variable is determined by its one-order lagged term and also the unobserved confounder; in addition, the unobserved confounder has a positive serial correlation and may also be influenced by the lagged explanatory variable.

With this general model, we discuss four scenarios. In Scenario 1, the lagged explanatory variable has no direct causal effect on the outcome variable or on the unobserved confounder. Therefore, Scenario 1 only violates the independence assumption. In Scenario 2, the lagged explanatory variable has a direct causal effect on the outcome variable. In Scenario 3, the lagged explanatory variable has a direct causal effect on the unobserved confounder. And in Scenario 4, the lagged explanatory variable has direct causal effects on both. Therefore, Scenarios 2, 3 and 4 violate both the independence assumption *and* the exclusion restriction.

In line with our analytical analysis, our simulation results show that in Scenario 1, (i) both the naïve OLS estimate and the lagged IV estimate are biased, but the bias of the lagged IV estimate is smaller than that of the naïve OLS estimate, (ii) both the naïve OLS estimate and the lagged IV estimate are consistent, (iii) the larger the extent to which the independence assumption is violated, the higher bias of the lagged IV estimate, (iv) root mean squared errors (RMSEs) show similar patterns as the biases, and (v) the likelihood that the lagged IV estimate suffers from the Type I error is very high close to 1. In other words, when only the independence assumption is violated, the lagged IV method is acceptable as its estimate is consistent, it mitigates bias from endogeneity some relative to the naïve OLS estimate, yet it remains problematic because of its very high likelihood of Type I error.

In Scenarios 2, 3 and 4, our simulation results show that, (i) both the naïve OLS estimate and the lagged IV estimate are biased, and the bias of the lagged IV estimate is smaller than that of the naïve OLS estimate, and (ii) both the naïve OLS estimate and the lagged IV estimate are inconsistent. In Scenarios 2 and 4, the lagged IV estimate is more inconsistent than the naïve OLS estimate; in Scenario 3, which estimator is the most consistent is ambiguous. Further, (iii) the larger the extent to which the exclusion restriction is violated, the higher the bias of the lagged IV, (iv) the RMSEs show similar patterns as the biases, and (v) the likelihood that the lagged IV estimate suffers from a Type I error is very high, and very close to 1. In others word, when violating *both* the independence assumption and the exclusion restriction, the lagged IV method yields an inconsistent estimate, and it can even aggravate endogeneity issues by increasing bias relative to a naïve OLS, and it also suffers from a very high likelihood of a Type I error.

Blundell and Bond (1998, 2000) have argued that because lagged explanatory variables are weakly correlated with the endogenous explanatory variable's first difference, GMM combined with lagged explanatory variables may not solve endogeneity problems. Our analysis focuses on using a one-order lagged explanatory variable as a single IV in estimation, a strategy that is is commonly used in empirical studies. Rossi (2014) argues in passing against using lagged explanatory variable as IV, but he does not provide precise derivations for why that is. Our findings, based on analytical and simulation results, are consistent with the previous literature. For applied researchers in the social sciences, our findings provide show that lagged IV method cannot obviate (and are likely to worsen) the consequences of endogeneity.

Before anything else, to see how common the lagged IV method is, we examine all articles published in the top general academic journals in economics and political science. We identify those articles using the lagged IV method by searching the text of each paper for the key words such as "lag," "lagged," or "lagging," and then seeing whether those papers used lagged endogenous variables as instrumental variables for those same endogenous variables. We do not discriminate between those papers where lagged explanatory variables are used as instrumental variables either as the main method or as a robustness check for the main results.

Table 1 shows, for the period 2013-2018, the number of papers using the lagged IV method, published in economic journals including the *American Economic Review*, *Econometrica*, the *Journal of Political Economy*, the *Quarterly Journal of Economics*, the *Review of Economic Studies*, and the *Review of Economics and Statistics*, and in political science journals including the *American Political Science Review*, the *American Journal of Political Science*, the *British Journal of Political Science*, *Comparative Political Studies*, and the *Journal of Politics*. In total, we find 31 papers in 2013-2018 using the lagged IV method, of which 19 in economics and 12 in political science journals. Narrowing the time period down to 2015-2018, 15 papers use the lagged IV method, of which nine in economics and 6 in political science journals.

These papers all use at least one first-order lagged (or first-order *and* multi-ordered lagged) explanatory variable(s) as instrumental variables to alleviate endogeneity concerns. Most papers mention that the availability of lagged explanatory variables is one of the key reasons why they are used as the IV. Seldom do those papers discuss whether the lagged explanatory variable has an explicit direct causal effect on the outcome variable—that is, whether the lagged IV violates the exclusion restriction, which makes the use of lagged IV methods more questionable. This literature review shows that the use of lagged IVs is not uncommon in economics and political science, and that authors of those papers using lagged IVs (act as if they) believe that although the lagged IV method is not perfect, it may somewhat mitigate endogeneity concerns.

The rest of this paper is organized as follows. Section II discusses the treatment effects of both the naïve OLS estimation and lagged IV estimation. Section III derives the biasedness and inconsistency of both estimations analytically. Section IV presents simulation results which illustrate and support our analytical results. Section V summarizes with directions for future research and recommendations for applied research.

II. TREATMENT EFFECTS

This section discusses the treatment effects in lagged IV estimation and in OLS estimation, and upon which compares their sources of endogeneity. In OLS estimation, lacking the ideal randomized experiment yields the selection bias that obstructs the identification of the ATT that we are interested in. In lagged IV estimation, violating the LATE theorem (Angrist and Pischke, 2009) results in the estimate different from the LATE that we are interested in.

We find that due to the synchronous relationship between the lagged IV and the unobserved confounder, the lagged IV estimation violates the independence assumption in the LATE Theorem. As a result, the lagged IV estimate suffers from endogeneity, of which the extent is ambiguous to that of the naïve OLS estimate. We also find that if the lagged IV causes the outcome variable not only through the explanatory variable but also through the unobserved confounder, the lagged IV violates the exclusion restriction in addition to the independence assumption in the LATE Theorem. As a result, the lagged IV estimate suffers more endogeneity than the naïve OLS estimate.

II.A. Setup

Empirically there are three reasons why the lagged explanatory variable may serve as a valid IV. Regarding the relevance restriction, autocorrelation in the explanatory variable implies that the endogenous variable is, to some extent, correlated with its lag.

Regarding the exclusion restriction, it is possible—at least in theory—that no causal relationship exists between the lagged explanatory variable and the outcome variable. Regarding data availability, the lagged IV method typically requires no additional data (at least not in longitudinal data sets), and the inherent statistical power issue that may arise from burning up one round of data due to lagging is less and less of a problem given the increasing availability of long panel data sets.[4]

In the case of the unobserved confounder, however, if there is autocorrelation both in the explanatory variable and in the unobserved confounder, the lagged explanatory variable could be correlated with the unobserved confounder in the current period through the lagged unobserved confounders, which would lead to biased estimates. To explain this, suppose that the structural model is such that

$$Y_{it} = \beta X_{it} + \xi X_{i,t-1} + \delta U_{it} + \epsilon_{it}, \tag{2.1}$$

where $Y_{it}, X_{it}, X_{i,t-1}, U_{it}$, and ϵ_{it} respectively denote the outcome variable, the explanatory variable, the lagged explanatory variable, the unobserved confounder, and an error term with mean zero, and where $Cov(X_{it}, U_{it}) \neq 0$, so that there is indeed an identification problem. If $\xi \neq 0$, the lagged explanatory variable has a direct impact on the outcome variable; otherwise it has obviously no such impact.

The autocorrelation function of explanatory variable is such that

$$X_{it} = \rho X_{i,t-1} + \kappa U_{it} + \eta_{it}. \tag{2.2}$$

The autocorrelation function of unobserved confounder is such that

$$U_{it} = \phi U_{i,t-1} + \psi X_{i,t-1} + \nu_{it}. \tag{2.3}$$

If $\psi \neq 0$, the lagged explanatory variable has a direct impact on unobserved confounder; otherwise there is no such impact.

We thus have four scenarios:

Scenario 1: $\xi = 0$, and $\psi = 0$. In this scenario, the lagged explanatory variable has no explicit impact on the outcome variable, nor does it have any explicit impact on the unobserved confounder.

Scenario 2: $\xi \neq 0$, while $\psi = 0$. In this scenario, the lagged explanatory variable has a direct impact on the outcome variable, but it has no explicit impact on the unobserved confounder.

Scenario 3: $\xi = 0$, while $\psi \neq 0$. In this scenario, the lagged explanatory variable

[4] For this reason, we ignore that issue in the remainder of this paper.

has no explicit impact on the outcome variable, but it has a direct impact on the unobserved confounder.

Scenario 4: $\xi \neq 0$, and $\psi \neq 0$. In this scenario, the lagged explanatory variable has a direct impact on the outcome variable, and it also has a direct impact on the unobserved confounder.

In light of the LATE theorem (Angrist and Pischke, 2009), we discuss the lagged IV estimate. For simplicity, and without any loss of generality, we assume a binary explanatory variable. We denote $Y_{it}(x, \tilde{x})$ as individual i's latent outcome when her treatment is $X_{it} = x$ and her lagged treatment, the lagged IV, is $X_{i,t-1} = \tilde{x}$. To specify the heterogeneous causal effect of the lagged IV, we denote X_{1it} as individual i's latent treatment state when $X_{i,t-1} = 1$, and X_{0it} as individual i's latent treatment state when $X_{i,t-1} = 0$. Thus, the observed treatment state is defined latently as

$$X_{it} = X_{0it} + (X_{1it} - X_{0it})X_{i,t-1} \tag{2.4}$$

in which either X_{1it} or X_{0it} can be observed, and $(X_{1it} - X_{0it})$ represents the heterogeneous causal effect of $X_{i,t-1}$. With this notation, we now state the independence assumption and the exclusion restriction of the lagged IV as follows.

1. The *independence assumption* implies that the instrumental variable should have no association with the latent outcome, nor should it have any association with the latent treatment state. Specifically, we have

$$[\{Y_{it}(x, \tilde{x}); \forall x, \tilde{x}\}, X_{1it}, X_{0it}\,] \perp\!\!\!\perp X_{i,t-1} \tag{2.5}$$

This implies that the lagged IV should have an effect similar to random assignment. In other words, the lagged IV should be uncorrelated with the outcome variable or with the latent treatment state by the explanatory variable.

Scenario 1 violates the independence assumption, because the lagged IV is synchronously correlated with the unobserved confounder. Specifically, because $U_{i,t-1}$ influences U_{it} by its marginal effect ϕ and influences $X_{i,t-1}$ by its marginal effect κ, $X_{i,t-1}$ and U_{it} have a simultaneous relationship. In other words, as $X_{i,t-1}$ changes, U_{it} changes—not causally but synchronously—and further causes a change in Y_{it} change. In other words, as $X_{i,t-1}$ changes by 1 unit, U_{it} changes synchronously by $\frac{\phi}{\kappa}$ unit. As a result, $X_{i,t-1}$ violates the independence assumption because it does not serve as a random exogenous shock. Only by assuming that there are no dynamics among unobserved confounders can the independence assumption be satisfied. Unfortunately, that assumption is unlikely to hold (Bellemare et al., 2017).

This implies that it is almost unavoidable for a lagged IV to be problematic.

2. The *exclusion restriction* implies that $Y_{it}(x,\tilde{x})$ is only a function of x. In other words, the lagged IV influence the outcome variable only through the explanatory variable. This is denoted as

$$Y_{it}(x,0) = Y_{it}(x,1), x = 0,1 \tag{2.6}$$

In Scenario 2, because $\xi \neq 0$, $X_{i,t-1}$ has a direct causal influence on Y_{it} by its marginal effect ξ. In Scenario 3, although $\xi = 0$, because $\psi \neq 0$, $X_{i,t-1}$ has a direct causal influence on Y_{it} by its marginal effect $\delta\psi$, derived from equations (2.1) and (2.3). As a result, both Scenarios 2 and 3 violate not only the independence assumption, but also the exclusion restriction. Scenario 4, which is a combination of Scenario 2 and 3, merely compounds the problem.

II.B. The Treatment Effects in Lagged IV and in OLS

The OLS estimate consists of the ATT that we are interested in to identify the treatment effect, as well as the selection bias, if no ideal randomized experiment exists. The selection bias is the source of endogeneity in OLS estimation. The lagged IV estimate consists of the LATE, which measures the treatment effect if the LATE Theorem theorem is satisfied. Yet the lagged IV estimate may suffer from endogeneity and fail to measure the LATE, when either the independence assumption or the exclusion restriction, or both, are violated.

To discuss the treatment effects and compare the endogeneity of lagged IV with that of naïve OLS, we first discuss the OLS estimate, which measures the observed difference between the participants and non-participants of the treatment, such that

$$\begin{aligned} &\mathbb{E}[Y_{it}|X_{it}=1] - \mathbb{E}[Y_{it}|X_{it}=0] \\ &= \mathbb{E}[Y_{1it}|X_{it}=1] - \mathbb{E}[Y_{0it}|X_{it}=1] + \mathbb{E}[Y_{0it}|X_{it}=1] - \mathbb{E}[Y_{0it}|X_{it}=0] \\ &= \mathbb{E}[Y_{1it} - Y_{0it}|X_{it}] + \mathbb{E}[Y_{0it}|X_{it}=1] - \mathbb{E}[Y_{0it}|X_{it}=0] \end{aligned} \tag{2.7}$$

where $\mathbb{E}[Y_{1it} - Y_{0it}|X_{it}]$ is the ATT, and $\mathbb{E}[Y_{0it}|X_{it}=1] - \mathbb{E}[Y_{0it}|X_{it}=0]$ is the selection bias that obstructs us from identifying the ATT.

In lagged IV estimation, , when both the exclusion restriction and the independence assumption are satisfied, the Wald estimate is $\frac{\mathbb{E}[Y_{it}|X_{i,t-1}=1] - \mathbb{E}[Y_{it}|X_{i,t-1}=0]}{\mathbb{E}[X_{it}|X_{i,t-1}=1] - \mathbb{E}[X_{it}|X_{i,t-1}=0]} = \mathbb{E}[Y_{1it} - Y_{0it}|X_{1it} > X_{0it}]$, which is also the LATE that we are interested in.

1. The lagged IV estimate in Scenario 1: In Scenario 1, where only the independence assumption is violated, we have

$$\mathbb{E}[Y_{it}|X_{i,t-1}=1]$$
$$=\mathbb{E}\left[Y_{it}(0,X_{i,t-1})+\left(Y_{it}(1,X_{i,t-1})-Y_{it}(0,X_{i,t-1})\right)X_{it}\middle|X_{i,t-1}=1\right] \quad (2.8)$$

Because the *exclusion restriction* is satisfied, we have $Y_{it}(0,X_{i,t-1})=Y_{0it}$, $Y_{it}(1,X_{i,t-1})=Y_{1it}$. Therefore,

$$\mathbb{E}[Y_{it}|X_{i,t-1}=1]$$
$$=\mathbb{E}[Y_{0it}+(Y_{1it}-Y_{0it})X_{1it}|X_{i,t-1}=1]$$
$$=\mathbb{E}[Y_{0it}|X_{i,t-1}=1]+\mathbb{E}[(Y_{1it}-Y_{0it})X_{1it}|X_{i,t-1}=1] \quad (2.9)$$

Similarly, we have

$$\mathbb{E}[Y_{it}|X_{i,t-1}=0]$$
$$=\mathbb{E}\left[Y_{it}(0,X_{i,t-1})+\left(Y_{it}(1,X_{i,t-1})-Y_{it}(0,X_{i,t-1})\right)X_{it}\middle|X_{i,t-1}=0\right]$$
$$=\mathbb{E}[Y_{0it}+(Y_{1it}-Y_{0it})X_{0it}|X_{i,t-1}=0]$$
$$=\mathbb{E}[Y_{0it}|X_{i,t-1}=0]+\mathbb{E}[(Y_{1it}-Y_{0it})X_{0it}|X_{i,t-1}=0] \quad (2.10)$$

If the exclusion restriction is satisfied, we have

$$\mathbb{E}[Y_{0it}|X_{i,t-1}=1]=\mathbb{E}[Y_{0it}|X_{i,t-1}=0] \quad (2.11)$$

Therefore, the Wald estimate is

$$\frac{\mathbb{E}[Y_{it}|X_{i,t-1}=1]-\mathbb{E}[Y_{it}|X_{i,t-1}=0]}{\mathbb{E}[X_{it}|X_{i,t-1}=1]-\mathbb{E}[X_{it}|X_{i,t-1}=0]}$$
$$=\frac{\mathbb{E}[(Y_{1it}-Y_{0it})X_{1it}|X_{i,t-1}=1]-\mathbb{E}[(Y_{1it}-Y_{0it})X_{0it}|X_{i,t-1}=0]}{\mathbb{E}[X_{1it}|X_{i,t-1}=1]-\mathbb{E}[X_{0it}|X_{i,t-1}=0]} \quad (2.12)$$

which is named as the "restricted local ATT", and is different from the LATE that we are interested in.

Compared with the naïve OLS estimate, it is easy to see that the lagged IV estimate in Scenario 1 does not have a selection bias, implying that the extent of the endogeneity problem in the Scenario 1 is smaller than the extent of the endogeneity problem in a naïve OLS. Furthermore, it is also easy to see that the lagged IV estimate in the Scenario 1 is still different from the LATE. This is because of (i) the lagged IV's dependency on the latent treatment, scaled by ρ, the marginal causal effect of the lagged IV on the treatment variable, and (ii) the lagged IV's dependency on the latent outcome, scaled by $\frac{\phi}{\kappa}$, the synchronous relationship between the lagged IV and the unobserved confounder. Because the unobserved confounder's marginal causal effect on the outcome variable is δ, we could initially predict that the key parameters for the extent of endogeneity of the Scenario 1 of lagged IV estimation are ρ, ϕ, κ and δ.

2. The lagged IV estimates in Scenario 2, 3 and 4: In Scenario 2, 3 and 4, both the exclusion restriction and the independence assumption are violated, to derive $\frac{\mathbb{E}[Y_{it}|X_{i,t-1}=1]-\mathbb{E}[Y_{it}|X_{i,t-1}=0]}{\mathbb{E}[X_{it}|X_{i,t-1}=1]-\mathbb{E}[X_{it}|X_{i,t-1}=0]}$, the LATE, we have

$$\begin{aligned}
&\mathbb{E}\left[Y_{it}\middle|X_{i,t-1}=1\right] \\
&=\mathbb{E}\left[Y_{it}(0,X_{i,t-1})+\left(Y_{it}(1,X_{i,t-1})-Y_{it}(0,X_{i,t-1})\right)X_{it}\middle|X_{i,t-1}=1\right] \\
&=\mathbb{E}\left[Y_{it}(0,X_{i,t-1})\middle|X_{i,t-1}=1\right] \\
&\quad+\mathbb{E}\left[\left(Y_{it}(1,X_{i,t-1})-Y_{it}(0,X_{i,t-1})\right)X_{it}\middle|X_{i,t-1}=1\right]
\end{aligned} \tag{2.13}$$

and similarly,

$$\begin{aligned}
&\mathbb{E}\left[Y_{it}\middle|X_{i,t-1}=0\right] \\
&=\mathbb{E}\left[Y_{it}(0,X_{i,t-1})+\left(Y_{it}(1,X_{i,t-1})-Y_{it}(0,X_{i,t-1})\right)X_{it}\middle|X_{i,t-1}=0\right] \\
&=\mathbb{E}\left[Y_{it}(0,X_{i,t-1})\middle|X_{i,t-1}=0\right] \\
&\quad+\mathbb{E}\left[\left(Y_{it}(1,X_{i,t-1})-Y_{it}(0,X_{i,t-1})\right)X_{it}\middle|X_{i,t-1}=0\right]
\end{aligned} \tag{2.14}$$

Therefore, the estimate becomes the sum of the "relaxed local ATT",

$$\frac{\mathbb{E}\left[\left(Y_{it}(1)-Y_{it}(0)\right)X_{it}\middle|X_{i,t-1}=1\right]-\mathbb{E}\left[\left(Y_{it}(1)-Y_{it}(0)\right)X_{it}\middle|X_{i,t-1}=0\right]}{\mathbb{E}\left[X_{it}\middle|X_{i,t-1}=1\right]-\mathbb{E}\left[X_{it}\middle|X_{i,t-1}=0\right]}$$

and the "local selection bias", $\frac{\mathbb{E}[Y_{it}(0)|X_{i,t-1}=1]-\mathbb{E}[Y_{it}(0)|X_{i,t-1}=0]}{\mathbb{E}[X_{it}|X_{i,t-1}=1]-\mathbb{E}[X_{it}|X_{i,t-1}=0]}$, where $Y_{it}(1)\equiv Y_{it}(1,X_{i,t-1})$, $Y_{it}(0)\equiv Y_{it}(0,X_{i,t-1})$. As a result, the estimates in Scenario 2, 3 and 4 consist of the "local selection bias" and the "relaxed local ATT", which are quite different from the LATEs that we are interested in.

Compared with the OLS estimate, it is easy to see that the lagged IV estimates in the Scenario 2, 3 and 4 include a "local selection bias", which could be greater than the selection bias in the naïve OLS estimate. Moreover, it is also easy to see that the lagged IV estimates in the Scenario 2, 3 and 4 also include the "relaxed local ATTs", which are different from the "restricted local ATT" in Scenario 1. These imply that the extent of endogeneity in Scenario 2, 3 and 4 are greater than that in Scenario 1, and could be greater than that in the naïve OLS.

To sum up, the naïve OLS estimate suffers from endogeneity because its estimate includes a selection bias. When the lagged IV estimate only violates the independence assumption, it suffers from the endogeneity, because the "restricted local ATT" in its estimate is different from the ATT in the naïve OLS estimate. When the lagged IV estimate violates both the exclusion restriction and the independence assumption, it suffers from endogeneity because, on the one hand, the estimate includes "local selection bias" and, on the other hand, the "relaxed local ATT", which is different from the ATT in the naïve OLS estimate.

III. BIASEDNESS AND INCONSISTENCY

The discussion of treatment effects in the last section demonstrates why using a lagged explanatory variable as an IV is unlikely to mitigate endogeneity issues. In this section, we characterize the treatment effect of lagged IV estimation analytically, and compare it with the analytical treatment effect of naïve OLS estimation. These analytical results are in line with what we find in last section.

For simplicity, we set up a bivariate regression scenario, and discuss the $AR(1)$ process in the data generation process for both the endogenous explanatory variable and the unobserved confounder.

III.A. Comparing Lagged IV and OLS Estimates

Scenario 1. We first quantitatively discuss Scenario 1, which violates the independence assumption but not the exclusion restriction. Following Bellemare et. al (2017), we consider the following setup:

$$Y_{it} = \beta X_{it} + \delta U_{it} + \epsilon_{it}, \tag{3.1}$$

$$X_{it} = \rho X_{i,t-1} + \kappa U_{it} + \eta_{it}, \text{and} \tag{3.2}$$

$$U_{it} = \phi U_{i,t-1} + \nu_{it}, \tag{3.3}$$

where i and t respectively denote units of observation and time, and where $i = 1, 2, \dots N$, $t = 1, 2, \dots, T$.

For simplicity, we drop i for the remainder of this section. As before, Y_t is the outcome variable, and X_t represents the explanatory variable. The $AR(1)$ process implies that X_t is determined both by its lagged value and by the unobserved confounder, and that U_t is determined by its one-order lagged value. For coefficients we assume that $\rho, \phi \in (0,1)$; for the error terms we assume that $\eta_t \sim N(0, \sigma_\eta^2)$, $\epsilon_t \sim N(0, \sigma_\epsilon^2)$, and $\nu_t \sim N(0, \sigma_\nu^2)$.

Since U_t, the unobserved confounder, is unobserved, a naïve OLS estimation of (3.1) obviously suffers from endogeneity, and the coefficient β is not identified. More formally, $\hat{\beta}_{OLS} = \frac{Cov(X_t, Y_t)}{Var(X_t)}$, and so

$$\hat{\beta}_{OLS} = \frac{Cov(X_t, \beta X_t + \delta U_t + \epsilon_t)}{Var(X_t)}$$

$$= \beta + \frac{\delta Cov(X_t, U_t)}{Var(X_t)} \tag{3.4}$$

Equation (3.4) implies that in Scenario 1, the OLS estimate is biased, and the bias $\frac{\delta Cov(X_t,U_t)}{Var(X_t)}$ is in line with the selection bias.

To discuss the consistency of the OLS estimate, we need to use equation (A.5) in the Appendix, which allows deriving the following expression:

$$Cov(X_t, U_t) = \frac{\kappa Var(U_t)}{1 - \phi\rho} \tag{3.5}$$

Thus, plugging equation (3.5) into (3.4) yields

$$\hat{\beta}_{OLS} = \beta + \frac{\delta\kappa Var(U_t)}{(1-\phi\rho)Var(X_t)}$$

$$= \beta + \frac{\delta\kappa \sum_{t=1}^{T} U_t^2}{(1-\phi\rho)\sum_{t=1}^{T} X_t^2} \tag{3.6}$$

Using the Slutsky theorem, (3.6) becomes

$$p\text{lim}\hat{\beta}_{OLS} = \beta + \frac{\delta\kappa[p\text{lim}\left(\frac{1}{T}\right)\sum_{t=1}^{T} U_t^2]}{(1-\phi\rho)[p\text{lim}\left(\frac{1}{T}\right)\sum_{t=1}^{T} X_t^2]} \tag{3.7}$$

where $\frac{\delta\kappa[p\text{lim}\left(\frac{1}{T}\right)\sum_{t=1}^{T} U_t^2]}{(1-\phi\rho)[p\text{lim}\left(\frac{1}{T}\right)\sum_{t=1}^{T} X_t^2]}$ is the estimation inconsistency. Because $\phi \in (0,1)$, (3.2) and (3.3) imply that as $T \to \infty$, $p\text{lim}\left(\frac{1}{T}\right)\sum_{t=1}^{T} U_t^2 \ll p\text{lim}\left(\frac{1}{T}\right)\sum_{t=1}^{T} X_t^2$. As a result, $p\lim_{T\to\infty} \hat{\beta}_{OLS} \to \beta$; in other words, the OLS estimate in Scenario 1 is consistent.

Now consider an IV estimation using X_{t-1} as the instrumental variable for X_t, the IV estimates expression implies that

$$\hat{\beta}_{IV,1} = \frac{Cov(X_{t-1}, Y_t)}{Cov(X_{t-1}, X_t)} \tag{3.8}$$

Plugging equation (3.1) into (3.8), we have

$$\hat{\beta}_{IV,1} = \frac{Cov(X_{t-1}, \beta X_t + \delta U_t + \epsilon_t)}{Cov(X_{t-1}, X_t)} \tag{3.9}$$

and so

$$\hat{\beta}_{IV,1} = \beta + \frac{\delta Cov(X_{t-1}, U_t)}{Cov(X_{t-1}, X_t)}$$

$$= \beta + \delta \frac{\frac{Cov(X_{t-1}, U_t)}{Var(X_{t-1})}}{\rho + \kappa \frac{Cov(X_{t-1}, U_t)}{Var(X_{t-1})}} \tag{3.10}$$

Therefore, (3.10) implies that in Scenario 1, the lagged IV estimate is biased, and that the bias is $\delta\frac{\frac{Cov(X_{t-1},U_t)}{Var(X_{t-1})}}{\rho+\kappa\frac{Cov(X_{t-1},U_t)}{Var(X_{t-1})}}$, which is in line with the "restricted local ATT". κ is the key parameter to determine to what extent the lagged IV estimate is biased. This is because the extent, to which the lagged IV violates the independence assumption is measured by κ.

To discuss the consistency of the lagged IV estimate in Scenario 1, we use equation (A.5) in the Online Appendix to derive

$$\begin{aligned}\frac{Cov(X_{t-1},U_t)}{Var(X_{t-1})} &= \frac{Cov(X_{t-1},\phi U_{t-1}+\nu_t)}{Var(X_{t-1})}\\ &= \frac{\phi Cov(X_{t-1},U_{t-1})}{Var(X_{t-1})}\\ &= \frac{\phi\kappa Var(U_t)}{(1-\phi\rho)Var(X_t)}\end{aligned} \quad (3.11)$$

Therefore, we have

$$\begin{aligned}\hat{\beta}_{IV,1} &= \beta+\frac{\delta\phi\kappa Var(U_t)}{\rho(1-\phi\rho)Var(X_t)+\phi\kappa^2 Var(U_t)}\\ &= \beta+\frac{\delta\kappa Var(U_t)}{\frac{\rho}{\phi}(1-\phi\rho)Var(X_t)+\kappa^2 Var(U_t)}\end{aligned} \quad (3.12)$$

Using the Slutsky theorem, (3.12) becomes

$$p\mathrm{lim}\,\hat{\beta}_{IV,1} = \beta+\frac{\delta\kappa[p\mathrm{lim}\left(\frac{1}{T}\right)\sum_{t=1}^{T}U_t^2]}{\frac{\rho}{\phi}(1-\phi\rho)\left[p\mathrm{lim}\left(\frac{1}{T}\right)\sum_{t=1}^{T}X_t^2\right]+\kappa^2[p\mathrm{lim}\left(\frac{1}{T}\right)\sum_{t=1}^{T}U_t^2]} \quad (3.13)$$

Because $\phi\in(0,1)$, (3.2) and (3.3) imply that as $T\rightarrow\infty$, $p\mathrm{lim}\left(\frac{1}{T}\right)\sum_{t=1}^{T}U_t^2 \ll p\mathrm{lim}\left(\frac{1}{T}\right)\sum_{t=1}^{T}X_t^2$. As a result, $p\lim_{T\rightarrow\infty}\hat{\beta}_{IV,1}\rightarrow\beta$; in other words, the lagged IV estimate in Scenario 1 is consistent.

Scenario 2. This scenario not only violates the independence assumption but also the exclusion restriction. We consider the following model

$$Y_{it} = \beta X_{it}+\xi X_{i,t-1}+\delta U_{it}+\epsilon_{it} \quad (3.14)$$

$$X_{it} = \rho X_{i,t-1}+\kappa U_{it}+\eta_{it} \quad (3.15)$$

$$U_{it} = \phi U_{i,t-1}+\nu_{it} \quad (3.16)$$

For simplicity, we drop i for the remainder of this section. Consider the OLS estimate in Scenario 2, which is such that

$$\hat{\beta}_{OLS} = \frac{Cov(X_t, Y_t)}{Var(X_t)}$$
$$= \frac{Cov(X_t, \beta X_t + \xi X_{t-1} + \delta U_t + \epsilon_t)}{Var(X_t)}$$
$$= \beta + \frac{\delta Cov(X_t, U_t)}{Var(X_t)} + \frac{\xi Cov(X_t, X_{t-1})}{Var(X_t)} \tag{3.17}$$

Therefore, (3.17) implies that in Scenario 2, the OLS estimate is biased, in which $\frac{\delta Cov(X_t, U_t)}{Var(X_t)} + \frac{\xi Cov(X_t, X_{t-1})}{Var(X_t)}$, the bias, is in line with the selection bias.

To discuss the consistency of the OLS estimate, we need to using equation (A.5) in the Online Appendix, and then we could derive the following expression that

$$Cov(X_t, U_t) = \frac{\kappa Var(U_t)}{1-\phi\rho} \tag{3.18}$$

and that

$$\frac{Cov(X_t, X_{t-1})}{Var(X_t)} = \rho + \frac{\phi\kappa^2 Var(U_t)}{(1-\phi\rho)Var(X_t)} \tag{3.19}$$

Therefore, we have an expression that

$$\hat{\beta}_{OLS} = \beta + \frac{\delta\kappa Var(U_t)}{(1-\phi\rho)Var(X_t)} + \xi\rho + \frac{\phi\xi\kappa^2 Var(U_t)}{(1-\phi\rho)Var(X_t)}$$
$$= \beta + \frac{\delta\kappa \sum_{t=1}^{T} U_t^2}{(1-\phi\rho)\sum_{t=1}^{T} X_t^2} + \frac{\phi\xi\kappa^2 \sum_{t=1}^{T} U_t^2}{(1-\phi\rho)\sum_{t=1}^{T} X_t^2} \tag{3.20}$$

Using the Slutsky theorem, (3.20) becomes

$$p\text{lim}\hat{\beta}_{OLS} = \beta + \xi\rho + \frac{(\delta\kappa + \phi\xi\kappa^2)[p\text{lim}\left(\frac{1}{T}\right)\sum_{t=1}^{T} U_t^2]}{(1-\phi\rho)[p\text{lim}\left(\frac{1}{T}\right)\sum_{t=1}^{T} X_t^2]} \tag{3.21}$$

Because $\phi \in (0,1)$, (3.15) and (3.16) imply that as $T \to \infty$, $p\text{lim}\left(\frac{1}{T}\right)\sum_{t=1}^{T} U_t^2 \ll p\text{lim}\left(\frac{1}{T}\right)\sum_{t=1}^{T} X_t^2$. As a result, $p\lim_{T\to\infty} \hat{\beta}_{OLS} \to \beta + \xi\rho$; in other words, the OLS estimate in Scenario 2 is inconsistent.

Consider an IV estimation using X_{t-1} as the instrumental variable for X_t, the IV estimates expression implies that

$$\hat{\beta}_{IV,2} = \frac{Cov(X_{t-1}, Y_t)}{Cov(X_{t-1}, X_t)} \tag{3.22}$$

Plugging equation (3.14) into (3.22), we have

$$\hat{\beta}_{IV,2} = \frac{Cov(X_{t-1}, \beta X_t + \xi X_{t-1} + \delta U_t + \epsilon_t)}{Cov(X_{t-1}, X_t)} \tag{3.23}$$

and then

$$\hat{\beta}_{IV,2} = \beta + \frac{\xi Var(X_{t-1})}{Cov(X_{t-1}, X_t)} + \frac{\delta Cov(X_{t-1}, U_t)}{Cov(X_{t-1}, X_t)}$$

$$= \beta + \xi \frac{1}{\rho + \kappa \frac{Cov(X_{t-1}, U_t)}{Var(X_{t-1})}} + \delta \frac{\frac{Cov(X_{t-1}, U_t)}{Var(X_{t-1})}}{\rho + \kappa \frac{Cov(X_{t-1}, U_t)}{Var(X_{t-1})}} \quad (3.24)$$

Therefore, (3.24) implies that in Scenario 2, the lagged IV estimate is biased, in which $\xi \frac{1}{\rho+\kappa\frac{Cov(X_{t-1},U_t)}{Var(X_{t-1})}}$ is in line with the "local selection bias", and $\beta + \delta \frac{\frac{Cov(X_{t-1},U_t)}{Var(X_{t-1})}}{\rho+\kappa\frac{Cov(X_{t-1},U_t)}{Var(X_{t-1})}}$ is in line with the "relaxed local ATT" in Scenario 2. κ and ξ are the key parameter determining to what extent the lagged IV estimate is biased. This is because the extent, to which the exclusion restriction of the lagged IV violates, is measured by ξ.

Then we discuss the consistency of the lagged IV estimate in Scenario 2. We have already known, from the Online Appendix, that

$$p\lim_{t\to\infty} \frac{Cov(X_{t-1}, U_t)}{Var(X_{t-1})} = \frac{\phi\kappa Var(U_t)}{(1-\phi\rho)Var(X_t)} \quad (3.25)$$

Therefore, we have

$$\hat{\beta}_{IV,2} = \beta + \frac{\xi(1-\phi\rho)Var(X_t)+\delta\phi\kappa Var(U_t)}{\rho(1-\phi\rho)Var(X_t)+\phi\kappa^2 Var(U_t)}$$

$$= \beta + \frac{\xi\left(\frac{1}{\phi}-\rho\right)Var(X_t) + \delta\kappa Var(U_t)}{\frac{\rho}{\phi}(1-\phi\rho)Var(X_t) + \kappa^2 Var(U_t)} \quad (3.26)$$

Using the Slutsky theorem, we have

$$\text{plim}\hat{\beta}_{IV,2} = \beta + \frac{\xi\left(\frac{1}{\phi}-\rho\right)\left[\text{plim}\left(\frac{1}{T}\right)\sum_{t=1}^{T} X_t^2\right]+\delta\kappa\left[\text{plim}\left(\frac{1}{T}\right)\sum_{t=1}^{T} U_t^2\right]}{\rho\left(\frac{1}{\phi}-\rho\right)\left[\text{plim}\left(\frac{1}{T}\right)\sum_{t=1}^{T} X_t^2\right]+\frac{\kappa^2}{\phi}\left[\text{plim}\left(\frac{1}{T}\right)\sum_{t=1}^{T} U_t^2\right]} \quad (3.27)$$

Because $\phi \in (0,1)$, (3.15) and (3.16) imply that as $T \to \infty$, $p\lim_{T\to\infty}\left(\frac{1}{T}\right)\sum_{t=1}^{T} U_t^2 \ll p\lim_{T\to\infty}\left(\frac{1}{T}\right)\sum_{t=1}^{T} X_t^2$. As a result, $p\lim_{T\to\infty}\hat{\beta}_{IV,2} \to \beta + \frac{\xi}{\rho}$; in other words, the lagged IV estimate in Scenario 2 is inconsistent. We could also derive that in Scenario 2, $p\lim_{T\to\infty}\hat{\beta}_{OLS} \to \beta + \xi\rho$; in other words, the OLS estimate in Scenario 2 is inconsistent. As $\frac{\xi}{\rho} > \xi\rho$, we know that the lagged IV estimate has significantly larger extent of inconsistency that the OLS estimate.

Scenario 3. This scenario violates both the *independence assumption* and the *exclusion restriction*. We consider the following model

$$Y_{it} = \beta X_{it} + \delta U_{it} + \epsilon_{it} \tag{3.28}$$

$$X_{it} = \rho X_{i,t-1} + \kappa U_{it} + \eta_{it} \tag{3.29}$$

$$U_{it} = \phi U_{i,t-1} + \psi X_{i,t-1} + \nu_{it} \tag{3.30}$$

For simplicity, we drop i for the reminder of this session, and everything is similar to those in Section III.A.

Consider the OLS estimate in Scenario 3, such that

$$\begin{aligned}\hat{\beta}_{OLS} &= \frac{Cov(X_t, Y_t)}{Var(X_t)} \\ &= \frac{Cov(X_t, \beta X_t + \delta U_t + \epsilon_t)}{Var(X_t)} \\ &= \beta + \frac{\delta Cov(X_t, U_t)}{Var(X_t)}\end{aligned} \tag{3.31}$$

Therefore, (3.31) implies that in Scenario 3, the OLS estimate is biased, in which $\frac{\delta Cov(X_t,U_t)}{Var(X_t)}$, the bias, is in line with the selection bias.

To discuss the consistency of the OLS estimate, we need to using equation (A.12) in the Online Appendix. Therefore, we know that using the Slutsky theorem, we have

$$\text{plim}\hat{\beta}_{OLS} = \beta + \frac{\psi\rho}{(1-\phi\rho)} + \frac{\kappa[\text{plim}\left(\frac{1}{T}\right)\sum_{t=1}^{T} U_t^2]}{(1-\phi\rho)[\text{plim}\left(\frac{1}{T}\right)\sum_{t=1}^{T} X_t^2]} \tag{3.32}$$

Because $\frac{\psi\rho}{(1-\phi\rho)} \neq 0$, in Scenario 3, the OLS estimate is inconsistent.

Consider an IV estimation using X_{t-1} as the instrumental variable for X_t, the lagged IV estimates expression implies that

$$\hat{\beta}_{IV,3} = \frac{Cov(X_{t-1}, Y_t)}{Cov(X_{t-1}, X_t)} \tag{3.33}$$

Plugging equation (3.27) into (3.32), we have

$$\hat{\beta}_{IV,3} = \frac{Cov(X_{t-1}, \beta X_t + \delta U_t + \epsilon_t)}{Cov(X_{t-1}, X_t)} \tag{3.34}$$

and then

$$\begin{aligned}\hat{\beta}_{IV,3} &= \beta + \frac{\delta Cov(X_{t-1}, U_t)}{Cov(X_{t-1}, X_t)} \\ &= \beta + \delta \frac{\frac{Cov(X_{t-1}, U_t)}{Var(X_{t-1})}}{\rho + \kappa \frac{Cov(X_{t-1}, U_t)}{Var(X_{t-1})}}\end{aligned} \tag{3.35}$$

Therefore, (3.35) implies that in Scenario 3, the lagged IV estimate is biased, in which $\beta + \delta \frac{\frac{Cov(X_{t-1},U_t)}{Var(X_{t-1})}}{\rho+\kappa\frac{Cov(X_{t-1},U_t)}{Var(X_{t-1})}}$ is in line with the "local selection bias" and the "relaxed local ATT", in Scenario 3. κ and ψ are the key parameter determining to what extent the lagged IV estimate is biased. This is because the extent, to which the exclusion restriction of the lagged IV violates, is measured by ψ.

Then we discuss the consistency of the lagged IV estimate in Scenario 3. We know from the appendix that

$$p\lim_{t\to\infty} \frac{Cov(X_{t-1},U_t)}{Var(X_{t-1})} = \frac{\phi\kappa Var(U_t)}{(1-\phi\rho)Var(X_t)} + \frac{\psi}{1-\phi\rho}$$

Therefore, we have

$$\text{plim}\hat{\beta}_{IV,3} = \beta + \frac{\delta\kappa Var(U_t) + \frac{\delta\psi}{\phi} Var(X_t)}{[\frac{\rho}{\phi}(1-\phi\rho) + \psi\kappa]Var(X_t) + \kappa^2 Var(U_t)} \tag{3.36}$$

Using the Slutsky theorem, we have

$$p\text{lim}\hat{\beta}_{IV,3} = \beta + \frac{\delta\kappa[p\text{lim}\left(\frac{1}{T}\right)\sum_{t=1}^{T} U_t^2] + \frac{\delta\psi}{\phi}\left[p\text{lim}\left(\frac{1}{T}\right)\sum_{t=1}^{T} X_t^2\right]}{[\frac{\rho}{\phi}(1-\phi\rho)+\psi\kappa]\left[p\text{lim}\left(\frac{1}{T}\right)\sum_{t=1}^{T} X_t^2\right] + \kappa^2[p\text{lim}\left(\frac{1}{T}\right)\sum_{t=1}^{T} U_t^2]} \tag{3.37}$$

It is easy to see that the "relaxed local ATT" in Scenario 3 is smaller than the "restrict local ATT" in Scenario 1; however, due to the "local selection bias" in Scenario 3, $\hat{\beta}_{IV,3}$ in Scenario 3 has greater extent of inconsistency than $\hat{\beta}_{IV,1}$ in Scenario 1.

When comparing with OLS, we know that in Scenario 3, $p\text{lim}\hat{\beta}_{OLS} = \beta + \frac{\psi\rho}{(1-\phi\rho)} + \frac{\kappa[p\text{lim}\left(\frac{1}{T}\right)\sum_{t=1}^{T} U_t^2]}{(1-\phi\rho)[p\text{lim}\left(\frac{1}{T}\right)\sum_{t=1}^{T} X_t^2]}$. Therefore, in Scenario 3, it is ambiguous whether the lagged IV estimate has larger extent of inconsistency than the OLS estimate.

Scenario 4. This scenario combines Scenarios 2 and 3, and so we know that the lagged IV estimate in this scenario could be more inconsistent than the OLS estimate.

III.B. Implications

The implications of the foregoing for empirical research are thus

(1) If both $\xi = 0$ and $\psi = 0$, the lagged IV satisfies the *exclusion restriction*, but violates the *independence assumption*. In this scenario, the lagged IV is safe, because its estimate is consistent.

(2) If $\xi \neq 0$ but $\psi = 0$, the lagged IV violates both the *exclusion restriction* and the *independence assumption*. In this scenario, the lagged IV is unambiguously bad,

because its estimate is inconsistent—more so than the OLS estimate.

(3) If $\xi = 0$ but $\psi \neq 0$, the lagged IV violates both the *exclusion restriction* and the *independence assumption*. In this scenario, the lagged IV is also unambiguously bad, because its estimate is inconsistent. In addition, it is unclear whether the lagged IV estimate is more inconsistent than the OLS estimate.

(4) If $\xi \neq 0$ and $\psi \neq 0$, the lagged IV violates both the *exclusion restriction* and the *independence assumption*. In this scenario, the lagged IV is once again unambiguously bad, because its estimate is inconsistent—again more so than the OLS estimate.

IV. SIMULATION ANALYSIS

In this section, we use Monte Carlo methods to create a simulation of the theoretical setups of our four scenarios discussed before, to quantitatively discuss the biases of both the lagged IV estimates and the OLS estimates, together with the root mean squared errors (RMSE) and the likelihoods of Type I errors of the lagged IV and the OLS estimation.

IV.A. Setup

We start with Scenario 1, which only violates the independence assumption but not the exclusion restriction. Figure 1 parameterizes the relations between the outcome variable, the explanatory variable and the unobserved confounders in Scenario 1. As is shown, the unobserved confounder, regarded as a general representation of endogeneity source, is correlated both with Y_t and with X_t. The parameter δ, the direct marginal effect of U_t on Y_t, is normalized as 1. The parameter β, the direct marginal effect of X_t on Y_t, is assigned a value of either 0 or 2.

The first key parameter in our simulation is κ, the marginal effect of U_t on X_t in our setup, which measures the magnitude of the endogeneity at the violation of the independence assumption. The value of κ is assigned a value of 0.5 or 2 to represent the attenuated and the amplified marginal effect of U on X, respectively. The second and the third key parameters are the autocorrelation parameters ρ and ϕ. They are set at 0.5 and {0, 0.1, 0.2, …, 0.9}, respectively, to represent the relevance of X_t, the endogenous variable, and X_{t-1}, the lagged IV, relative to the relevance of the current and the lagged unobserved confounder. In each simulation, we generate a panel with

$T = 50$ periods and $N = 100$ cross-section units, for a total of 5,000 observations.

Our simulation follows the same data generating process (DGPs) as in section III. Each set of parameter values, shown in Table 2, are simulated 100 times. Then three estimators of β are illustrated: (1) the "naïve" estimator ($\hat{\beta}_{NAIVE}$), or the OLS estimator, that regress Y_t on X_t and ignores the unobserved confounder, (2) the "lagged IV" estimator ($\hat{\beta}_{LAGIV}$) that regress Y_t on X_t and use X_{t-1} as the IV for X_t, and (3) the "correct" estimator ($\hat{\beta}_{CORRECT}$) that regress Y_t on X_t and also the unobserved confounder. Here the "correct" estimator is the counterfactual, and since researchers cannot observe the confounders in their applied studies, our DGPs provides the tests of the performance of both the OLS estimates and the lagged IV estimates, by comparing each of their bias with the "correct" estimator, of which the bias is zero. To make our analysis simple and straightforward, we just use the one-period autocorrelation.

Three criteria are used to evaluate the performance of the lagged IV estimates: (1) bias, (2) root mean squared error (RMSE), and (3) likelihood of Type I error, which tells researchers the extent to which they could make false inference on the estimates, rejecting the true null hypotheses that $\beta = 0$.

We then discuss Scenario 2, which violates not only the independence assumption, but also the exclusion restriction directly. Figure 4 parameterizes the relations between the outcome variable, the explanatory variable and the unobserved confounders in Scenario 2. In this scenario, the first key parameter in our simulation is ξ, the marginal effect of X_{t-1} on Y_t in our setup, which measures the magnitude of the endogeneity at the violation of the exclusion restriction. The value of ξ is set at 0.5 or 2, to represent the attenuated and the amplified marginal effect of X_{t-1} on Y_t, respectively.

After those, we discuss Scenario 3, which violates not only the independence assumption, but also the exclusion restriction indirectly. Figure 7 parameterizes the relations between the outcome variable, the explanatory variable and the unobserved confounders in Scenario 3. In this scenario, the first key parameter in our simulation is ψ, the marginal effect of X_{t-1} on U_t in our setup, which measures the magnitude of the endogeneity at the violation of the exclusion restriction. The value of ψ is set at 0.5 or 2, to represent the attenuated and the amplified marginal effect of X_{t-1} on U_t, respectively.

IV.B. Monte Carlo Simulation Results

Figure 2 summarizes the simulation results when κ=0.5 and 2, $\rho = 0.5$, and ϕ ranges from 0 to 0.9. The simulation results show that

(1) both $\hat{\beta}_{NAIVE}$ and $\hat{\beta}_{LAGIV}$ are biased, and the bias of the lagged IV estimate is smaller than that of the OLS estimate. This is consistent with our theoretical prediction that as the lagged IV only violates the independence assumption in Scenario 1, it is less problematic than the OLS estimate, and

(2) As ϕ increases, the bias of the lagged IV estimate also increases; as κ increases, the bias of the lagged IV estimate decreases. This is also consistent with our theoretical prediction that the lagged IV estimate's violation of the independence assumption is quantified with $\frac{\phi}{\kappa}$, the synchronous change of U_t by X_t; as $\frac{\phi}{\kappa}$ increases, the independence assumption is violated to a larger extent and as a result, the lagged IV estimate suffers from higher bias. (3) The RMSEs show similar patterns as the biases.

Our simulations also show what happens when the null hypothesis (i.e., $\beta = 0$) is true and an applied researcher uses the lagged IV method to test the alternative hypothesis that $\beta \neq 0$. Here we use the 95% confidence levels.

Our simulation results imply that when $\kappa > 0$ and as ϕ ranges from 0 to 1, the likelihood of a Type I error rises dramatically. The reason is that lagged IV identification will lead to nonzero estimates of β even with $\beta = 0$, because δ, the marginal effect of unobserved confounder on the outcome variable, and κ, the marginal effect of unobserved confounder on the explanatory variable, are both nonzero. In addition, similar to the magnitude of estimation bias, the likelihood of rejecting the true null hypothesis rises dramatically and becomes close to 1, as ϕ goes up.

Further, Figure 3 presents simulation results when $\phi = 0.5$, ρ ranges from 0 to 1, and κ=0.5 and 2. These results show that

(1) Both $\hat{\beta}_{NAIVE}$ and $\hat{\beta}_{LAGIV}$ are biased, and the bias of the lagged IV estimate is smaller than that of the OLS estimate. This is consistent with our theoretical prediction that as the lagged IV only violates the independence assumption in Scenario 1, it is less problematic than the OLS estimate.

(2) As ρ increases, the bias of the lagged IV estimate decreases. This shows that as the relevance of the lagged IV and the endogenous variable goes up, the validity of the lagged IV also goes up.

(3) As κ increases, the bias of the lagged IV estimate decreases. This is also

consistent with our theoretical prediction that the lagged IV estimate's violation of the independence assumption is quantified with $\frac{\phi}{\kappa}$, the synchronous change of U_t by X_t; as $\frac{\phi}{\kappa}$ increases, the independence assumption is violated to a larger extent and as a result, the lagged IV estimate suffers from higher bias.

(4) The RMSEs show similar patterns as the biases. (5) The likelihood of the Type I error is very high.

In sum, our simulation results convey an unambiguous message: If the lagged explanatory variable has neither a direct causal effect on the outcome variable or on the unobserved confounder, using lagged explanatory variable as the IV in instrumental estimation can mitigate both bias and RMSE. The likelihood of a Type I error, however, can hardly be mitigated by the lagged IV method. These results imply that even if the exclusion restriction is satisfied, the lagged IV method is still problematic.

We also discuss the case in which the lagged explanatory variable has a direct causal effect on the outcome variable, the case in which the lagged explanatory variable has a direct causal effect on the unobserved confounder, and the case in which the lagged explanatory variable has direct causal effects both on the outcome variable and on the unobserved confounder. These cases coincide with Scenario 2, 3 and 4 discussed in our conceptual framework. These three cases yield much different results regarding estimation bias and RMSE, that both bias and RMSE in lagged IV estimation are significantly larger than those in OLS; besides, in these three cases the likelihood of a Type I error is close—and sometimes equal—to one, and significantly higher than those in OLS. These results imply that when lagged IV estimation violates both the exclusion restriction and the independence assumption, it even aggravates the endogeneity.

Figure 5 summarizes simulation results where ξ=0.5 and 2, $\rho = 0.5$, and ϕ ranges from 0 to 0.9. These simulation results show that

(1) both $\hat{\beta}_{NAIVE}$ and $\hat{\beta}_{LAGIV}$ are biased, and the bias of the lagged IV estimate is much larger than that of the OLS estimate. This is consistent with our theoretical prediction that as the lagged IV violates both the independence assumption and the exclusion restriction in Scenario 2, it is much more problematic than the OLS estimate.

(2) As ϕ increases, the bias of the lagged IV estimate also increases. This is also consistent with our theoretical prediction that the lagged IV estimate's violation of the independence assumption is quantified with $\frac{\phi}{\kappa}$, the synchronous change of U_t by X_t;

as $\frac{\phi}{\kappa}$ increases, the independence assumption is violated to a larger extent and as a result, the lagged IV estimate suffers from higher bias.

(3) As ξ increases, the bias of the lagged IV estimate also increases. This is also consistent with our theoretical prediction that the lagged IV estimate's violation of the exclusion restriction in Scenario 2 is quantified with ξ, the marginal effect of X_{t-1} on Y_t; as ξ increases, the exclusion restriction is violated to a larger extent and as a result, the lagged IV estimate suffers from higher bias.

(4) The RMSEs show similar patterns as the biases.

(5) The likelihood of the Type I error is very high, and close to 1.

Figure 6 summarizes simulation results where ξ=0.5 and 2, $\phi = 0.5$, and ρ ranges from 0 to 0.9. These simulation results show that

(1) Both $\hat{\beta}_{NAIVE}$ and $\hat{\beta}_{LAGIV}$ are biased, and the bias of the lagged IV estimate is much larger than that of the OLS estimate. This is consistent with our theoretical prediction that as the lagged IV violates both the independence assumption and the exclusion restriction in Scenario 2, it is much more problematic than the OLS estimate.

(2) As ρ increases, the bias of the lagged IV estimate decreases. This shows that as the relevance of the lagged IV and the endogenous variable goes up, the validity of the lagged IV also goes up.

(3) As ξ increases, the bias of the lagged IV estimate also increases. This is also consistent with our theoretical prediction that the lagged IV estimate's violation of the exclusion restriction in Scenario 2 is quantified with ξ, the marginal effect of X_{t-1} on Y_t; as ξ increases, the exclusion restriction is violated to a larger extent and as a result, the lagged IV estimate suffers from higher bias.

(4) The RMSEs show similar patterns as the biases.

(5) The likelihood of the Type I error is very high, and close to 1.

Figure 8 shows simulation results where ψ=0.5 and 2, $\rho = 0.5$, and ϕ ranges from 0 to 0.9. These results show that

(1) both $\hat{\beta}_{NAIVE}$ and $\hat{\beta}_{LAGIV}$ are biased, and the bias of the lagged IV estimate is much larger than that of the OLS estimate. This is consistent with our theoretical prediction that as the lagged IV violates both the independence assumption and the exclusion restriction in Scenario 3, it is much more problematic than the OLS estimate.

(2) As ϕ increases, the bias of the lagged IV estimate also increases. This is also consistent with our theoretical prediction that the lagged IV estimate's violation of the

independence assumption is quantified with $\frac{\phi}{\kappa}$, the synchronous change of U_t by X_t; as $\frac{\phi}{\kappa}$ increases, the independence assumption is violated to a larger extent and as a result, the lagged IV estimate suffers from higher bias.

(3) As ψ increases, the bias of the lagged IV estimate also increases. This is also consistent with our theoretical prediction that the lagged IV estimate's violation of the exclusion restriction in Scenario 3 is quantified with ψ, the marginal effect of X_{t-1} on U_t; as ψ increases, the exclusion restriction is violated to a larger extent and as a result, the lagged IV estimate suffers from higher bias.

(4) The RMSEs show similar patterns as the biases. (5) The likelihood of the Type I error is very high, and close to 1.

Finally, Figure 9 shows simulation results where ψ=0.5 and 2, $\phi = 0.5$, and ρ ranges from 0 to 0.9. These results show that

(1) both $\hat{\beta}_{NAIVE}$ and $\hat{\beta}_{LAGIV}$ are biased, and the bias of the lagged IV estimate is much larger than that of the OLS estimate. This is consistent with our theoretical prediction that as the lagged IV violates both the independence assumption and the exclusion restriction in Scenario 3, it is much more problematic than the OLS estimate.

(2) As ρ increases, the bias of the lagged IV estimate decreases. This shows that as the relevance of the lagged IV and the endogenous variable goes up, the validity of the lagged IV also goes up.

(3) As ψ increases, the bias of the lagged IV estimate also increases. This is also consistent with our theoretical prediction that the lagged IV estimate's violation of the exclusion restriction in Scenario 3 is quantified with ψ, the marginal effect of X_{t-1} on U_t; as ψ increases, the exclusion restriction is violated to a larger extent and as a result, the lagged IV estimate suffers from higher bias.

(4) The RMSEs show similar patterns as the biases.

(5) The likelihood of the Type I error is very high, and close to 1.

V. CONCLUSION

We have looked at the practice of using a lagged endogenous variable to use it as an instrumental variable for the same endogenous variable—a practice we have dubbed "lagged IV" for brevity. Given our discussion of the independence assumption and exclusion restriction, our main result is that if the lagged IV satisfies the exclusion restriction (which assumes no causal influence), then the lagged IV method is acceptable and helpful, as its estimate is consistent and is less biased than the OLS estimate. A violation of independence assumption, however, results in a high likelihood that the lagged IV will lead to a Type I error. If the lagged IV violates both the independence assumption and the exclusion restriction, the resulting estimate is unambiguously inconsistent, and is much more biased than the OLS estimate.

When using a lagged IV, most applied researchers fail to discuss the independence assumption and the exclusion restriction in details, assuming empirically that the lagged IV method can at the very least yield estimates whose bias is less than that of OLS. Our simulation results show that this only obtains in a very narrow range of cases. Worse, a lagged IV always significantly increases the likelihood of a Type I error.

In the absence of experimental data, causal inference requires either (i) that the assumption of selection on observables be satisfied (i.e., that all back-door paths between the treatment and outcome variable are successfully blocked), (ii) that the data include a mediator variable between the treatment and outcome variables which satisfy the requirements of the front-door criterion (Pearl 1995, 2000, 2009; Bellemare, Bloem, and Wexler 2020), or (iii) that the data include a valid IV (i.e., a variable that satisfies the requirements we have laid out at the very beginning of this paper). The results in this paper show that in the latter case, using a lagged endogenous variable is unlikely to lead to credible estimates, and so the practice of using lagged IVs should be discontinued.

Appendix

A.I. Derivation of $Cov(X,U)$

In Scenario 1 and 2, following the appendix of Bellemare et al. (2017), we have, given the equations (3.2) and (3.3), or (3.17) and (3.18), the expression that

$$Cov(X_{t-1},U_{t-1}) = Cov\left(\frac{1}{\rho}X_t - \frac{\kappa}{\rho}U_t - \frac{1}{\rho}\eta_t, \frac{1}{\phi}U_t - \frac{1}{\phi}\upsilon_t\right) \quad (A.1)$$

Then we have

$$Cov(X_{t-1},U_{t-1}) = \frac{1}{\phi\rho}[Cov(X_t,U_t) - \kappa Var(U_t)]$$

which yields

$$Cov(X_t,U_t) - \kappa Var(U_t) = \phi\rho Cov(X_{t-1},U_{t-1}) \quad (A.2)$$

Since $\rho,\phi \in (0,1)$, both X and U are mean-reverting series, that is, the covariance between X and U does not depend on t. In other words, asymptotically, we have

$$p\lim_{t\to\infty} Cov(X_t,U_t) = p\lim_{t\to\infty} Cov(X_{t-1},U_{t-1}) = Cov(X,U) \quad (A.3)$$

Therefore, (A.2) becomes

$$Cov(X,U) - \kappa Var(U) = \phi\rho Cov(X,U) \quad (A.4)$$

implying that

$$Cov(X,U) = \frac{\kappa Var(U)}{1-\phi\rho} \quad (A.5)$$

In Scenario 3, similarly, we have, given the equations (3.26) and (3.27), the expression that

$$Cov(X_{t-1},U_{t-1}) = Cov\left(\frac{1}{\rho}X_t - \frac{\kappa}{\rho}U_t - \frac{1}{\rho}\eta_t, \frac{1}{\phi}U_t - \frac{\psi}{\phi}X_{t-1} - \frac{1}{\phi}\upsilon_t\right) \quad (A.6)$$

Then we have

$$Cov(X_{t-1},U_{t-1}) \\ = \frac{1}{\phi\rho}[Cov(X_t,U_t) - \psi Cov(X_t,X_{t-1}) - \kappa Var(U_t) + \kappa\psi Cov(X_{t-1},U_t)] \quad (A.7)$$

which yields

$$Cov(X_t,U_t) - \psi Cov(X_t,X_{t-1}) - \kappa Var(U_t) + \kappa\psi Cov(X_{t-1},U_t) \\ = \phi\rho Cov(X_{t-1},U_{t-1}) \quad (A.8)$$

Similarly, since $\rho,\phi \in (0,1)$, both X and U are mean-reverting series, that is, the covariance between X and U does not depend on t. In other words, asymptotically, we have

$$p\lim_{t\to\infty} Cov(X_t, U_t) = p\lim_{t\to\infty} Cov(X_{t-1}, U_{t-1}) = Cov(X, U) \quad \text{(A.9)}$$

We also know that

$$Cov(X_t, X_{t-1}) = Cov(\rho X_{t-1} + \kappa U_t + \eta_t, X_{t-1})$$
$$= \rho Var(X) + \kappa Cov(X_{t-1}, U_t)$$

and that

$$Cov(X_{t-1}, U_t) = Cov(X_{t-1}, \phi U_{t-1} + \psi X_{t-1} + \upsilon_t)$$
$$= \phi\, Cov(X, U) + \psi Var(X)$$

Therefore, (A.8) becomes

$$Cov(X, U) - \kappa Var(U) - \psi\rho Var(X) = \phi\rho Cov(X, U) \quad \text{(A.10)}$$

implying that

$$Cov(X, U) = \frac{\kappa Var(U) + \psi\rho Var(X)}{1 - \phi\rho} \quad \text{(A.11)}$$

Therefore,

$$\frac{Cov(X_{t-1}, U_t)}{Var(X_{t-1})} = \frac{\phi\kappa Var(U) + \phi\psi\rho Var(X)}{(1 - \phi\rho)Var(X)} + \psi$$
$$= \frac{\phi\kappa Var(U)}{(1 - \phi\rho)Var(X)} + \frac{\psi}{1 - \phi\rho} \quad \text{(A.12)}$$

REFERENCES

Anderson, Theodore Wilbur, and Cheng Hsiao. "Estimation of Dynamic Models with Error Components." *Journal of the American Statistical Association* 76, no. 375 (1981): 598-606.

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91, no. 434 (1996): 444-455.

Angrist, Joshua, and Jörn-Steffen Pischke. "Mostly Harmless Econometrics: An Empiricists Guide." Princeton: Princeton University Press, 2009.

Bellemare, Marc F., Jeffrey R. Bloem and Noah Wexler. *The Paper of How: Estimating Treatment Effects Using the Front-Door Criterion*. Working Paper, 2020.

Bellemare, Marc F., Takaaki Masaki, and Thomas B. Pepinsky. "Lagged Explanatory Variables and the Estimation of Causal Effect." *The Journal of Politics* 79, no. 3 (2017): 949-963.

Blundell, Richard, and Stephen Bond. "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models." *Journal of Econometrics* 87, no. 1 (1998): 115-143.

Blundell, Richard, and Stephen Bond. "GMM Estimation with Persistent Panel Data: An Application to Production Functions." *Econometric Reviews* 19, no. 3 (2000): 321-340.

Imbens, Guido W., and Joshua D. Angrist. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62, no. 2 (1994): 467-475.

Pearl, Judea. "Causal Diagrams for Empirical Research." *Biometrika* 82, no. 4 (1995): 669-688.

Pearl, Judea. "Causal Inference without Counterfactuals: Comment." *Journal of the American Statistical Association* 95, no. 450 (2000): 428-431.

Pearl, Judea. *Causality*. Cambridge University Press, 2009.

Rossi, Peter E. "Even the Rich Can Make Themselves Poor: A Critical Examination of IV Methods in Marketing Applications." *Marketing Science* 33, no. 5 (2014): 655-672.

Todd, Petra E., and Kenneth I. Wolpin. "On the Specification and Estimation of the Production Function for Cognitive Achievement." *The Economic Journal* 113, no. 485 (2003).

Table 1. Reviewed Journals Published in 2013-2018, Using Lagged IV Methods

Journal Name	Discipline	2013-2018	2015-2018
American Economic Review	Economics	5	3
Econometrica	Economics	0	0
Journal of Political Economy	Economics	1	0
Quarterly Journal of Economics	Economics	3	2
Review of Economic Studies	Economics	3	1
Review of Economics & Statistics	Economics	7	2
American Political Science Review	Political Science	1	0
American Journal of Political Science	Political Science	1	1
British Journal of Political Science	Political Science	6	4
Comparative Political Studies	Political Science	3	1
Journal of Politics	Political Science	1	0

Table 2. Simulation Parameters

Parameters	Causal Pathway	Simulation Values
Basic Parameters		
β	$X_t \rightarrow Y_t$	{0, 2}
δ	$U_t \rightarrow Y_t$	{1}
Key Parameters		
ϕ	$U_{t-1} \rightarrow U_t$	{0, 0.1, 0.2,…,0.9}, {0.5}
ρ	$X_{t-1} \rightarrow X_t$	{0.5}, {0, 0.1, 0.2,…,0.9}
κ	$U_t \rightarrow X_t, U_{t-1} \rightarrow X_{t-1}$	{0.5, 2}
ξ	$X_{t-1} \rightarrow Y_t$	{0.5, 2}
ψ	$X_{t-1} \rightarrow U_t$	{0.5, 2}

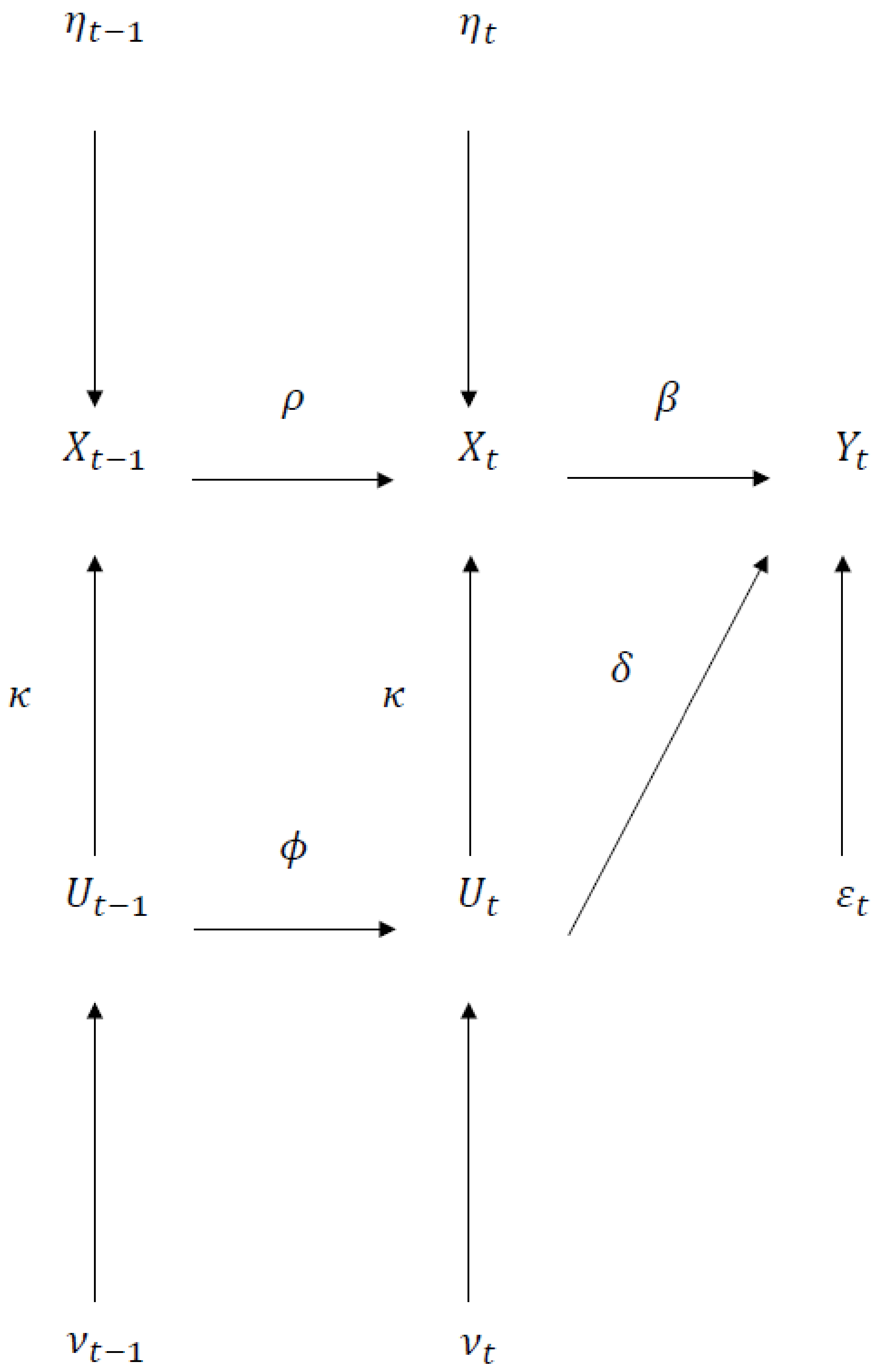

Figure 1. Representation of Monte Carlo Simulation Setup

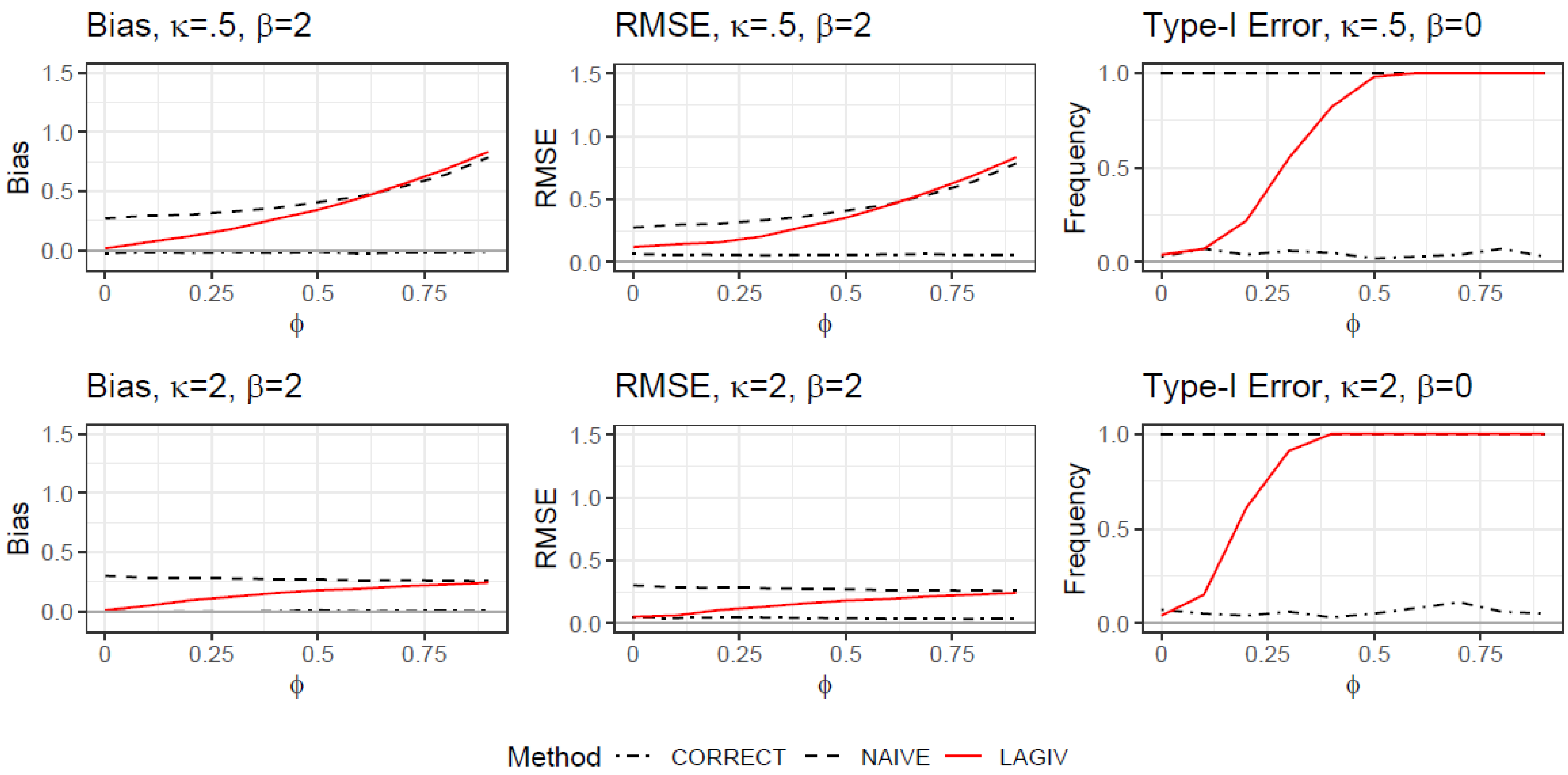

Figure 2. Monte Carlo Results: κ=0.5 and 2, ϕ ranges from 0 to 1, $\rho = 0.5$

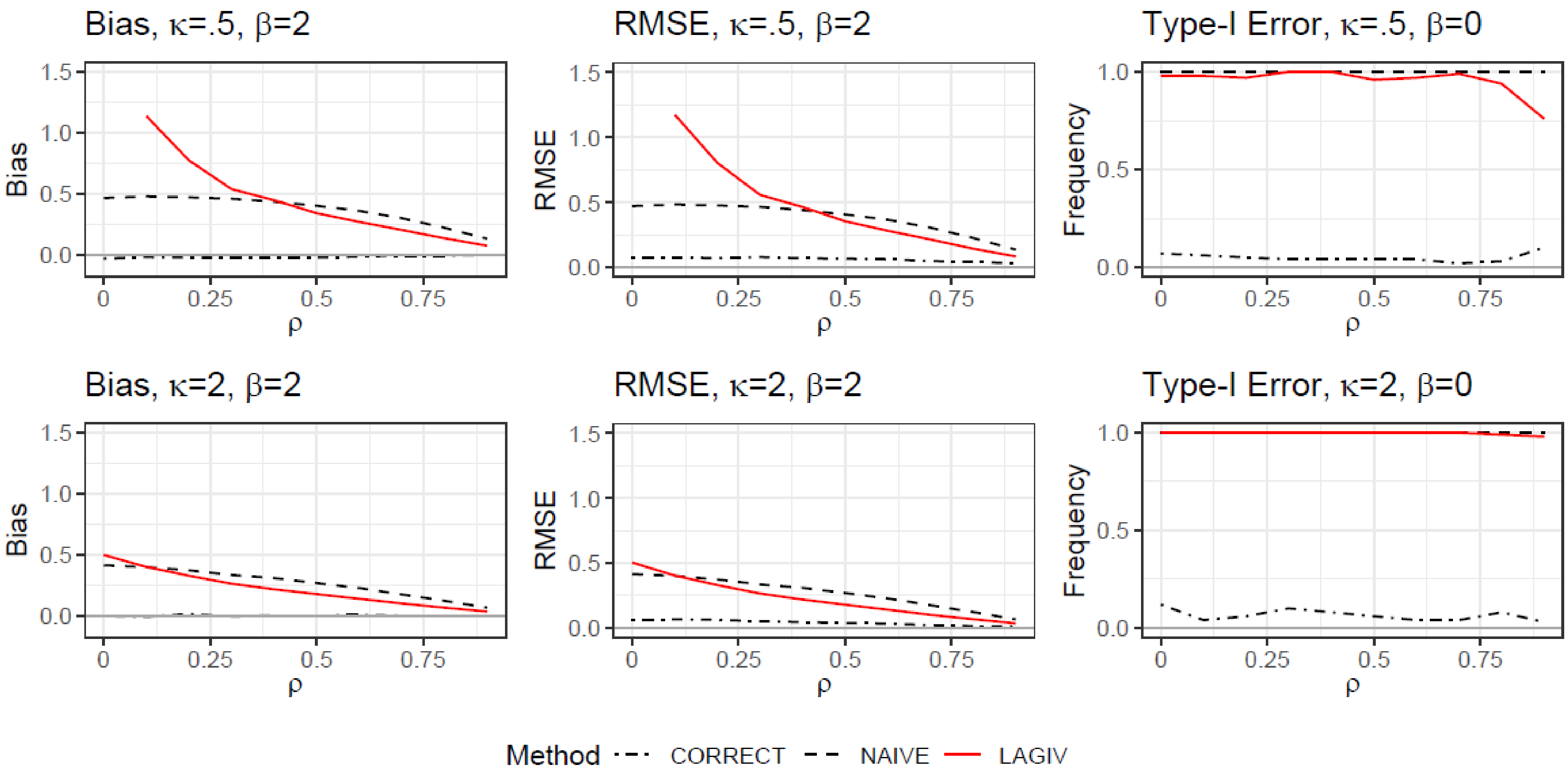

Figure 3. Monte Carlo Results: κ=0.5 and 2, ρ ranges from 0 to 1, $\phi = 0.5$

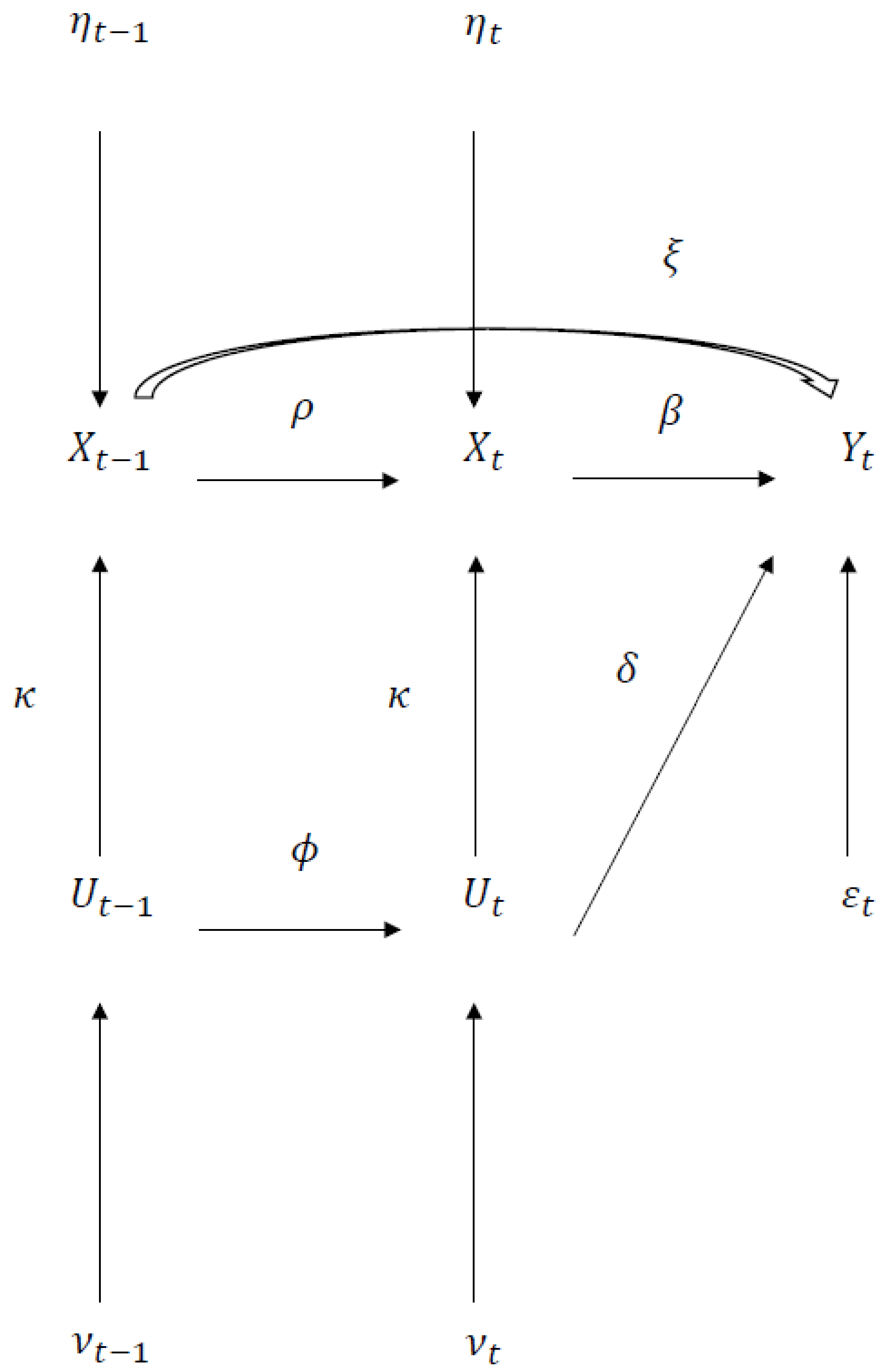

Figure 4. Representation of Monte Carlo Simulation Setup: X_{t-1} Also Has Causal Effects on Y_t

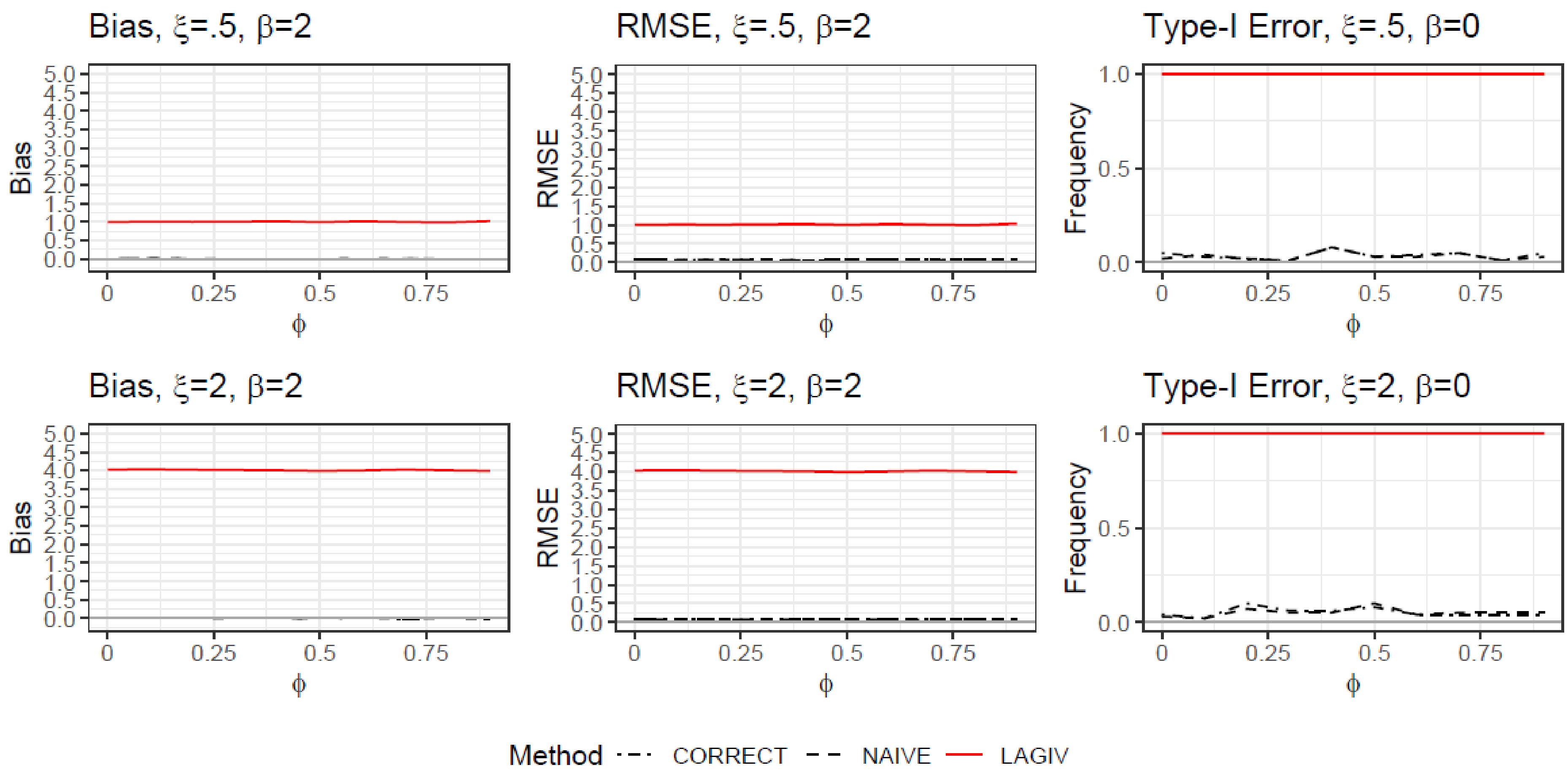

Figure 5. Monte Carlo Results: κ=0.5 and 2, ϕ ranges from 0 to 1; Lagged Causality on Outcome variable

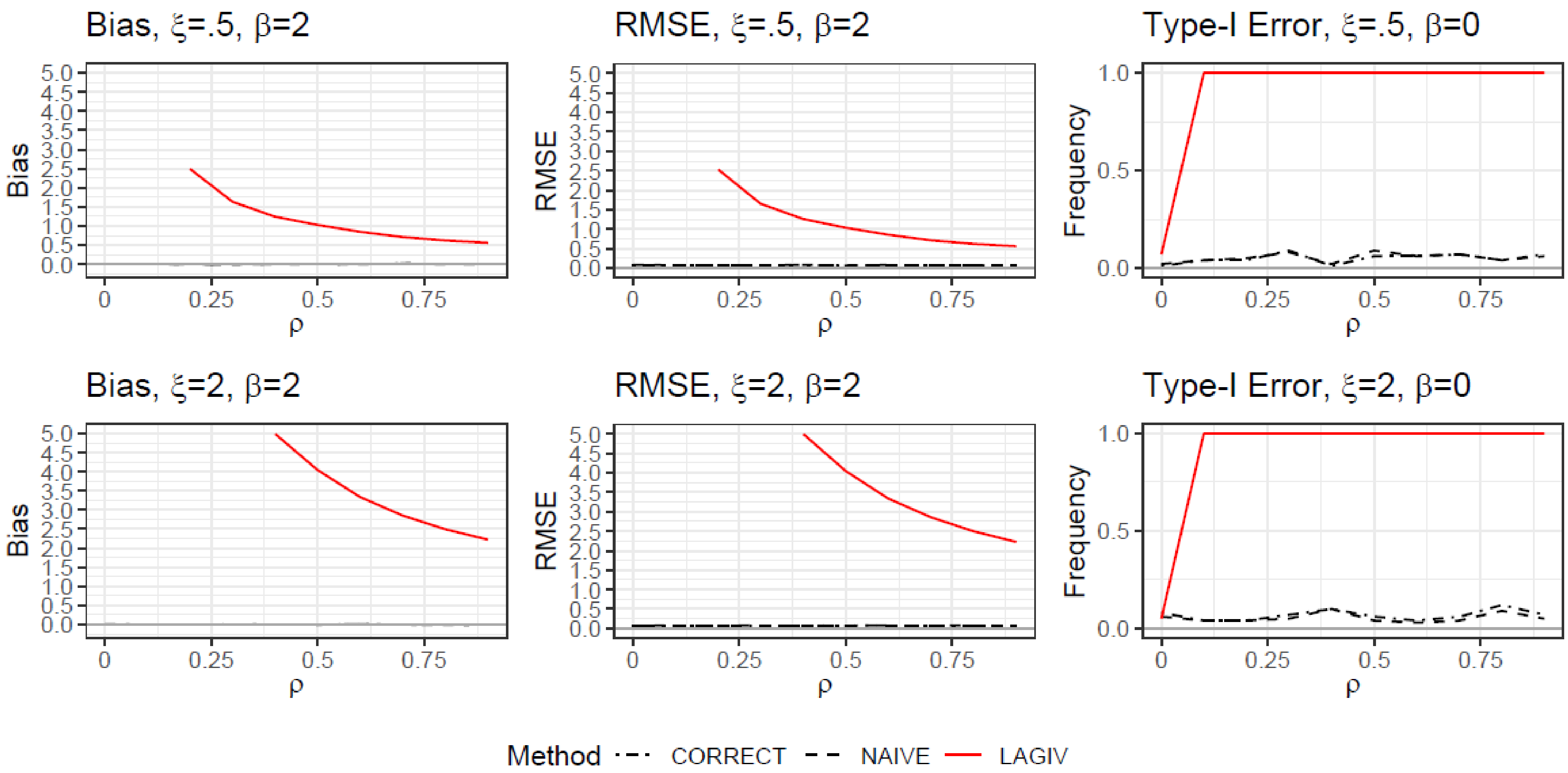

Figure 6. Monte Carlo Results: κ=0.5 and 2, ρ ranges from 0 to 1; Lagged Causality on Outcome variable

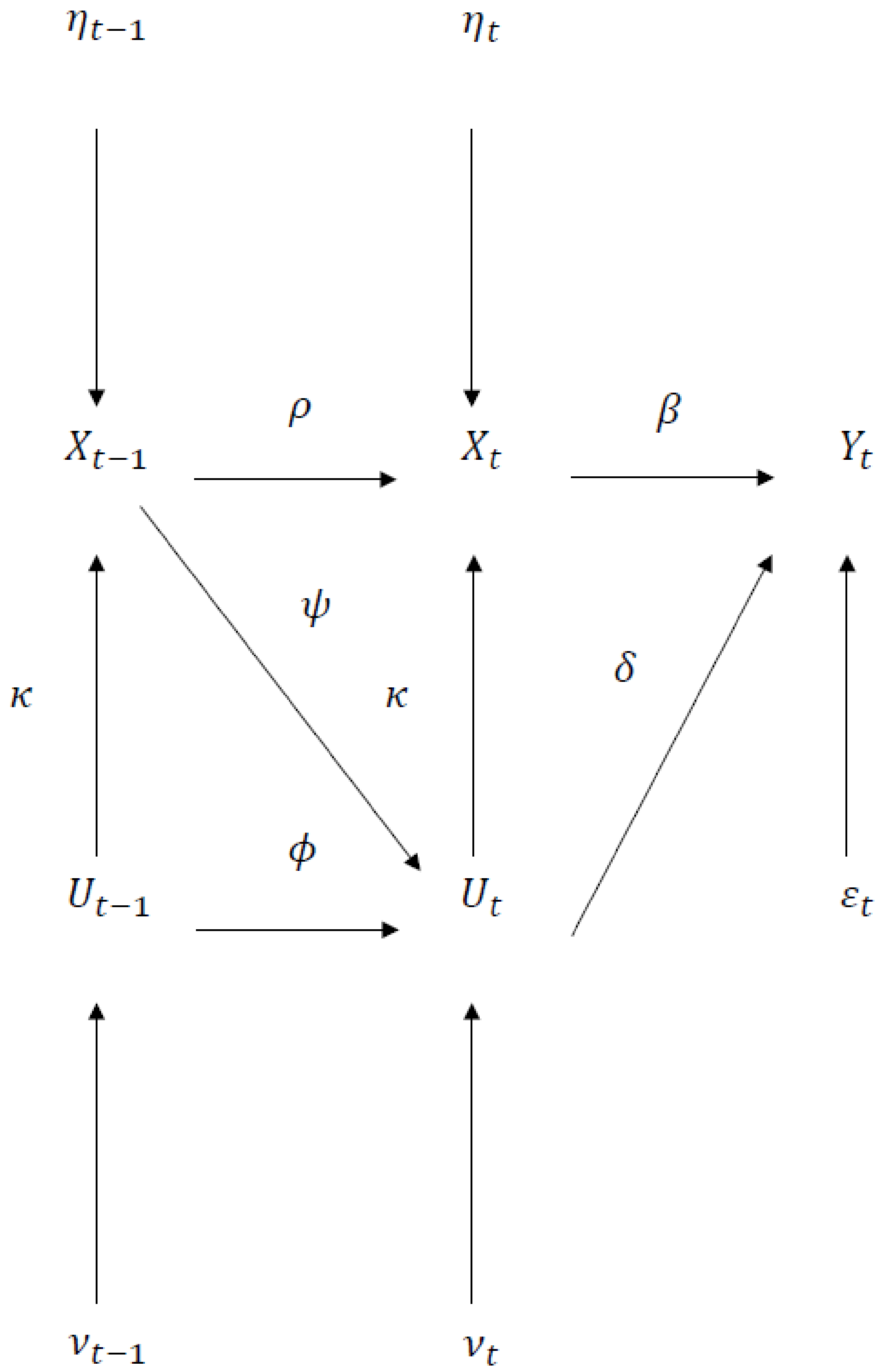

Figure 7. Representation of Monte Carlo Simulation Setup: X_{t-1} Also Has Causal Effects on U_t

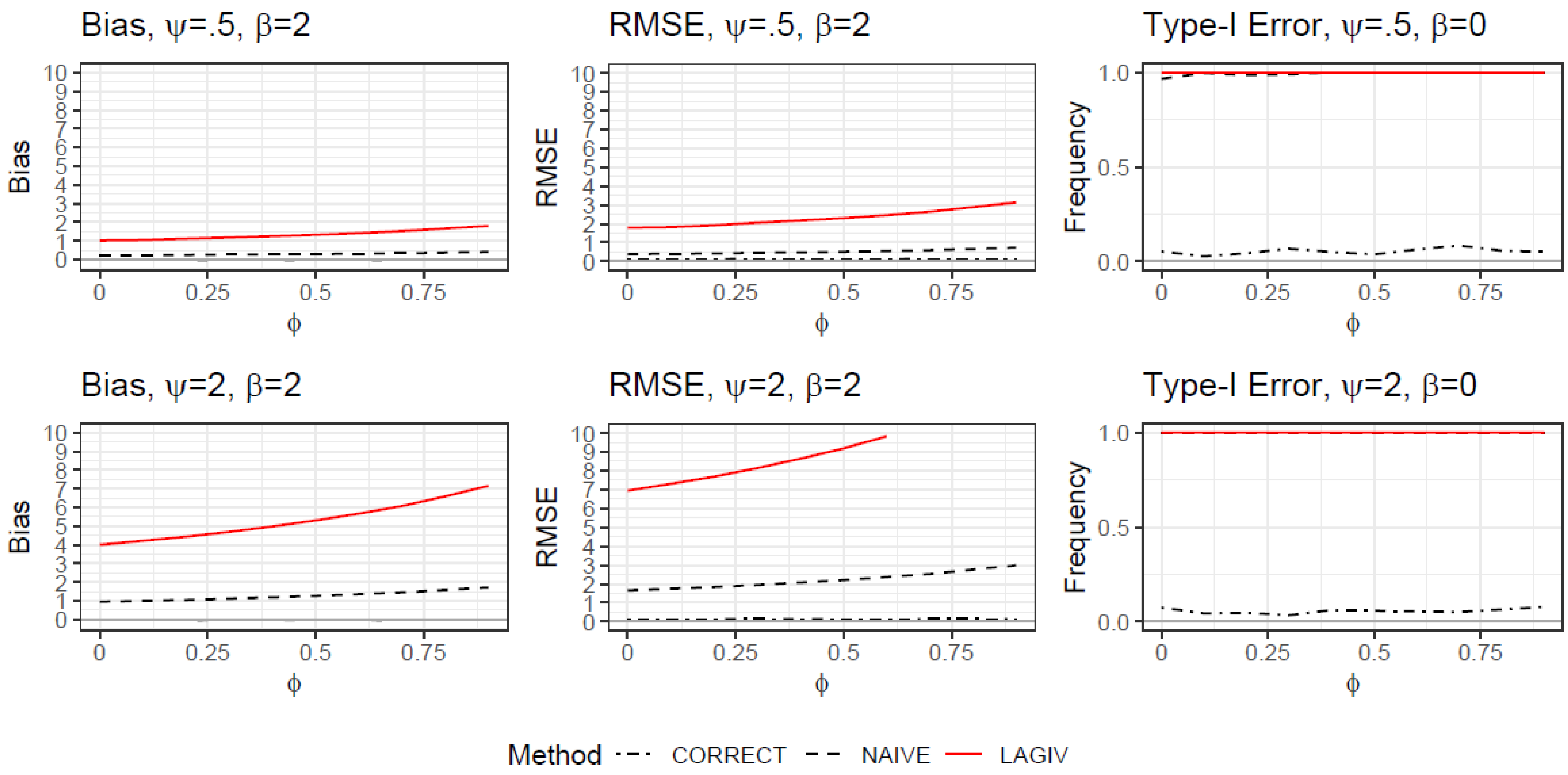

Figure 8. Monte Carlo Results: κ=0.5 and 2, ϕ ranges from 0 to 1; Lagged Causality on Unobserved Confounder

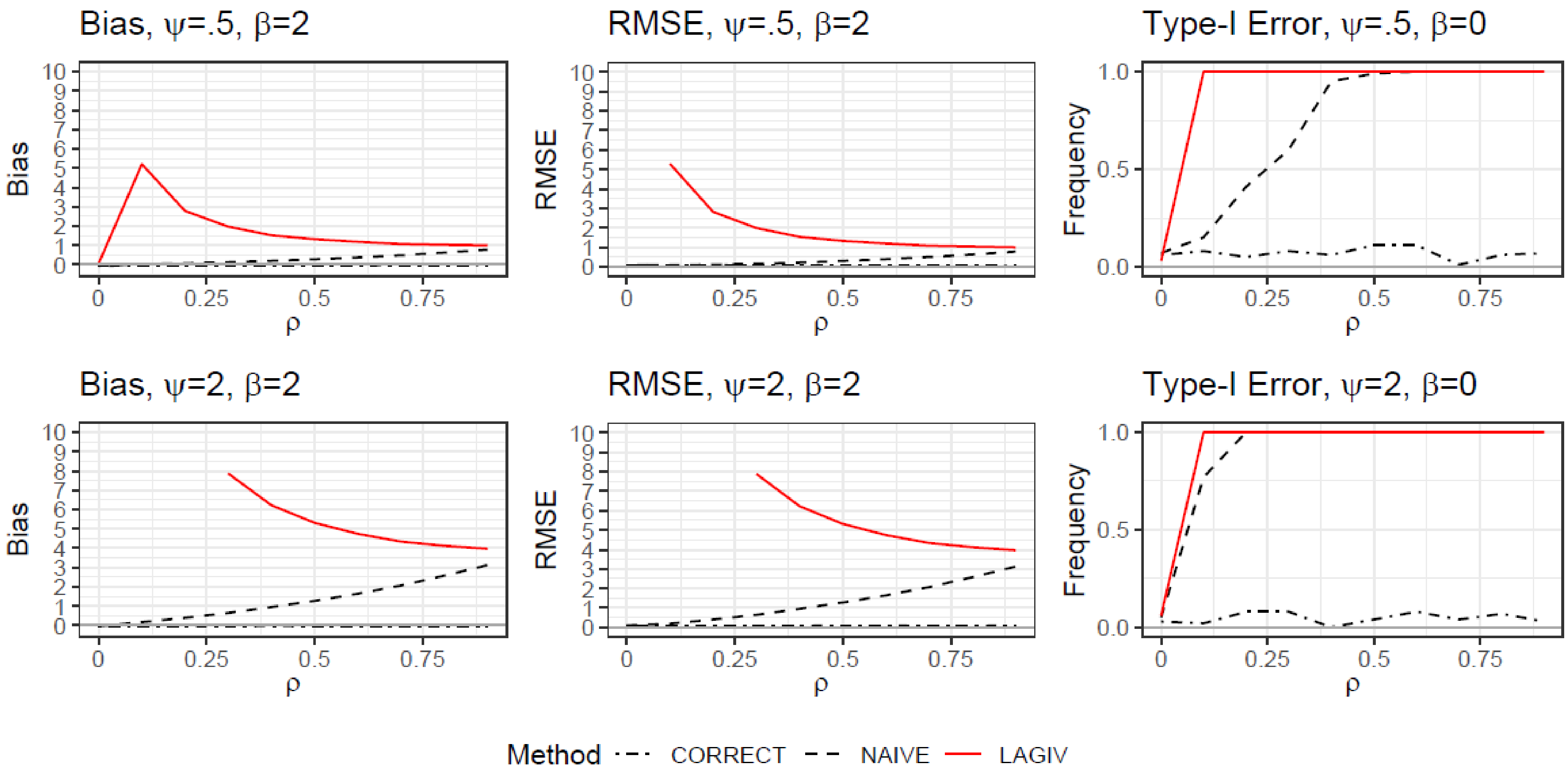

Figure 9. Monte Carlo Results: κ=0.5 and 2, ρ ranges from 0 to 1; Lagged Causality on Unobserved Confounder