Lagged Variables as Instruments*

Yu Wang† Marc F. Bellemare‡

May 31, 2019

*PLEASE DO NOT CITE WITHOUT PERMISSION FROM THE AUTHORS. We thank Jeff Bloem, Jay Coggins, John Freeman, and Paul Glewwe for useful comments and suggestions. All remaining errors are ours.

†Corresponding Author. PhD Candidate, Department of Applied Economics, University of Minnesota, Email: wang5979@umn.edu.

‡Northrop Professor, Department of Applied Economics, University of Minnesota, Email: mbellema@umn.edu.

LAGS AS INSTRUMENTS

Abstract

Lagged explanatory variable remain commonly used as instrumental variables (IVs) to address endogeneity concerns in empirical studies with observational data. Few theoretical studies, however, address whether "lagged IVs" mitigate endogeneity. We develop a theoretical setup in which dynamics among the endogenous explanatory variable and the unobserved confounders cannot be ruled out and look at the consequences of lagged IVs for bias and the root mean square error (RMSE). We then use Monte Carlo simulations to illustrate our analytical findings. We show that when lagged explanatory variables have no direct causal effect on the dependent variable or on the unobserved confounders, the "lagged IV" method mitigates the endogeneity problem by reducing both bias and the root mean squared errors given specific parameter values relative to the naïve OLS case. If either or both of the causal relationships above are present, however, lagged IVs increase both bias and the RMSE relative to OLS, and they virtually blow up the likelihood of a Type I error to one.

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

JEL Classification Codes: C13, C15, C21

1 Introduction

The credibility revolution of the last few decades in applied economics has brought about a greater emphasis on causal inference (Angrist and Pischke, 2010). One of the oldest of all working tools available to an applied econometrician working with observational data and interested in the estimation of a causal relationship is the use of instrumental variables (IVs).

To identify the causal relationship between an endogenous variable of interest x and a dependent variable y in such a setup, one uses an instrument z, on which x is conditioned before y is regressed on $\widehat{x}$, the predicted value of x when conditioned on z (Angrist and Pischke, 2014).[1] Under the right conditions—namely that the instrumental (IV) be relevant, and that it satisfies both the independence assumption and the exclusion restriction—the coefficient on $\widehat{x}$ is the local average treatment effect (LATE) of x on y (Imbens and Angrist, 1994), i.e., the effect of x on y for the subset of compliers. In a setup where x and z are both dichotomous and respectively refer to the uptake of a treatment and the assignment to the same treatment, the subset of compliers is the subset of observations for which $x = 1$ because $z = 1$.

But as every econometrics textbook is quick to emphasize, a good instrument—one that satisfies the right conditions above—is hard to find. As a result, applied econometricians often settle on less-than-ideal IVs in an effort to "exogenize" x, doing so without carefully thinking through the consequences

[1] We focus in this paper on the case where there is a single endogenous variable of interest, and for which there is a single IV.

of their choices for bias and inference.

One such less-than-ideal identification strategy is the use of what we refer to throughout this paper as a "lagged IV." In the context above, a lagged IV entails using a lag x_{it-1} of the variable of interest x_{it} as an IV for x_{it}.[2] The argument that is typically (and often implicitly) made in such cases is that since x_{it-1} precedes x_{it} in time, the causality runs entirely from x_{it-1} to x_{it}, and since there is presumably a high degree of autocorrelation in x, x_{it-1} should be a valid IV for x_{it-1}. Though the latter is testable and tends to be true in observational data, the former often confounds the notion of statistical endogeneity with that of theoretical endogeneity.

Still, there might be cases where the use of a lagged IV might be warranted. In this paper, we look at the consequences of a lagged IV on the bias of the estimated coefficient on $\widehat{x}$, its root mean squared error (RMSE), and on the likelihood of making a Type I error, i.e., on the likelihood of rejecting the null hypothesis that the estimated coefficient on $\widehat{x}$ is equal to zero when it is equal to zero. We first do so analytically, which allows identifying the precise conditions under which one can use a lagged IV. Though we maintain the assumption of IV relevance throughout, we are careful to characterize what happens when (i) both the independence assumption (i.e., the assumption that the IV should have no association with latent outcomes

[2]Though our notation implies the use of longitudinal data, it is not uncommon for applied econometricians to use cross-sectional data where both the contemporaneous and the lagged value of a specific variables was collected. This is usually done with the use of a lagged IV identification strategy in mind.

or with latent variables of interest) and the exclusion restriction (i.e., the lagged IV influences the dependent variable only through the variable of interest) hold, (ii) the independence assumption is violated but the exclusion restriction holds, (ii) the independence assumption holds but the exclusion restriction is violated, and (iv) both the independence assumption and the exclusion restriction are violated. We next use Monte Carlo simulations to show what happens to bias, RMSE, and the likelihood of a Type I error for a broad range of the relevant parameters.

There is a small literature on the use of lagged variables for identification. Blundell and Bond (1998, 2000) argue that since lagged explanatory variables tend to only be weakly correlated with the first difference of the endogenous explanatory variable, GMM using lagged explanatory variables may not solve the endogeneity problem. Similarly, in his discussion of poor practices surrounding the use of IVs in the marketing literature, Rossi (2014) briefly touches upon how lagged IVs tend not to be valid. Reed (2015) and Bellemare et al. (2017) both look at the practice of replacing an endogenous variable with its lag, and Reed concludes that instead of using a lag as a control variable, one should use it as an IV, adding that a lagged IV identification strategy only works if x_{t-1} satisfies both the exclusion restriction and is a strong enough instrument. Reed, however, only considers situations where endogeneity stems from simultaneity between y and x.

We build on Reed's (2015) result and generalize it by bringing in the independence assumption in addition to the exclusion restriction and relevance of the IV as well as by considering any and all sources of statistical endogeneity, captured here by the presence of unobserved confounders affecting both x and y. We then build on Bellemare et al.'s (2017) Monte Carlo framework to explore the consequences of lagged IVs for bias, RMSE, and the likelihood of making a Type I error.

On the one hand, we find that if the lagged IV x_{it-1} has no direct causal impact (i) on the dependent variable nor (ii) on the unobserved confounder, it violates the independence assumption, but not the exclusion restriction. In this case, a lagged IV can mitigate the endogeneity problems by reducing bias and the root mean square error (RMSE) relative to OLS for common ranges of parameter values. Even in this case, however, the likelihood of a Type I error remains large. On the other hand, we find that if the lagged IV x_{it-1} has a direct causal impact (i) on the dependent variable, on (ii) on the unobserved confounder, or both, then it violates the exclusion restriction in addition to the independence assumption. In such cases, a lagged IV worsens the endogeneity problem by increasing bias as well as the RMSE relative to OLS. Moreover, in such cases, the likelihood of Type I error is almost always equal to one for common ranges of parameter values. In practical terms, this means that the use of a lagged IV often leads one to report coefficients estimates of questionable economic significance (because of the increased bias) and statistical significance (because of the greater likelihood of a Type I error). Worse, the use of lagged IVs will tend to lead one to conclude that a causal relationship exists where it does not.

The remainder of this paper is organized as follows. In section 2, we present our conceptual framework and discuss the independence and exclusion restriction assumptions made when using IVs, and how the violation of either or both of those assumptions by a lagged IV can compromise identification. We then present some evidence about how common the use of lagged IVs is in economics. Section 3 derives our analytical results and show the bias of lagged IV compared to that of OLS. In section 4, we conduct Monte Carlo simulations aimed at exploring the properties of lagged IV compared to OLS, and show the consequences of using a lagged IV for identification on the bias, RMSE, and likelihood of a Type I error. Section 5 summarizes our main findings and concludes with a list of recommendations for applied econometricians.

2 The Problem with Lagged IVs

In this section, we first lay out our conceptual framework, which allows us to state the problem with using a lagged IV to identify a causal relationship of interest. We then discuss how widespread the use of lagged IVs remains is in the literature, focusing on the economics literature of the past five years.

2.1 Nature of the Problem

Suppose we have the structural equation

$$y_{it} = \beta x_{it} + \theta x_{it-1} + \delta u_{it} + \epsilon_{it}, \tag{1}$$

where y, x, and ϵ respectively denote the dependent variable, the variable of interest, and an error term with mean zero, as usual, but where u denotes confounders. We refer to equation (1) as "structural" because it summarizes the exact causal relationships between y and the variables on the right-hand side.

Suppose that $Cov(x_{it}, u_{it}) \neq 0$ and that u_{it} is unobserved, i.e., there is an endogeneity problem. If $\theta \neq 0$, the lagged variable of interest x_{it-1} has a direct causal impact on the dependent variable y_{it}; similarly, if $\theta = 0$, x_{it-1} has no direct causal impact on y_{it}.

We specify two autocorrelation functions: one for x, and one for u, such that

$$x_{it} = \rho x_{it-1} + \kappa u_{it} + \eta_{it}, \tag{2}$$

and

$$u_{it} = \phi u_{it-1} + \psi x_{it-1} + \upsilon_{it}. \tag{3}$$

Using the framework laid out in equations (1), (2), and (3), we can explore four distinct endogeneity scenarios:

1. $\theta = 0$ and $\psi = 0$, i.e., the lagged variable of interest has no direct causal impact on the dependent variable, nor does it have a causal impact on the unobserved confounder.

2. $\theta \neq 0$ and $\psi = 0$, i.e., the lagged variable of interest has a direct causal impact on the dependent variable, but it does not have a causal impact on the unobserved confounder.

3. $\theta = 0$ and $\psi \neq 0$, i.e., the lagged variable of interest has no direct causal impact on the dependent variable, but it has a causal impact on the unobserved confounder.

4. $\theta \neq 0$ and $\psi \neq 0$, i.e., the lagged variable of interest has a direct causal impact on the dependent variable, and it has a causal impact on the unobserved confounder.

Recall that one of the key assumption of the local average treatment effect (LATE) theorem is the independence assumption, which implies that the IV should have no association with latent outcomes or with latent variables of interest (Angrist and Pischke 2009). Here, this means that

$$[\{y_{it}(x_{it}, x_{it-1}); \forall x_{it}, x_{it-1}\}, \{x_{it}(x_{it-1}) \forall x_{it-1}\}] \perp x_{it-1}, \tag{4}$$

which implies that the lagged IV should have an effect that is similar effect to what random assignment of x_{it} does.

Another key assumption of the LATE theorem is the exclusion restriction, i.e., that $y_{it}(x_{it}, x_{it-1})$ only be a function of x_{it}. Generally speaking, this assumption—the exclusion restriction—requires that the lagged IV influence the dependent variable only through the variable of interest. This is denoted as

$$y_{it}(x_{it}, x_{it-1}) = y_{it}(x_{it}, x^{'}_{it-1}) \forall x^{'}_{it-1} \neq x_{it-1}. \tag{5}$$

We now turn to discussing the four scenarios above, beginning with those scenarios that feature endogeneity, viz. scenarios 2 to 4.

In scenario 2, since $\theta \neq 0$, x_{it-1} directly influences y_{it} via its marginal effect θ. In Scenario 3, although $\theta = 0$, $\psi \neq 0$, and x_{it-1} still influences y_{it} via marginal effect $\theta\psi$, derived from equations (2) and (3). Therefore, both scenario 2 and 3 violate not only the independence assumption, but also the exclusion restriction, and so they will result in biased estimates. A similar result obtains for scenario 4, which is just a combination of the undesirable features (i.e., $\theta \neq 0$ and $\psi \neq 0$) in scenarios 2 and 3.

Since in scenario 3, x_{it-1} has a direct impact on u_{it}, which could include more than one unobserved covariate, it implies that x_{it-1} could have more than one causal path whereby it influences y_{it}. Accordingly, even if theoretical arguments state that the lagged variable of interest has no direct impact on the dependent variable—in other words, even if those arguments make the case that scenario 2 does not hold—it is difficult to argue against the possible existence of scenarios 3 and 4, which both results in a violation of the exclusion restriction.

Turning to scenario 1, although the lagged IV in this case has neither a direct causal impact on the dependent variable nor on the unobserved confounder, the lagged IV may still indirectly be correlated with the dependent variable. Specifically, since u_{it-1} influences both u_{it} and u_{it-1}, x_{it-1} and u_{it} have a simultaneous relationship. In other words, as x_{it-1} changes, u_{it} changes contemporaneously (albeit not causally), and so y_{it} changes as well. In this case although x_{it-1} influences y_{it} only through x_{it}, which satisfies the exclusion restriction, the IV x_{it-1} violates the independence assumption because it does not serve as a random exogenous shock. In other words, the independence assumption can only be satisfied by assuming that there are no dynamics among unobserved confounders (Bellemare et al. 2017). Therefore, even if the dynamic causal impacts are restricted and thus exclusion restriction is satisfied, a lagged IV can still be problematic because of the unavoidable violation of independence assumption that it entails.

To understand how a violation of either the independence assumption or of the exclusion restriction influences bias, we rely on conditional expectations. Without any loss of generality, suppose that x_{it} is a dichotomous variable. For OLS, the average treatment effect (ATE) is such that

$$
\begin{aligned}
E(y_{it}|x_{it} &= 1) - E(y_{it}|x_{it} = 0) = \\
&E(y_{1it} - y_{0it}|x_{it}) + \frac{E(y_{0it}|x_{it} = 1) - E(y_{0it}|x_{it} = 0)}{E(x_{it}|x_{it} = 1) - E(x_{it}|x_{it} = 0)}, \quad (6)
\end{aligned}
$$

where the estimation bias stems from selection bias, or the second term on the RHS of equation (6), whose denominator we know to be equal to 1.

Under lagged IV, assuming that the independence assumption and exclusion restriction both hold, and further assuming monotonicity of the effect of x_{it-1} on x_{it} as well as the existence of the first-stage regression, per the LATE theorem we know that the local average treatment effect (LATE) is such that

$$\frac{E(y_{it}|x_{it-1}=1)-E(y_{it}|x_{it-1}=0)}{E(x_{it}|x_{it-1}=1)-E(x_{it}|x_{it-1}=0)} = E(y_{1it}-y_{0it}|x_{1it}>x_{0it}). \tag{7}$$

But if neither the independence assumption nor the exclusion restriction are satisfied, the LATE becomes

$$\frac{E[y_{0it}+(y_{1it}-y_{0it})x_{1it}|x_{it-1}=1]-E[y_{0it}+(y_{1it}-y_{0it})x_{0it}|x_{it-1}=0]}{E(x_{it}|x_{it}=1)-E(x_{it}|x_{it}=0)}+ \\ \frac{E(y_{0it}|x_{it-1}=1)-E(y_{0it}|x_{it-1}=0)}{E(x_{it}|x_{it-1}=1)-E(x_{it}|x_{it-1}=0)}. \tag{8}$$

In this case, because the exclusion restriction is not satisfied, the second term appears, which could be even larger than $\frac{E(y_{0it}|x_{it}=1)-E(y_{0it}|x_{it}=0)}{E(x_{it}|x_{it}=1)-E(x_{it}|x_{it}=0)}$. This implies that the bias from lagged IV estimation may be even larger than the bias from naïve OLS estimation.

Even if the exclusion restriction is satisfied and the second item in equation (8) collapses to zero, the violation of the independence assumption implies that we simply cannot have

$$E[y_{0it} + (y_{1it} - y_{0it})x_{1it}|x_{it-1} = 1] = E[y_{0it} + (y_{1it} - y_{0it})x_{1it}], \tag{9}$$

nor can we have

$$E[y_{0it} + (y_{1it} - y_{0it})x_{0it}|x_{it-1} = 0] = E[y_{0it} + (y_{1it} - y_{0it})x_{0it}]. \tag{10}$$

Thus we cannot derive the LATE in equation (8) similarly as in equation (7), such that

$$\begin{aligned} & \frac{E[y_{0it} + (y_{1it} - y_{0it})x_{1it}|x_{it-1} = 1] - E[y_{0it} + (y_{1it} - y_{0it})x_{0it}|x_{it-1} = 0]}{E(x_{it}|x_{it-1} = 1) - E(x_{it}|x_{it-1} = 0)} \\ = \; & E(y_{1it} - y_{0it}|x_{1it} > x_{0it}) \end{aligned} \tag{11}$$

which makes ambiguous whether the estimation bias in OLS is smaller than that in lagged IV estimation ambiguous. This implies that even in scenario 1, lagged IV estimation remains problematic. What's more, the assumption that there are no dynamics among unobserved confounders, which has to hold in order for the independence assumption to be satisfied, is unlikely to hold in most—if not all—cases.

2.2 Extent of the Problem

To see how common the practice of lagged IV is, we examine the articles published in the top journals in economics from 2013 to 2018. We identify the articles using lagged IV, either as the core identification strategy or as a robustness check on the core identification strategy.

Table 1 shows the number of papers using lagged IV published in 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* from 2013 to 2018.

In total, we identify 21 papers using a lagged IV. All these papers use one or more lagged explanatory variables as IVs to mitigate endogeneity concerns. Doraszelski et al. (2018), for instance, instrument a firms' bid for a balancing mechanism in month t with the same firm's bid in month $t-1$. Most of those lagged IV papers mention that the availability of lagged explanatory variables is one of the key reasons why those same variables are used as IVs. Those papers, however, do not address the difference between the bias of the lagged IV method and that of OLS in details. In Bøler et al. (2015), for instance, one cannot rule out that the lagged explanatory variable could have a direct causal effect on the dependent variable, as in scenarios 2 and 4 above, which casts doubt on those results of theirs that rely on lagged IV as an identification strategy.

Our review of the literature shows that lagged IVs remain common in economics. Moreover, it shows that several authors believe that lagging endogenous variable is a somewhat valid identification strategy, because it is at least exogenous to some extent and satisfies the relevance restriction. In the analysis that follows, we explore contrast lagged IV with OLS, and we specifically look at what lagged IV does to bias.

3 Analytical Derivations

The conceptual framework above demonstrates why lagged IV estimation is unlikely to make endogeneity problems go away, and why it is likely to aggravate them. To characterize the LATE under lagged IV quantitatively, we compare the bias under OLS and under lagged IV. For simplicity, we set up a bivariate regression scenario, and discuss the AR(1) process in the data generation process both for the endogenous explanatory variable and the unobserved confounder.

Since the causal relationships of interest in scenarios 2 to 4 suffer from endogeneity, we quantitatively discuss the estimation bias only for scenario 1, viz. the simplest case. Following Bellemare et. al. (2017), we consider the following setup:

$$y_{it} = \beta x_{it} + \delta u_{it} + \epsilon_{it}, \tag{12}$$

$$x_{it} = \rho x_{it-1} + \kappa u_{it} + \eta_{it}, \tag{13}$$

and

$$u_{it} = \phi u_{it-1} + \omega_{it}, \tag{14}$$

where y is the dependent variable, x is the variable of interest, and u represents all unobserved confounders. Since the unobserved confounder is omitted from equation (12), estimating it by OLS will return a biased estimate $\widehat{\beta}_{OLS}$ of β. The AR(1) process in equation (13) implies that x_{it} determined both by its lagged value x_{it-1} and by the contemporaneous value u_{it} unobserved confounder. Similarly, the AR(1) process in equation (14) implies that u_{it} is only determined by its lagged value. In what follows, we assume that $\rho, \phi \in (0,1)$, that $\epsilon_{it} \sim N(0, \sigma_\epsilon^2)$, $\eta_{it} \sim N(0, \sigma_\eta^2)$, and $\omega_{it} \sim N(0, \sigma_\omega^2)$.

3.1 Bias of OLS Estimation

Without unobserved confounder, OLS yields unbiased estimates. Given the fact that the unobserved confounders are almost always present in observational data, the magnitude of the bias from OLS estimation is a function of $Cov(x_{it}, u_{it})$, $Var(x_{it})$, as well as δ, i.e., the magnitude of the causal effect of the unobserved confounder on y_{it}:

$$\begin{aligned} E(\widehat{\beta}_{OLS}|x_{it}) &= \frac{Cov(x_{it}, y_{it})}{Var(x_{it})} \\ &= \frac{Cov(x_{it}, \beta x_{it} + \delta u_{it} + \epsilon_{it})}{Var(x_{it})} \\ &= \beta + \delta \frac{Cov(x_{it}, u_{it})}{Var(x_{it})}. \end{aligned} \tag{15}$$

If either $\delta = 0$ or $Cov(x_{it}, y_{it}) = 0$, i.e., if u_{it} has no impact on y_{it}, or u_{it} is uncorrelated with x_{it}, then endogeneity is not an issue, and we get the standard result that $E(\widehat{\beta}_{OLS}|x_{it}) = \beta$.

Using the derivations in Appendix A, we have that

$$Cov(x_{it}, u_{it}) = \kappa \frac{Var(u_{it})}{1 - \phi\rho}. \tag{16}$$

Plugging this last equation into equation (15), we get an expression for the bias of OLS that is such that

$$E(\widehat{\beta}_{OLS}|x_{it}) = \beta + \frac{\delta\kappa Var(u_{it})}{(1 - \phi\rho)Var(x_{it})}, \tag{17}$$

which shows the bias arising from OLS.

3.2 Bias of Lagged IV Estimation

Now consider what happens when using x_{t-1} as an IV for x_t. This implies that

$$E(\widehat{\beta}_{IV}|x_{it}) = \frac{Cov(x_{it-1}, y_{it})}{Cov(x_{it-1}x_{it})}. \tag{18}$$

Plugging equation (12) into equation (18), we get

$$E(\widehat{\beta}_{IV}|x_{it}) = \frac{Cov(x_{it-1}, \beta x_{it} + \delta u_{it} + \epsilon_{it})}{Cov(x_{it-1}, x_{it})}, \tag{19}$$

and so

$$
\begin{aligned}
E(\widehat{\beta}_{IV}|x_{it}) &= \beta + \delta\frac{Cov(x_{it-1}, u_{it})}{Var(x_{it-1})} \\
&= \beta + \delta\frac{\frac{Cov(x_{it-1}, u_{it})}{Cov(x_{it-1}, x_{it})}}{\rho + \kappa\frac{Cov(x_{it-1}, u_{it})}{Cov(x_{it-1}, x_{it})}}. \qquad (20)
\end{aligned}
$$

Using the derivations in the Appendix, we also have

$$
\begin{aligned}
\frac{Cov(x_{it-1}, u_{it})}{Var(x_{it-1})} &= \frac{Cov(x_{it-1}, \phi u_{it-1} + \omega_{it})}{Var(x_{it-1})} \\
&= \phi\frac{Cov(x_{it-1}, u_{it-1})}{Var(x_{it-1})} \\
&= \frac{\phi\kappa Var(u_{it})}{(1-\phi\rho)Var(x_{it})}. \qquad (21)
\end{aligned}
$$

Therefore, we have

$$
\begin{aligned}
E(\widehat{\beta}_{IV}|x_{it}) &= \beta + \frac{\delta\phi\kappa Var(u_{it})}{\rho(1-\phi\rho)Var(x_{it}) + \phi\kappa^2 Var(u_{it})} \qquad (22) \\
&= \beta + \frac{\delta\kappa Var(u_{it})}{\frac{\rho}{\phi}(1-\phi\rho)Var(x_{it}) + \kappa^2 Var(u_{it})}. \qquad (23)
\end{aligned}
$$

Equation (23) shows the bias of the lagged IV method.

3.3 Comparing the Bias of OLS with that of Lagged IV

We can now compare the bias of OLS with that of lagged IV, i.e., equations (17) and (23), or

$$E(\widehat{\beta}_{OLS}|x_{it}) = \beta + \frac{\delta\kappa Var(u_{it})}{(1-\phi\rho)Var(x_{it})} \tag{24}$$

with

$$E(\widehat{\beta}_{IV}|x_{it}) = \beta + \frac{\delta\kappa Var(u_{it})}{\frac{\rho}{\phi}(1-\phi\rho)Var(x_{it}) + \kappa^2 Var(u_{it})}. \tag{25}$$

Setting up an inequality between the two yields

$$\beta + \frac{\delta\kappa Var(u_{it})}{(1-\phi\rho)Var(x_{it})} \gtreqless \beta + \frac{\delta\kappa Var(u_{it})}{\frac{\rho}{\phi}(1-\phi\rho)Var(x_{it}) + \kappa^2 Var(u_{it})} \tag{26}$$

or

$$\frac{\delta\kappa Var(u_{it})}{(1-\phi\rho)Var(x_{it})} \gtreqless \frac{\delta\kappa Var(u_{it})}{\frac{\rho}{\phi}(1-\phi\rho)Var(x_{it}) + \kappa^2 Var(u_{it})}, \tag{27}$$

This is equivalent to

$$\frac{\rho}{\phi}(1-\phi\rho)Var(x_{it}) + \kappa^2 Var(u_{it}) \gtreqless (1-\phi\rho)Var(x_{it}), \tag{28}$$

so that we need to consider how the parameter ratio $\frac{\rho}{\phi}$ and the parameter κ affect this comparison.

By comparing $\frac{\rho}{\phi}(1-\phi\rho)$ with $(1-\phi\rho)$, we know on the one hand that because $\rho, \phi \in (0,1)$, if $\frac{\rho}{\phi} < 1$, $\frac{\rho}{\phi}(1-\phi\rho) < (1-\phi\rho)$. On the other hand, if $\frac{\rho}{\phi} > 1$, then $\frac{\rho}{\phi}(1-\phi\rho) > (1-\phi\rho)$.

If $\frac{\rho}{\phi} < 1$, then, $\frac{\rho}{\phi}(1-\phi\rho)Var(x_{it}) < (1-\phi\rho)Var(x_{it})$. Thus, if κ is small enough, $\kappa^2 Var(u_{it})$, the bias of lagged IV is always larger than that of OLS. With κ large enough, however, the bias of lagged IV is actually smaller than the bias of OLS.

If $\frac{\rho}{\phi} > 1$, then, $\frac{\rho}{\phi}(1-\phi\rho)Var(x_{it}) > (1-\phi\rho)Var(x_{it})$. Thus, we always have that $\frac{\rho}{\phi}(1-\phi\rho)Var(x_{it}) + \kappa^2 Var(u_{it}) > (1-\phi\rho)Var(x_{it})$, in which case the bias of lagged IV is always smaller than that of OLS.

To sum up, κ and $\frac{\rho}{\phi}$ are the key quantities, because they determine whether the bias of lagged IV exceeds the bias of OLS. According to our derivations, it is only in cases where κ is small and $\frac{\rho}{\phi} < 1$ that the bias of lagged IV is larger than the bias of OLS. In other cases, the bias in lagged IV is smaller than that of OLS. Therefore, even in scenario 1—the case with the least amount of endogeneity—using a lagged IV does not guarantee unbiasedness. In section 4, we show simulation results that test these theoretical findings for a range of credible values of κ and $\frac{\rho}{\phi}$.

3.4 Root Mean Square Error of OLS and Lagged IV

We continue with scenario 1 and derive the RMSE for both OLS and lagged IV. The expression for the RMSE under OLS estimation is such that

$$RMSE(\widehat{\beta}_{OLS}) = \sqrt{\frac{\sigma^2}{SST_x} + \left[\frac{\delta\kappa Var(u_{it})}{(1-\phi\rho)Var(x_{it})}\right]^2} \tag{29}$$

where SST_x is the total sum of squares of x_{it}.

Similarly, the expression for RMSE under lagged IV estimation is such that

$$\begin{aligned} RMSE(\widehat{\beta}_{IV}) &= \sqrt{MSE(\widehat{\beta}_{IV})} \\ &= \sqrt{Var(\widehat{\beta}_{IV}) + Bias(\widehat{\beta}_{IV})^2}, \end{aligned} \tag{30}$$

which is equivalent to

$$RMSE(\widehat{\beta}_{IV}) = \sqrt{Var(\widehat{\beta}_{IV}) + \left[\frac{\delta\kappa Var(u_{it})}{\frac{\rho}{\phi}(1-\phi\rho)Var(x_{it}) + \kappa^2 Var(u_{it})}\right]^2}, \tag{31}$$

or more explicitly to

$$RMSE(\widehat{\beta}_{IV}) = \sqrt{\begin{array}{c} Var\left(\beta + \frac{\delta\kappa Var(u_{it})}{\frac{\rho}{\phi}(1-\phi\rho)Var(x_{it})+\kappa^2 Var(u_{it})}\right) \\ + \left[\frac{\delta\kappa Var(u_{it})}{\frac{\rho}{\phi}(1-\phi\rho)Var(x_{it})+\kappa^2 Var(u_{it})}\right]^2 \end{array}}. \tag{32}$$

We know that the first term under the square root is zero, since the expression in parentheses is constant. Therefore,

$$RMSE(\widehat{\beta}_{IV}) = \frac{\delta\kappa Var(u_{it})}{\frac{\rho}{\phi}(1-\phi\rho)Var(x_{it}) + \kappa^2 Var(u_{it})}. \tag{33}$$

Note that $RMSE(\widehat{\beta}_{OLS}) > \frac{\delta\kappa Var(u_{it})}{(1-\phi\rho)Var(x_{it})}$. From the last section, we know that if the bias from lagged IV estimates is smaller than that of OLS, it has to be the case that

$$\frac{\delta\kappa Var(u_{it})}{(1-\phi\rho)Var(x_{it})} > \frac{\delta\kappa Var(u_{it})}{\frac{\rho}{\phi}(1-\phi\rho)Var(x_{it}) + \kappa^2 Var(u_{it})}, \tag{34}$$

which leads to the conclusion that $RMSE(\widehat{\beta}_{OLS}) > RMSE(\widehat{\beta}_{IV})$. Conversely, if the bias in lagged IV is larger than that of OLS, then the inequality flips and $RMSE(\widehat{\beta}_{OLS}) < RMSE(\widehat{\beta}_{IV})$. Whether RMSE is larger under one estimator versus the other depends directly on whether the bias of the former exceeds that of the latter.

4 Monte Carlo Simulations

We have shown analytically that lagged a variable of interest as an IV for itself can either alleviate or aggravate the issues deriving from the presence of endogeneity. In this section, we build on the Monte Carlo setup developed by Bellemare et al. (2017) to simulate the four scenarios outlined in our conceptual framework so as to quantitatively discuss bias of both the lagged IV and OLS estimates, along with the RMSE and the likelihood of a Type I error under both methods.

4.1 Setup

We start with scenario 1. Figure 1 illustrates our theoretical framework and parameterizes the relations between the dependent variable, the variable of interest, and the unobserved confounder. As shown, the unobserved confounder is correlated both with y and with x. The effect δ of u on y, is normalized to 1. The effect β of x on y is assigned a value ranging from 0 to 2.

Apart from those, the effect κ of u on x, is assigned a value of 0.5 and 2 to see its impact on the two biases. In addition, autocorrelation parameters ρ and ϕ are set respectively at 0.5 and $\{0, 0.1, 0.2, ..., 0.9\}$, so that when ρ ranges from 0 to 1 and $\phi = 0.5$, we can study what happens both when $\frac{\rho}{\phi} < 1$ and when $\frac{\rho}{\phi} > 1$, and likewise when $\rho = 0.5$ and ϕ ranges from 0 to 1.

Following Bellemare et al. (2017), for each simulation, we generate a panel in which $N = 100$ units are observed $T = 50$ periods each, for a total of $NT = 5,000$ observations. Our simulation follows the same data generating process (DGPs) as in section 3. Each set of parameter values, shown in Table 2, are simulated 100 times. Then three estimators of β are illustrated: (i) the naïve estimator ($\widehat{\beta}_{NAIVE}$), or a regression of y_{it} on x_{it} that ignores unobserved confounders, (ii) the lagged IV estimator ($\widehat{\beta}_{LAGIV}$), or a regression of y_{it} on x_{it} which relies on x_{it-1} as an IV for x_{it}, and (iii) the "correct" estimator ($\widehat{\beta}_{CORRECT}$), which regresses y_{it} on x_{it} and u_{it}.

Here, the correct estimator is the counterfactual, and since applied researchers cannot observe confounders, our DGPs test the performance of both OLS and lagged IV by comparing each of their bias with the correct estimator, whose bias is zero. To make our analysis simple and straightforward, we only consider one-period autocorrelation.

As in Bellemare et al. (2017), three criteria are used to evaluate the performance of the lagged IV method: (i) bias, (ii) RMSE, and (iii) the likelihood of Type I error, i.e., the odds of rejecting the null hypothesis that $\beta = 0$ when it is, in fact, true.

4.2 Scenario 1

For each combination of parameter values, $\widehat{\beta}_{NAIVE}$, $\widehat{\beta}_{LAGIV}$, and $\widehat{\beta}_{CORRECT}$ are saved from each of the 100 simulations, then the average bias as well as the corresponding RMSE and likelihood of Type I error are plotted.

Figure 2 summarizes the simulation results when $\kappa = 0.5$ and 2, $\rho = 0.5$, and ϕ ranges from 0 to 1. The simulation results show that both $\widehat{\beta}_{NAIVE}$ and $\widehat{\beta}_{LAGIV}$ are biased because of the unobserved confounder. As ϕ increases, the bias of both OLS and lagged IV estimates also increases, which is consistent with the theoretical predictions in sections 3.1 and 3.2. More importantly, as predicted in section 3.3, the bias of lagged IV is smaller than the bias of OLS when $\kappa = 0.5$, and $\phi < 0.5$. That is, when $\frac{\rho}{\phi} > 1$, no matter the value of κ, the bias of lagged IV is always smaller than that of OLS.

When $\kappa = 0.5$ and $\phi > 0.5$, that is, when $\frac{\rho}{\phi} < 1$, the bias in lagged IV is still smaller than that of OLS. As $\frac{\rho}{\phi}$ gets smaller, however, the bias of lagged IV becomes larger than the bias of OLS. When κ gets larger and goes to 2, the bias of lagged IV is always smaller than the bias of OLS.

In line with our derivations of RMSE, the respective RMSEs for OLS and lagged IV show patterns similar to their respective biases. These results provide supportive evidence for our mathematical argument that using a lagged variable of interest as an IV could mitigate the endogeneity problem. For that to be true, however, it has to be the case that the autocorrelation in the unobserved confounders is small and that the impact of unobserved confounders on y is small—two untestable conditions.

Similar to our results for bias and RMSE, when $\kappa > 0$, and as ϕ ranges from 0 to 1, the likelihood of making a Type I error rises dramatically. The reason is that lagged IV identification will lead to nonzero estimates of β even if $\beta = 0$ because the effect δ of the unobserved confounder on the dependent variable and the effect κ of the unobserved confounders on the variable of interest are both nonzero. In addition, the likelihood of rejecting a true null rises dramatically and gets close to 1 as ϕ increases. Accordingly, these results and interpretations suggest that using lagged explanatory variable as instruments for themselves can hardly help mitigate the likelihood of Type I error.

Figure 3 shows simulation results when $\phi = 0.5$, ρ ranges from 0 to 1, and $\kappa = 0.5$ or 2. These results show that both $\widehat{\beta}_{NAIVE}$ and $\widehat{\beta}_{LAGIV}$ are still biased. As ρ ranges from 0.5 to 1, bias decreases both for $\widehat{\beta}_{NAIVE}$ and

$\widehat{\beta}_{LAGIV}$. More importantly, when $\kappa = 0.5$ and $\rho > 0.5$ (i.e., when $\frac{\rho}{\phi} > 1$), no matter the value of κ, the bias of lagged IV is always smaller than that of OLS. When $\kappa = 0.5$ and $\rho < 0.5$ (i.e., when $\frac{\rho}{\phi} < 1$), the bias in lagged IV is still smaller than that of OLS. As $\frac{\rho}{\phi}$ gets smaller, however, the bias of lagged IV exceeds that of OLS. Here, too, RMSEs exhibit a pattern that is in line with that of the biases. As regards the likelihood of Type I error, Figure 3 shows that when $\phi = 0.5$, and ρ ranges from 0 to 1, the likelihood of rejecting a true null is almost always equal to 1. This echoes the findings in Figure 2.

In sum, our simulation results convey an unambiguous message: If the lagged variable of interest has no direct effect on the dependent variable or on the unobserved confounder, a lagged IV can reduce bias and the RMSE. This is only true, however, for specific parameter values such that the ratio $\frac{\rho}{\phi}$ of the explanatory variable's autocorrelation parameter to the unobserved confounder's autocorrelation parameter is small, and the impact of the unobserved confounder on the explanatory variable is also small. Still, our results leave little hope for the likelihood of Type I error, which is always high under either OLS or lagged IV.

These results imply that even if the exclusion restriction holds, the lagged IV method remains problematic. Since lagged IV estimation violates the independence assumption because of the lagged explanatory variable's simultaneous relationship with the unobserved confounder and the resulting indirect relationship with the dependent variable, it only mitigates endogeneity to a limited extent, and it may even aggravate the problem for some values of the parameters.[3]

We also discuss the cases in which (i) the lagged explanatory variable has a direct effect on the dependent variable, (ii) the lagged explanatory variable has a direct effect on the unobserved confounder, and (iii) the lagged explanatory variable has direct effects on both the dependent variable and on the unobserved confounder. These cases coincide with scenarios 2 to 4 in our conceptual framework. These three cases yield different results regarding estimation bias and RMSE, in that both bias and RMSE from lagged IV estimation are significantly larger than those from OLS. Moreover, in these three cases, the likelihood of a Type I error is either close to or equal to 1, and significantly under lagged IV than under OLS. These results imply that when lagged IV estimation violates both the exclusion restriction and the independence assumption, it makes the endogeneity problem worse than under naïve OLS estimation.

4.3 Scenario 2

As is discussed in scenario 2 in conceptual framework, when lagged endogenous explanatory variable has direct causal impact on the explained variable, our model setup becomes

$$y_{it} = \beta x_{it} + \theta x_{it-1} + \delta u_{it} + \epsilon_{it}, \tag{35}$$

[3]In the appendix, we also discuss the case in which fixed effects are included in estimation. That analysis yields results that are similar to our main results.

$$x_{it} = \rho x_{it-1} + \kappa u_{it} + \eta_{it}, \tag{36}$$

and

$$u_{it} = \phi u_{it-1} + \upsilon_{it}, \tag{37}$$

where x_{it-1} has a direct effect on y_{it} with marginal effect θ, which we normal to 1 in our simulations. Figure 4 shows the causal relationships in our simulations.

Figure 5 and 6 show simulation results for this case. In contrast with results in Figures 2 and 3, using x_{it-1} as an IV for x_{it} introduces more bias and a larger RMSE than OLS, and the likelihood of a Type I error gets very close to one for a wide range of parameter values. These results imply that since lagged IVs violate not only the independence assumption but also the exclusion restriction, they cannot mitigate the endogeneity problem, and may even aggravate it.

4.4 Scenario 3

We now add the lagged endogenous explanatory variable as another determinant of the dependent variable, as in scenario 3. In this case, our model setup becomes

$$y_{it} = \beta x_{it} + \delta u_{it} + \epsilon_{it}, \tag{38}$$

$$x_{it} = \rho x_{it-1} + \kappa u_{it} + \eta_{it}, \tag{39}$$

and

$$u_{it} = \phi u_{it-1} + \psi x_{it-1} + \upsilon_{it}, \tag{40}$$

where x_{it-1} also has a causal effect on u_{it} with marginal effect ψ, which is also normalized to 1 in our simulations. Figure 7 shows the causal relationships in our simulations. Figures 8 and 9 show that using x_{it-1} as an IV for x_{it} increases bias and RMSE relative to OLS, and once again, the likelihood of a Type I error gets very close to 1. These results imply that since using the lagged explanatory variable as an IV violates both the independence assumption and the exclusion restriction, the lagged IV method cannot mitigate endogeneity, and it may even aggravate the problem. What's worse, as was explained in the conceptual framework, since there could exist more than one potential causal path from x_{it-1} to u_{it}, it is difficult to argue that the exclusion restriction is satisfied by assuming x_{it-1} influences y_{it} only through x_{it}.

4.5 Scenario 4

We now add the lagged endogenous explanatory variable's causal impact both on the dependent variable and on the unobserved confounder, as is discussed in scenario 4 in conceptual framework. Specifically, our model setup becomes

$$y_{it} = \beta x_{it} + \theta x_{it-1} + \delta u_{it} + \epsilon_{it}, \tag{41}$$

$$x_{it} = \rho x_{it-1} + \kappa u_{it} + \eta_{it}, \tag{42}$$

and

$$u_{it} = \phi u_{it-1} + \psi x_{it-1} + \upsilon_{it}, \tag{43}$$

where x_{it-1} has a direct causal effect on y_{it} with the marginal effect θ, which is normalized to 1, and also has a causal effect on u_{it} with the marginal effect of ψ, which is also normalized as 1 in our simulation. Figure 10 shows the causal relationships in our simulations.

Figures 11 and 12 show that using x_{it-1} as an IV for x_{it} increases the bias and RMSE relative to OLS; exceptionally in this case, the likelihood of a Type I error is always equal to 1 given a wide range of parameter values. These results also imply that a lagged IV violates not only the independence assumption, but also the exclusion restriction. So once again, the endogeneity problem is not mitigated, and it can even be made worse.

5 Summary and Concluding Remarks

In this paper, we have looked at the consequences of using the lagged value of an endogenous variable of interet as an instrument for that same variable—what we have dubbed "lagged IV"—for bias, the RMSE, and the likelihood of a Type I error. We have done so both analytically as well as with the aid of Monte Carlo simulations.

We find on the one hand that if the lagged IV has no direct causal impact on (i) the dependent variable nor (ii) the unobserved confounder, it violates the independence assumption, but not the exclusion restriction. In this case, a lagged IV can mitigate the endogeneity problems by reducing bias and the root mean squared error (RMSE) for common ranges of parameter values. But even in such a case, the likelihood of a Type I error remains large. On the other hand, we find that if the lagged IV has a direct causal impact on (i) the dependent variable, (ii) the unobserved confounder, or (iii) both, it violates the exclusion restriction as well as the independence assumption. In such cases, a lagged IV worsens the endogeneity problem by increasing bias as well as the RMSE. Additionally the likelihood of a Type I error in such cases is almost always equal to one for common ranges of parameter values. In practical terms, this means that the use of a lagged IV often leads one to report coefficients estimates of questionable economic and statistical significance. Worse, the use of lagged IVs will tend to lead one to conclude that a statistically significant relationship exists where it in fact does not.

The implications of our findings for the practice of applied econometrics are obvious. Unless one can make the claim that both the independence assumption and the exclusion restriction hold, lagged IVs should be avoided in the name of bias, RMSE, and the likelihood of a Type I error. But given that the independence assumption requires that one make the dubious assumption that there are no dynamics among unobserved confounders, this essentially means that lagged IVs should be avoided entirely.

Our review of the recent literature in economics shows that the practice of using lagged IVs remains common. The implications of our findings for editors and reviewers of journals in the social sciences are thus that they should be especially skeptical of any finding that involves a lagged IV. In most cases, estimation results relying on a lagged IV identification strategy are no better than a naïve OLS, and so the latter should be favored over the former. At the very least, lagged IV results should be subjected to a battery of robustness checks. Better yet, they should only be presented when a viable alternative identification strategy is shown or comparison.

Causal inference usually requires experimental data to identify the treatment effect of variables of interest. With observational data, natural experiments are usually indispensable to provide with an exogenous shock in causal identification (Angrist and Krueger, 2001, Freedman 2005), although they lack underlying theoretical relationships (Rosenzweig and Wolpin, 2000). Therefore, valid instrumental variables are likely from natural experiments because in this sense, they are very likely to be exogenous and satisfy both the independence assumption and the exclusion restriction. Lagged explanatory variables, on the contrary, have simultaneous relationship with the unobserved confounder that influences the dependent variable, lacking the exogeneity as natural experiments do, thus can hardly provide with additional information in causal inference.

References

[1] Angrist, Joshua D., and Alan B. Krueger (2001), "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments," *Journal of Economic Perspectives* 15(4): 69-85.

[2] Angrist, Joshua, and Jörn-Steffen Pischke (2009), *Mostly Harmless Econometrics*, Princeton: Princeton University Press.

[3] Angrist, Joshua, and Jörn-Steffen Pischke (2010), "The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics," *Journal of Economic Perspectives* 24(2): 3-30.

[4] Angrist, Joshua, and Jörn-Steffen Pischke (2014), *Mastering 'Metrics*, Princeton: Princeton University Press.

[5] Bellemare, Marc F., Takaaki Masaki, and Thomas B. Pepinsky (2017), "Lagged Explanatory Variables and the Estimation of Causal Effects," *Journal of Politics* 79(3): 949-963.

[6] Blundell, Richard, and Stephen Bond (1998), "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics* 87(1): 115-143.

[7] Blundell, Richard, and Stephen Bond (2000), "GMM Estimation with Persistent Panel Data: An Application to Production Functions," *Econometric Reviews* 19(3): 321-340.

[8] Bøler, Esther Ann, Andreas Moxnes, and Karen Helene Ulltveit-Moe (2015), "R&D, International Sourcing, and the Joint Impact on Firm Performance," *American Economic Review* 105(12): 3704-3739.

[9] Doraszelski, Ulrich, Gregory Lewis, and Ariel Pakes (2018), "Just Starting out: Learning and Equilibrium in a New Market," *American Economic Review* 108(3): 565-615.

[10] Freedman, David A. (2005), *Linear Statistical Models for Causation.* Wiley StatsRef: Statistics Reference Online

[11] Imbens, Guido W., and Joshua D. Angrist (1994), "Identification and Estimation of Local Average Treatment Effects," *Econometrica* 62(2): 467-475.

[12] Reed, William Robert (2015), "On the Practice of Lagging Variables to Avoid Simultaneity," *Oxford Bulletin of Economics and Statistics* 77(6): 897-905.

[13] Rosenzweig, Mark R., and Kenneth I. Wolpin (2000), "'Natural' Natural Experiments" in Economics," *Journal of Economic Literature* 38(4): 827-874.

[14] Rossi, Peter E. (2014), "Even the Rich Can Make Themselves Poor: A Critical Examination of IV Methods in Marketing Applications," *Marketing Science* 33(5)): 655-672.

Appendix

Derivation of $Cov(x, u)$

Following Bellemare et al. (2017), given equations (2) and (3), we have

$$Cov(x_{it-1}, u_{it-1}) = Cov\left(\frac{1}{\rho}x_{it} - \frac{\kappa}{\rho}u_{it} - \frac{1}{\rho}\eta_t, \frac{1}{\phi}u_t - \frac{1}{\phi}\nu_t\right). \tag{44}$$

We then have

$$Cov(x_{it-1}, u_{it-1}) = \frac{1}{\phi\rho}[Cov(x_{it}, u_{it}) - \kappa Var(u_{it})], \tag{45}$$

which yields

$$Cov(x_{it}, u_{it}) - \kappa Var(u_{it}) = \phi\rho Cov(x_{it-1}, u_{it-1}). \tag{46}$$

Since $\rho, \phi \in (0, 1)$, both x and u revert to their means, i.e., $Cov(x_{it}, u_{it})$ does not depend on t, and so we have

$$Cov(x_{it}, u_{it}) = Cov(x_{it-1}, u_{it-1}). \tag{47}$$

From equation (46), we thus know that

$$Cov(x_{it}, u_{it}) = Cov(x_{it-1}, u_{it-1}) = \frac{\kappa Var(u_{it})}{1 - \phi\rho}. \tag{48}$$

Extensions to the Monte Carlo Simulations

In this section, we discuss controlling fixed effects to address unobserved heterogeneity in estimation using scenario 1 as example. So far, our DGPs incorporate no unit-level unobserved heterogeneity, in other words, we are using the pooled estimator for β. Fixed effects, however, are commonly used to account for unobserved heterogeneity.

Introducing unobserved unit-level heterogeneity to account for unit fixed effects, our simulation results remain identical to previous ones. As Figures 13 and 14 show, using a lagged explanatory variable as an IV yields smaller bias and RMSE in fixed effect estimates than in OLS estimates at first; but as ϕ goes up, or as ρ goes down, bias and RMSE in fixed effect estimates gets larger and become gradually slightly larger than those of the naïve OLS estimates. Similar to the simulation results in section 4, the likelihood a Type I error rises dramatically and becomes close to 1 as ϕ goes up from 0 to 1; when ρ ranges from 0 to 1, the likelihood of rejecting the true null hypothesis is mostly equal to 1.

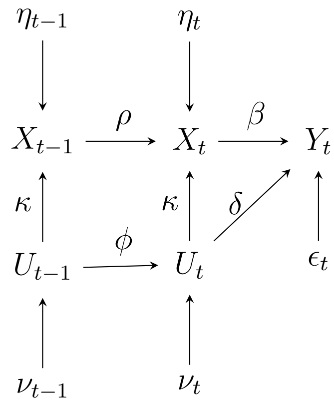
Table 1: Number of Articles Using Lagged IVs, 2013-2018

Journal	Number of Articles
American Economic Review	5
Econometrica	1
Journal of Political Economy	1
Quarterly Journal of Economics	4
Review of Economic Studies	3
Review of Economics and Statistics	7

Table 2: Simulation Parameters

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

Figure 1: Monte Carlo Simulations



Notes: Visual representation of the Monte Carlo simulations setup. Greek letters denote parameters; X denotes the variable of interest; U denotes unobserved confounders; Y is the dependent variable.

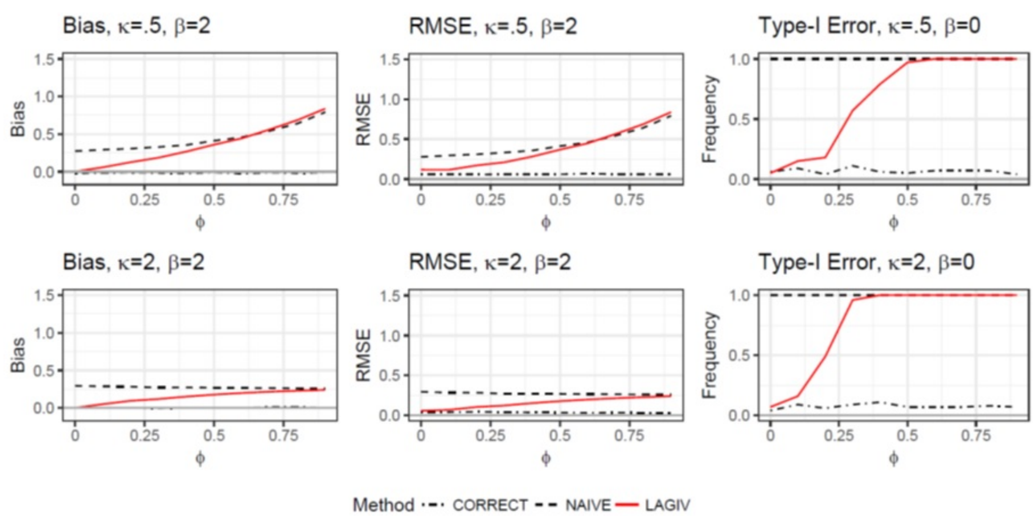


Figure 2: Monte Carlo Results with $\kappa = 0.5$ and 2 and ϕ ranging from 0 to 1.

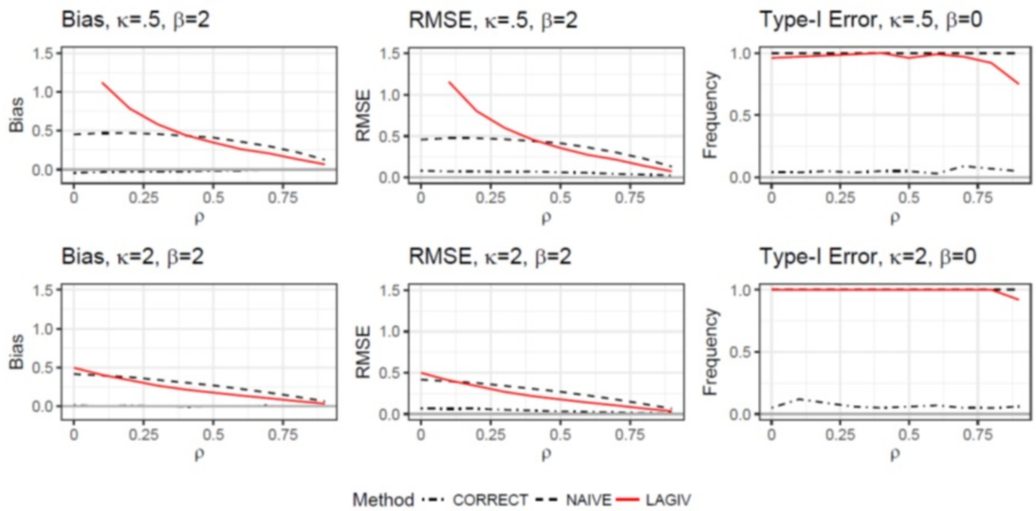
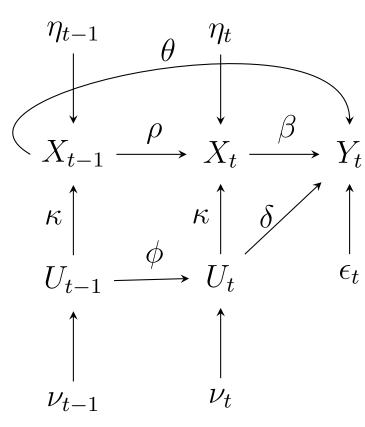


Figure 3: Monte Carlo Results with $\kappa = 0.5$ and 2 and ρ ranging from 0 to 1.

Figure 4: Monte Carlo Simulations



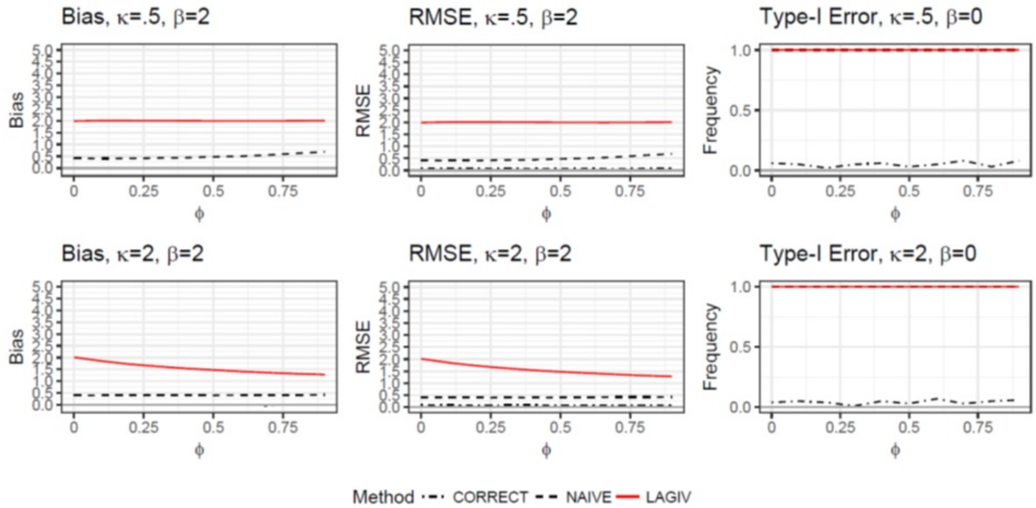


Figure 5: Monte Carlo Results with $\kappa = 0.5$ and 2 and ϕ ranging from 0 to 1. Lagged Causality in Dependent Variable.

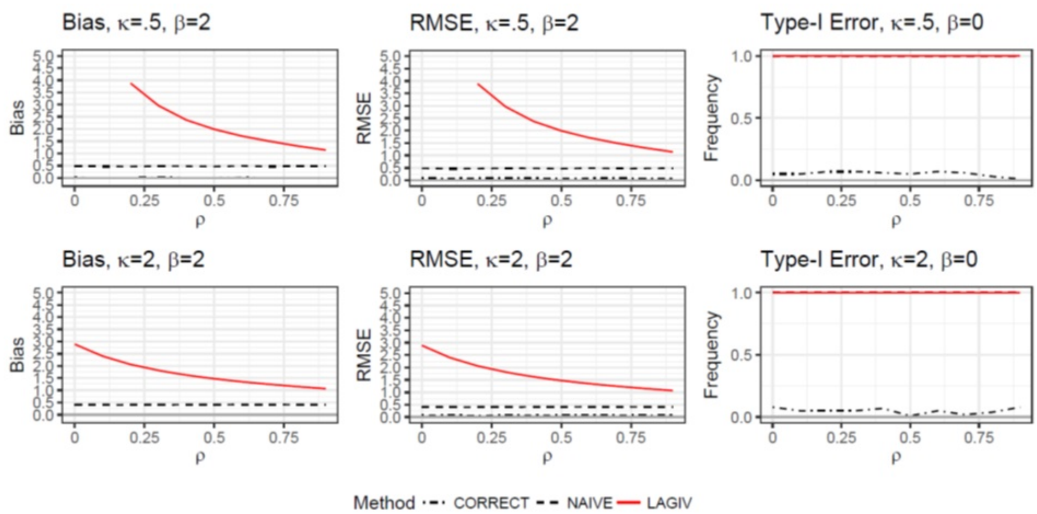
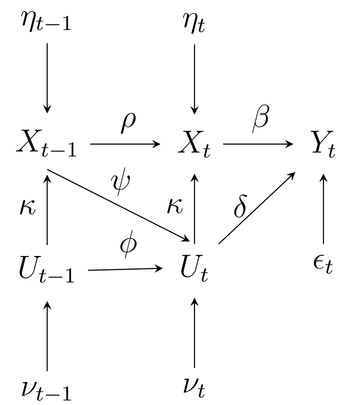


Figure 6: Monte Carlo Results with $\kappa = 0.5$ and 2 and ρ ranging from 0 to 1. Lagged Causality in Dependent Variable.

Figure 7: Monte Carlo Simulations



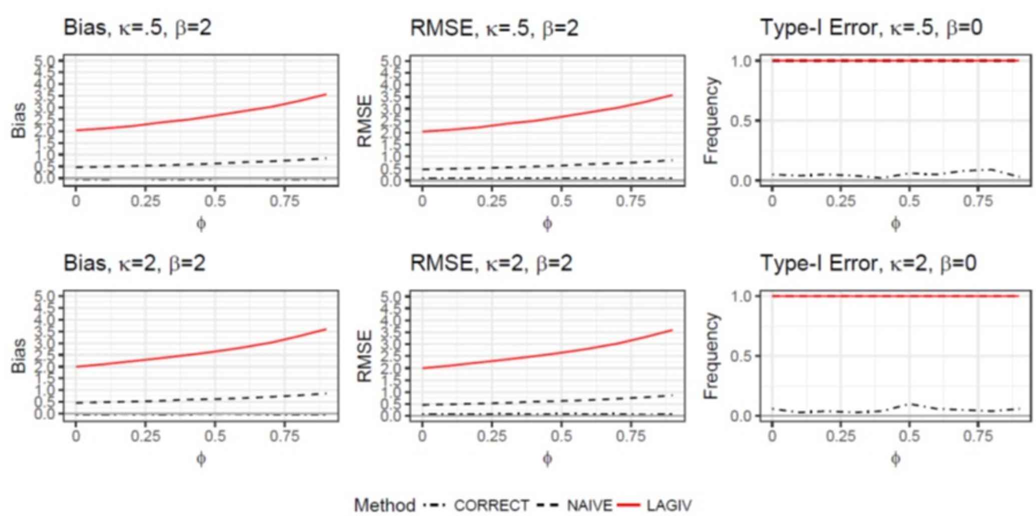


Figure 8: Monte Carlo Results with $\kappa = 0.5$ and 2 and ϕ ranging from 0 to 1. Lagged Causality in Unobserved Confounders.

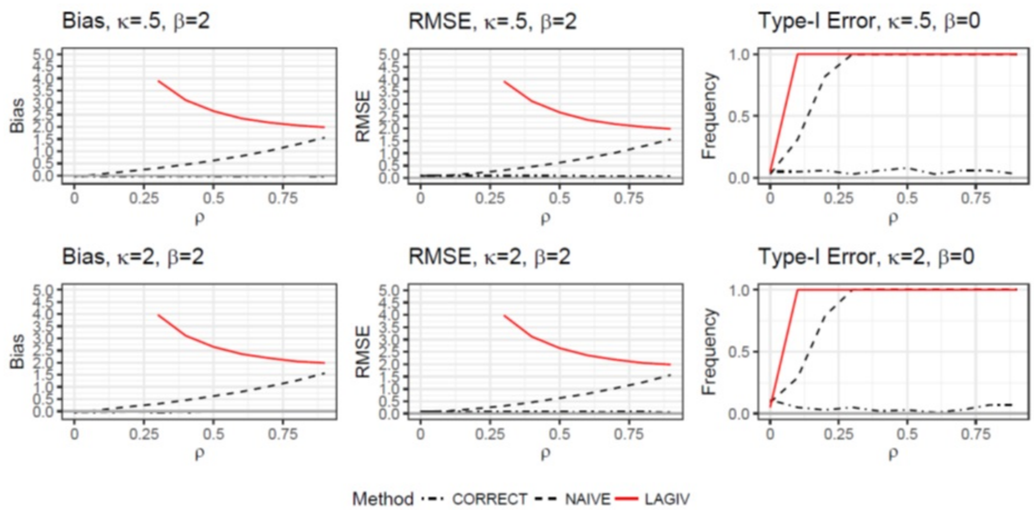
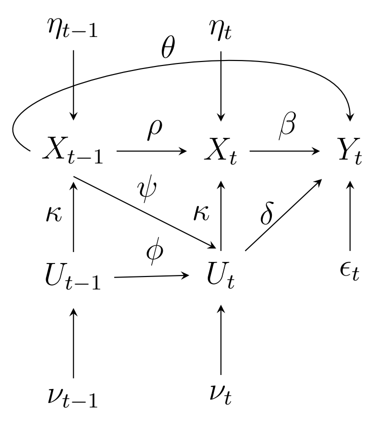


Figure 9: Monte Carlo Results with $\kappa = 0.5$ and 2 and ρ ranging from 0 to 1. Lagged Causality in Unobserved Confounders.

Figure 10: Monte Carlo Simulations



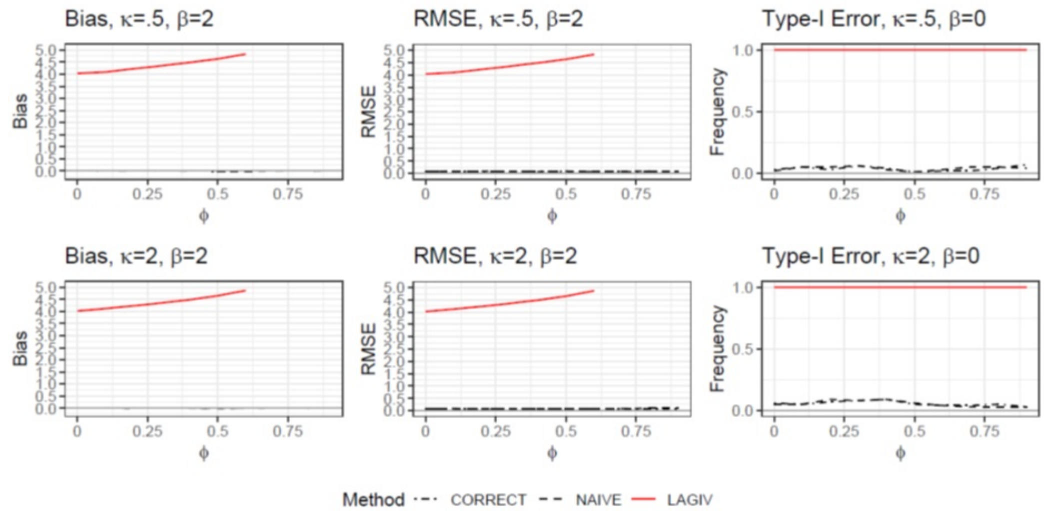


Figure 11: Monte Carlo Results with $\kappa = 0.5$ and 2 and ϕ ranging from 0 to 1. Lagged Causality in Both the Dependent Variable and the Unobserved Confounders.

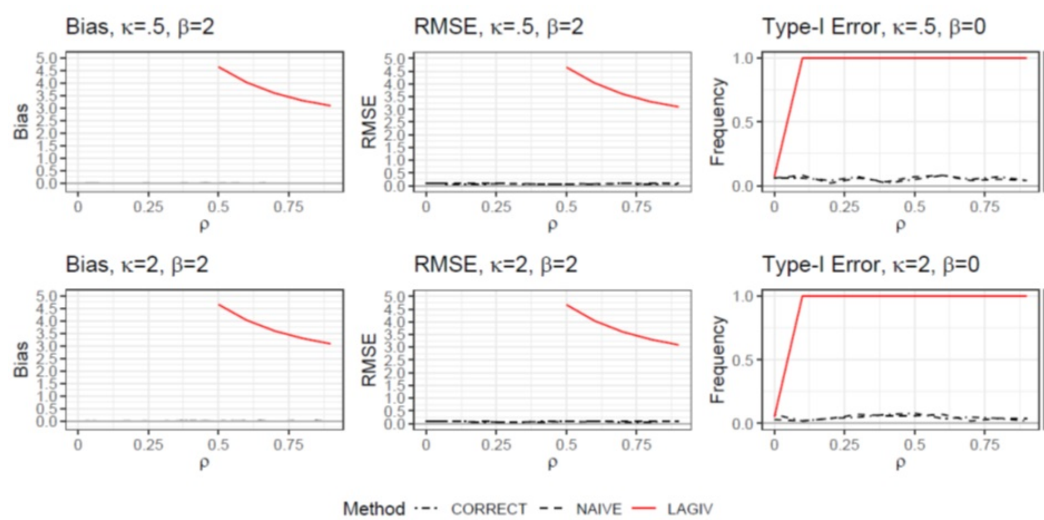


Figure 12: Monte Carlo Results with $\kappa = 0.5$ and 2 and ρ ranging from 0 to 1. Lagged Causality in Both the Dependent Variable and the Unobserved Confounders.

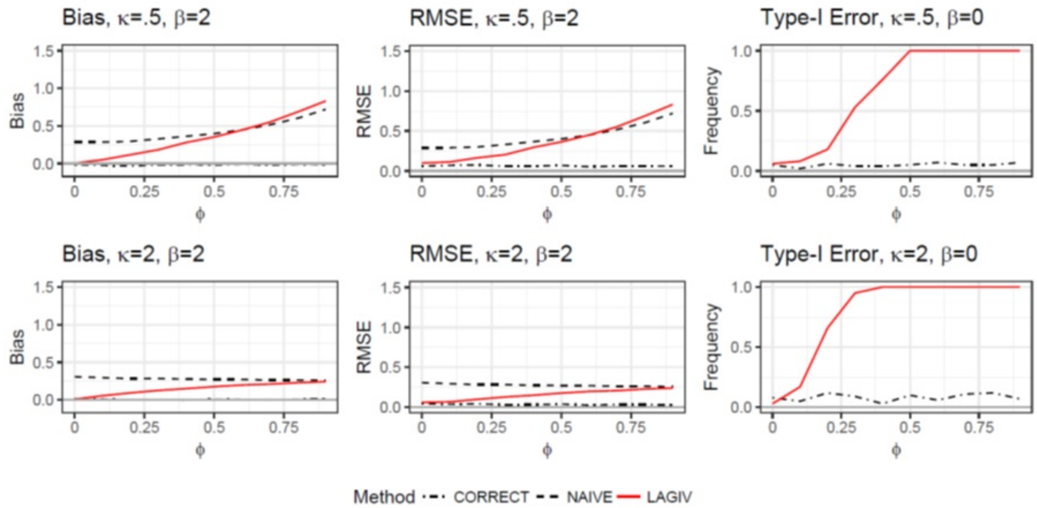


Figure A1: Monte Carlo Results with $\kappa = 0.5$ and 2 and ϕ ranging from 0 to 1. Fixed Effects Used in Naïve Specification.

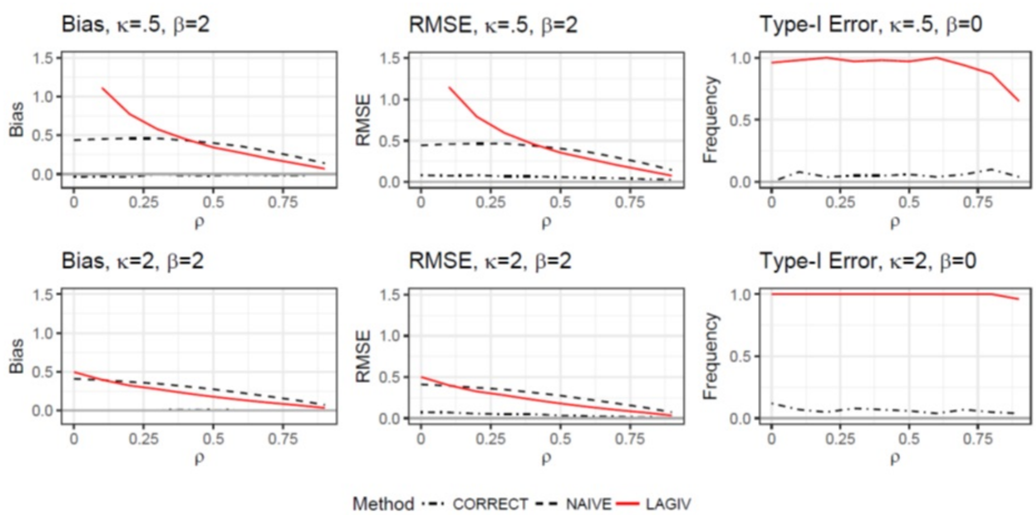


Figure A2: Monte Carlo Results with $\kappa = 0.5$ and 2 and ρ ranging from 0 to 1. Fixed Effects Used in Naïve Specification.