Lagged Explanatory Variables and the Estimation of Causal Effects

Marc F. Bellemare* Takaaki Masaki† Thomas B. Pepinsky‡

September 26, 2016

Abstract

Lagged explanatory variables are commonly used in political science in response to endogeneity concerns in observational data. There exist surprisingly few formal analyses or theoretical results, however, that establish whether lagged explanatory variables are effective in surmounting endogeneity concerns and, if so, under what conditions. We show that lagging explanatory variables as a response to endogeneity moves the channel through which endogeneity biases parameter estimates, supplementing a "selection on observables" assumption with an equally untestable "no dynamics among unobservables" assumption. We build our argument intuitively using directed acyclic graphs and then provide analytical results on the bias of lag identification in a simple linear regression framework. We then use Monte Carlo simulations to show how, even under favorable conditions, lag identification leads to incorrect inferences. We conclude by specifying the conditions under which lagged explanatory variables are appropriate responses to endogeneity concerns.

Keywords: Endogeneity, Causal Identification, Lagged Variables, Time-Series Cross-Sectional Data

WORD COUNT: 9113 (excluding appendices)

---

*Associate Professor and Director, Center for International Food and Agricultural Policy, Department of Applied Economics, University of Minnesota, Saint Paul, MN 55108, mbellema@umn.edu.

†Postdoctoral Fellow, Institute for the Theory & Practice of International Relations, College of William & Mary, VA 23185, tmasaki@wm.edu.

‡Associate Professor, Department of Government, Cornell University, Ithaca, NY 14853, pepinsky@cornell.edu.

Lagged Explanatory Variables and the Estimation of Causal Effects

1 Introduction

Lagged explanatory variables are a common strategy used in political science in response to endogeneity concerns in observational data. This strategy—what we term "lag identification"—is particularly attractive because it purports to alleviate threats to causal identification without requiring any other data than that available in the dataset. Surprisingly, however, there exists no formal analysis of lagged explanatory variables in the context of endogeneity. Despite the popularity of lag identification in contemporary political science, researchers have few theoretical results to guide them about whether lagged explanatory variables are actually effective in surmounting endogeneity concerns, under what conditions, or whether lagged explanatory variables might generate even more misleading results than if researchers were simply to ignore endogeneity altogether.

In this paper we provide such an analysis. Focusing on what we consider to be the best-case scenarios for lag identification to be effective, we find that lag identification is almost never a solution to endogeneity problems in observational data. Rather than mitigating endogeneity threats, or facilitating the identification of causal effects, lag identification merely moves the channel through which endogeneity affects the estimates of parameters of interest. We characterize precisely the conditions under which lagging an explanatory variable purges an estimate of endogeneity: these are (i) serial correlation in the potentially endogenous explanatory variable and (ii) no serial correlation among the unobserved sources of endogeneity. This supplements the selection on observables assumption that motivates the regression with a new identification assumption of "no dynamics among unobservables," which may be problematic insofar as it places substantive restrictions on the properties of a variable that is not observed. Our analysis allows schol-

ars to appreciate just what it is at stake when lagging explanatory variables in the context of endogeneity, and provides design-based reasoning that will enable more considered choices between different statistical approaches to this problem.

Our argument is most closely related to concurrent research by Reed (2015), who studies the use of lagged explanatory variables for causal inference in economics, but focuses on simultaneity and proposes the use of lagged explanatory variables as instruments for endogenous explanatory variables. In contrast, our work focuses on more general forms of endogeneity, and our results imply that Reed's recommendations are unlikely to represent a valid solution to the identification problem. Our work is also related to Blackwell and Glynn (2014), who are broadly concerned with establishing theoretical results about causal inference using time-series cross-sectional (i.e., large-T and large-N panel) data. All of our arguments are consistent with theirs, and in fact, our results emerge as a special and yet particularly important case of their more general analysis with only two time periods (see also Robins et al. 2000, p. 551). Our contribution is more focused, and designed to identify a specific practice in political science research whose consequences are not properly understood. As we demonstrate in Section 2 below, lag identification is a particularly common statistical practice that is pervasive in contemporary political science research, counting for dozens upon dozens of articles published in prominent political science outlets in the year 2014 alone. Finally, our work is also closely related to Cranmer et al. (forthcoming), who propose a Bayesian methodology for identifying lag structures in the face of theoretical uncertainty about how long lags should be.

Our analysis applies to any statistical problem in the social sciences in which endogeneity threatens the identification of a parameter of interest in time-series cross-sectional data, including cases where researchers do not claim to be estimating treatment effects. However, our contribution is also motivated by the same concern for credible estimates of causal effects that has motivated recent advances in randomized controlled trials (Duflo et al. 2007; Glennerster and Takavarasha 2013), field experiments (Gerber and Green

2012; Harrison and List 2004; Teele 2014), instrumental variables (Angrist et al. 1996; Imbens 2014; Sovey and Green 2011), regression discontinuity (Imbens and Lemieux 2008), differences-in-differences estimation (Bertrand et al. 2004), marginal structural models (Robins et al. 2000), and other design-based statistical techniques in the social sciences (see Dunning 2012; Samii 2016). The common theme uniting this literature is the importance of design; in the words of Sekhon (2009, 503), "without an experiment, a natural experiment, a discontinuity, or some other strong design, no amount of econometric or statistical modeling can make the move from correlation to causation persuasive." We echo this conclusion, showing that one popular statistical "fix" does not rescue credible estimates of causal effects in the presence of endogeneity.

The rest of this paper is organized as follows. In section 2, we discuss the general problem posed by the use of lagged variables as regressors using directed acyclic graphs (Pearl 2009). We also present an overview of recent articles in the most prominent political science journals which rely on lagged explanatory variables as a source of exogenous variation. Section 3 derives analytical results for the biases of lag identification in a common parametric setting: an ordinary least squares (OLS) regression. In this parametric case, we provide a formal result for the "no dynamics among unobservables" condition that allows for conservative estimates of causal effects using lagged explanatory variables in the presence of endogeneity. Section 4 presents Monte Carlo results showing that the use of lagged explanatory variables can worsen the identification problem, with consequences for inference that are often worse than simply ignoring endogeneity altogether. These results also allow us to describe the tradeoffs between ignoring endogeneity and lagging explanatory variables, which reveals how design-based reasoning informs the choice between alternative specifications. Section 5 entertains a wide range of extensions: alternative lag structures, generalized method of moments (GMM) estimation, classic time series applications, and different sources of endogeneity, demonstrating that our findings emerge across many data structures and estimation strategies. Section 6 outlines a set of

guidelines for researchers to follow when using lagged explanatory variables to surmount endogeneity problems in observational data, and Section 7 concludes.

2 Problem Definition

There are three reasons why a lagged value of an independent variable might appear on the right hand side of a regression.

1. *Theoretical*: In some contexts, there are clear theoretical reasons to expect that the effect of an explanatory variable only operates with a one-period lag. Such is the case, for example, when economists estimate Euler equations in order to study intertemporal substitution behaviors, or when considering the efficient market hypothesis in its random walk version, wherein p_t, the price of an asset today, is a function of the price of the same asset yesterday, p_{t-1}, and an error term e_t. It could also be the case that the analysis is directly interested in lagged effects conditional on contemporaneous effects, in which both current and lagged values of the independent variable would appear on the right hand side of a regression.

2. *Statistical*: In other contexts, lagged independent variables serve a statistical function. Examples include dynamic panel data analysis (Arellano and Bond 1991) as well as distributed lag, error correction, and related families of dynamic statistical models (see De Boef and Keele 2008). In such cases, a general model is required and lagged explanatory variables enable the calculation of both short- and long-term effects, but all parameters are still weakly exogenous by assumption.

3. *Identification*: Frequently, applied researchers propose to use a lagged value of an explanatory variable X in order to "exogenize" it when estimating the effect of X on Y. Since Y_t cannot possibly cause X_{t-1}, the (often implicitly made) argument goes, replacing X_t with X_{t-1} obviates concerns that X is endogenous to Y.

Our focus in this paper is on the use of lagged explanatory variables for identification purposes. None of our critiques of lag identification apply to theoretical or statistical motivations for including lagged values of independent variables on the right hand side of a regression, although we will touch briefly on both of these in our Monte Carlo analysis in section 4. If theory indicates that causal effects operate with a one period lag, or if the research design implies a dynamic panel approach, then lagged explanatory variables are wholly appropriate. But as we now discuss, in much applied research, neither a theoretical nor a statistical argument is provided when researchers lag their independent variables.

How common is the practice of lagging explanatory variables for identification purposes? To answer this question, we examined all articles published in the top general journals in political science, economics, and sociology, as well as several top journals in the political science subfields of comparative politics and international relations (country-year data structures are common in these subfields, so we anticipate that problems would be particularly acute here). We identified articles that used lagged explanatory variables by searching the full text of each for the word "lag," and then discarding articles that used lags purely for the purposes of forecasting, or that used the word "lag" in some other context, including articles that lagged only their dependent variable, or included only spatial lags. We also looked closely at the justifications that authors provided for including lagged explanatory variables.

In the *American Political Science Review*, between 2010 and 2014, we uncovered twenty-three published articles in which authors employed a lagged independent variable in a regression analysis. In fifteen of these articles—65%—we found that authors either had explicitly argued that lagged explanatory variables were used to alleviate endogeneity concerns, or provided no justification at all for lagging an explanatory variable. The picture that emerges from other major journals in political science is even worse (see Table 1). In 2014 alone, we count a total of seventy-seven published articles in prominent political

Table 1: Reviewed Journals Published in 2014

Journal Name	Discipline	Lag Articles	Lag "Identified"
American Political Science Review	Political Science	3	1
American Journal of Political Science	Political Science	10	6
Journal of Politics	Political Science	10	8
British Journal of Political Science	Political Science	10	8
Comparative Political Studies	Political Science	14	7
International Organization	Political Science	8	8
International Studies Quarterly	Political Science	15	10
World Politics	Political Science	7	6
American Economic Review	Economics	4	2
Econometrica	Economics	1	1
Journal of Political Economy	Economics	1	1
Quarterly Journal of Economics	Economics	2	0
Review of Economic Studies	Economics	1	1
Review of Economics and Statistics	Economics	8	6
American Sociological Review	Sociology	1	1
American Journal of Sociology	Sociology	0	0
European Sociological Review	Sociology	1	1

Notes: Lag Articles is a raw count of the number of articles published in 2014 that employed a lagged explanatory variable. Lag"Identified" is the number of Lag Articles that either involved endogeneity as a justification for lagging an explanatory variable, or contained no justification at all for lagging an explanatory variable.

science journals employing lagged explanatory variables, fifty-four of which we coded as cases of lag identification.

Comparing across disciplines using the data in Table 1 also suggests that this practice is much more common in political science relative to economics or sociology. Moreover, articles in political science journals frequently invoked "simultaneity" or "reverse causality" explicitly as the sole motivation for lagging explanatory variables.[1] Somewhat more

[1]Some examples are as follows: Baccini and Urpelainen (2014, 205) write "Most of these variables are lagged by one year to avoid endogeneity problems." Lehoucq and Perez-Linan (2014, 1113) write "We lag both economic variables one year to minimize problems of endogeneity." Steinberg and Malhotra (2014, 513) write "All independent and control

concerning, a substantial minority of articles that we identified in this survey contained no justification whatsoever for their lag choice. We did identify a number of cases where authors employed lagged explanatory variables as part of an error correction or distributed lag model, but these remain the minority of the articles that we identified. As Table 1 shows, in 2014, across a range of journals, more than half of the articles that employed lagged exogenous variables either explicitly invoked endogeneity, or contained no justification at all.

This review of recent scholarship reveals that the practice of lagging explanatory variables for identification purposes remains common in the most prominent political science journals. We acknowledge that authors who use lags for identification purposes almost certainly would not hold that lags are a "true solution" to identification problems. Still, this review of recent literature and the examples that we have cited above reveal that authors justify this choice on the presumption that lag identification somehow *mitigates* these problems. Our analysis allows us to characterize this claim with greater precision. We now turn to closer examination of the conceptual problems that lag identification creates, which is a foundation for better understanding the tradeoffs between different estimating strategies in the presence of endogeneity.

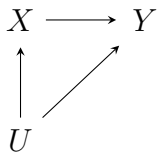
2.1 Directed Acyclic Graphs

Following Pearl (2009), we begin with an intuitive discussion of the problem which relies on directed acyclic graph (DAGs). The DAG in Figure 1 shows the fundamental identification problem in observational data: the identification of the causal relationship flowing from X to Y is compromised by the presence of unobservable factors U which are correlated with both X and Y.

Figure 2, wherein we add subscripts t to clarify temporal ordering, illustrates the lag

variables are lagged by one year to mitigate the possibility of simultaneity or reverse causality bias."

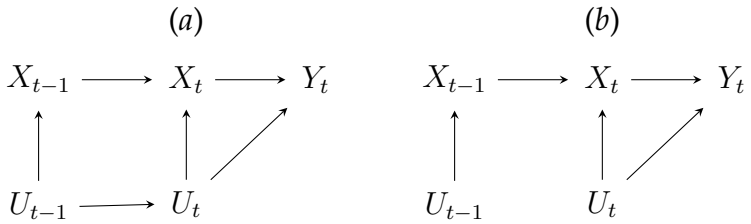
Figure 1: The Identification Problem



Notes: This is a representation of a causal relationship from X to Y where identification is compromised by unobservables U.

identification approach that we study in this paper. Figure 2 is a special case of Figure 1(a) in Robins et al. (2000, p. 551), which we believe is the first published example of a DAG representing our data structure, but which does not consider lag identification in any way. Lag identification means replacing X_t with its lagged value, X_{t-1} in a regression of X on Y. The DAG representation in Figure 2 clarifies the logic behind this approach, which is formally a *selection on observables* identification strategy (see Keele 2015: 321-322). It must be the case that there is a causal pathway from $X_{t-1} \rightarrow X_t$, or else X_{t-1} could not be related to Y. However, the fact that there is no direct causal link running from U_t to X_{t-1} means that there is no possibility that this particular unobserved confounder U_t threatens causal identification. But Figure 2(a) also shows that replacing X_t with X_{t-1} merely

Figure 2: Lagged Independent Variable as a Solution?



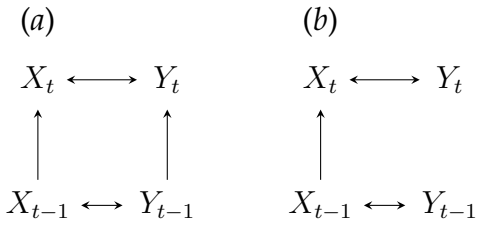
Notes: This is a representation of the causal relationship from X to Y that is implied when using a lagged value of X to overcome the identification problem in figure 1. In (*a*), U_t depends on its previous value U_{t-1}. In (*b*) the two are independent.

moves the endogeneity problem back one time period. It is true that X_{t-1} is unaffected by U_t, but it *is* affected by U_{t-1} for the same reason that $U_t \rightarrow X_t$. As a result, if there are any temporal dynamics in the unobservables, then the causal pathways $U_{t-1} \rightarrow U_t \rightarrow Y_t$

and $U_{t-1} \rightarrow X_{t-1} \rightarrow X_t \rightarrow Y_t$ prevent causal identification using X_{t-1}. The critical identification assumption in lag identification, therefore, is that there are no temporal dynamics among the unobservables, as shown in Figure 2(b). This assumption is not testable: doing so would require observing U, the unobservable confounder that motivates lagging X on identification grounds. Interestingly, if $U_{t-1} \not\rightarrow U_t$ then X_{t-1} is actually an *instrument* for X_t; that is, if lagged explanatory variables are valid exogenous proxies for contemporary explanatory variables, they are also valid instruments as well.

Our discussion thus far has focused on endogeneity in the form of unobserved heterogeneity. In many applications, however, lag identification is justified on "reverse causality" grounds rather than unobserved heterogeneity grounds. The classic example of simultaneous causation is Haavelmo's (1943) treatment of the joint determination of consumption and investment. This causal process is depicted in Figure 3, which shows that if Y_t causes X_t, Y_{t-1} also causes X_{t-1}. Note that Figure 3 is not a DAG because the causal relations that it depicts are cyclic. Nevertheless as Pearl (2009, pp. 27-28, 215-217) illustrates, a graphical approach facilitates the exposition of identifiable causal relationships.

Figure 3: Lagged Independent Variable with Reverse Causality



Notes: This is a representation of simultaneous causation with no unobservables. In (*a*), Y_t depends on its previous value Y_{t-1}. In (*b*) the two are independent conditional on X.

The argument that temporal ordering prevents current realizations of the dependent variable from affecting past values of a causal variable may be more reasonable as a defense against simultaneous or reverse causation. The identification assumption, however, is now that there are dynamics in X but not Y (Figure 3(b)). We will analyze a system of this

sort in section 5.4 below, but to preview our results, we find that simultaneous causation without unobserved confounders generates similar pathologies for lag identification.[2]

3 Analytical Results

The DAGs in the preceding section are useful for clarifying the intuition behind lagged independent variables, and also for demonstrating why they are unlikely to sidestep problems of endogeneity. To characterize precisely the consequences of lagged independent variables in the context of endogeneity, in this section we analyze formally the consequences of lag identification in a bivariate OLS regression setup. We reiterate that results apply to *any* situation in which endogeneity affects estimates of the relationship between X and Y using observational data, not only to analyses that explicitly seek to identify casual effects.[3]

[2]In some cases, it is possible to reformulate reverse causality as problems of unobserved heterogeneity in which a latent variable representing the "likelihood" or "propensity" of Y is an unobserved confounder that causes both Y and X. See Pearl (2009, 145-49) for a related argument on the observational equivalence of structural equation models.

[3] Our analysis applies directly to the estimation of treatment effects if we assume that the OLS regression framework is the correct functional form for the estimation of the causal effect of X on Y. Of course, if the correct functional form is unknown, then a non-parametric approach such as those offered by Pearl or Rubin, as well as precise assumptions about counterfactual outcomes, are necessary to define estimators that estimate causal effects.

Consider the model

$$Y_{it} = \beta X_{it} + \delta U_{it} + \epsilon_{it} \tag{1}$$

$$X_{it} = \rho X_{it-1} + \kappa U_{it} + \eta_{it} \tag{2}$$

$$U_{it} = \phi U_{it-1} + \nu_{it} \tag{3}$$

where i and t index units and time, respectively; $0 \leq \rho < 1$; $0 \leq \phi < 1$; and $\epsilon_{it} \sim N(0, \sigma_\epsilon^2)$, $\eta_{it} \sim N(0, \sigma_\eta^2)$, and $\nu_{it} \sim N(0, \sigma_\nu^2)$. Dropping i for the remainder of this section (it will reappear in the next section), it is well known that if we estimate

$$Y_t = bX_t + e_t \tag{4}$$

then the resulting estimate of β is biased because the unobserved confounder U is omitted.[4] The magnitude of the bias is a function of the variances and covariances of X and U as well as magnitude of the causal effect of the unobserved confounder:

$$\mathrm{E}[\hat{b}_{X_t}] = \beta + \delta \cdot \frac{\mathrm{Cov}(X, U)}{\mathrm{V}(X)} \tag{5}$$

If either δ or $\mathrm{Cov}(X, U) = 0$—if U has no effect on Y, or if U is uncorrelated with X—then endogeneity is not a problem, and $\mathrm{E}[\hat{b}_{X_t}] = \beta$.

Now consider a regression that replaces X with X_{t-1}.[5] This means estimating the following equation:

$$Y_t = bX_{t-1} + e_t \tag{6}$$

[4]We use Greek letters for population coefficients and Latin letters for sample coefficients.

[5]For expositional purposes we do not consider here more complicated models that condition on past values of Y. For a contemporary analysis of when to include lagged dependent variables to estimate causal effects, see Dafoe forthcoming.

While this is plainly not an unbiased estimate of β, one hope is that lag identification will estimate a function of β and the autocorrelation in X, or ρ—a moderated, or "conservative," estimate of β. However, by expression (2) we know that X_t is also a function of U_t. We may therefore rewrite equation (1) as follows:

$$\begin{aligned} Y_t &= \beta(\rho X_{t-1} + \kappa U_t + \eta_t) + \delta U_t + \epsilon_t \\ &= \beta\rho X_{t-1} + (\beta\kappa + \delta)U_t + \beta\eta_t + \epsilon_t \end{aligned} \tag{7}$$

This immediately makes clear that the error term e_t in (6) now contains $(\beta\kappa+\delta)U_t+\beta\eta_t+\epsilon_t$. Therefore, $\hat{b}_{X_{t-1}}$ is not a consistent estimate of either β or the conservative $\beta\rho$. To see what exactly $\hat{b}_{X_{t-1}}$ does estimate, recall that $\hat{b}_{X_{t-1}} = \frac{Cov(X_{t-1},Y_t)}{V(X_{t-1})}$. We will assume here that κ represents the only source of endogeneity of X, so we impose that $\sigma_{\eta,\nu} = 0$. We may therefore write

$$\begin{aligned} \text{plim}_{n\to\infty}\, \hat{b}_{X_{t-1}} = &\ \frac{\text{Cov}(X_{t-1}, \beta\rho X_{t-1} + (\beta\kappa + \delta)U_t + \beta\eta_t + \epsilon_t)}{\text{V}(X_{t-1})} \\ = &\ \frac{\text{Cov}(X_{t-1}, \beta\rho X_{t-1})}{\text{V}(X_{t-1})} + \frac{\text{Cov}(X_{t-1}, (\beta\kappa + \delta)U_t)}{\text{V}(X_{t-1})} + \\ &\ \frac{\text{Cov}(X_{t-1}, \beta\eta_t)}{\text{V}(X_{t-1})} + \frac{\text{Cov}(X_{t-1}, \epsilon_t)}{\text{V}(X_{t-1})} \end{aligned} \tag{8}$$

We know that by design, given expressions (1– 3), the third and fourth terms in (8) reduce to zero. In Appendix 1 we show how to simplify the remaining terms to produce a final expression for what lagged independent variables estimate.

$$\text{plim}_{n\to\infty}\, \hat{b}_{X_{t-1}} = \beta\rho + \frac{\phi\kappa(\beta\kappa + \delta)\text{V}(U)}{(1 - \phi\rho)\text{V}(X)} \tag{9}$$

Contrasting lag identification bias in (9) with the standard result for omitted variable bias in (5) usefully highlights the troublesome properties of lagged independent variables for estimating β, which is the true contemporaneous effect of X on Y, or $\beta\rho$, which is the

true lagged effect of X on Y. The difference between $\hat{b}_{X_{t-1}}$ and $\beta\rho$ increases with both the degree of endogeneity, κ, and the serial correlation in U, ϕ. Furthermore, when (5) does *not* yield a biased estimate of β because $\delta = 0$ but $\phi \neq 0$ or $\rho \neq 0$, (16) does. In general, lagging X_t and using it as a regressor creates a "back-door channel" through $U_{t-1} \rightarrow X_{t-1}$ and $U_{t-1} \rightarrow U_t \rightarrow Y_t$. Expression (16) also establishes that one of the following conditions must hold for lag identification to produce a consistent estimate of $\beta\rho$, the "conservative" estimate of the effect of X on Y, attenuated by ρ.

1. No serial autocorrelation in U ($\phi = 0$), i.e., no dynamics among unobservables.

2. There is no endogeneity, which means that $\kappa = 0$ or $\delta = 0$.

The first case is precisely the condition identified in Section 2.1 above. In that case, the second term reduces to zero, and $\text{plim}_{n\rightarrow\infty}\ \hat{b}_{X_{t-1}} = \beta\rho$. In the second case, there is obviously no identification problem, and thus no need for lag identification to begin with.

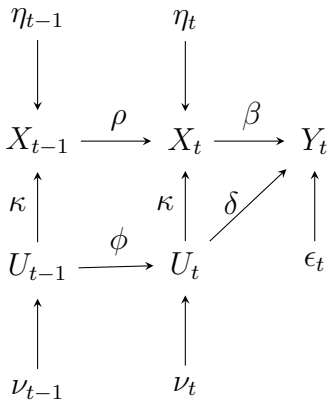
4 Monte Carlo Analysis

We have shown so far that lagging independent variables does not produce statistical estimates that are free from endogeneity, except under the specific conditions that there is no endogeneity, or there are no dynamics among the unobserved sources of endogeneity. We have also characterized analytically the source and magnitude of this bias in a simple OLS regression setup. In this section, we use Monte Carlo experiments to study the consequences of lagged explanatory variables in common time-series cross-sectional applications. Doing so allows us to illustrate tradeoffs that arise when researchers do not have access to a perfect research design: is it better to ignore endogeneity rather than to lag explanatory variables in the presence of endogeneity?

4.1 Setup

Our task is to estimate β, the effect of X on Y.[6] Figure 4 is an extension of our earlier analysis which parameterizes the relations of interest. As above, the source of endogeneity

Figure 4: Monte Carlo Simulations



Notes: This is a schematic representation of our Monte Carlo simulations, with Greek letters representing the parameters that we vary in our simulations. X is the explanatory variable of interest, represented here as a function of a random variable ν and its own past value. U is an unbserved source of endogeneity, and is itself a function of a random variable ν and its own past value. Y is the dependent variable, and is a function of observed X, unobserved U, and a random error term ϵ. β is the parameter to be estimated, κ measures the size of the endogeneity problem, and ρ and ϕ capture dynamics in X and U, respectively.

is the unobserved confounder U, which is correlated with both X and Y. In all simulations, we set the direct effect of U on Y (which we called δ above) equal to 1, and explore the consequences of endogeneity by varying κ, the pathway that makes X endogenous to Y by forcing $\text{Cov}(X, U) \neq 0$. The remaining two parameters are the autocorrelation parameters ρ and ϕ, which capture serial correlation in X and U, respectively. When either of the autocorrelation parameters is zero, then the value of each variable is statistically independent of its own lag. In our simulations, we set ρ equal to .5, and then vary ϕ. A summary of the parameters that we vary in our simulations in Table 2. For each simulation, we generate a panel with $N = 100$ units and $T = 50$ periods, for a total of 5,000

[6]This is the *causal* effect of X on Y under the assumption in footnote 3.

Table 2: Simulation Parameters

Parameter	Causal Pathway	Simulation Values
β	$X_t \rightarrow Y_t$	{0, 2}
κ	$U_t \rightarrow X_t$, $U_{t-1} \rightarrow X_{t-1}$	{.5, 2}
ϕ	$U_{t-1} \rightarrow U_t$	{0, .1, .2, ..., .9}
δ	$U_t \rightarrow Y_t$	{1}
ρ	$X_{t-1} \rightarrow X_t$	{.5}

unit-period observations.[7]

Our simulations adopt the same DGPs as expressed in equations (1-3). We simulate each combination of parameter values in Table 2 a total of 100 times, and then test the performance of the following three estimators in estimating β (the true contemporaneous effect of X on Y) or $\beta\rho$ (the true lagged effect of X on Y): (i) the "naïve" estimator ($\hat{\beta}_{NAIVE}$) that regresses Y_t on X_t and ignores endogeneity , (ii) the "lag explanatory variables" estimator ($\hat{\beta}_{LAGID}$) that regresses Y_t on X_{t-1} in an attempt to avoid endogeneity problems, and (iii) a "correct" estimator ($\hat{\beta}_{CORRECT}$) that regresses Y_t on both X_t *and* the unobservable U_t.[8] The "true" estimator is, of course, counterfactual: we presume that the researcher does not observe U, or else she would condition on it. The estimates obtained from a regression model that correctly follows the data generating process, however, will serve as our empirical benchmark against which to gauge the performance of the other two estimators.

Our simulations have many moving parts, but the underlying DGP is still simple in terms of the dynamics that it allows. Among many other simplifications, we assume that

[7]In each simulation, variables U, X and errors η, ν are each drawn independently from a standard normal distribution $N(0, 1)$. ϵ is drawn from $N(0, 5)$. We set the variance of ϵ at 5 in order to allow for a realistic amount of model uncertainty. Most estimates from our simulations have an overall R^2 between 0.05 and 0.1, which is comparable to R^2 measures in much applied political science research using panel data.

[8]We estimate each using the "pooled" estimator implemented in the `plm` library in R.

there are no dynamic relationships among unobservables and observables. For example, lagged omitted variables U are not direct causes of current values of Y—they only affect Y through the pathway $U_{t-1} \rightarrow U_t \rightarrow Y_t$. Moreover, there are no complex temporal dynamics in X or U, just simple one-period autocorrelation. We view this relatively straightforward setup as a conservative way to explore the performance of lagged independent variables in the most favorable cases.

We evaluate the consequences of lag identification according to three criteria: (i) bias, (ii) root mean squared error (RMSE), and (iii) the likelihood of Type 1. The last of these is perhaps the most important from the perspective of applied researchers, as it tells us the extent to which researchers will make faulty inferences—rejecting true null hypotheses that $\beta = 0$, or failing to reject the null hypothesis when the true $\beta \neq 0$—when using lagged independent variables.

4.2 Results

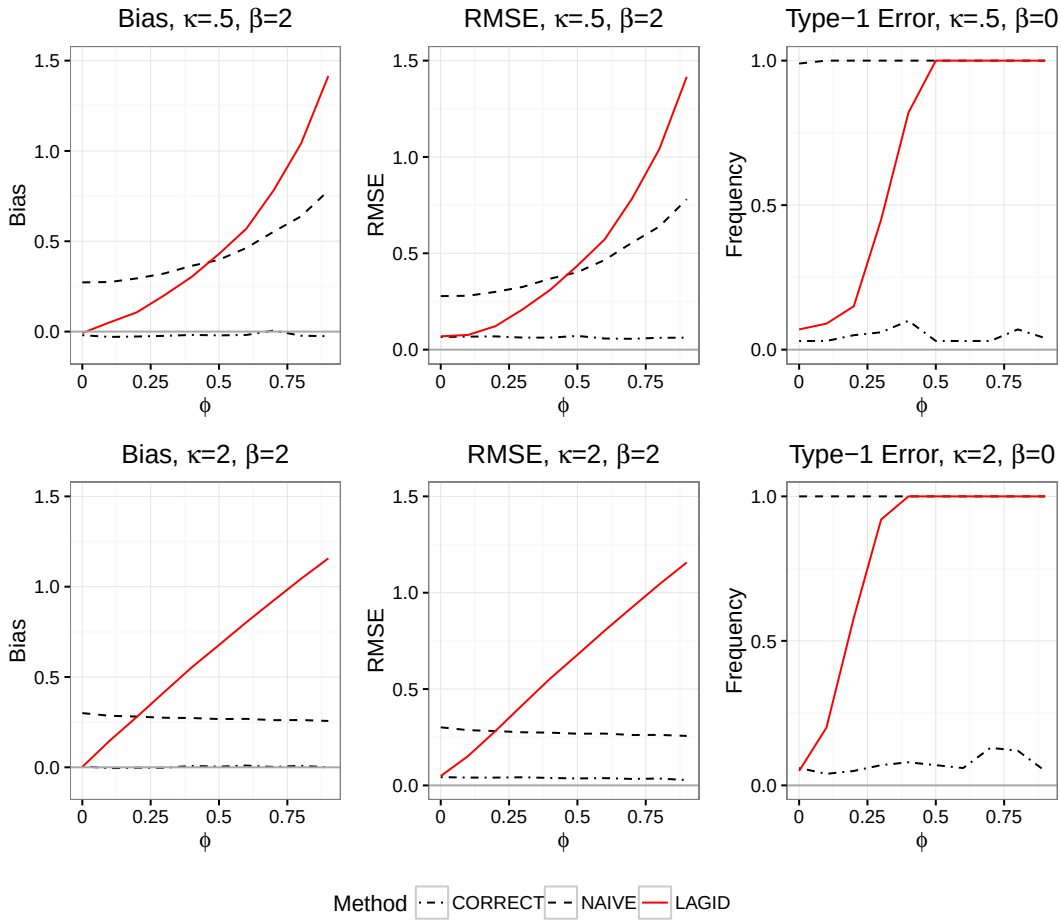
We begin by comparing bias across the three estimators. For each combination of parameter values, we save the estimated parameters $\hat{\beta}_{NAIVE}$, $\hat{\beta}_{LAGID}$, and $\hat{\beta}_{CORRECT}$ from each of the 100 simulations,[9] and then plot average levels of bias, which is the deviation of $\hat{\beta}_{NAIVE}$ and $\hat{\beta}_{CORRECT}$ from the true contemporaneous effect of X on Y (β). For the lag identification model, we define bias as the deviation between $\hat{\beta}_{LAGID}$ and $\beta\rho$, which is the conservative estimate of the effect of X on Y. Figure 5 summarizes our main results.

The results from these simulations are clear. When X is endogenous to Y ($\kappa \neq 0$), $\hat{\beta}_{NAIVE}$ is biased (regardless of the value of ϕ).[10] More importantly, as equation (9) sug-

[9]We find that 100 simulations is sufficient to produce stable results across our models; increasing the simulations further offers no noticeable improvement in precision and makes no difference for our conclusions.

[10]Needless to say, when there is no endogeneity (κ=0), both $\hat{\beta}_{NAIVE}$ and $\hat{\beta}_{LAGID}$ serve as unbiased estimators for β and $\beta\rho$, respectively.

Figure 5: Monte Carlo Results



Notes: The first two rows of this figure display the bias and RMSE for $\hat{\beta}_{CORRECT}$ and $\hat{\beta}_{NAIVE}$ relative to the true value of β, and bias and RMSE $\hat{\beta}_{LAGID}$ relative to $\beta\rho$, based on 100 simulations of the data generating process in Figure 4 with $\beta = 2$ and $\rho = .5$. The third row plots the frequency of Type-1 error when $\beta = 0$ and $\rho = .5$.

gests, as the degree of endogeneity (κ) and the strength of the first-order autoregressive parameter in the unobservables (ϕ) increase, the degree of bias associated with $\hat{\beta}_{LAGID}$ also increases. Observe in Figure 5 that the bias in $\hat{\beta}_{LAGID}$ is even *greater than* the degree of bias in $\hat{\beta}_{NAIVE}$ at higher values of ϕ. A similar pattern emerges for RMSE. Our results in Figure 5 show that the RMSE of the lag explanatory variable estimator is *larger than* that of the naïve estimator when there exists endogeneity ($\kappa \neq 0$) and the unobservable is

sufficiently persistent ($\phi > 0.5$ at $\kappa = 0.5$ and $\phi > 0.25$ at $\kappa = 2$).[11]

Of course, the finding that an incorrect regression specification generates biased parameter estimates is not surprising. In fact, for most applied researchers, bias may not matter because—all too commonly, in our view—the size of the estimate of β or $\beta\rho$ is not necessarily of direct interest, but rather its p-value. That is, researchers are less interested in whether the sizes of their estimates are upward or downward biased, but whether the associated p-value from their t-test leads them to reject the null that $\beta = 0$ or $\beta\rho = 0$ at some level of significance. We think that the overwhelming focus placed on statistical significance in political science research is a major problem, but this is nonetheless an accurate description of current practice in political science and many other social sciences. And so in the third column of Figure 5, we ask what would happen if an applied researcher were to use a lagged independent variable in the standard fashion to test the alternative hypothesis that $\beta \neq 0$ when the null hypothesis is true ($\beta = 0$), using the standard 95% confidence threshold.

Our results are troubling. The likelihood of Type 1 error increases dramatically when $\kappa > 0$ and ϕ increases. The reason for this is apparent in equation (15), which shows that $\hat{b}_{X_{t-1}}$ is a function of the causal effect of the unobserved confounder U, κ, as well as ρ and ϕ. Unless there is no endogeneity ($\kappa = 0$), which again obviates the need for lagging X in the first place, lag identification will produce non-zero estimates of $\beta\rho$ even when $\beta = 0$. Substantively, this means that lagging independent variables in response to concerns about endogeneity will lead analysts working within the mainstream approach to hypothesis testing to reject null hypotheses that are true, and to find too many estimates of causal effects that are spurious.

The summary message from these Monte Carlo simulations is unambiguous. Under conditions where there exists endogeneity and dynamics among unobservables, lag-

[11]The only exceptions are purely incidental: when there is no endogeneity ($\kappa = 0$), $\hat{\beta}_{NAIVE}$ and $\hat{\beta}_{LAGID}$ are as efficient as the estimator that conditions on U.

ging independent variables may generate estimates that are more biased, and with higher RMSE, than *simply ignoring endogeneity altogether*. Worst of all, such estimates are more likely to produce a Type 1 error when endogeneity actually does threaten identification and the true effect of X on Y is zero.[12]

These results suggest difficult tradeoffs. On one hand, Type 1 error is nearly certain when using the naive estimator that ignores endogeneity. On the other hand, Type 1 error is also nearly certain with the lag estimator unless the autocorrelation parameter in the unobservables is small, and the larger the endogeneity problem, the more demanding this requirement. The bias and error of the lag estimator is more likely to exceed that of the naive estimator under similar conditions. When choosing between lagging explanatory variables to mitigate endogeneity and just ignoring endogeneity altogether, the latter approach is superior only in cases when endogeneity threats and persistence in the unobservables are both small.

5 Extensions

In this section we entertain several potential objections to our simulation results, focusing on temporal sequencing of causal effects, unobserved heterogeneity and GMM, and cases of "pure" simultaneous causality.

5.1 Lagged Effects

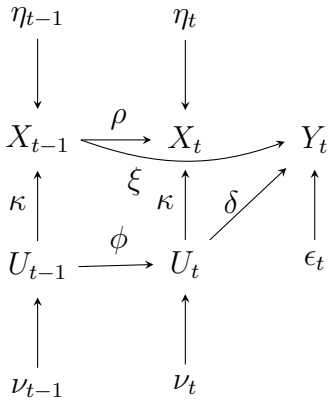
One criticism of our baseline results is that they do not realistically reflect the kinds of data generating processes that scholars mean to capture when lagging independent variables to confront endogeneity problems. If theory suggests that causal effects operate with a one-period time lag, for example, then lag identification is not just a way to

[12]In Appendix 2, we present the cases where $\rho = .1$ and $\rho = .9$, which confirm that these results hold regardless of the degree to which X_t and X_{t-1} are correlated.

avoid endogeneity, it is also the natural way to estimate the parameter of interest. Such an objection might suggest that our results in the previous subsection are simply a consequence of proposing a different data generating process than the one that might justify lag identification.

Attuned to such concerns, in Figure 6 we propose a different model for Monte Carlo analysis. Here, as before X is endogenous to Y through U, but the parameter of interest

Figure 6: Monte Carlo Simulations: X_{t-1} as the Causal Variable



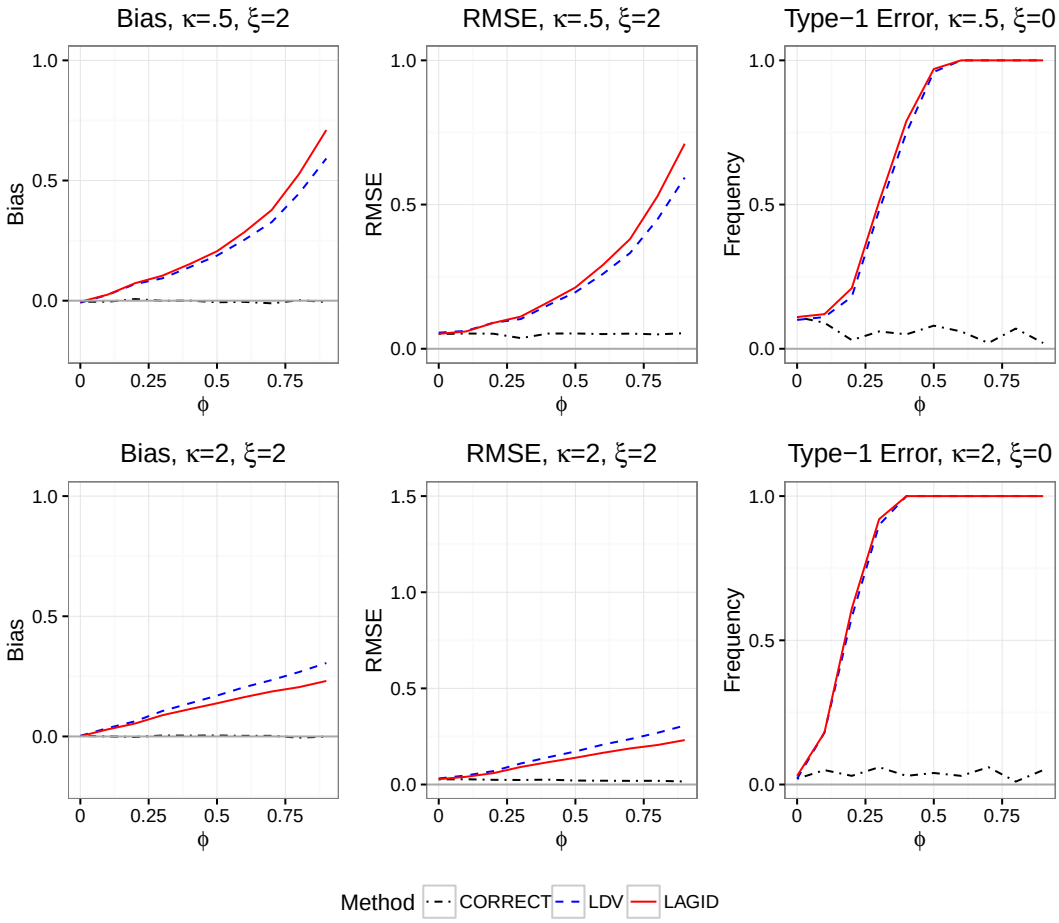
Notes: This is a schematic representation of our Monte Carlo simulations where X_{t-1} is the true causal variable (the causal effect of X_t is, by assumption, 0). It is otherwise identical to Figure 4.

ξ is the one-period lagged effect of X on Y. We therefore assume that the effect operates with a one-period lag, that the empirical specification is designed to estimate that quantity, and also that the contemporaneous effect of X_t on Y_t is exactly zero. This reflects perhaps the most favorable case for lag identification, one in which causal effects operate over time and in which there is no direct pathway that runs from the unobserved U_t to the variable of interest, in this case X_{t-1}.

The DGPs for simulations based on Figure 6 can be found in Appendix 3. In Figure 7 we compare estimates of ξ from the lag explanatory variable estimator ($\hat{\xi}_{LAGID}$), an extended version of the lag explanatory variable estimator that also conditions on Y_{t-1} in an attempt to capture temporal dynamics in the unobservables ($\hat{\xi}_{LDV}$), and the "correct" model that conditions both on X_{t-1} and U_t ($\hat{\xi}_{CORRECT}$), once again as an empirical benchmark against

which to judge the others. Our results show that even under the favorable assumption that X_{t-1} is the causal variable of interest, lagged independent variables generate biased estimates of ξ, even when also including a lagged dependent variable. As above, when $\kappa > 0$, however, lagging independent variables generates biased and inefficient estimates of ξ, with bias increasing in ϕ.

Figure 7: Monte Carlo Results, X_{t-1} as the Causal Variable



Notes: The first two rows of this figure display the bias and RMSE for $\hat{\xi}_{CORRECT}$ and $\hat{\xi}_{LDV}$, and $\hat{\xi}_{LAGID}$ relative to the true value of ξ based on 100 simulations of the data generating process in Figure 6 with $\xi = 2$ and $\rho = .5$. The third row plots the frequency of Type-1 error when $\xi = 0$ and $\rho = .5$.

We also find similar results for Type 1 error, which appear in the third column of Figure 7. As before, these results indicate that with any amount of endogeneity, t-statistics from lagged explanatory variables (either $\hat{\xi}_{LAGID}$ or $\hat{\xi}_{LDV}$) are likely to lead applied researchers

to reject the null that $\xi = 0$ when the null is true. The implication of this analysis is that even if a strong theory dictates that the causal process linking X to Y operates with exactly and exclusively a one-period lag, lagged independent variables do not avoid problems of endogeneity.

5.2 Fixed Effects and GMM Estimation

The data generating processes we have entertained so far contain no sources of unobserved unit-level heterogeneity, so we can safely use the pooled estimator to estimate β. However, in most observational data contexts, fixed effects are necessary to account for unobserved heterogeneity. In separate results that we do not report here, all of our findings remain identical in Monte Carlo simulations that introduce unobserved unit-level heterogeneity, accounted for using unit fixed effects. We also explore whether standard dynamic panel data models (Arellano and Bond 1991; Blundell and Bond 1998), which use higher order lags and differences of both X and Y as instruments for X, X_{t-1}, and Y_{t-1}, yield better results in simulations with unobserved unit-level heterogeneity. The results of these analyses are available in Appendix 4. The summary finding is straightforward: GMM estimation fares better than does lag identification with fixed effects in the context of unobserved heterogeneity, but results remain biased away from zero and Type 1 errors remain very likely.

5.3 Advanced Time-Series Estimation

There are many possible extensions that are possible in the context of time-series data: ARIMA modeling, VAR estimation, and others. However, without either observing U or finding an instrument for X, then estimates of β will always be biased in a lagged explanatory variable model. However, our unambiguous results from above, that lagged explanatory variables are superior to ignoring endogeneity only in cases when endogeneity threats and persistence in the unobservables are both small, may not hold in under

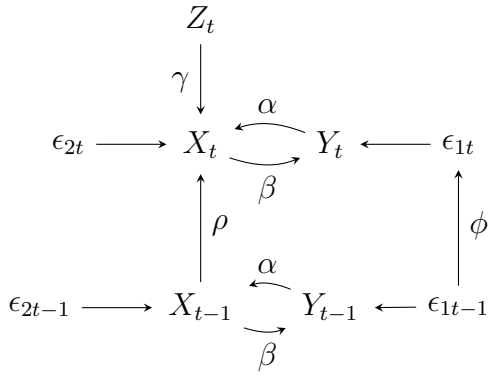
such models. This recommends further care when using such statistical models to model data generating processes such as those that we have considered here.

5.4 True Simultaneous Causality

Our final extension returns to the problem of simultaneous causality. Here, we show that true simultaneous causality has different implications for the estimation of causal effects.

Specifically, we consider the DGP in Figure 8, which is an extension of Haavelmo (1943) (for a discussion of this model in the context of causal analysis, see Pearl forthcoming). We incorporate into this model an instrument for X, denoted Z. Endogeneity in Figure 8 is not

Figure 8: Monte Carlo Simulations: Simultaneous Causality



Notes: This is a schematic representation of our Monte Carlo simulations where X and Y are truly "simultaneous" equations. Z serves as an instrument for X whenever $\gamma \neq 0$. ϵ_{1t} follows an autoregressive process in our simulations: $\epsilon_{1t} = \phi\epsilon_{1t-1} + \eta$.

a function of unobserved confounders, but rather of a simultaneous causal relationship in which Y and X directly cause one another. We use the following system of equations to

represent this causal structure:

$$Y_{it} = \beta X_{it} + \epsilon_{1it} \tag{10}$$

$$X_{it} = \rho X_{it-1} + \alpha Y_{it} + \gamma Z_{it} + \epsilon_{2it} \tag{11}$$

$$\epsilon_{1it} = \phi \epsilon_{1it-1} + \eta_{it} \tag{12}$$

where $0 \leq \rho < 1$, $0 \leq \phi < 1$, $\epsilon_{2it} \sim N(0, 5)$, and $\eta_{it} \sim N(0, 1)$.

By substituting equations (10) and (12) into (11), we can express X solely in terms of its own lag, model parameters and errors:

$$X_{it} = \frac{\rho X_{it-1} + \alpha\phi\epsilon_{1it-1} + \alpha\eta_{it} + \gamma Z_{it} + \epsilon_{2it}}{1 - \alpha\beta} \tag{13}$$

This expression reveals an important difference between the simultaneous causality case and our results above for unobserved heterogeneity. Because X_t is a function of Y_t, it is no longer true that $\hat{b}_{X_{t-1}}$ will estimate $\beta\rho$ when $\phi = 0$, which was true in our discussion of unobserved heterogeneity. We can show this by substituting (13) in (10), which yields:

$$Y_{it} = \frac{\beta\rho}{1 - \alpha\beta} X_{it-1} + \frac{\beta(\alpha\phi\epsilon_{1it-1} + \alpha\eta_{it} + \gamma Z_{it} + \epsilon_{2it})}{1 - \alpha\beta} + \epsilon_{1it} \tag{14}$$

The first term will reduce to $\beta\rho$ only when $\alpha = 0$, meaning that there is no simultaneous causation at all.

We test the performance of the following three estimators in estimating β: (i) the "naïve" estimator ($\hat{\beta}_{NAIVE}$) that regresses Y_t on X_t and ignores simultaneity, (ii) the "lag explanatory variables" estimator ($\hat{\beta}_{LAGID}$) that regresses Y_t on X_{t-1}, and (iii) a instrumental variable estimator ($\hat{\beta}_{IV}$) that employs a two-stage least squares model by using Z_t as an instrument for X_t. The parameters for this final set of simulations are summarized in Table 3.

Our main results where $\gamma = 10$ appear in Figure 9. Instrumental variables reliably

Table 3: Simulation Parameters: Simultaneous Causality

Parameter	Causal Pathway	Simulation Values
β	$X \rightarrow Y$	{3}
α	$Y \rightarrow X$	{1,10}
ϕ	$\epsilon_{1t-1} \rightarrow \epsilon_{1t}$	{0, .1, .2,..., .9}
ρ	$X_{t-1} \rightarrow X_t$	{.5}
γ	$Z_t \rightarrow X_t$	{0,10}

estimate β in all simulations. As predicted, $\hat{\beta}_{NAIVE}$ and $\hat{\beta}_{LAGID}$ are consistently downward biased unless there is no simultaneity (α=0); estimates of $\beta\rho$ are biased even when $\phi = 0$. (In Appendix 5 we present the case where $\gamma = 0$, in which case lag identification only identifies $\beta\rho$ when $\phi = 0$ and there is no endogeneity anyway.) The third column in Figure 9 shows once again that lag identification leads to Type 1 error.

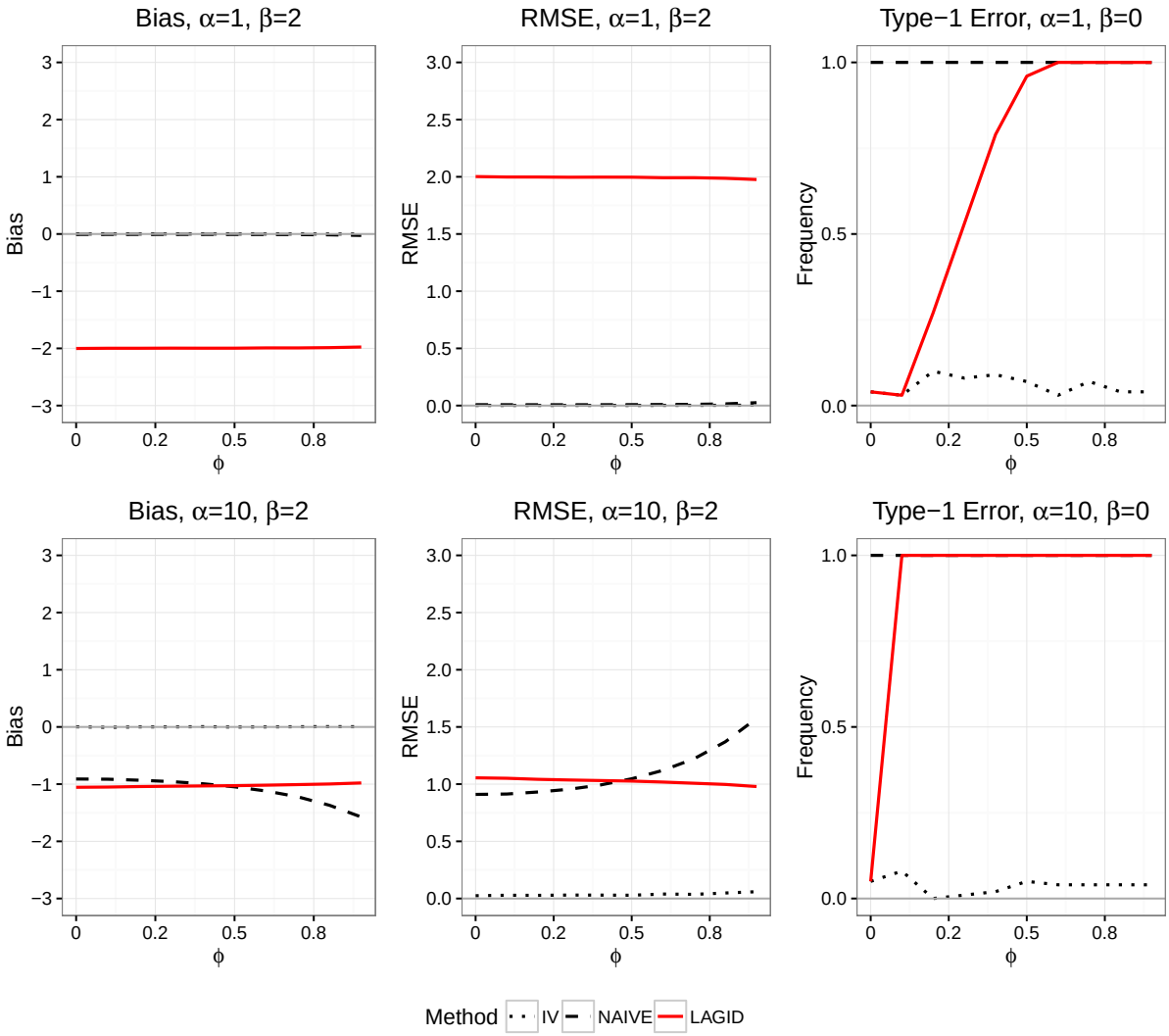
When endogeneity takes the form of simultaneous causality, then, the tradeoffs are somewhat different than those described above with the unobserved confounder. Now, when simultaneity problems are small relative to the casual effect of interest ($\alpha < \beta$), $\hat{\beta}_{NAIVE}$ consistently outperforms $\hat{\beta}_{LAGID}$ in terms of bias and efficiency (See Figure Figure 5 in comparison where the degree of bias and efficiency in $\hat{\beta}_{LAGID}$ is less than $\hat{\beta}_{NAIVE}$ depending on the persistence of the unobservable ϕ). The bias and error of the lag estimator are more comparable to those of the naive estimator when simultaneity is large.

6 Summary Recommendations

Applied researchers will find our conclusions troubling. We have shown using both directed graphs and Monte Carlo simulations that lag identification not only fails to avoid the identification problem without adding new assumptions, it will also lead to misleading inferences under the null hypothesis significance testing paradigm. Given the prevalence of lag identification in applied work, our results call for a shift in current practice.

However, our results *do not imply that lagged explanatory variables are always and every-*

Figure 9: Monte Carlo Results, Simultaneous Causality



Notes: The first two rows of this figure display the bias and RMSE for $\hat{\beta}_{IV}$ and $\hat{\beta}_{NAIVE}$ relative to the true value of β, and $\hat{\beta}_{LAGID}$ relative to the true value of $\beta\rho$, based on 100 simulations of the data generating process in Figure 8 with $\beta = 2$ and $\rho = .5$. The third row plots the frequency of Type-1 error when $\beta = 0$.

where inappropriate. In this section we provide some simple guidelines for researchers seeking to use them. There are several kinds of data generating processes in which lagged explanatory variables are appropriate:

1. In the context of unobserved confounding, in which case we have shown that the following two auxiliary assumptions are necessary:

(a) No dynamics among unobservables U.

(b) The lagged endogenous variable X is a stationary autoregressive process.

2. In the context of no unobserved confounding, in which case one of the following DGPs must be assumed:

   (a) There is no reverse causality ($Y \not\rightarrow X$) and the causal effect operates with a one period lag only ($X_{t-1} \rightarrow Y$ but $X_t \not\rightarrow Y_t$)

   (b) There is reverse causality ($Y \rightarrow X$), but reverse causality is contemporaneous only, and the causal effect of X on Y operates with a one period lag only (see Appendix 6 Figure A5 for one example).

   (c) There is reverse causality, and the causal effect of X on Y is contemporaneous, there are no dynamics in Y ($Y_{t-1} \not\rightarrow Y_t$), but there are dynamics in X ($X_{t-1} \rightarrow X_t$) (see the Appendix Figure A6 for one example).

Our focus in this paper has been on showing why Scenario 1 does not improve estimates of causal effects. But Scenarios 2(a), 2(b), and 2(c) are valid justifications for lagging explanatory variables.[13]

We can look at existing research to show these insights might affect statistical models in existing research. For example, Kelley and Simmons (2015) study "the effect of monitoring and ranking on state behavior" (62), arguing that U.S. human rights reports shame countries into criminalizing human trafficking. They model their dependent variable Y (a dummy for "whether countries criminalize human trafficking in their domestic legislation") as a function of several key explanatory variables, including whether a country is named in the U.S. annual Trafficking in Persons Report. They are explicitly concerned

[13]The causal effect of X_{t-1} is identified under Scenarios 2(a) and 2(b) only. See Appendix Figure A6 for a discussion of why Scenario 2(c) does not allow for identification of causal effects of either X_t or X_{t-1} on Y.

about reverse causality: "All explanatory and control variables are lagged to help address reverse causality and selection issues" (62), and ask "Does the United States strategically shame countries that are likely to criminalize anyway?" (63). This articulation of the inferential threat facing their analysis is illuminating: the identification problem is not that criminalizing human trafficking causes countries to be named in the Trafficking in Persons Report, which would be a case of reverse causality. Rather, it is that strategic dynamics not captured in the observables determine both criminalization and being included in the report. In this case, the unobservables are whatever propensity to criminalize human trafficking is not captured in the explanatory or control variables, but which also drives U.S. scrutiny of a country's human trafficking problem. Substantively, this may be something like activism and political pressure by D.C.-linked activists in trafficking countries. Our design based approach suggests that lagging their key explanatory variable "to address...selection issues" will only do so if they are willing to assume that there is a relationship between being named in the U.S. annual Trafficking in Persons Report across years, but not in those unobservables that drive U.S. scrutiny of a country's human trafficking problem.

For a different example, Warren (2014) tests the hypothesis that "states with high levels of media accessibility will be less likely to experience the onset of civil war" (123). The independent variable of interest is a media density index. Identification is a problem, however: "to guard against spurious results due to reverse causation, all independent variables are lagged by one year" (126). In this case, it is theoretically possible that the onset of war directly reduces the density of countries' media markets. If so, then (in addition to the assumption of selection on observables) for this lag structure to "guard against spurious results" it must be the case either that

1. war onset only affects the density of media markets in the same year, and media density only affect war onset with a one-year lag (Scenario 2(b)), or

2. war onset only affects the density of media markets in a single year, there are no

dynamics in war onset, but there are dynamics in media density (Scenario 2(c))

As these examples illustrate, our theoretical results provide scholars the ability to assess more clearly the conditions under which their choice of lagging explanatory variables will have the effects that they desire.

Any use of lagged explanatory variables requires assumptions about the data generating process. As assumptions, these are ultimately untestable. But the two groups of scenarios do differ in one respect: Scenarios 2(b) and 2(c), because they imply selection on observables, do suggest falsification tests that can be used to *rule out* each. Each requires only data that is already available to the analyst.

Those tests are as follows:

1. Under Scenario 2(b), it must be the case that there there is no contemporary correlation between X and Y. If a regression of $Y_t = b_1 X_t + b_2 X_{t-1}$ uncovers a non-zero coefficient on b_1, then the data reject Scenario 2(b) as the data generating process.

2. Under Scenario 2(c), it must be the case that there are no dynamics in Y. If a regression of $Y_t = bX_t + \lambda Y_{t-1}$ uncovers a non-zero coefficient on λ, then the data reject Scenario 2(c) as the data generating process.

When employing lagged explanatory variables in the context of endogeneity, following these guidelines will help researchers to make explicit the assumptions about the data generating process that underlies their identification strategy. Doing so, in turn, will ensure that research designs with lagged explanatory variables are credible.

7 Conclusion

We conclude by stepping back from the problem of statistical modeling and considering best practices in research design. When scholars suspect that endogeneity may bias their estimates, the solution cannot arrive at the analysis stage. It must come earlier, at

the research design stage. As Sekhon (2009) notes in a critical discussion of matching and causal inference,

> for causal inference, issues of design are of utmost importance; a lot more is needed than just an algorithm. Like other methods, matching algorithms can always be used, and they usually are, even when design issues are ignored in order to obtain a nonparametric estimate from the data. Of course, in such cases, what exactly has been estimated is unclear.

This point applies equally to our results for lagged explanatory variables. As we have stressed, lag identification supplements the assumption of "selection on observables" with the assumption of "no dynamics among unobservables." We stress that this assumption of no dynamics among unobservables could *in principle* be defensible. But we conclude that without careful arguments on substantive grounds, lagged explanatory variables should not be used for identification purposes. This may be a hard conclusion, and it may be distressing to applied researchers, but we follow Samii (2016) in emphasizing that credible causal claims require political scientists to "unlearn" many statistical practices that have long been popular.

However, we recognize that tradeoffs will inevitably be made between ignoring endogeneity threats and employing lagged explanatory variables in an attempt to circumvent them. We always recommend design-based inference, but in many applied contexts researchers face a tradeoff between two imperfect statistical fixes. Under these circumstances, when is lag identification preferable to ignoring endogeneity? The answer depends on the relative size of the endogeneity threat (greater or less than the causal effect of interest), its source (unobserved confounding or simultaneity), and in the case of unobserved confounding, persistence in the unobservables. Our discussion in this manuscript shows that even when design-based identification of causal effects is impossible, design-based *reasoning* about identification can help scholars to make informed choices about the benefits of lag identification versus alternatives.

References

[1] Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91(434): 444-55.

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

[3] Angrist, Joshua D., 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 D., and Jörn-Steffen Pischke. 2014. *Mastering "Metrics": The Path from Cause to Effect*. Princeton: Princeton University Press.

[5] Arellano, Manuel, and Stephen Bond. 1991. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *Review of Economic Studies* 58(2): 277-97.

[6] Baccini, Leonardo, and Johannes Urpelainen. 2014. "International Institutions and Domestic Politics: Can Preferential Trading Agreements Help Leaders Promote Economic Reform?" *Journal of Politics* 76(1): 195-214.

[7] Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics* 119(1): 249-75.

[8] Blackwell, Matthew, and Adam Glynn. 2014. "How to Make Causal Inferences with Time-Series Cross-Sectional Data." Working Paper, Harvard University. http://www.mattblackwell.org/files/papers/causal-tscs.pdf.

[9] Blundell, Richard, and Stephen Bond. 1998. "Initial conditions and moment restrictions in dynamic panel data models." *Journal of Econometrics* 87(1): 115-43.

[10] Cranmer, Skyler J., Douglas R. Rice, and Randolph M. Siverson. forthcoming. "What To Do About Atheoretic Lags." *Political Science Research and Methods*.

[11] Dafoe, Allan. forthcoming. "Nonparametric Identification of Causal Effects under Temporal Dependence." *Sociological Methods and Research*.

[12] De Boef, Suzanna, and Luke Keele. 2008. "Taking Time Seriously." *American Journal of Political Science* 52(1): 184-200.

[13] Dunning, Thad. 2012. *Natural Experiments in the Social Sciences: A Design-Based Approach.* New York: Cambridge University Press.

[14] Duflo, Esther, Rachel Glennerster, and Michael Kremer. 2007. "Using Randomization in Development Economics Research: A Toolkit." In Vol. 4 of *Handbook of Development Economics*, eds. T. Paul Schultz and John A. Strauss. Amsterdam: Elsevier-North Holland.

[15] Gerber, Alan S., and Donald P. Green. 2012. *Field Experiments: Design, Analysis, and Interpretation*. New York: W.W. Norton & Company.

[16] Glennerster, Rachel, and Kudzai Takavarasha. 2013. *Running Randomized Evaluations: A Practical Guide*. Princeton, NJ: Princeton University Press.

[17] Haavelmo, Trygve. 1943. "The Statistical Implications of a System of Simultaneous Equations." *Econometrica* 11(1): 1-12.

[18] Harrison, Glenn W., and John A. List. 2004. "Field Experiments." *Journal of Economic Literature* 42(4): 1009-55.

[19] Imbens, Guido W. 2014. "Instrumental Variables: An Econometrician's Perspective." Working Paper, National Bureau of Economic Research. http://ftp.iza.org/dp8048.pdf.

[20] Imbens, Guido W., and Thomas Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics* 142(2): 615-35.

[21] Lehoucq, Fabrice, and Anibal Pérez-Liñán. 2014. " Breaking Out of the Coup Trap: Political Competition and Military Coups in Latin America." *Comparative Political Studies* 47(8): 1105-29.

[22] Keele, Luke. 2015. "The Statistics of Causal Inference: A View from Political Methodology." *Political Analysis* 23(3): 313-335.

[23] Kelley, Judith, and Beth A. Simmons. 2015. "Politics by Number: Indicators as Social Pressure in Inter national Relations." *American Journal of Political Science* 59(1): 55-70.

[24] Pearl, Judea. 2009. *Causality: Models, Reasoning, and Inference*. 2nd ed. Cambridge: Cambridge University Press.

[25] Pearl, Judea. 2015. "Trygve Haavelmo and the Emergence of Causal Calculus," *Econometric Theory* 31(1): 152-179.

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

[27] Robins, James M., Miguel Ángel Hernán, and Babette Brumback. 2000. "Marginal Structural Models and Causal Inference in Epidemiology," *Epidemiology* 11(5): 550-560.

[28] Samii, Cyrus. 2016. "Causal Empiricism in Quantitative Research." *Journal of Politics* 78(3): 941-955

[29] Sekhon, Jasjeet S. 2009. "Opiates for the Matches: Matching Methods for Causal Inference." *Annual Review of Political Science* 12: 487-508.

[30] Steinberg, David A., and Krishna Malhotra. 2014. "The Effect of Authoritarian Regime Type on Exchange Rate Policy." *World Politics* 66(3): 491-529.

[31] Sovey, Allison J., and Donald P. Green. 2011. "Instrumental Variables Estimation in Political Science: A Readers' Guide." *American Journal of Political Science* 55(1): 188-200.

[32] Teele, Dawn Langan. 2014. *Field Experiments and Their Critics: Essays on the Uses and Abuses of Experimentation in the Social Sciences*. New Haven: Yale University Press.

[33] Warren, T. Camber. 2014. "Not by the Sword Alone: Soft Power, Mass Media, and the Production of State Sovereignty." *International Organization* 68(1): 111-41.

Appendix 1 Complete Derivation of $\hat{b}_{X_{t-1}}$

In this appendix we show how to derive equation (9) in the main text.

Equation (8) reduces to

$$\begin{aligned}
\text{plim}_{n\to\infty}\,\hat{b}_{X_{t-1}} =& \beta\rho + \frac{(\beta\kappa+\delta)\text{Cov}(X_{t-1},U_t)}{\text{V}(X_{t-1})} \qquad (15)\\
=& \beta\rho + \frac{(\beta\kappa+\delta)\text{Cov}(X_{t-1},\phi U_{t-1}+\nu_t)}{\text{V}(X_{t-1})}\\
=& \beta\rho + \frac{\phi(\beta\kappa+\delta)\text{Cov}(X_{t-1},U_{t-1})}{\text{V}(X_{t-1})} + \frac{\phi(\beta\kappa+\delta)\text{Cov}(X_{t-1},\nu_t)}{\text{V}(X_{t-1})}\\
=& \beta\rho + \frac{\phi(\beta\kappa+\delta)\text{Cov}(X_{t-1},U_{t-1})}{\text{V}(X_{t-1})} \qquad (16)
\end{aligned}$$

We found this derivation by replacing U_t with $\phi U_{t-1} + \nu_t$ in (15). However, we can also rearrange (2) to express X_{t-1} in terms of contemporary values X_t and U_t. Doing so, and then substituting the resulting expression in for X_{t-1} in (15), produces

$$\begin{aligned}
\text{plim}_{n\to\infty}\,\hat{b}_{X_{t-1}} =& \beta\rho + \frac{\phi(\beta\kappa+\delta)\text{Cov}(\frac{1}{\rho}X_t - \frac{\kappa}{\rho}U_t - \frac{1}{\rho}\eta_t, U_t)}{\text{V}(X_{t-1})}\\
=& \beta\rho + (\beta\kappa+\delta)\Big[\frac{\frac{1}{\rho}\text{Cov}(X_t,U_t)}{\text{V}(X_{t-1})} + \frac{\frac{\kappa}{\rho}\text{Cov}(U_t,U_t)}{\text{V}(X_{t-1})} + \frac{\frac{1}{\rho}\text{Cov}(\eta_t,U_t)}{\text{V}(X_{t-1})}\Big]\\
=& \beta\rho + \frac{(\beta\kappa+\delta)\big[\text{Cov}(X_t,U_t) - \kappa\text{V}(U_t)\big]}{\rho\text{V}(X_{t-1})} \qquad (17)
\end{aligned}$$

Setting the two expressions for $\text{plim}_{n\to\infty}\,\hat{b}_{X_{t-1}}$ in (16) and (17) equal to each other produces

$$\text{Cov}(X_t,U_t) - \kappa\text{V}(U_t) = \phi\rho\text{Cov}(X_{t-1},U_{t-1}) \qquad (18)$$

By assumption, $0 \leq \rho < 1$ and $0 \leq \phi < 1$. This means that both X and U are mean-reverting series, which in turn entails that the covariance between X and U does not depend on the period, or $\text{Cov}(X_t,U_t) = \text{Cov}(X_{t-1},U_{t-1}) = \text{Cov}(X,U)$. We can therefore solve

for $\mathrm{Cov}(X, U)$:

$$\begin{aligned}
&\mathrm{Cov}(X, U) - \kappa \mathrm{V}(U) = \phi\rho \mathrm{Cov}(X, U) \\
&\mathrm{Cov}(X, U) = \frac{\kappa \mathrm{V}(U)}{1 - \phi\rho} \qquad (19)
\end{aligned}$$

We can now simplify (16) further by replacing $\mathrm{Cov}(X_{t-1}, U_{t-1})$ with $\kappa V(U)/(1-\phi\rho)$, which yields

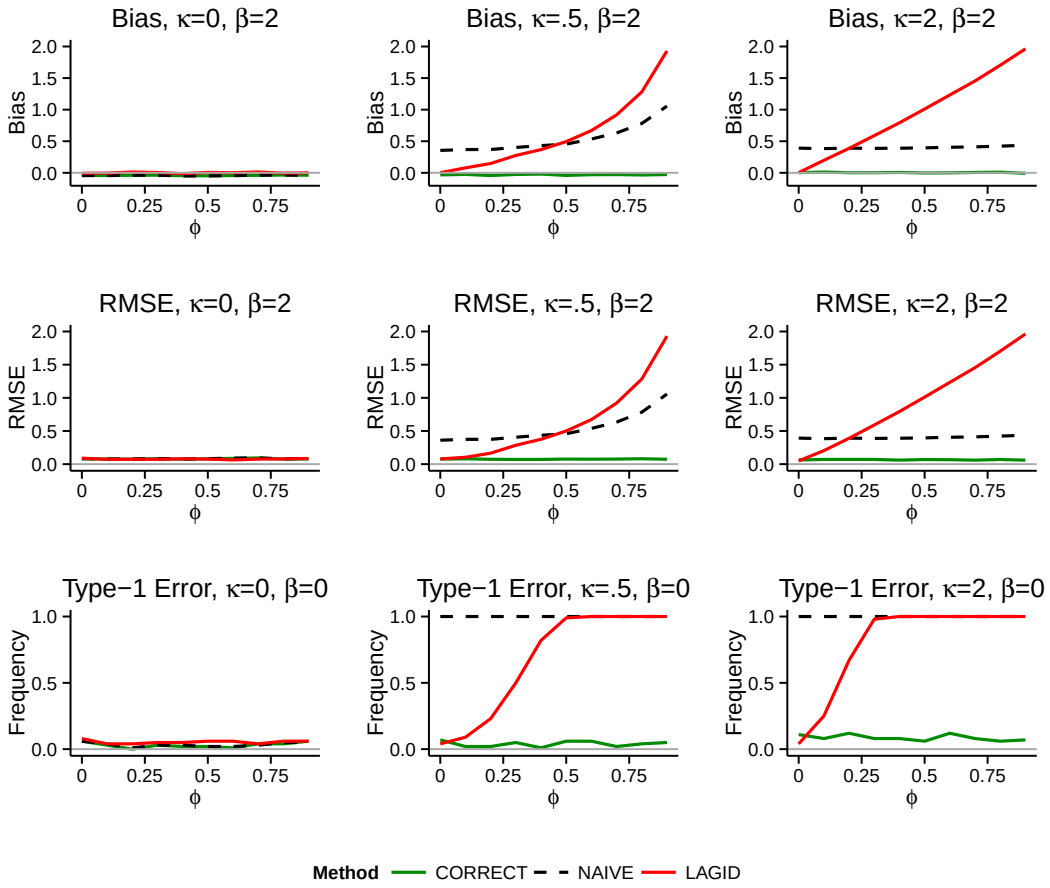
$$\mathrm{plim}_{n\to\infty}\, \hat{b}_{X_{t-1}} = \beta\rho + \frac{\phi\kappa(\beta\kappa + \delta)\mathrm{V}(U)}{(1 - \phi\rho)\mathrm{V}(X)} \qquad (20)$$

This result is (9) in the main text.

Appendix 2 Additional Results, Baseline Simulations

In this appendix we provide additional simulation results for alternative values of ρ. We begin first with the case of $\rho = .1$. The bias arising from lagged explanatory variables

Figure A1: Monte Carlo Results, $\rho = .1$

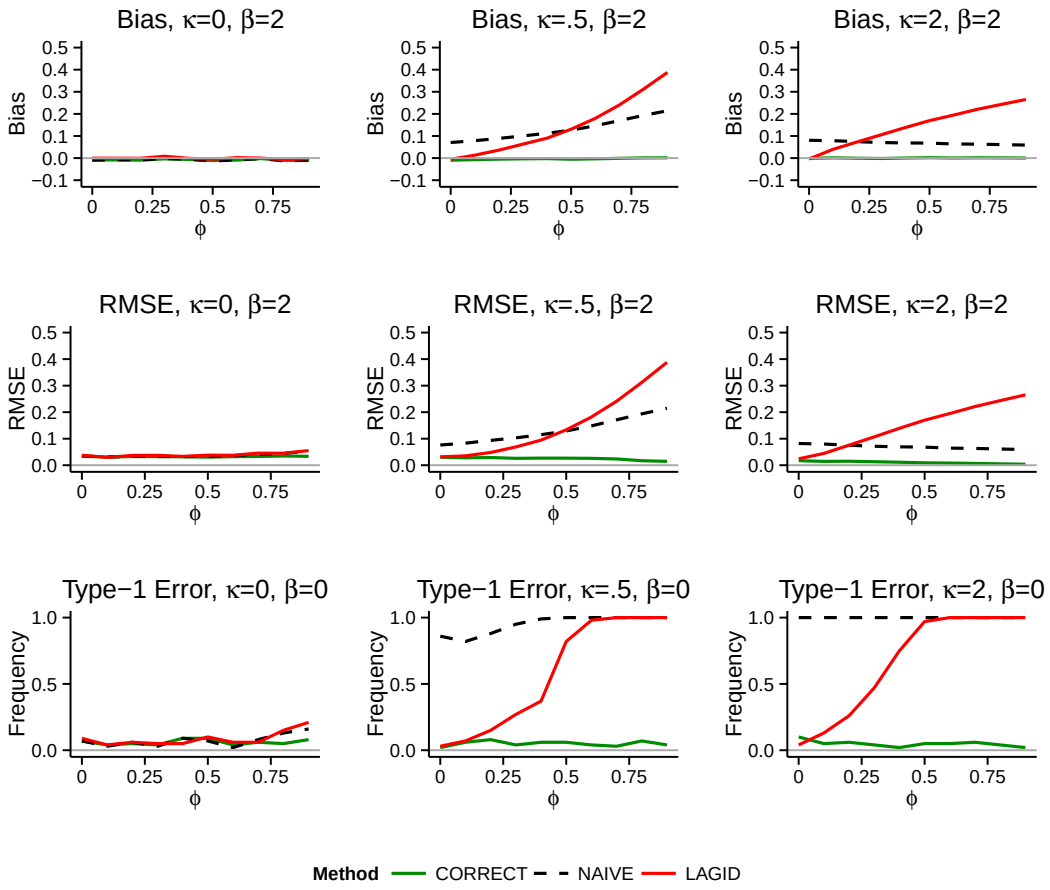


Notes: The first two rows of this figure display the bias and RMSE for $\hat{\beta}_{CORRECT}$ and $\hat{\beta}_{NAIVE}$ relative to the true value of β, and bias and RMSE $\hat{\beta}_{LAGID}$ relative to $\beta\rho$, based on 100 simulations of the data generating process in Figure 4 with $\beta = 2$ and $\rho = .1$. The third row plots the frequency of Type-1 error when $\beta = 0$ and $\rho = .1$.

is even more pronounced in these simulations.

Now consider the case where $\rho = .9$. As before, we continue to find that the bias of the lagged explanatory variables estimates is still significant, and still greater than the naive estimator that ignores endogeneity altogether for sufficiently high values of ϕ. Type-1 error

Figure A2: Monte Carlo Results, $\rho = .9$



Notes: The first two rows of this figure display the bias and RMSE for $\hat{\beta}_{CORRECT}$ and $\hat{\beta}_{NAIVE}$ relative to the true value of β, and bias and RMSE $\hat{\beta}_{LAGID}$ relative to $\beta\rho$, based on 100 simulations of the data generating process in Figure 4 with $\beta = 2$ and $\rho = .9$. The third row plots the frequency of Type-1 error when $\beta = 0$ and $\rho = .9$.

is still a major issue. Together with the results in the main text, these results demonstrate that lagged explanatory variables avoid endogeneity problems only in the absence of dynamics among the unobservables, or when endogeneity is not actually a problem anyway.

Appendix 3 DGP and Simulation Parameters, X_{t-1} as the Causal Variable

$$Y_{it} = \xi X_{it-1} + \delta U_{it} + \epsilon_{it} \tag{21}$$

$$X_{it} = \rho X_{it-1} + \kappa U_{it} + \eta_{it} \tag{22}$$

$$U_{it} = \phi U_{it-1} + \nu_{it} \tag{23}$$

where we assume $0 \leq \rho < 1$, $0 \leq \phi < 1$, $\epsilon_{it} \sim N(0, 5)$, $\eta_{it} \sim N(0, 1)$, and $\nu_{it} \sim N(0, 1)$. We vary the parameters of ξ, δ, ρ, κ, and ϕ according to Table A-1.

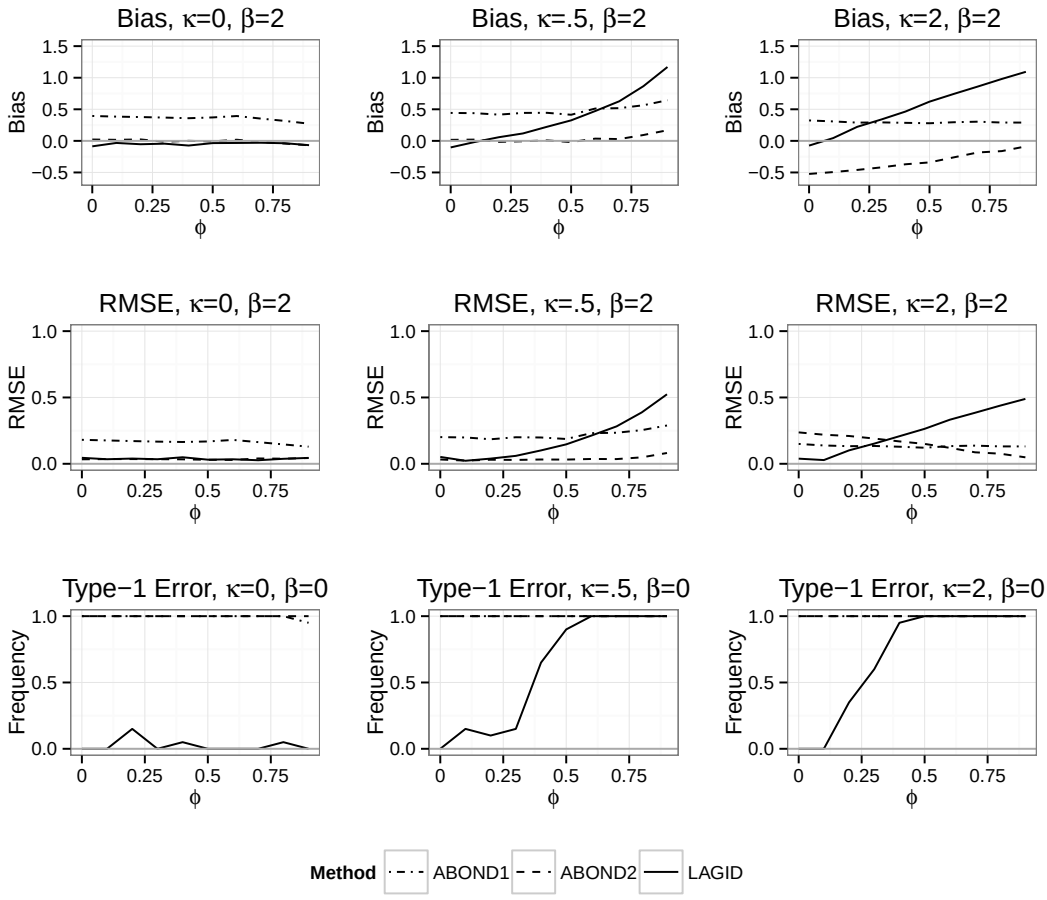
Table A-1: Simulation Parameters, X_{t-1} as the Causal Variable

Parameter	Causal Pathway	Simulation Values
ξ	$X_{t-1} \rightarrow Y_t$	{0, 2}
κ	$U_t \rightarrow X_t$, $U_{t-1} \rightarrow X_{t-1}$	{.5, 2}
ϕ	$U_{t-1} \rightarrow U_t$	{0, .1, .2,..., .9}
δ	$U_t \rightarrow Y_t$	{1}
ρ	$X_{t-1} \rightarrow X_t$	{.5}

Appendix 4 Fixed Effects and GMM Simulations

We estimate these models using the `pgmm` estimator implemented in the `plm` library in R. We use the second and third lags of X and Y as instruments, and estimate models in which we condition on X_t ($\hat{\beta}_{ABOND1}$) as well as models in which we condition on X_{t-1} ($\hat{\beta}_{ABOND2}$). We compare these results to a lag identification model with fixed effects, implemented using the `"within"` estimator in `plm`. Note that when $k = 2$, $\hat{\beta}_{ABOND2}$ approaches the true

Figure A3: Monte Carlo Results, Fixed Effects and GMM



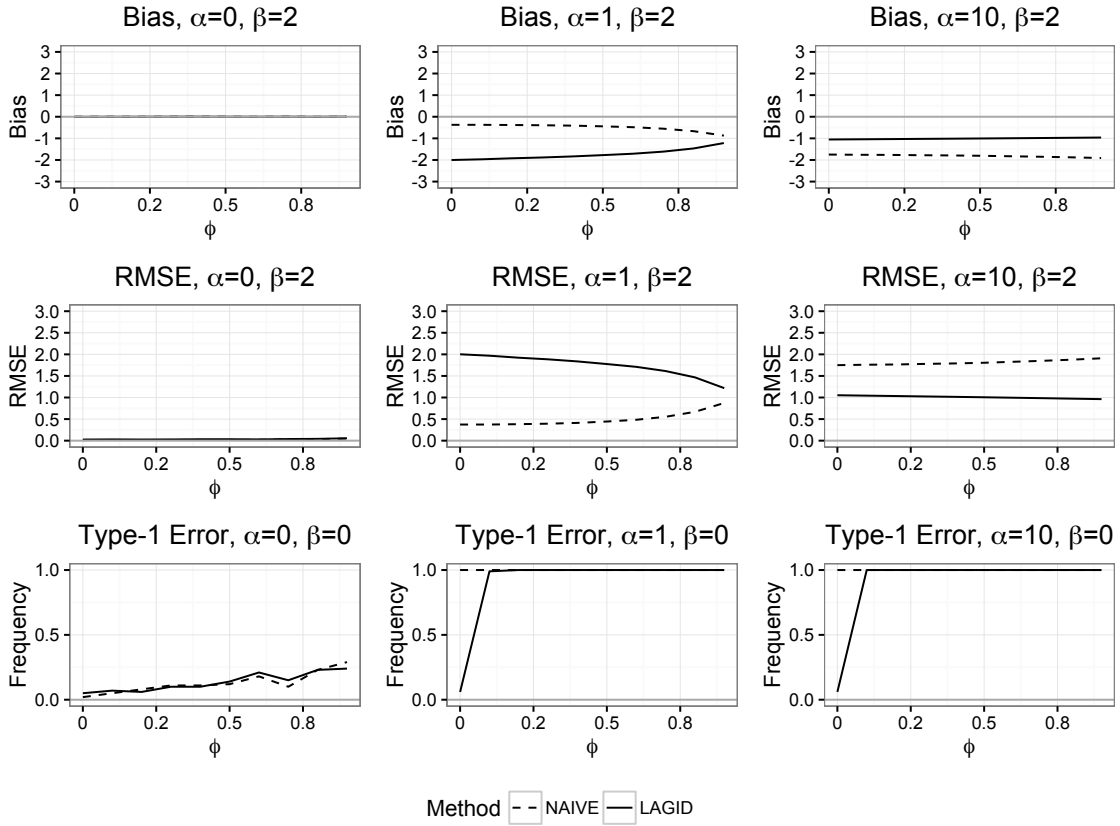
Notes: The first two rows of this figure display the bias and RMSE for $\hat{\beta}_{ABOND1}$ relative to the true value of β, and $\hat{\beta}_{ABOND2}$, and $\hat{\beta}_{LAGID}$ relative to the true value of $\beta\rho$ based on 100 simulations with $\beta = 2$ and $\rho = .5$. The third row plots the frequency of Type-1 error when $\beta = 0$ and $\rho = .5$.

value of $\beta\rho$ as ϕ grows larger. This is an incidental result stemming from the fact that $\beta = \kappa$ in this model, and it does not appear for larger values of κ.

Appendix 5 Simultaneous Causation when $\gamma = 0$

When $\gamma = 0$, there is no instrumental variables strategy available to estimate β. Lag identification only performs well in this case when $\phi = 0$ and there is no reverse causality. In other cases, whether or not lag identification performs better than the naive estimator that ignores endogeneity altogether depends on the relative size of α and β.

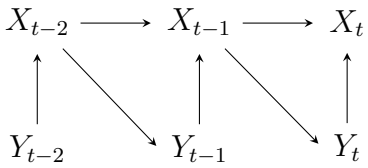
Figure A4: Monte Carlo Results, Simultaneous Causality



Notes: The first two rows of this figure display the bias and RMSE for $\hat{\beta}_{NAIVE}$ relative to the true value of β, and $\hat{\beta}_{LAGID}$ relative to the true value of $\beta\rho$, based on 100 simulations of the data generating process in Figure 8 with $\beta = 2$ and $\rho = .5$. The third row plots the frequency of Type-1 error when $\beta = 0$.

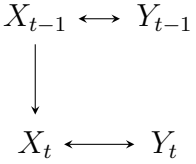
Appendix 6 Valid Data Generating Processes

Figure A5: Scenario 2(b)



Notes: This is a representation of a data generating process in which reverse causality exists ($Y_t \rightarrow X_t$) but the causal effect of X_{t-1} on Y is identified because $X_t \not\rightarrow Y_t$.

Figure A6: Scenario 2(c)



Notes: This is a representation of a data generating process in which reverse causality exists ($Y_t \rightarrow X_t$) but a "conservative" effect of X_{t-1} on Y is identified because $Y_{t-1} \not\rightarrow Y_t$. To be clear, however, *neither* the direct effect of X_t on Y_t *nor* the total effect of X_{t-1} on Y_t is identified because of the simultaneous relationship between X and Y.