PA

Not so Harmless After All: The Fixed-Effects Model

Thomas Plümper[1] and Vera E. Troeger[2]

[1] Vienna University of Economics, Department of Socioeconomics, Welthandelsplatz 1, 1020 Vienna, Austria.
Email: thomas.pluemper@wu.ac.at
[2] University of Warwick, Department of Economics and CAGE, Coventry CV4 7AL, UK. Email: v.e.troeger@warwick.ac.uk

Abstract

The fixed-effects estimator is biased in the presence of dynamic misspecification and omitted within variation correlated with one of the regressors. We argue and demonstrate that fixed-effects estimates can amplify the bias from dynamic misspecification and that with omitted time-invariant variables *and* dynamic misspecifications, the fixed-effects estimator can be more biased than the 'naïve' OLS model. We also demonstrate that the Hausman test does not reliably identify the least biased estimator when time-invariant and time-varying omitted variables or dynamic misspecifications exist. Accordingly, empirical researchers are ill-advised to rely on the Hausman test for model selection or use the fixed-effects model as default unless they can convincingly justify the assumption of correctly specified dynamics. Our findings caution applied researchers to not overlook the potential drawbacks of relying on the fixed-effects estimator as a default. The results presented here also call upon methodologists to study the properties of estimators in the presence of multiple model misspecifications. Our results suggest that scholars ought to devote much more attention to modeling dynamics appropriately instead of relying on a default solution before they control for potentially omitted variables with constant effects using a fixed-effects specification.

Keywords: consistency, efficiency, misspecification, omitted variable bias, panel data, Monte Carlo simulation

1 Introduction

The reputation of the fixed-effects estimator[1] is better than its finite sample properties. Among the panel and pooled analysis textbooks that we are aware of, Wooldridge has perhaps the most precise description of the conditions under which fixed-effects models are unbiased: "Under a strict exogeneity assumption on the explanatory variables, the fixed-effects estimator is unbiased." (Wooldridge 2002, 442) This in turn means that if the variables included in the model are correlated with a model misspecification other than omitted variables with constant effects, the fixed-effects model is not unbiased. For example, recent research has argued that in the presence of dynamic misspecification, fixed-effects estimates are biased and inconsistent (Harris *et al.* 2009; Lee 2012; Ahn, Lee, and Schmidt 2013; see also Nickell 1981).[2]

Authors' note: We thank Jonathan Kropko and the participants of the workshop "Modeling Politics & Policy in Time and Space" organized by Guy Whitten and Scott Cook at Texas A&M for helpful comments and input.
The replication files for the MC analysis can be found on the PA dataverse: Troeger and Pluemper (2017), "Replication Data for: Not so Harmless After All: The Fixed-Effects Model", doi:10.7910/DVN/RAUIHG, Harvard Dataverse.

1 The good reputation the fixed-effects model enjoys among econometricians and increasingly among applied researchers, is perhaps best summarized with the following claim: "With panel data, always model the fixed effects using dummy variables (. . .). Do not estimate random-effects models without ensuring that the estimator is consistent with respect to the fixed-effects estimator (using a Hausman test)" (Antonakis *et al.* 2010, 1113). This quote demonstrates a common misperception of the Hausman test (Hausman (1978), see also Ahn and Low 1996; Frondel and Vance 2010). The Hausman test does not test the consistency of the random-effects model, it tests whether the random-effects model generates estimates that significantly differ from the fixed-effects model. This would be an indirect test of the random-effects model's consistency if and only if omitted time-invariant variables were the only reason that could produce such a significant difference in estimates. We later demonstrate that with more than one model misspecification the Hausman test does not reliably identify the less biased estimator.

2 Bell and Jones (2015) discuss the possibility of different effect strengths for level and changes, which can also be interpreted as dynamic misspecification (Bell and Jones 2015). For a discussion of fixed versus random effects see also Clark and Linzer (2015). Note that for our specification of the data-generating process in the Monte Carlo analyses, random effects and pooled-OLS give identical point estimates and very similar standard errors. We therefore do not report random-effects results but everything we say about pooled-OLS also applies to random effects.

PA

We take this literature one step further and demonstrate that fixed-effects estimates amplify the bias from dynamic misspecification[3] relative to pooled-OLS estimates. This finding has two implications: In the absence of omitted time-invariant variables and the presence of dynamic misspecification, the pooled-OLS model is strictly less biased than the fixed-effects model. And second, in the simultaneous presence of omitted variables with both constant and time-varying effects, the fixed-effects model is *more* biased than the pooled-OLS model (and the random-effects model) if the correlation between the variable of interest and the omitted time-invariant variance is smaller than the correlation between the variable of interest and the omitted time-varying variance. Accordingly, relative to the naïve pooled-OLS benchmark, the fixed-effects model solves the problem of omitted variables with constant effects at the expense of rendering other problems worse.[4]

We use the pooled-OLS estimator as a benchmark for the fixed-effects model in the following sense: The properties of the pooled-OLS estimator in the presence of omitted time-invariant variables, omitted time-varying variables, and dynamic misspecifications are known to be poor. Pooled-OLS gives biased estimates in the presence of omitted time-invariant variables, omitted time-varying variables and other dynamic misspecifications.

By demonstrating that the fixed-effects model often performs worse than the pooled-OLS estimator when dynamic misspecifications exist, we try to alert applied researchers about the importance of choosing the correct dynamic specification when relying on fixed-effects estimates. Of course, we do not argue that ignoring dynamics and using the pooled-OLS model is an appropriate alternative. Optimally, scholars would use the correct dynamic specification for their model. However, in many applications, the chances of getting the dynamic specification right remain slim (Adolph, Butler, and Wilson 2005; Plümper, Troeger, and Manow 2005; Wilson and Butler 2007; DeBoef and Keele 2008). Our findings also suggest that a difference in pooled-OLS and fixed-effects estimates cannot with certainty be attributed to time constant unit heterogeneity.[5] It may equally be caused by, inter alia, omitted time-varying variables, wrong assumptions about the functional forms of the treatment effect, and misspecified lag structures.

The article studies the consequences of dynamic misspecification that occur in static models or when applied researchers use simple econometric patches instead of a correct dynamic specification. Our formal and simulation analyses support previous arguments that fixed-effects estimates *are biased* if the model suffers from excluded time-varying variables and if trends and dynamics are not correctly modeled. We also demonstrate that the widely shared assumption, the fixed-effects model is superior to our naïve benchmark, pooled-OLS, does not necessarily hold in the presence of dynamic misspecifications. We provide evidence that under identifiable and plausible conditions the fixed-effects estimator may actually exacerbate the bias in comparison to a naïve estimator even in presence of omitted time-invariant variables, because dropping the between variation increases the influence of dynamic misspecifications on parameter estimates.

We examine the overall logic of dynamic misspecifications based on three simple examples: the existence of omitted time-invariant and time-varying variance (experiment 1), trends in both an

---

3 Note that we solely discuss omitting important dynamics such as trends or time-varying variables or lags of RHS variables. We do not analyze the effect of including unnecessary dynamics directly. However our MC analysis include an element of including dynamic components that are not necessarily in the DGP. For example, many fixes that are used to control for serial correlation in the error term are not part of the DGP. Including a Lagged Dependent Variable (LDV) or time fixed effects into the RHS of the model is an example for this. These dynamic specifications may generate additional bias because these elements can pick up variation that should be attributed to other elements in the DGP.

4 These statements are based on the assumption that within and between effects are the same. If this is not the case it depends whether the researcher is interested in between or within or average effects across time and space. We discuss this issue in more detail later on.

5 Some authors (i.e. Gamm and Kousser 2010) demonstrate that their estimates are robust to a change from fixed-effects estimates to pooled-OLS. In the light of our results, we believe this is a useful research strategy.

omitted variable and the variable of interest (experiment 2), and a misspecified lag structure of the explanatory variable of interest (experiment 3). Our results confirm that the fixed-effects estimator is biased in the presence of omitted variables which either vary over time or exert a time-lagged effect on the outcome—a result known from theoretical work (Lee 2012; Ahn, Lee, and Schmidt 2013). We go one step further and demonstrate that using fixed effects in the presence of time-varying and time-invariant omitted variables can under plausible assumptions increase the bias relative to a naïve estimation with pooled-OLS or the random-effects estimator. Our results also invalidate the common interpretation of the Hausman test, namely that if the fixed and random-effects (or pooled-OLS) estimates significantly differ, then researchers should use the (consistent) fixed-effects model (Hausman 1978; Baltagi 2001, 65–70). This interpretation of the Hausman test assumes the absence of any other model misspecification that influences fixed effects and pooled-OLS estimates differently.[6] The results of our analyses bring a common problem to attention: econometric solutions to a single specification issue can impede the accuracy of estimates even though the econometric patch solves the problem it has been invented for. For example, the fixed-effects estimator has been developed to eliminate bias from 'unobserved heterogeneity'[7] due to constant unit-specific effects, but by doing so it can amplify the bias resulting from un-modeled dynamics.[8] Our findings stress the importance of developing model specifications for multiple simultaneous model misspecifications. Biases generated by different model misspecifications are often not additive, which implies that solving one problem can exacerbate the bias emanating from another misspecification.[9]

2 The Sources of and Potential for Dynamic Misspecification

Applied researchers often perceive serially correlated errors as noise rather than information (DeBoef and Keele 2008). Yet, serially correlated errors clearly indicate a potentially severe model misspecification, which can result from various sources (Neumayer and Plümper 2017). Perhaps most obviously, serially correlated errors are caused by incompletely or incorrectly modeled persistency in the dependent variable, time-varying omitted variables or changes in the effect strengths of time-invariant variables, or misspecified lagged effects of explanatory variables. Conditionality makes modeling dynamics more complicated (Franzese 2003a,b; Franzese and Kam 2009). Few empirical analyses model all potential conditioning factors of the variables of interest. If, however, treatment effects are conditioned by unobserved time-varying factors—as for example the effect of higher education on income is conditioned by structural change of the economy and ruptures in economic policies—then treatment effects vary over time, and the strength of these effects also changes over time as un-modeled conditioning factors change. Finally, serially correlated errors may result from misspecification that at first sight have little to do with dynamics, for example from spatial dependence. Yet, spatial effects are certainly misunderstood if they are perceived as time-invariant, ignoring spatial dependence causes errors to be serially correlated (Franzese and Hays 2007). Virtually all of these complications depend on

---

6 The poor performance of the Hausman test for different misspecifications including serial correlation, non-stationarity, and heteroscedasticity is known (Arellano 1993; Ahn and Low 1996; Bole and Rebec 2013).

7 It also does not help that econometric textbooks usually do not define the term 'unobserved heterogeneity', tend to be imprecise about the conditions under which the fixed-effects estimator is consistent, and hardly ever discuss the conditions under which the fixed-effects model generates biased and inconsistent estimates—at least not in a way that non-econometricians understand easily (Hendry 1995; Baltagi 2001; Wooldridge 2002; Hsiao 2014). Interestingly, identification textbooks discuss the FE model's properties in greater detail, see Angrist and Pischke (2009) and the excellent discussion in Morgan and Winship (2007).

8 Variables that are usually treated as time-invariant including culture (Kayser and Satyanath 2014), distance (Wegener 1912), institutions (North 1990), genetic markers (Hedrick 2005), tend to vary at least slowly over time. The only truly time-invariant variable is 'inheritance'. However, still in this case the effects of inherited factors are not likely to be constant over time. Park (2012) develops a procedure that allows testing the assumption that unobserved heterogeneity is indeed time-invariant.

9 Pickup (2017) suggests a general-to-specific approach to dynamics for 'short panels' and argues that researchers should first find a plausible dynamic specification before dealing with unobserved heterogeneity.

an arbitrary decision that no researcher can avoid: the periodization of continuous time that is a necessary condition if researchers wishing to study 'periods'. If researchers choose relatively short periods, effects do no longer necessarily occur in the same period as the treatment. If periods cover a long stretch of time, the probability that estimates are biased by confounders rises quickly. In the social sciences, the lengths of a period is rarely chosen to optimize the analysis. Instead, social scientists often have to accept data that is collected on a daily, monthly or—most often—annual basis.

At least in an optimal world these model misspecifications should be avoided: dynamics should be directly modeled to obtain unbiased estimates. This proves to be difficult. Since dynamic misspecifications are manifold and complex, econometric tests for 'dynamics' at best reveal serially correlated errors, but they are usually unable to identify the underlying root causes of autocorrelation. Often, these tests are also weak and do not reveal the true dynamic structure of the data-generating process, which may lead to overfitting of the data (Keele, Linn, and Webb 2016). Thus, empirical researchers are probably best advised to simplify their empirical model and to treat problems such as serially correlated errors with straightforward econometric patches such as lagged dependent variables, period dummies, and simple homogeneous lag structures.

Yet, using misspecification patches should not mislead researchers into believing that the dynamics of their model are correctly specified. Econometric fixes are not correct per se because they are usually not modeling the true dynamic process in the underlying data-generating process. For example, periods do not exert a direct effect on the dependent variable but period dummies capture variation over time which can help to "clean" residuals. Perhaps even worse, more than one econometric specification allows eliminating serial correlation, and there is no guarantee that different models lead to identical or at least sufficiently similar results. Empirical researchers should also not expect that so called 'dynamic econometric models', e.g. the Arellano–Bond (A–B) estimator (Arellano and Bond 1991), solve various problems of dynamics. Dynamic panel models only eliminate Nickell-bias. Period dummies control for common trends, common shocks, and common breaks, but they do not perfectly account for unit-specific, heterogeneous trends, shocks, and breaks. Including a lagged dependent variable to the right hand side of the estimation equation without including lags of the explanatory variables (x) assumes that the dynamics of all independent variables are identical. These assumptions are convenient, but not always plausible.

Still, the vast majority of panel data analyses pushes serially correlated errors into uninformative econometric patches: lagged dependent variables and period fixed effects appear to be the most common solutions, but they are not the only ones. More often than not analysts seem to "adopt restrictive dynamic specifications on the basis of limited theoretical guidance and without empirical evidence that restrictions are valid, potentially biasing inferences and invalidating hypothesis tests" (DeBoef and Keele 2008, 184).[10] A review of recent political science publications reveals that a large majority of panel data analyses rely on of the following four strategies: do nothing and ignore the potential for dynamics (Humphreys and Weinstein 2006; Ross 2008), assume that all dynamics are captured by period fixed effects (Egorov, Guriev, and Sonin 2009; Besley and Reynal-Querol 2011; Menaldo 2012 among many others), try to capture dynamics by a lagged dependent variable (e.g. Acemoglu *et al.* 2008; Guisinger and Singer 2010; Lupu and Pontusson 2011; Kogan, Lavertu, and Peskowitz 2016), or finally follow Beck and Katz (1995) and model dynamics by a combination of period fixed effects and a lagged dependent variable (e.g. Lipsmeyer and Zhu 2011; Getmansky and Zeitzoff 2014). Significantly fewer authors

---

10 "Substantive theory, then, typically does not provide enough guidance for precise dynamic specifications" (DeBoef and Keele 2008, 196).

rely on GLS estimators (Mukherjee, Smith, and Li 2009; Lupu and Pontusson 2011), distributed lag models (e.g. Gerber *et al.* 2011) or error correction models (Lebo, McGlynn, and Koger 2007; Kayser 2009; Soroka, Stecula, and Wlezien 2015).[11] Overall, the vast majority of panel data analyses in political science assumes rather simple dynamics.[12] This finding is consistent with DeBoef and Keele (2008, 185), who also conclude that the vast majority of authors do not test for the underlying dynamic structure. Thus, social scientists often model the dynamic aspects with very little theoretical guidance (Keele and Kelly 2006; DeBoef and Keele 2008),[13] use ad hoc econometric solutions, which make rather rigid assumptions, do not try to model the true data-generating process, and do not report results of minimal tests for serial correlation.

One strategy that may reduce the size of the problem is to use less constrained econometric solutions. Distributed lag models, models with a unit-specific lagged dependent variable, panel co-integration models, models with heterogeneous lag structure (Plümper, Troeger, and Manow 2005), more attention to periodization (Franzese 2003a), better specified spatial models (Franzese and Hays 2007; Neumayer and Plümper 2016) may all reduce the size of the problem. However, as the number of possible dynamic specifications increases, a higher order problem of model selection arises: since all these different models likely generate different estimates and often demand different inferences, the question becomes how empirical researchers select their preferred model. To eliminate or at least reduce the arbitrariness of model selection, DeBoef and Keele (2008, 187) suggest a testing-down approach, starting with a full autoregressive distributive lag model and stepwise removing parameters according to pre-determined criteria, often the significance of parameters. This procedure will result in a dynamic specification that maximizes the variance absorbed by the minimum number of parameters. As with all testing-down approaches, this approach suffers from the arbitrariness in the choice of a start model because we do not have an infinite number of degrees of freedom. DeBoef and Keele recommend starting with a general autoregressive distributed lag (ADL) model. They argue that this model has 'no constraints'. Yet, the model still assumes a homogeneous lag structure and it will run quickly out of degrees of freedom if the number of controls is large because a finite number of distributed lag parameters has to be estimated for each regressor. Accordingly, these models only work if the number of periods is much larger than the number of variables—a criterion that is not necessarily met in panel data analyses. Since the specification includes a lagged dependent variable, the estimator is inconsistent when unit fixed effects are included, though the bias declines if the number of periods increases (Nickell 1981; Kiviet 1995). Gerber *et al.* (2011) thus prefer an alternative strategy. Rather than relying on a single 'best' dynamic specification, they report the result of various different dynamic specification and they demonstrate that their results are robust "for varying lag lengths and polynomial orders" (Gerber *et al.* 2011, 143). Relying on robustness tests has at least two advantages: first, it largely reduces the necessity to make arbitrary dynamic modeling assumptions, and, second, it helps identifying possible relevant model uncertainties (Neumayer and Plümper 2017).

For our purposes and in the remainder of this article, the problem is not so much which techniques minimizes the potential for dynamic misspecification. Instead, we assume that

---

11 When political scientists employ distributed lag or error correction models, they often do not include unit dummies, and when they use fixed effects, they rarely control for complex dynamics. Exceptions exist, e.g. Haber and Menaldo (2011) and Treisman (2015) combine unit and period fixed effects with an error correction model.

12 We do not wish to suggest here that error correction models and distributed lag models allow social scientists to model dynamics correctly. These models do assume homogeneous dynamic processes which neither capture omitted time-varying variables nor unobserved time-varying conditionality of the variable of interest, they remain limited in their ability to capture functional forms of effects which do not simply diminish at a constant rate, and they de facto rely on homogeneous lag structures.

13 The absence of theoretical guidance may be caused by theories which "typically tell us only generally how inputs relate to processes we care about. They are nearly always silent on which lags matter, (. . .), what characterizes equilibrium behavior, or what effects are likely to be biggest in the long run" (DeBoef and Keele 2008, 186).

dynamic misspecifications exist and analyze the performance of the fixed-effects model in the presence of various dynamic misspecifications—some of which could be dealt with easily, others are more difficult though not impossible to eliminate if only researchers knew the true data-generating process. But of course the whole point of estimation is that researchers do not know the true data-generating process and that theory, econometric tests, and testing-down procedures cannot identify the optimal model beyond reasonable doubt. Having said this, we do not claim that social scientists inevitably misspecify dynamics, but we emphasize that in the presence of dynamic misspecifications, the fixed-effects model has problematic properties. Needless to say that modeling dynamics correctly is always preferable.

3 The Bias of the Fixed-Effects Estimator with Dynamic Misspecification

This section analyzes how dynamic misspecifications cause fixed-effects estimates to be biased and we demonstrate that the bias of the fixed-effects estimator can exceed the bias of the naïve pooled-OLS estimator under plausible assumptions. We are not the first to do so. Lee (2012) analytically demonstrates that the fixed-effects estimator is biased when the lag order is not correctly chosen and stresses that "existing bias corrections would not work properly because the correction formulas assume correct model specification. In fact, attempts to adjust for the bias using formulas that correct for AR(1) dynamics would be wrong and may even exacerbate the bias when the true lag order is larger than one" (Lee 2012, 57).

Misspecified lag structures are clearly not the only dynamic misspecification that biases fixed-effects estimates. Rather, fixed-effects estimates are likely to be biased in the presence of any dynamic misspecification or omitted time-varying variables.[14] Ahn, Lee, and Schmidt (2013) argue that the fixed-effects estimator is biased when omitted variables vary over time and develop a generalized method of moments procedure that accounts for multiple factorial time-varying fixed effects. This estimator, however, requires the existence of instruments which are correlated with the dynamic fixed effects but not with the errors—an assumption that is unlikely to be satisfied and that cannot be tested properly since errors remain unobserved. Finally, Park (2012) at least implicitly confirms the existence of bias in fixed-effects models with structural breaks and develops a Bayesian estimator that seeks to identify these structural breaks. As one would expect, a model that corrects for 'turning points' fits the data better than the classical fixed-effects estimator. We build on these contributions and prove that the bias from dynamic misspecifications can be larger for the fixed-effects estimator than for pooled-OLS. As we have already mentioned, we make this comparison not so much because we intend to rehabilitate the pooled-OLS estimator. Rather, we use this comparison to demonstrate how poorly the fixed-effects estimator performs in the presence of dynamic misspecifications.

3.1 Bias of the fixed-effects estimator induced by correlated within variance

Fixed-effects estimation accounts for potential bias from unobserved time-invariant variables by eliminating all between variation from the estimation. Obviously, the effect of variance that is dropped from the estimation cannot be biased by correlated confounders. And the correlation between the remaining within variation and omitted time-invariant cross-sectional variation is zero. Therefore, if the effect of omitted variables is really exclusively time-invariant, the estimates which rely on an analysis of the remaining time-varying variance does not suffer from omitted

---

14 Fixed-effects estimates are also biased by what is known as the incidental parameter problem (Neyman and Scott 1948; Lancester 2000; Hahn and Kuersteiner 2011). This incidental parameter problem for fixed-effects estimation of pooled data is insofar interesting for our argument because it implies that the fixed-effects estimator is consistent when T approaches infinity. From this perspective using fixed effects becomes a catch 22 because as the number of periods increases, fixed-effects estimates become more precise but the probability of dynamic misspecification bias increases as well.

variable bias. This, of course, immediately changes when the aggregate effect of omitted variables is not strictly time-invariant.

In this section, we derive the causes of the bias of the dynamic fixed-effects estimator using a time-varying omitted variable as an example. We demonstrate that the bias of the fixed-effects estimate of β exceeds the bias of the pooled-OLS estimate of β when the correlation between x and omitted within variation is larger than the correlation between x and omitted between variation.

Assume that

$$y_{it} = \alpha + \beta x_{it} + u_i + \varepsilon_{it} \tag{1}$$

is the true data-generating process with x_{it} a time-varying observed variable, u_i a vector of time-invariant unobserved variables, and ε_{it} an i.i.d. error component. Note that the data-generating process is static. Estimating this model by a 'naïve' OLS estimator leads to bias if $\text{var}(\bar{x}_i, u_i) \neq 0$, or $\text{var}(x_{it}, \varepsilon_{it}) \neq 0$.

The fixed-effects estimator eliminates the between variation from Equation (1) so that

$$y_{it} - \bar{y}_i = \beta\left(x_{it} - \bar{x}_i\right) + u_i - \bar{u}_i + \varepsilon_{it} - \bar{\varepsilon}_i \tag{2}$$

which is equivalent to

$$y_{it} - \bar{y}_i = \beta\left(x_{it} - \bar{x}_i\right) + \varepsilon_{it} - \bar{\varepsilon}_i, \tag{3}$$

because $u_i - \bar{u}_i = 0$.

Assume now the following data-generating process

$$y_{it} = \alpha_1 x_{it} + \alpha_2 w_{it} + u_i + \varepsilon_{it}; \quad \text{with } \alpha_2 = 1 \tag{4}$$

where x_{it} and w_{it} are time-varying right-hand-side variables and u_i is a unit-specific effect.

The omitted variable w_{it} is correlated with the included right-hand-side variable x_{it}. γ_1 and γ_2 indicate the strength of the correlation between w_{it} and the within variance of x_{it} and w_{it} and the between variation of x_{it}, respectively:

$$w_{it} = \gamma_1 \ddot{x}_{it} + \gamma_2 \bar{x}_i + \omega_{it} \tag{5}$$

with $\bar{x}_i = (1/T) \sum_{t=1}^{T} x_{it}$, $\ddot{x}_{it} = (x_{it} - \bar{x}_i)$.

Finally, the unit-specific effect u_i covaries with the between variance of x_{it} to a degree of delta.

$$u_i = \delta_1 \bar{x}_i + v_i. \tag{6}$$

We omit w_{it} from the estimation and can easily derive the biases for the fixed effects and the pooled-OLS estimators ($\hat{\alpha}_{1,FE}$ and $\hat{\alpha}_{1,OLS}$) under the assumptions in (4)–(6). We also can demonstrate that under certain conditions the bias of fixed-effects estimates exceeds that of pooled-OLS estimates. Needless to say that neither of these two estimators is unbiased in case of time-varying omitted variables.[15]

---

15 If bias exclusively results from correlation of the between variation of x_{it} and w_{it}, it is of course possible to throw away all the between variation and regress the $\ddot{y}$ on $\ddot{x}$—the within variation of y on the within variation of x—for an unbiased estimate.

Conditional on all of the x_{it}, Equation (7) derives the bias for the pooled-OLS estimator:

$$
\text{Bias}\,(\hat{\alpha}_{1,OLS}) = \frac{\gamma_1 \sum\limits_{i=1}^{N}\sum\limits_{t=1}^{T} \ddot{x}_{it}^2 + T\gamma_2 \sum\limits_{i=1}^{N} \bar{x}_i^2}{\sum\limits_{i=1}^{N}\sum\limits_{t=1}^{T} \ddot{x}_{it}^2 + T \sum\limits_{i=1}^{N} \bar{x}_i^2} + \frac{\delta_1 T \sum\limits_{i=1}^{N} \bar{x}_i^2}{\sum\limits_{i=1}^{N}\sum\limits_{t=1}^{T} \ddot{x}_{it}^2 + T \sum\limits_{i=1}^{N} \bar{x}_i^2}.
\tag{7}
$$

Equation (7) indicates that the OLS bias depends both on the correlation between x_{it} and w_{it} as well as the correlation between u_i and x_{it}.

As usual the bias for the FE estimator is given by:

$$
\text{Bias}\,(\hat{\alpha}_{1,FE}) = \gamma_1.
\tag{8}
$$

The bias of the fixed-effects estimator depends on the correlation between x_{it} and w_{it}, but not on the covariance between the unit-specific effects u_i and x_{it}, because the within-transformation on which FE estimation relies, effectively eliminates all, endogenous and exogenous, between variation from the estimation.

If we assume that $\delta_1 = 0$ (no correlation between u_i and $\bar{x}_i$) and $\gamma_2 = 0$ (no correlation between w_{it} and $\bar{x}_i$), then,

$$
\text{Bias}\,(\hat{\alpha}_{1,OLS}) = \gamma_1 \frac{\sum\limits_{i=1}^{N}\sum\limits_{t=1}^{T} \ddot{x}_{it}^2}{\sum\limits_{i=1}^{N}\sum\limits_{t=1}^{T} \ddot{x}_{it}^2 + T \sum\limits_{i=1}^{N} \bar{x}_i^2} < \gamma_1 = \text{Bias}\,(\hat{\alpha}_{1,FE}).
\tag{9}
$$

In this case, for any given $T < \infty$, the bias of the FE estimator that results from the omission of w_{it} is larger than that of OLS. This is so because the fraction term of the OLS bias in Equation (9) is always smaller than 1.

This case might seem rare in real data but can emerge when neither x_{it} nor w_{it} have a specific dynamic structure (autocorrelation or trends) but only the variation over time and not across units of these two variables is related. An exogenous shock could have this property. Alternatively, w_{it} has no between variation and represents an omitted common trend. More often, however, applied researchers specify empirical models that suffer from both, omitted between variation correlated with the regressors *and* omitted within variation correlated with the regressors. In these cases, one cannot say whether the fixed effects or the pooled-OLS estimator gives less biased estimates. One can ex ante know that both estimators give biased results, but which one is more reliable (or less unreliable) depends on the relative strengths of the correlations with the omitted variance. Unfortunately, these correlations cannot be observed. Researchers may often know that a relevant variable has been omitted, but one cannot know with certainty that no relevant variable has been omitted. Still, one can evaluate whether omitted variables are potentially more problematic for the included within or between variation by estimating how much of the within and between variance of the dependent variable remains unexplained.

Now consider a situation where x_{it} (and w_{it}) follows a deterministic trend so that its within variance grows with increasing number of time periods, and approaches infinity as $T \to \infty$. In this case, even if $\delta_1 \neq 0$ (non-zero correlation between u_i and $\bar{x}_i$) the second term of the OLS bias Equation (7) approaches zero because the within variation ($\ddot{x}_{it}^2$) grows but only appears in the denominator while the between variance ($\bar{x}_i^2$) does not change. The bias that is caused by the

correlation between w_{it} and x_{it} increases with T because of the trend and will outweigh the bias induced by omitted time-invariant variables if T grows large enough.

3.2 Bias from dynamic misspecification

Correlated within variation and common trends of included and excluded explanatory variables are obvious sources of omitted variable bias occurring in fixed-effects estimates. Yet, there are many examples of dynamic misspecifications that can cause bias. Assume a data-generating process representing the simplest form of dynamic misspecification, an explanatory variable that does not exert a contemporaneous effect on the dependent variable but a one period lagged effect:

$$y_{it} = \beta x_{it-1} + u_i + \varepsilon_{it}. \tag{10}$$

If we estimate Equation (10) ignoring the lagged effect of x_{it}, the probability limit (plim) of the OLS estimator of β in the regression $y_{it} = \beta x_{it} + \varepsilon_{it}$ is given by:

$$\frac{\text{Cov}(y_{it}, x_{it})}{\text{Var}(x_{it})} = \beta \frac{\text{Cov}(x_{it-1}, x_{it})}{\text{Var}(x_{it})} + \frac{\text{Cov}(u_i, x_{it})}{\text{Var}(x_{it})}. \tag{11}$$

The second term of Equation (11) is similar to the bias of estimating a model without fixed effects while the true DGP has correlated unit effects: the estimated β wrongly captures the unit-specific effects (unless $\bar{x}_i = 0$).

The probability limit of the fixed-effects estimator equals:

$$\frac{\text{Cov}(y_{it}, \ddot{x}_{it})}{\text{Var}(x_{it})} = \beta \frac{\text{Cov}(x_{it-1}, \ddot{x}_{it})}{\text{Var}(x_{it})} + \frac{\text{Cov}(u_i, \ddot{x}_{it})}{\text{Var}(x_{it})}. \tag{12}$$

The second term now vanishes since $\ddot{x}_{it}$ has no unit-specific mean. We can rewrite Equation (12) so that

$$\frac{\text{Cov}(y_{it}, \ddot{x}_{it})}{\text{Var}(x_{it})} = \beta \frac{\text{Cov}(x_{it-1}, x_{it})}{\text{Var}(x_{it})} - \beta \frac{\text{Cov}(x_{it-1}, \bar{x}_i)}{\text{Var}(x_{it})} \tag{13}$$

$$\Leftrightarrow \frac{\text{Cov}(y_{it}, \ddot{x}_{it})}{\text{Var}(x_{it})} = \beta \frac{\text{Cov}(x_{it-1}, x_{it})}{\text{Var}(x_{it})} - \beta \frac{\text{Var}(\bar{x}_i)}{\text{Var}(x_{it})}. \tag{14}$$

The second term of Equation (14) equals β multiplied by the between variance of x_{it} divided by its total variance. The result will fall between 0 and 1. It follows that the probability limit of the within estimator (FE) is smaller than β. The estimate will thus be downward biased and this bias increases as the share of ignored between variation in x_{it} increases.

The total bias of the OLS estimator depends on the autocorrelation of x_{it} and the bias induced by the omission of the unit-specific effects. If the majority of autocorrelations in real world data generation processes is positive (which seems to be the case), the bias of a fixed-effects estimator exceeds the bias of pooled-OLS. It is of course possible to estimate whether x_{it-1} and x_{it} are positively correlated and how strong this correlation is. However, it is much more complicated to identify the correct lag structure of explanatory variables (Adolph, Butler, and Wilson 2005; Plümper, Troeger, and Manow 2005). Time series tests such as information criteria (BIC, AIC etc.) have low power in complex models and usually predict diverging lag lengths depending on the number of lags and right-hand-side variables included. The problem of misspecified lag length is exacerbated if the lag length is not uniform but varies across units which can occur frequently in political science data, for example because institutional settings will usually influence responsiveness of actors (Plümper, Troeger, and Manow 2005). We analyze the effect of unit-specific lag length in the Monte Carlo experiments below.

3.3 Discussion

We have demonstrated that biases from two different sources of model misspecification are not simply additive. Rather, the solution to one problem, time-invariant omitted variables, can easily make another problem, say omitted time-varying variables, worse. In the following section we use Monte Carlo analyses to compare the bias of the fixed-effects estimator to the bias of the estimator that econometricians call naïve, pooled-OLS.[16] We do so to identify some of the conditions under which the fixed-effects estimator has poor properties. As we have mentioned before, we use pooled-OLS to have a benchmark for 'poor properties'—and not to recommend the choice of the pooled-OLS estimator in applied research.

4 Design of the Monte Carlo Experiments[17]

Bias in fixed-effects estimation can result, inter alia, from omitted time-varying variables, from omitted trends, a misspecified lag structure, and other—more complex—dynamic misspecifications. Since social scientists often rely on standard dynamic specifications rather than on explicitly modeling the dynamics, bias may be reduced, but is unlikely to disappear. As we have shown in the previous section, the existence of any form of unaccounted within variation correlated with the regressors biases fixed-effects estimates. The Monte Carlo analyses in this section aim at exploring the relevance of the problem. To benchmark the bias of the fixed-effects estimator, we use the pooled-OLS estimator which is known to have poor properties in the presence of omitted time-invariant variables and dynamic misspecifications. Naïvely, one could expect that, since pooled-OLS suffers from (at least) two problems while the fixed-effects estimator solves the problem of omitted time-invariant variables, the bias of the fixed-effects estimator is always strictly smaller than the bias of pooled-OLS. However, this perspective ignores the fact that the fixed-effects estimator solely uses the within variation and is therefore more vulnerable to dynamic misspecification than pooled-OLS that uses both, the within- and the between variation. As we demonstrate analytically, it is thus possible that fixed-effects estimates are more biased than the pooled-OLS estimates under identifiable conditions. To study the properties of the fixed-effects estimator with potential dynamic misspecification and to reveal the conditions under which the use of fixed effects produces larger bias than the naïve pooled-OLS estimator, we employ a set of Monte Carlo experiments.

Our data-generating process follows a straightforward set-up:

$$y_{it} = x_{it}^1 + (x_{it}^2) + u_i + \varepsilon_{it} \tag{15}$$

with $x_{it}^1, x_{it}^2, \varepsilon_{it}$, and u_i being drawn from a standard normal probability density function.

We use three rather straightforward types of model misspecifications as examples: an omitted time-varying variable, an omitted time trend when the variable of interest x_{it}^1 is trended, and the simple dynamic misspecification analyzed formally in the previous section—a one period lagged effect of x_{it}^1. We distinguish three levels of correlation between our variable of interest x_{it}^1 and an omitted strictly time-invariant, constant effect variable u_i. We set this correlation between x_{it}^1 and u_i to 0.0 (absent), 0.2 (weak), and 0.5 (substantive). Higher correlation between x_{it}^1 and u_i implies higher bias of the pooled-OLS estimator, while the correlation between x_{it}^1 and u_i does not bias the fixed-effects estimator. The higher the correlation between x_{it}^1 and u_i, the larger the bias advantage of the fixed-effects model before we consider a dynamic misspecification. Obviously, in the absence of dynamic misspecification the fixed-effects estimates are unbiased. Throughout

---

16 We could add here the analysis of the random-effects model but this should produce exactly the same average bias as an OLS model.

17 The replication files for the MC analysis can be found on the PA dataverse: Troeger, Vera; Pluemper, Thomas, 2017, "Replication Data for: Not so Harmless After All: The Fixed-Effects Model", doi:10.7910/DVN/RAUIHG, Harvard Dataverse.

all specifications we assume that between and within effects are equal. We acknowledge that this is a strong assumption and that pooled-OLS gives an average estimate of the two effects while the fixed-effects estimator provides a clean estimate for the within effect only. For a discussion of dealing with different within and between effects see Bell and Jones (2015). We refrain from adding a discussion of different effects across units and over time since it would distract from the focus on bias stemming from dynamic misspecification.

We are of course aware that social scientists could be able to correctly model these simple dynamic misspecifications. But this argument misses the point: we do not seek to identify dynamic misspecifications which are so difficult to model that social scientists probably fail to fully eliminate them. Instead, we are analyzing the consequences of dynamic misspecifications. The advantage of simple dynamic misspecifications, thus, is that it is easy to understand how they bias the fixed-effects estimator. Only in a second step will we generate complex data-generating processes for which simple solutions are not available. None of the data-generating processes we study here are likely to be as complex as true data-generating processes. Given that we include simple dynamics, we do not just use a simple fixed-effects specification, but rather compare fixed-effects estimation with dynamic specifications that applied researchers are likely to use as econometric solutions for potential dynamic misspecifications:[18] a lagged dependent variable (or Arellano–Bond dynamic panel model[19]), the Prais–Winsten transformation, or period fixed effects. This also allows us to demonstrate that these simple fixes, which are widely used in panel and pooled analyses, do not sufficiently eliminate simple dynamic misspecifications. In addition to these simple but commonly employed dynamic fixes, we use more general dynamic specifications as offered by ADL models and show that capturing the most salient dynamic elements of a DGP can reduce the bias considerably. This is consistent with Pickup (2017).

For our first two experiments, x_{it}^2 is the omitted part of the data-generating process. We first directly manipulate the correlation between the within variation of x_{it}^1 and x_{it}^2 and the unit heterogeneity, e.g. the covariance of the between variance of x_{it}^1 and the unobserved unit-specific effects u_i with $\mathrm{corr}\left(\ddot{x}_{it}^1, \ddot{x}_{it}^2\right) = \{0.2, 0.5, 0.8\}$.

The second set of experiments aims at demonstrating the logic of our argument without ex ante assuming that x_{it}^1 and x_{it}^2 are correlated. We generate a dynamic misspecification by merely trending both variables so that the correlation results from the trends only. We discuss two different variants of this second Monte Carlo experiment.[20] The first variant assumes common trends across all units: Both included and excluded right-hand-side (RHS) variables are continuous with a common trend of 0.1 increase per time period:

$$x_{it}^{1,2} = N \sim (0, 1) + 0.1 * t, \quad t = 1, \ldots, T. \tag{16}$$

The second variant relaxes this assumption and allows for unit-specific trends,[21] which merely means that trends are conditioned by other factors—a plausible assumption for social scientists,

---

18 See Acemoglu *et al.* (2008) for the choice of an Arellano–Bond model, Beck and Katz (1995) for the use of the lagged dependent variable (but see Achen 2000 and Keele and Kelly 2006), Huber and Stevens (2012) for the Prais–Winsten transformation, and Becker and Woessmann (2013) for the inclusion of period dummies. For a broader discussion see DeBoef and Keele (2008).

19 Since the combination of a LDV and unit-specific effects generates Nickell-bias (Nickell 1981) we also show results for the most common solution to Nickell-bias—and Arellano–Bond model (Arellano and Bond 1991).

20 We have conducted additional experiments. Since findings remain consistent with the results discussed here, we do not report additional findings.

21 In experiments not shown here we also studied bias of the fixed-effects model with a binary treatment variable (Beck and Katz 2001; Green, Kim, and Yoon 2001). Binary treatments can be trended if the probability of treatment increases or declines over time. Epidemics may serve as the most obvious example. Furthermore, most studies of treatment effects only observe two periods: pre-treatment and post-treatment. In this situation, the probability of treatment increases from zero to a probability determined by the share of the treated cases to the total cases. In such a case, every omitted trended variable will bias the results unless the effect of this variable is strictly identical for treatment and control group. Because of limited space we relegate the results for binary treatment variables to an online appendix.

---

since trends are unlikely to be homogeneous across units. Specifically, we randomly draw a third of the units that receives a positive trend of 0.1 per time period (see Equation (16)), a third of the units remains untrended ($x_{it}^{1,2} = N \sim (0, 1)$), and the last third of units has a negative trend of 0.1 per time period $x_{it}^{1,2} = N \sim (0, 1) - 0.1 * t, t = 1, \ldots, T$).[22]

The third experiment is based upon a slightly different DGP to account for a misspecified lag structure of x_{it}^1:

$$y_{it} = x_{it-1}^1 + u_i + \varepsilon_{it}. \tag{17}$$

We compare the bias generated by a static OLS estimator ($y_{it} = x_{it}^1 + \varepsilon_{it}$) to that of a static FE estimator ($\ddot{y}_{it} = \ddot{x}_{it}^1 + \ddot{\varepsilon}_{it}$) where the lagged effect of x_{it}^1 is not taken into account.

Finally, we also allow the lag length of x_{it}^1 to vary across units in the following way: for one randomly drawn third of the units x_{it}^1 exerts a one period lagged effect on y_{it} as in Equation (17), for the second randomly drawn third of units we observe a two period lagged effect $y_{it} = x_{it-2}^1 + u_i + \varepsilon_{it}$ and for the last third we model a three period lagged effect $y_{it} = x_{it-3}^1 + u_i + \varepsilon_{it}$.

We vary the number of periods [$T = \{10, 30, 50\}$] but we hold the number of units constant at 20 throughout all experiments. Note that increasing the number of units increases the between variation and favors pooled-OLS over fixed effects (Plümper and Troeger 2007, 2011). In each permutation of the experiments we estimate 500 models with independently drawn errors.

Since econometricians have developed different solutions for models with potential dynamic misspecifications, we incorporate these variants of the fixed effects and the pooled-OLS estimators into the simulation. The most commonly used 'solutions' to dynamic misspecification are the inclusion of the lagged dependent variable (LDV: $y_{it} = \alpha + \beta_1 y_{it-1} + \beta_2 x_{it}^1 + \varepsilon_{it}$),[23] and period fixed effects ($y_{it} = \alpha_t + \beta_2 x_{it}^1 + \varepsilon_{it}$) or a combination of the two ($y_{it} = \alpha_t + \beta_1 y_{it-1} + \beta_2 x_{it}^1 + \varepsilon_{it}$). Less often researchers employ a Prais–Winsten transformation (PW: ($y_{it} - \rho y_{it-1}$) = $\alpha + \beta_2 (x_{it}^1 - \rho x_{it-1}^1) + (\varepsilon_{it} - \rho \varepsilon_{it-1})$),[24] or an ADL ((1,1): $y_{it} = \alpha + \beta_1 y_{it-1} + \beta_2 x_{it}^1 + \beta_3 x_{it-1}^1 + \varepsilon_{it}$) model.

Though it seems to increasingly be the case that scholars estimate fixed-effects models without justification and thus as default, econometric textbooks suggest a variant of the Hausman specification test (Hausman 1978) to decide whether to estimate a fixed effects or a random-effects/ OLS specification. The Hausman test (and its variants) have been shown to be consistent (for a short overview see Baltagi 2001, 65–70), therefore, if fixed-effects estimates are significantly different from random effects or pooled-OLS estimates, the latter are biased because of unit heterogeneity.[25] However, the asymptotic properties of the Hausman test do not necessarily translate into favorable finite sample characteristics especially when other misspecifications do exist and are not accounted for. We also present Monte Carlo results for the performance of this test. This is related to our main research interest, because we intend to demonstrate that pooled-OLS may be less biased than the fixed-effects model in situations in which the Hausman test favors a fixed-effects specification. These instances may occur frequently and under plausible conditions.

5 Results

Applied researchers should select estimators according to their reliability for the sample at hand. The root mean squared error has been suggested as the appropriate criterion for selecting

---

22 This set-up might seem somewhat unrealistic but we run the same experiment with one half of the units positively trended and one half not trended and get similar results.

23 Since the combination of unit fixed effects and a lagged dependent variable induces additional bias, the so called Nickell-bias (Nickell 1981), we also run dynamic panel models that allow for the combination of unit-specific effects and a lagged dependent variable.

24 The OLS variant with Prais–Winsten transformation results in a GLS model.

25 Note that generalization from asymptotic properties to small sample properties are not valid. At the same time, this logic overlooks multiple other reasons for parameter heterogeneity.

Table 1. Bias over all Experiments.

estimate	Econometric Specification	Bias: pooled-OLS			Bias: FE		
		mean	min	max	mean	min	max
x^1_{it}	No dynamics	0.378	0.030	0.666	0.619	0.055	1.103
	LDV	0.315	0.000	0.738	0.580	0.044	1.127
	Arellano–Bond (A–B)				0.597	0.002	1.392
	Prais–Winsten GLS	0.549	0.015	1.324	0.612	0.037	1.189
	Period Fixed Effects	0.337	0.001	0.663	0.563	0.000	1.103
	LDV+ Period Fixed Effects	0.316	0.005	0.737	0.546	0.001	1.126
	A–B + Period Fixed Effects				0.568	0.000	1.385
	ADL	0.256	0.005	0.757	0.283	0.000	0.803
x^1_{it-1}	ADL	0.537	0.033	1.160	0.188	0.000	0.763
	LDV	0.335	0.184	0.464	0.066	0.005	0.229
y_{it-1}	LDV+ Period FE	0.338	0.192	0.472	0.051	0.005	0.210
	ADL	0.444	0.265	0.586	0.066	0.010	0.148

estimators in finite samples. The root mean squared error provides information on the average deviation of an estimator from the true relationship. This average deviation results from bias and sampling variation of an estimator. We show the bias for our MC experiments because whenever the bias of the fixed-effects estimates exceeds the bias of the benchmark, the pooled-OLS estimator, the root mean squared error is also larger. Since OLS is using both within and between variation for estimation it is the more efficient estimator as compared to Fixed Effects.

We run five sets of experiments that examine different dynamic misspecifications: (i) omitted time-varying variable, (ii) omitted common trend, (iii) unit-specific trend, (iv) misspecified common lag structure, and (v) misspecified unit-specific lag structure. For each of the misspecifications we estimate six different fixed effects and pooled-OLS models with different dynamic specification: no dynamics, lagged dependent variable (LDV, or Arellano–Bond (A–B) model), Prais–Winsten GLS transformation, period fixed effects, a combination of LDV/A–B and period fixed effects, and an ADL (1,1) model. Finally we vary the correlations between the unit-specific effects u_i and the interesting RHS variable x^1_{it} (as described above), as well as the number of periods.

Table 1 summarizes the findings of all conducted experiments. We show the average, minimum and maximum bias generated by pooled-OLS and the fixed-effects model for each dynamic specification.

Table 1 gives a first impression of the general performance of pooled-OLS and fixed-effects models with different econometric patches when dynamic misspecifications are present in the DGP but not necessarily properly accounted for in the specification of the estimation equation. Overall, the average bias of the coefficient for x^1_{it} (the RHS variable of interest) produced by pooled-OLS is up to 45 percent smaller than that generated by the fixed-effects estimator. In addition, the maximum bias of OLS is usually considerably smaller than the maximum bias of the fixed-effects estimates (except when a Prais–Winsten GLS transformation is applied). An ADL(1,1) model estimates coefficients for both x^1_{it} and the one period lagged x^1_{it-1}. The ADL model produces on average less biased estimates for x^1_{it-1} when unit fixed effects are included. However, the computed average bias for estimates of x^1_{it} and x^1_{it-1} in the ADL(1,1) model is somewhat misleading because in experiments 1 and 2 x^1_{it} should be included in the estimation but x^1_{it-1} is not part of the DGP, while in experiment 3 only x^1_{it-1} has an effect on the outcome. Three of the dynamic specifications we test (LDV, LDV+period FE, ADL) also estimate coefficients for y_{it-1}. This coefficient should be zero because y_{it-1} is never part of the DGP. Specifications that include unit fixed effects on average produce coefficients for the LDV that are closer to zero. In a

pooled-OLS specification the LDV on average seems to pick up potential unit-specific effects that remain un-modeled.[26] If researchers are interested in persistency of the dependent variable or long term effects and unit effects are indeed present, a fixed-effects specification produces less biased estimates of the LDV coefficient. This often comes at the expense of a more biased estimate for the interesting explanatory variables when dynamic misspecifications are present. To unpack the relative performance of both estimators in the presence of different dynamic misspecifications we present disaggregated results for each misspecification and different econometric controls for dynamics.

5.1 Experiment 1: omitted time-varying variable

We start with examining the effect of omitted time-varying variables for different levels of correlated unit-specific heterogeneity. The results confirm the theoretical results in Section 3. Table 2 depicts the bias of OLS (solid line) and FE (dashed line) with an assumed within correlation between included and omitted time-varying variables of 0.5. We include the results for eighteen combinations for the level of correlation of x_{it}^1 and u_i and a dynamic specification. Each single figure displays the bias for the OLS estimates and the bias for the fixed-effects estimates (right axis) plus the probability that the Hausman test finds a significant difference between the OLS and the FE estimates (at the 95 percent level—gray shaded area, left axis). The larger the gray shaded area, the higher the probability that the Hausman test recommends the FE model. We show results for each of the six specifications that political scientists frequently use to control for dynamics: no control for dynamics, lagged dependent variable (with Arellano–Bond estimator—dotted line[27]), Prais–Winsten transformation, period fixed effects, the combination of the LDV and period fixed effects, and an ADL specification. For the ADL(1,1) model (last specification in each table) we display bias for estimates of x_{it}^1 (black lines) and x_{it-1}^1 (gray lines). The columns depict these results for different levels of correlation between the unit-specific effects u_i and the included treatment x_{it}^1.

Table 2 illustrates that the bias of the fixed-effects model increases as the correlation between the variable of interest and an omitted time-varying variable increases (see tables A1 and A2 in appendix for comparison). The fixed-effects estimator is not immune against different sources of unobserved or omitted heterogeneity, it merely shelters estimates from omitted time-invariant variables with constant effects (which is referred to as 'unobserved heterogeneity' in most econometric textbooks). The bias of the fixed-effects estimates remains unaffected by changes in the correlation between the variable of interest and an omitted time-invariant variable.

The omission of a time-varying variable that is correlated with included right-hand-side variables may lead to serially correlated errors and it will induce bias. As we have explained in Section 2, social scientists use various econometric solutions to control for the serial correlation of errors potentially resulting from omitted time-varying variables. We find that these solutions have virtually no effect on the bias of the fixed-effects estimate in the presence of omitted time-varying variables. Yet, omitted time-varying variables are a common problem in the social sciences—arguably more common than the omission of variables with time-invariant effects that vary across units.

A comparison between the properties of the fixed-effects model and the benchmark pooled-OLS estimator reveals that the fixed-effects model is more (less) biased if the correlation of the variable of interest with the omitted within variation is larger (smaller) than the correlation with the omitted between variation. Of course, if no omitted time-invariant variable exists but the model is dynamically misspecified, pooled-OLS is strictly less biased than the fixed-effects estimator. This confirms the results from Section 3. The results for the ADL(1,1) model show the

---

26 We show detailed results for the bias of the coefficient of the LDV for all MC experiment in Appendix tables A3 to A9.
27 In some cases the dotted line for the bias of the A–B estimator cannot be seen because it is equal to the bias produced by the FE estimator and they completely overlap.

Table 2. Omitted Within Variance $corr(\ddot{x}_{it}^1, \ddot{x}_{it}^2) = 0.5$: Bias for Estimate of x_{it}^1 and x_{it-1}^1.
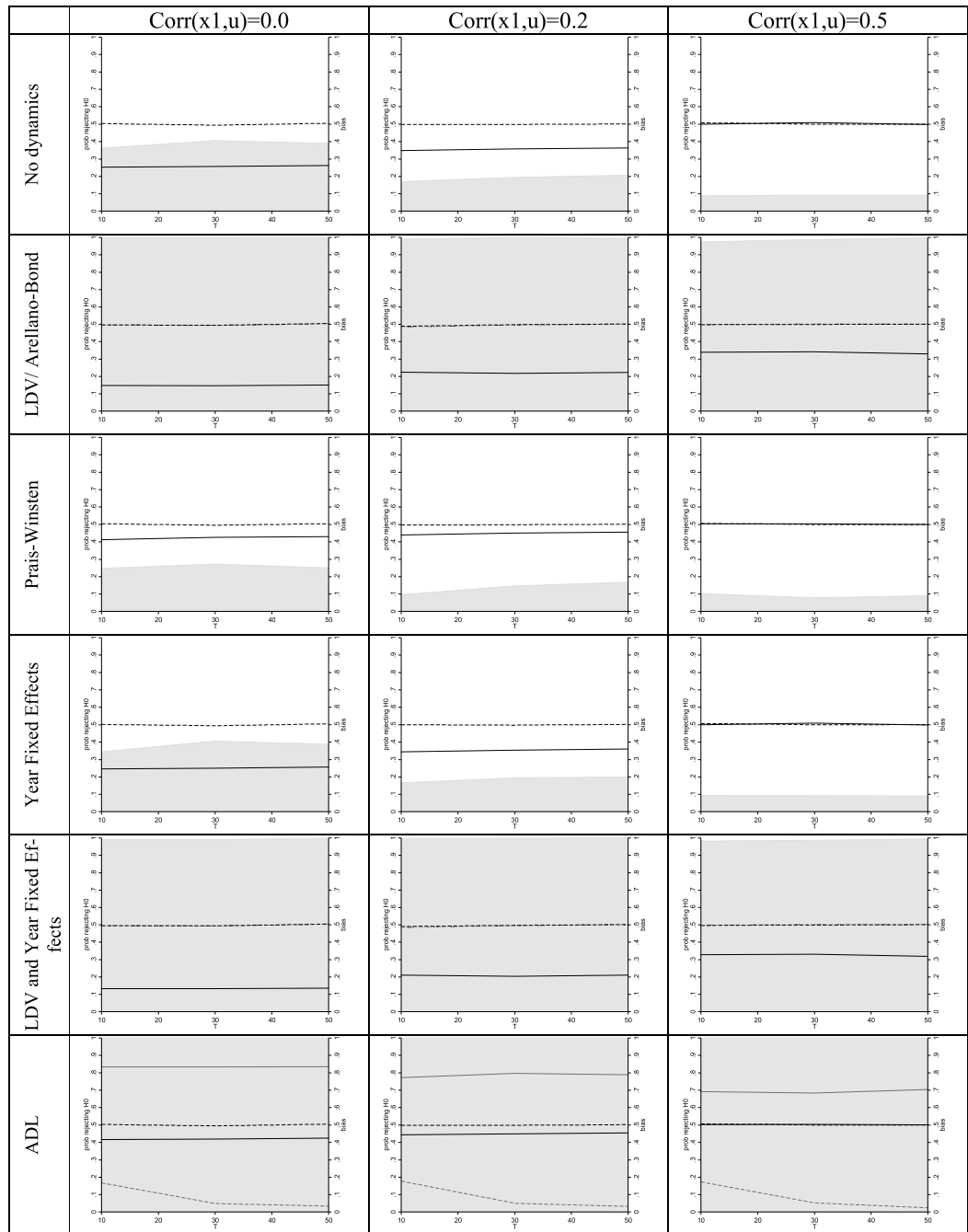
Note: Right Axis—Absolute Bias: —— OLS, - - - - - FE, · · · · · · A–B (ADL: gray lines = bias of coefficient for x_{it-1}^1); Left Axis—Probability of rejecting the H0 on the 5% level and thus suggesting FE: gray shaded area = Hausman Test.

same bias differential between OLS and FE estimates for x_{it}^1 (black lines), though the difference is smaller. However, a fixed-effects specification seems to be able to deal much better with elements that are not included in the DGP since it produces a much smaller bias for the unnecessarily added x_{it-1}^1 (gray lines).

Table 2 also reveals the low power of the Hausman test in the presence of dynamic misspecification. It gives erratic results and in the worst case with no omitted time-invariant variables but a correlated omitted time-varying variable, the Hausman test always suggests the

use of the fixed-effects model—even if no omitted between variation exists. We also find that the Hausman test is sensitive to the choice of dynamic specification. If applied researchers include a lagged dependent variable, the Hausman test is biased toward the fixed-effects model—a finding that confirms previous research (Arellano 1993; Ahn and Low 1996; Godfrey 1998; Baltagi 2001; Hoechle 2007, 66–69). In other words, the 'consistency' of the Hausman test is conditional on a perfectly specified model that suffers solely from omitted between variation with constant unit effects.

Finally, in our MC analyses all results are largely independent of the number of periods, because we hold the within correlation constant. If, in reality, adding periods leads to a change in the correlation, the bias will also change. As adding time periods increases the probability of correlated time-varying omitted variables, the bias will increase over-proportionally for fixed-effects estimates.

## 5.2 Experiment 2. correlated common and unit-specific trends

In the second experiment, we study the bias of the fixed-effects model and the pooled-OLS estimator when both the variable of interest x_{it}^1 and an omitted time-varying variable x_{it}^2 are trended. Two trended variables tend to be correlated even if they are independent of each other. Table 3a displays the results for an excluded trended variable, while Table 3b provides the results for experiments in which the trend is assumed to be unit-specific.

This experiment confirms that the static fixed-effects model is biased, but this bias—expectedly—disappears when scholars include period fixed effects in the presence of a true common trend. A similar result can be achieved by the inclusion of splines, but period fixed effects follow the functional form of the omitted trended variable more closely. Unfortunately, period fixed effects also capture the trend of other trended variables. Hence, if scholars aim at analyzing dynamic processes, period fixed effects only leave unit-specific deviations from the common trend for variables of interest, since period fixed effects account for all common trends. This does not mean that we suggest leaving out period fixed effects in general, we advocate a less ad hoc approach to modeling of the salient dynamic features of the data-generating process and a far more cautious interpretation of the estimation results.

With omitted trended variables and no period FE, pooled-OLS tends to outperform the fixed-effects model unless the number of periods remains small and the correlation between a time-invariant omitted variable u_i and the variable of interest x_{it}^1 is high. As we have demonstrated in Section 3, trends increase the within variation of included and omitted RHS variables when T grows larger. As a consequence, the bias resulting from omitted trends increases in T, which affects fixed-effects models more strongly than pooled-OLS models because FE solely relies on within variation for estimation. This observation holds true for the ADL (1,1) estimation of x_{it}^1. However, as in experiment 1, including unit fixed effects allows estimating zero effects of unnecessary components (x_{it-1}^1) more precisely though not without bias.

In the likely case that omitted trends are not common to all units (Table 3b), period dummies can no longer guarantee the unbiasedness of the fixed-effects model. In this case, the period fixed effects capture the mean of these unit-specific trends so that residuals for other units, units that follow a different trend, still show serial correlation, which of course can be correlated with the variable of interest and, indeed, will almost certainly be correlated if the variable of interest is also trended in a unit-specific fashion. Our results thus run directly counter to Allan and Scruggs (2004, 505) belief that "fixed effects do allow us to reduce the possibility that the substantive estimates are in fact attributable to country-specific trends." We find this statement unlikely to be correct. Instead, the presence of unit-specific trends that are not otherwise accounted for renders the choice of a fixed-effects model more problematic.

PA

Table 3a. Omitted common trend: bias for estimate of x_{it}^1 and x_{it-1}^1.
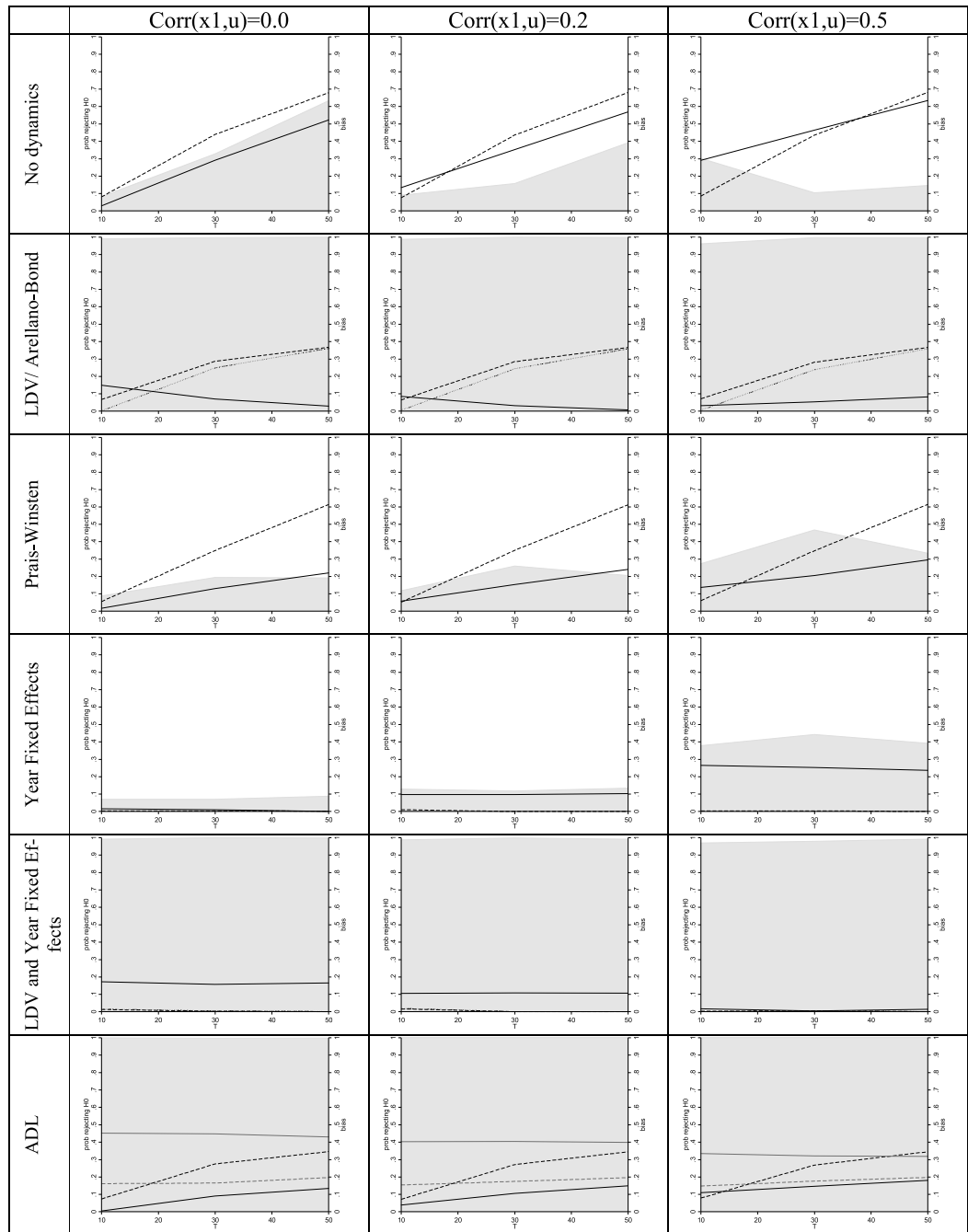
Note: Right Axis—Absolute Bias: —— OLS, - - - - - FE, · · · · · · A–B (ADL: gray lines = bias of coefficient for x_{it-1}^1); Left Axis—Probability of rejecting the H0 on the 5% level and thus suggesting FE: gray shaded area = Hausman Test.

In general, these Monte Carlo analyses provide ample evidence that the bias of the fixed-effects model depends on the existence of dynamic misspecifications *and* on the degree to which econometric solutions capture the dynamic misspecification. The inclusion of the correct dynamic model provides of course a solution but it is usually hard to test for the source of dynamic misspecifications, especially when different dynamic issues occur jointly. Different dynamic misspecifications can lead to similar manifestations in the residuals, e.g. serial correlation.

Table 3b. Omitted unit-specific trends: bias for Estimate of x_{it}^1 and x_{it-1}^1.
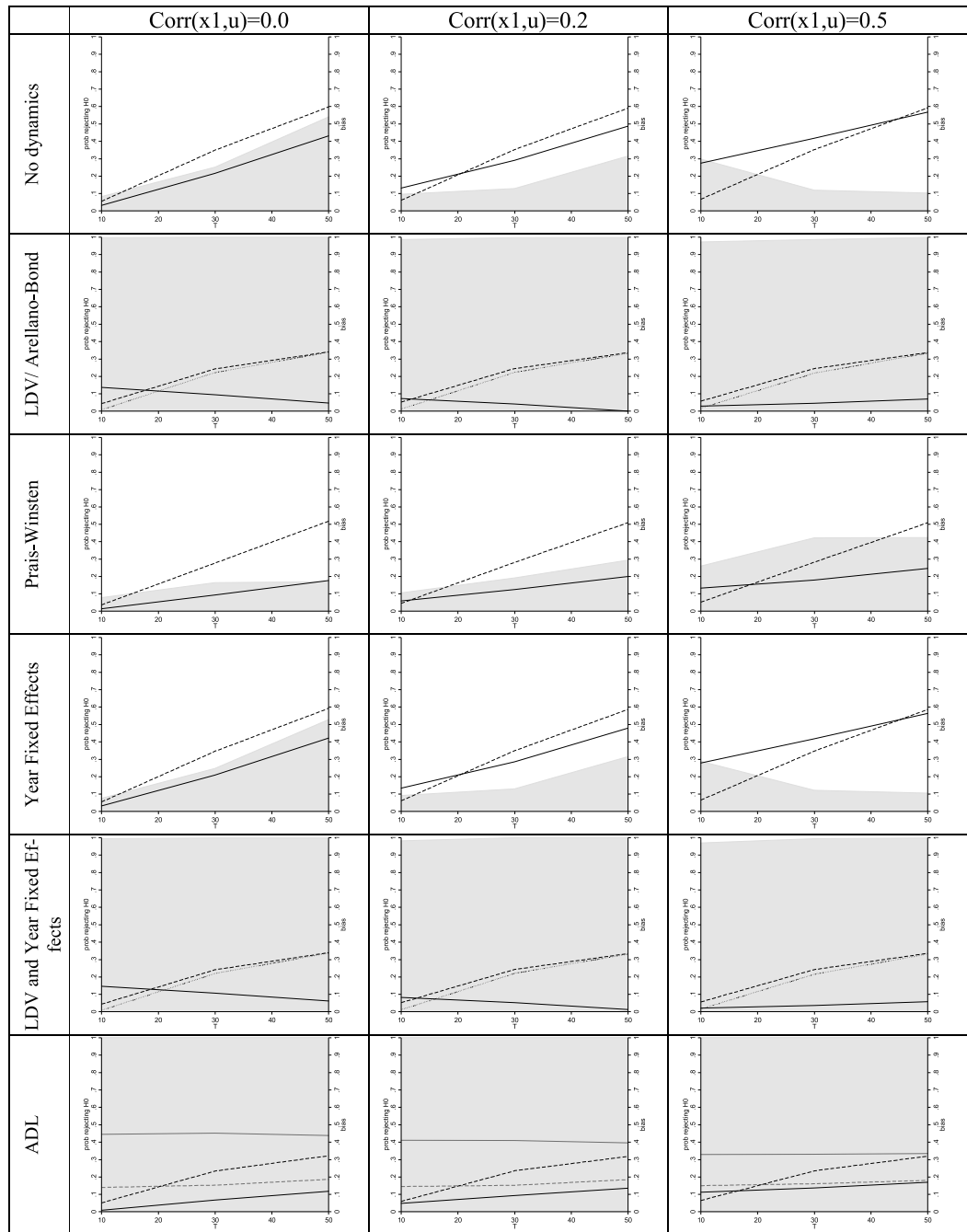
Note: Right axis—absolute bias: —— OLS, - - - - - FE, · · · · · · A–B (ADL: gray lines = bias of coefficient for x_{it-1}^1);
Left axis—probability of rejecting the H0 on the 5% level and thus suggesting FE: gray shaded area = Hausman Test.

However, not every econometric model controlling for autocorrelation (e.g. LDV, ADL, Prais–Winsten) will treat the source of the problem successfully and might exacerbate the bias.

5.3 Experiment 3. misspecified lag structure

In the final set of simulations we study the impact of a very common dynamic misspecification (Adolph, Butler, and Wilson 2005; Wilson and Butler 2007) on the performance of pooled-OLS and fixed-effects estimators. Many applied researchers do not sufficiently explore the potential

of lagged effects on the outcome. Often, ignoring lagged effects will result in not rejecting the Null Hypothesis, concluding that there are no effects from x on y (Plümper, Troeger, and Manow 2005). In models with several right-hand-side variables and complex dynamics, especially when analyzing pooled data, it becomes very difficult if not impossible to test for the correct lag length of right-hand-side variables.

In pooled social science data we also find very often that effects are delayed differently for different units. The lag length can vary because for example different electoral systems generate different political reaction functions. It is conceivable that changes in the political color of the executive have differently delayed effects on political outcomes in coalition vs. single party governments due to different bargaining situations (Plümper, Troeger, and Manow 2005). Table 4a presents the results for an un-modeled (except in the ADL(1,1) specification) one period lagged effect of x_{it}^1, while Table 4b presents MC findings for unit-specific lag length.

Experiment 3 adds further support to the notion that dynamic misspecification biases fixed-effects estimates and that this bias can outweigh the bias of pooled-OLS estimates facing the same dynamic problems. We also find evidence that common econometric solutions to dynamic misspecifications can exacerbate the bias. The results for experiment 4 are indeed staggering: All dynamic specifications except the ADL(1,1) model produce largely biased estimates when the correct lag length is ignored. The bias generated by including unit-specific effects in these cases exceeds 100 per cent. This, in our perspective, potentially provides the best argument for preferring pooled-OLS to the fixed-effects model when dynamics are not explicitly modeled by substantive variables or the correct dynamic specification but controlled away by econometric patches. However, in the presence of dynamic misspecification, neither fixed effects nor pooled-OLS will be unbiased.

Only econometric specifications that explicitly include a one period lagged right-hand-side variable (x_{it-1}^1) like the ADL (1,1) model can recover the true effect of x_{it-1}^1. In the simpler case where x exerts a uniform one period lagged effect on the outcome y (Table 4a) both OLS and FE estimation produces unbiased estimates of x_{it-1}^1 (gray lines) which is included in the DGP. The FE estimator also generates unbiased estimates for x_{it}^1 which is an unnecessary element while the OLS estimator produces slightly biased estimates of x_{it}^1 (black lines). In the more complex situation where lag structures are unit-specific (Table 4b), both estimators produce biased estimates for x_{it-1}^1 (gray lines), and this problem appears to affect the fixed-effects estimator more strongly than pooled-OLS. In comparison, either FE or pooled-OLS are able to recover the zero effect of the unnecessary component x_{it}^1 (black lines), with the FE estimator performing slightly better especially as T grows larger.

The poor performance of the Hausman test is starkest in this set of experiments. Independent of existing correlation between unit-specific effects and RHS variables and independent of whether the FE model generates a larger bias than an OLS or RE specification, the Hausman test indiscriminately and wrongly favors the FE estimator.

The first best strategy to estimate models with complex dynamics in the true data-generating process, heterogeneous lag structures, time-varying conditionality, trended regressors, and so on, is to actually try modeling these dynamics directly rather than eliminating serially correlated errors. This error structure exists not because nature invented a complex error process that ought to be controlled away, but because of a dynamic misspecification in the underlying data-generating process. A fixed-effects model with some added fixes for dynamics does not offer a valid strategy for analyzing dynamic phenomena in the social sciences. Our findings for pooled data with relatively large T are consistent with recent research on short dynamic panels and correlated unit-specific effects (Pickup 2017).

Our results also support findings by Adolph, Butler, and Wilson (2005) as well as Wilson and Butler (2007) who have demonstrated that the use of so called dynamic panel models

Table 4a. Misspecified lag of RHS variable: bias for estimate of x^1_{it} and x^1_{it-1}.
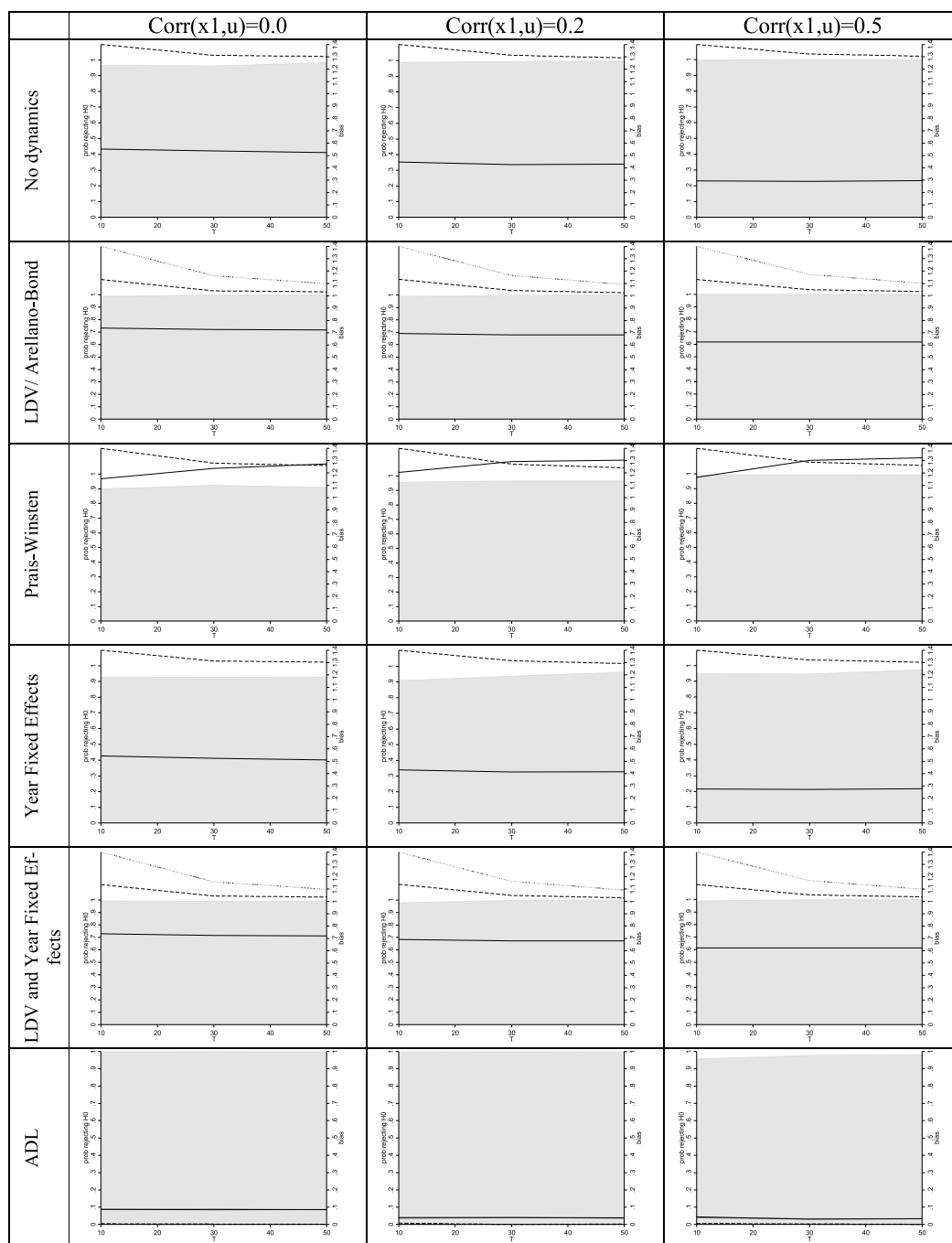
Note: Right axis—absolute bias: —— OLS, - - - - - FE, · · · · · · A–B (ADL: gray lines = bias of coefficient for x^1_{it-1}); Left axis—probability of rejecting the H0 on the 5% level and thus suggesting FE: gray shaded area = Hausman Test.

(Arellano and Bond 1991; Blundell and Bond 1998 etc.) does not alleviate bias from other dynamic misspecifications, even simple ones but only the Nickell-bias that stems from combining fixed effects with a lagged dependent variable. Even if the dynamics in the data-generating process remain fairly trivial, we find substantive bias in the Arellano–Bond model.

Table 4b. Misspecified unit-specific lag of RHS variable: Bias for Estimate of x^1_{it} and x^1_{it-1}.
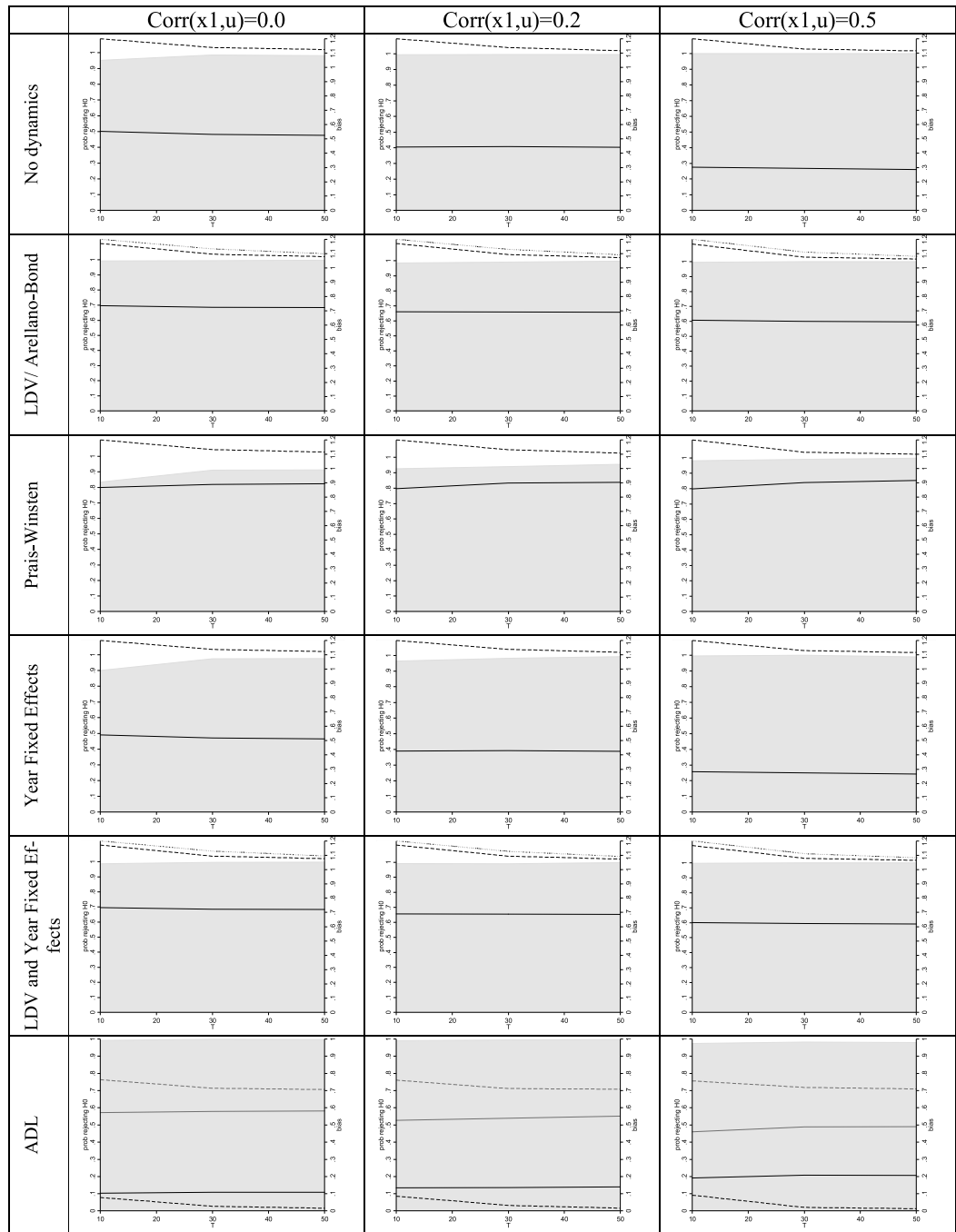
Note: Right Axis—Absolute Bias: —— OLS, - - - - - FE, · · · · · · A–B (ADL: gray lines = bias of coefficient for x^1_{it-1});
Left Axis—Probability of rejecting the H0 on the 5% level and thus suggesting FE: gray shaded area = Hausman Test.

5.4 Discussion

The MC analyses we conduct do not tackle the question whether social scientists can manage to model dynamics properly. Widely used 'from the shelf' model specifications such as the fixed-effects model with a lagged dependent variable, with period fixed effects, or the Arellano–Bond model reveal substantive bias if the data-generating process assumed in the simulations is not completely trivial. Yet, true data-generating processes usually tend to be much more complex than

the ones we design here. We further demonstrate that the widely employed fixed-effects estimator performs poorly, and often even worse than our benchmark, the naïve pooled-OLS model, which is widely criticized for its poor properties. We do not argue that our analyses rehabilitate the pooled-OLS model because its poor performance in the presence of unobserved unit-specific effects and other misspecifications is widely studied and known.

6 Conclusion

The fixed-effects estimator is consistent in the presence of omitted variables with time-invariant effects. It is not consistent in the presence of dynamic misspecification. The fixed-effects estimator deals with one problem and one problem only: its consistency depends on the strong assumption of the strict absence of any specification error other than omitted constant variables with effects that are entirely independent of time. These conditions are not likely to exist in real social sciences data, where few if any variables have constant effects over time.

Dynamic misspecification does not merely render the fixed-effects model biased. Instead we demonstrate in this article that the fixed-effects estimator amplifies the bias from dynamic misspecification relative to estimators that do not shelter the estimation from the between variation. The increase of bias from dynamic misspecification potentially reaches the point where the combined bias from omitted time-invariant variables and dynamic misspecification of OLS estimates becomes smaller than the bias of the fixed-effects model from dynamic misspecification alone.

One could feel tempted to argue that the fixed-effects model solves one particular problem perfectly and thus advise to use the fixed-effects estimator in the likely presence of this problem and deal with all other issues through other model specifications. However, this solution would only be convincing if researchers could eliminate all other model misspecification or if FE would not influence the bias that emanates from model misspecifications which FE do not treat. But as we have demonstrated: this latter assumption is wrong: the use of the fixed-effects model can increase the bias from dynamic misspecifications relative to the naïve pooled-OLS model. Therefore, the case for the FE estimator is limited to situations in which researchers are confident and can thus plausibly argue that they have gotten the dynamic specification of their empirical model correct. Our analyses suggest that simple econometric solutions for modeling dynamics are not very likely to guarantee a correct dynamic specification.[28] Our results demonstrate the importance of carefully modeling underlying dynamics before testing for the existence and potential correlation of unit-specific time-invariant heterogeneity.

These results have rather general implications for econometric research: Misspecifications of the empirical model are not necessarily additive so that solving one problem does not strictly improve the overall performance of the estimator. Quite the contrary is true: Model misspecifications interact with each other so that accounting for one problem by an econometric solution may actually exacerbate the overall bias and therefore increase the probability of wrong inferences. Model misspecifications are not likely to be independent of each other: empirical models suffer from numerous misspecifications (Box 1976; Plümper, Troeger, and Manow 2005; Neumayer and Plümper 2017) and the solution to one problem often renders another problem worse and more difficult to solve. In other words, our analysis casts some doubt on the usefulness of the econometric practice to 'solve' single model misspecifications in isolation. The proof that

---

28 The Fixed-Effects Estimator is of course the correct choice if researchers are theoretically and empirically only interested in within effects. In this case the fixed-effects estimator will give a more adequate econometric answer, though it will still suffer from bias induced by dynamic misspecifications. Throughout this paper we have assumed that within and between effects are the same. This assumption is essential for our conclusions because only if it is met, using between variation in addition to within variation to identify the effects will generate less biased and more reliable estimates. However, as mentioned before, we are not advocating using OLS over FE but are using OLS estimates as benchmark because the undesirable properties are known in the presence of misspecifications.

---

estimators are consistent in respect to a single model misspecification does not guarantee correct inferences if applied researchers cannot plausibly guarantee that their empirical model suffers from the treated misspecification alone.

Supplementary material

For supplementary material accompanying this paper, please visit https://doi.org/10.1017/pan.2018.17.

References

Acemoglu, D., S. Johnson, A. James, J. A. Robinson, and P. Yared. 2008. Income and democracy. *American Economic Review* 98(3):808–842.

Achen, C. H. 2000. Why lagged dependent variables can suppress the explanatory power of other independent variables. Presented at the Annual Meeting of the Political Methodology, Los Angeles.

Adolph, C., D. M. Butler, and S. E. Wilson. 2005. Like shoes and shirt, one size does not fit all: Evidence on time series cross-section estimators and specifications from Monte Carlo experiments. unpubl. Manuscript, Harvard University.

Ahn, S. C., and S. Low. 1996. A reformulation of the Hausman-test for regression models with pooled cross-section-time-series data. *Journal of Econometrics* 71:309–319.

Ahn, S. C., Y. H. Lee, and P. Schmidt. 2013. Panel data models with multiple time-varying individual effects. *Journal of Econometrics* 174:1–14.

Allan, J. P., and L. Scruggs. 2004. Political partisanship and welfare state reform in advanced industrial societies. *American Journal of Political Science* 48(3):496–512.

Angrist, J. D., and J. S. Pischke. 2009. *Mostly harmless econometrics. An empiricists' companion*. Princeton: Princeton University Press.

Antonakis, J., S. Bendahan, P. Jacquart, and R. Lalive. 2010. On making causal claims: a review and recommendations. *Leadership Quarterly* 21:1086–1120.

Arellano, M. 1993. On the testing of correlated effects with panel data. *Journal of Econometrics* 59(1–2):87–97.

Arellano, M., and S. Bond. 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies* 58:277–297.

Baltagi, B. 2001. *Econometric analysis of panel data*. John Wiley & Sons.

Beck, N., and J. N. Katz. 1995. What to do (and not to do) with time-series cross-section data. *American Political Science Review* 89(03):634–647.

Beck, N., and J. N. Katz. 2001. Throwing out the Baby with the Bath Water: A comment on Green, Kim, and Yoon. *International Organization* 55:487-+.

Becker, S. O., and L. Woessmann. 2013. Not the opium of the people: Income and secularization in a panel of Prussian counties. *American Economic Review* 103(3):539–544.

Bell, A., and K. Jones. 2015. Explaining fixed effects: Random effects modeling of time-series cross-sectional and panel data. *Political Science Research and Methods* 3(01):133–153.

Besley, T., and M. Reynal-Querol. 2011. Do democracies select more educated leaders? *American Political Science Review* 105(3):552–566.

Blundell, R., and S. Bond. 1998. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics* 87(1):115–143.

Bole, V., and P. Rebec. 2013. Bootstrapping the Hausman test in panel data models. *Communications in Statistics-Simulation and Computation* 42(3):650–670.

Box, G. E. P. 1976. Science and Statistics. *Journal of the American Statistical Association* 71:791–799.

Clark, T. S., and D. A. Linzerx. 2015. Should I use fixed or random effects? *Political Science Research and Methods* 3(02):399–408.

DeBoef, S., and L. J. Keele. 2008. Taking time seriously: Dynamic regression. *American Journal of Political Science* 52(1):184–200.

Egorov, G., S. Guriev, and K. Sonin. 2009. Why resource-poor dictators allow freer media: A theory and evidence from panel data. *American Political Science Review* 103(4):645–668.

Franzese, R. J. Jr. 2003a. Multiple hands on the wheel: empirically modeling partial delegation and shared policy control in the open and institutionalized economy. *Political Analysis* 445–474.

Franzese, R.J. 2003b. Quantitative empirical methods and the context-conditionality of classic and modern comparative politics. *CP: Newsletter of the Comparative Politics Organized Section of the American Political Science Association* 14(1):20–24.

Franzese, R. J. Jr, and J. C. Hays. 2007. Spatial econometric models of cross-sectional interdependence in political science panel and time-series-cross-section data. *Political Analysis* 140–164.

Franzese, R., and C. Kam. 2009. *Modeling and interpreting interactive hypotheses in regression analysis*. University of Michigan Press.

Frondel, M., and C. Vance. 2010. Fixed, random, or something in between? A variant of Hausman's specification test for panel data estimators. *Economics Letters* 107:327–329.

Gamm, G., and T. Kousser. 2010. Broad bills or particularistic policy? Historical patterns in American state legislatures. *American Political Science Review* 104(1):151–170.

Gerber, A. S., J. G. Gimpel, D. P. Green, and D. R. Shaw. 2011. How large and long-lasting are the persuasive effects of televised campaign ads? Results from a randomized field experiment. *American Political Science Review* 105(01):135–150.

Getmansky, A., and T. Zeitzoff. 2014. Terrorism and voting: The effect of rocket threat on voting in Israeli elections. *American Political Science Review* 108(3):588–604.

Godfrey, L. G. 1998. Hausman tests for autocorrelation in the presence of lagged dependent variables Some further results. *Journal of econometrics* 82(2):197–207.

Green, D. P., S. Y. Kim, and D. H. Yoon. 2001. Dirty Pool. *International Organization* 55:441-+.

Guisinger, A., and D. A. Singer. 2010. Exchange Rate proclamations and inflation-fighting credibility. *International Organization* 64(2):313–337.

Haber, S., and V. Menaldo. 2011. Do natural resources fuel authoritarianism? A reappraisal of the resource curse. *American Political Science Review* 105(01):1–26.

Hahn, J., and G. Kuersteiner. 2011. Bias reduction for dynamic nonlinear panel models with fixed effects. *Econometric Theory* 27:1152–1191.

Harris, M. N., W. Kostenko, L. Matyas, and I. Timol. 2009. The robustness of estimators for dynamic panel data models to misspecification. *Singapore Economic Review* 54:399–426.

Hausman, J. A. 1978. Specification tests in econometrics. *Econometrica* 46(6):1251–1271.

Hedrick, P. W. 2005. A standardized genetic differentiation measure. *Evolution* 59:1633–1638.

Hendry, D. F. 1995. *Dynamic econometrics*. Oxford: Oxford University Press.

Hoechle, D. 2007. Robust standard errors for panel regressions with cross-sectional dependence. *Stata Journal* 7(3):281.

Hsiao, C. 2014. *Analysis of panel data*. Cambridge: Cambridge University Press.

Huber, J. D., and E. Stevens. 2012. Democracy and the left: Social policy and inequality in Latin America. *Journal of Social Policy* 42(3):660–661.

Humphreys, M., and J. M. Weinstein. 2006. Handling and manhandling civilians in civil war. *American Political Science Review* 100(3):429.

Kayser, M. A., and S. Satyanath. 2014. Fairytale Growth. Unp. Manuscript, Hertie School of Governance, Berlin.

Kayser, M. A. 2009. Partisan waves: International business cycles and electoral choice. *American Journal of Political Science* 53(4):950–970.

Keele, L. J., and N. J. Kelly. 2006. Dynamic models for dynamic theories: The ins and outs of LDVs. *Political Analysis* 14(2):186–205.

Keele, L., S. Linn, and C. M. Webb. 2016. Treating time with all due seriousness. *Political Analysis* 24(1):31–41.

Kiviet, J. F. 1995. On bias, inconsistency, and efficiency of various estimators in dynamic panel data models. *Journal of Econometrics* 68(1):53–78.

Kogan, V., S. Lavertu, and Z. Peskowitz. 2016. Performance federalism and local democracy: Theory and evidence from school tax referenda. *American Journal of Political Science* 60(2):418–435.

Lancester, T. 2000. The Incidental Parameter Problem since 1948. *Journal of Econometrics* 95:391–413.

Lebo, M. J., A. J. McGlynn, and G. Koger. 2007. Strategic party government: Party influence in congress, 1789–2000. *American Journal of Political Science* 51(3):464–481.

Lee, Y. 2012. Bias in dynamic panel models under time series misspecification. *Journal of Econometrics* 169:54–60.

Lipsmeyer, C. S., and L. Zhu. 2011. Immigration, globalization, and unemployment benefits in developed EU states. *American Journal of Political Science* 55(3):647–664.

Lupu, N., and J. Pontusson. 2011. The structure of inequality and the politics of redistribution. *American Political Science Review* 105(02):316–336.

Menaldo, V. 2012. The middle east and north Africa's resilient monarchs. *The Journal of Politics* 74(3):707–722.

Morgan, S. L., and C. Winship. 2007. *Counterfactuals and causal inference. Methods and principles for social research*. Cambridge: Cambridge University Press.

Mukherjee, B., D. L. Smith, and Q. Li. 2009. Labor (im) mobility and the politics of trade protection in majoritarian democracies. *Journal of Politics* 71(1):291–308.

Neumayer, E., and T. Plümper. 2017. *Robustness tests for quantitative research*. Cambridge University Press.

Neumayer, E., and T. Plümper. 2016. W. *Political Science Research and Methods* 4(1):175–193.

Neyman, J., and E. Scott. 1948. Consistent estimates based on partially consistent observations. *Econometrica* 16:1–32.

Nickell, S. 1981. Biases in dynamic models with fixed effects. *Econometrica* 49:1417–1426.

North, D. C. 1990. *Institutions, institutional change, and economic performance*. Cambridge: Cambridge University Press.

Park, J. H. 2012. A unified method for dynamic and cross-sectional heterogeneity: Introducing hidden markov panel models. *American Political Science Review* 56:1040–1054.

Pickup, M. 2017. A general-to-specific approach to dynamic panel models with a very small T. Presented to the 2017 Meeting of the Midwest Political Science Association, Chicago Illinois.

Plümper, T., V. E. Troeger, and P. Manow. 2005. Panel data analysis in comparative politics: Linking method to theory. *European Journal of Political Research* 44(2):327–354.

Plümper, T., and V. E. Troeger. 2007. Efficient estimation of time-invariant and rarely changing variables in finite sample panel analyses with unit fixed effects. *Political Analysis* 15(2):124–139.

Plümper, T., and V. E. Troeger. 2011. Fixed-effects vector decomposition: properties, reliability, and instruments. *Political Analysis* 19(2):147–164.

Ross, M. L. 2008. Oil, Islam, and women. *American Political Science Review* 102(1):107–123.

Soroka, S. N., D. A. Stecula, and C. Wlezien. 2015. It's (change in) the (future) economy, stupid: economic indicators, the media, and public opinion. *American Journal of Political Science* 59(2):457–474.

Treisman, D. 2015. Income, democracy, and leader turnover. *American Journal of Political Science* 59(4):927–942.

Troeger, Vera, and Thomas Pluemper. 2017. Replication Data for: Not so Harmless After All: The Fixed-Effects Model. doi:10.7910/DVN/RAUIHG, Harvard Dataverse.

Wegener, A. 1912. Die Entstehung der Kontinente. *Geologische Rundschau* 3:276–292.

Wilson, S. E., and D. M. Butler. 2007. A lot more to do: The sensitivity of time-series cross-section analyses to simple alternative specifications. *Political Analysis* 15(2):101–123.

Wooldridge, J. M. 2002. *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.