

Fractionally integrated data and the autodistributed lag model: results from a simulation study

Justin Esarey*

July 1, 2015

Abstract

Two contributions in this issue, Grant and Lebo (2015) and Keele, Linn and Webb (2015), recommend using an ARFIMA model to diagnose the presence of and estimate the degree of fractional integration, then either (i) fractionally differencing the data before analysis or, (ii) for cointegrated variables, estimating a fractional error correction model. But Keele, Linn and Webb (2015) also presents evidence that ARFIMA models yield misleading indicators of the presence and degree of fractional integration in a series with fewer than 1000 observations. In a simulation study, I find evidence that the simple autodistributed lag model (ADL) or equivalent error correction model (ECM) can, without first testing or correcting for fractional integration, provide a useful estimate of the immediate and long-run effects of weakly exogenous variables in fractionally integrated (but stationary) data.

Introduction

Political scientists strive to study the substantive meaning of temporal dynamics rather than to treat them as nuisances to be fixed in the error term (Beck and Katz, 1996). If the autodistributed lag (ADL) and mathematically equivalent error correction models (ECM) described by De Boef and Keele (2008) have been inappropriately used in some instances, I suspect that it is out of a laudable desire to pursue this agenda. This issue's article by Grant and Lebo (2015) reminds us that how we handle the nuisances is still important, and that doing so improperly can lead to "discovering" non-existent relationships. In particular, I concur with Grant and Lebo's advice to carefully assess the stationarity of variables prior

*Assistant Professor of Political Science, Rice University. E-mail: justin@justinesarey.com.

to analysis: time series analysis of non-stationary variables without appropriate differencing (or an appropriate ECM, if variables are cointegrated) can yield misleading results. This advice is also encapsulated in Keele, Linn, and Webb's (2015) very helpful Table 5.

But the analysis guidelines produced by Grant and Lebo (2015) and Keele, Linn and Webb (2015) raise a significant question about the handling of fractionally integrated data. When fractional integration is suspected in a time series, both articles recommend measuring the degree of fractional integration d with an ARFIMA model, then using d to either (i) fractionally difference the data before analysis, or (ii) when co-integration is present, estimate a fractional ECM model. But Keele, Linn and Webb (2015) are *also* justifiably skeptical of the ARFIMA model's ability to recover d in a data set with a small number of temporal observations T . First, they present simulation evidence (in Figure 1) that ARFIMA estimates of d are both noisy and biased when $T < 1000$. They also show that the ARFIMA model frequently produces false positive results for d in non-fractionally-integrated data under the same circumstances. They cite prior studies indicating that these limitations of the ARFIMA model have been observed by other researchers. Finally, they express concern that complex models in very short data sets (with many parameters per observation) are likely to be overfitted. If fractionally integrated data are both ubiquitous (as Grant and Lebo (2015) suggest they are) and difficult to study in short data sets, the evidence offered by Keele, Linn and Webb (2015) suggests that a large number of important questions in political science may be very difficult to answer using the recommended methodology.

I examine an alternative: if data might be fractionally integrated but are stationary (and the independent variables are weakly exogenous), estimate an ADL/ECM model on the data without first estimating d and fractionally differencing. Although this model is undoubtedly misspecified, it may nevertheless provide an accurate approximation of important dynamic relationships in the data. Moreover, it is considerably simpler than the ARFIMA model and perhaps less-susceptible to over-fitting. In a simulation study, I find evidence that an ADL/ECM can accurately detect and recover immediate and long-run relationships in this

setting while avoiding false positives. Consequently, the ADL/ECM appears to be a valid option for studying short T data sets with fractional integration. The results suggest that dealing with the non-stationarity and/or co-integration of time series data is a methodologically higher priority compared to correcting for possible fractional integration, and that researchers may generally trust the results of ADL/ECM models in this environment.

Approximating ARFIMA data with ADL/ECM models

The general form of an ARFIMA process for a variable y , as defined by Shumway and Stoffer (2010, p. 272), is written as:

$$\phi(L)(1 - L)^d(y_t - \mu_t) = \theta(L)\varepsilon_t \quad (1)$$

where $t = 1 \dots T$ indexes time, a non-integer d indicates a degree of fractional differencing, $\phi(L)$ is an autoregressive function of lag operators¹ L which acts on the fractionally-differenced and de-measured y ,² and $\theta(L)$ is a moving average function lag operators which acts on the white noise term ε .³ When y is suspected to be fractionally integrated, Grant and Lebo (2015) and Keele, Linn and Webb (2015) recommend estimating d using an ARFIMA model that matches this process, then fractionally differencing y before determining its relationship with other variables like x . x may also need to be differenced prior to analysis as well; a fractional ECM may be possible if x and y share the same d .

Given the objections to ARFIMA modeling in short T data sets raised by Keele, Linn and Webb (2015), it may be possible to use the ADL/ECM to *approximate* a fractionally integrated data generating process in order to recover both immediate and long-run relationships dy/dx when x is weakly exogenous with respect to y .⁴ As long as $d \in (-\frac{1}{2}, \frac{1}{2})$, a series

¹The lag operator is: $L(x_t) = x_{t-1}$, with higher multiples of L leading to deeper lags (e.g., $L^2(x_t) = x_{t-2}$).

²Specifically, $\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p$ with p the order of the AR process.

³Specifically, $\theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q$ with q the order of the MA process. For more details on ARFIMA modeling, see Shumway and Stoffer (2010) pp. 12, 85, and 90.

⁴Weak exogeneity implies that y does not have a direct or indirect (error-mediated) causal impact on x ;

like equation 1 is stationary (Shumway and Stoffer, 2010, p. 269); De Boef and Keele (2008) demonstrated that the ADL/ECM can be very useful for studying dynamic relationships in stationary data.

The ADL/ECM is obviously misspecified for the data generating process in equation (1). The intent is not to precisely mirror the data generating process, but to approximate immediate and long-run relationships between x and y within an acceptable degree of error. Given the problems that Keele, Linn and Webb (2015) identify in estimating d in short panels (and the complexity of time series analysis in general), some form of misspecification may be inevitable. Moreover, because the ADL/ECM is a relatively simple model, the risk of overfitting may be reduced compared to the more complex ARFIMA model.

Grant and Lebo (2015) is primarily motivated by imprudent use of the ADL/ECM model that often finds relationships among variables where none exist. Thus, to recommend the use of ADL/ECM models to study fractionally integrated data without first fractionally differencing, it is imperative to demonstrate that this use does not encounter the problems that they identify. The key issue is whether the long-run memory present in a series with fractional integration creates spurious or severely biased estimates of short- or long-run relationships between x and y . I answer this question using a simulation study.

Monte Carlo evidence

Consonant with the concerns of Grant and Lebo (2015), my simulation study is designed to answer four questions:

1. Do ADL/ECM models find immediate or instantaneous relationships in fractionally integrated data where they do not exist?
2. Do ADL/ECM models find long-term or dynamic relationships in fractionally integrated data where they do not exist?

see Enders (2015, pp. 394-395) for details.

3. Can ADL/ECM models accurately recover the magnitude and direction of immediate/instantaneous relationships in fractionally integrated data where they do exist?
4. Can ADL/ECM models accurately recover the magnitude and direction of long term relationships in fractionally integrated data where they do exist?

To answer these questions, I create three simulated data generating processes with different characteristics. Specifically, I vary the relationship between x and y : the possibilities are that (1) changes in x have no impact on y , (2) permanent changes in x at time t^* have an immediate, permanent impact on y at time t^* but no further impact, and (3) permanent changes in x have an immediate impact on y that continues to increase over time as a long-run adjustment in y . Because Keele, Linn and Webb (2015) note that difficulties in estimating d are most acute in short data sets, I assess the ADL's suitability for time series with $T = 100$; this is sufficiently short that we might be skeptical of estimates of d from an ARFIMA model.

Fractionally-integrated DGPs without long-term adjustment

To simulate fractionally integrated data for y and x , I create data using an ARFIMA process with the form:

$$(1 - 0.5L)(1 - L)^d(y_t - \mu_t) = \varepsilon_t \quad (2)$$

$$\mu_t = \gamma_x x_t \quad (3)$$

$$(1 - 0.5L)(1 - L)^{0.3}x_t = \psi_t \quad (4)$$

$\gamma_x = 0$ when there is no relationship between x and y ; I set $\gamma_x = 0.5$ when there is such a relationship. In this ARFIMA process, the short-run and long-run impacts of x are identical; the mean of y shifts immediately to reflect a change in x , and fluctuations around this mean (which are subject to long memory) are unrelated to the value of x . I vary $d \in$

$\{0, 0.1, 0.2, 0.3, 0.4, 0.45\}$, similar to Keele, Linn and Webb (2015). The noise terms ε and ψ are $\sim N(0, 1)$. Series for y and x of length $T = 100$ out of this process are generated using the `arfima.sim` function in the `arfima` package for R. I draw 1000 data sets for each Monte Carlo study.

For each data set, I estimate an ADL model⁵ of the form:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 \Delta x_t + \beta_3 x_{t-1} + \zeta_t \quad (5)$$

where $\Delta x_t = x_t - x_{t-1}$ and its coefficient β_2 shows the immediate impact of a change in x on y at the time of the change, t^* . The long run impact of a permanent change in x on y is given by $LRM = \beta_3 / (1 - \beta_1)$; I estimate the Bewley transformation of the model (described in De Boef and Keele, 2008) in order to measure this impact and its variance.⁶

False positive rates for immediate and long-run impacts

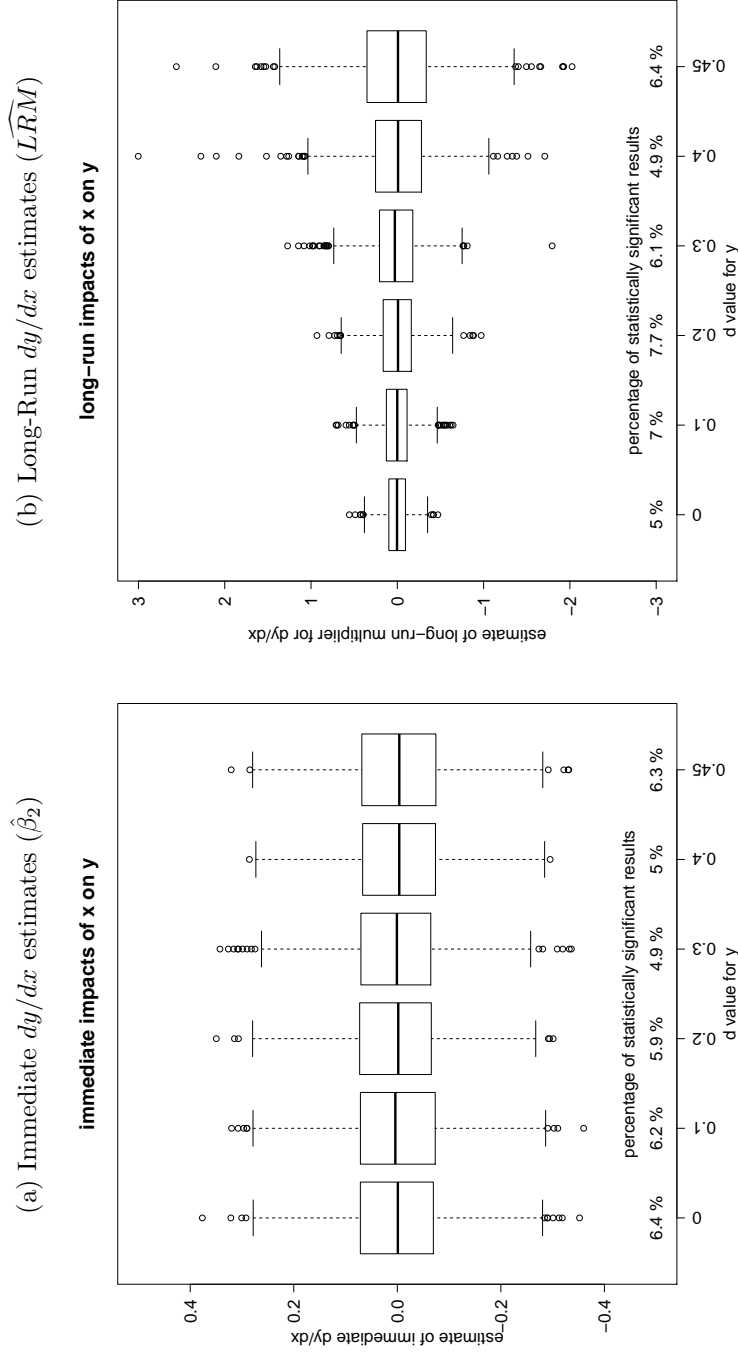
When $\gamma_x = 0$ in equation (3), there is no immediate or long-term impact of x on y . In this environment, I designate a false positive immediate impact as a statistically significant value of $\hat{\beta}_2$ using a two-tailed t -test, $\alpha = 0.05$. I similarly designate a false positive long run impact as a statistically significant \widehat{LRM} using the same test.

Consider Figure 1, which shows the estimated values of $\hat{\beta}_2$ and \widehat{LRM} for each of the 1000 simulated data sets and the percentage of estimates that are statistically significant. As the figures make clear, both the immediate impacts (estimates of $\hat{\beta}_2$ from equation 5) and long run impacts (estimates of \widehat{LRM} using the Bewley method) are consistently centered on zero with false positive rates near the nominal $\alpha = 0.05$ value of the statistical significance test. In other words, in fractionally integrated data, the ADL is resistant to finding immediate and long-run relationships between y and x where they do not exist.

⁵There is a mathematically equivalent ECM for this model, as laid out in De Boef and Keele (2008) and Keele, Linn and Webb (2015); I focus on the ADL formulation for ease of interpretation.

⁶The Bewley model estimates $y_t = \alpha_0 + \alpha_1 \Delta y_t + \alpha_2 x_t + \alpha_3 \Delta x_t + \eta_t$, using y_{t-1} as an instrument for Δy_t . The coefficient α_2 is the estimate of LRM , with its estimated variance as the variance of this impact.

Figure 1: ADL approximation of null ARFIMA models



These figures depict the result of analyzing 1000 simulated data sets from the ARFIMA process in (equations 2-4) with $\gamma_x = 0$ using the ADL model in equation (5). The left figure depicts the estimates of $\hat{\beta}_2$ from the model; these should be centered on 0 if the estimates are accurate. The line of numbers at the bottom of the left figure depicts the proportion of $\hat{\beta}_2$ estimates that are statistically significant ($\alpha = 0.05$, two-tailed) for each value of d . The right figure depicts the estimates of \widehat{LRM} from the Bewley transformation of the model; these should also be centered on 0 if the estimates are accurate. The line of numbers at the bottom of the right figure depicts the proportion of \widehat{LRM} estimates that are statistically significant ($\alpha = 0.05$, two-tailed) for each value of d .

True positive rates and accuracy estimates

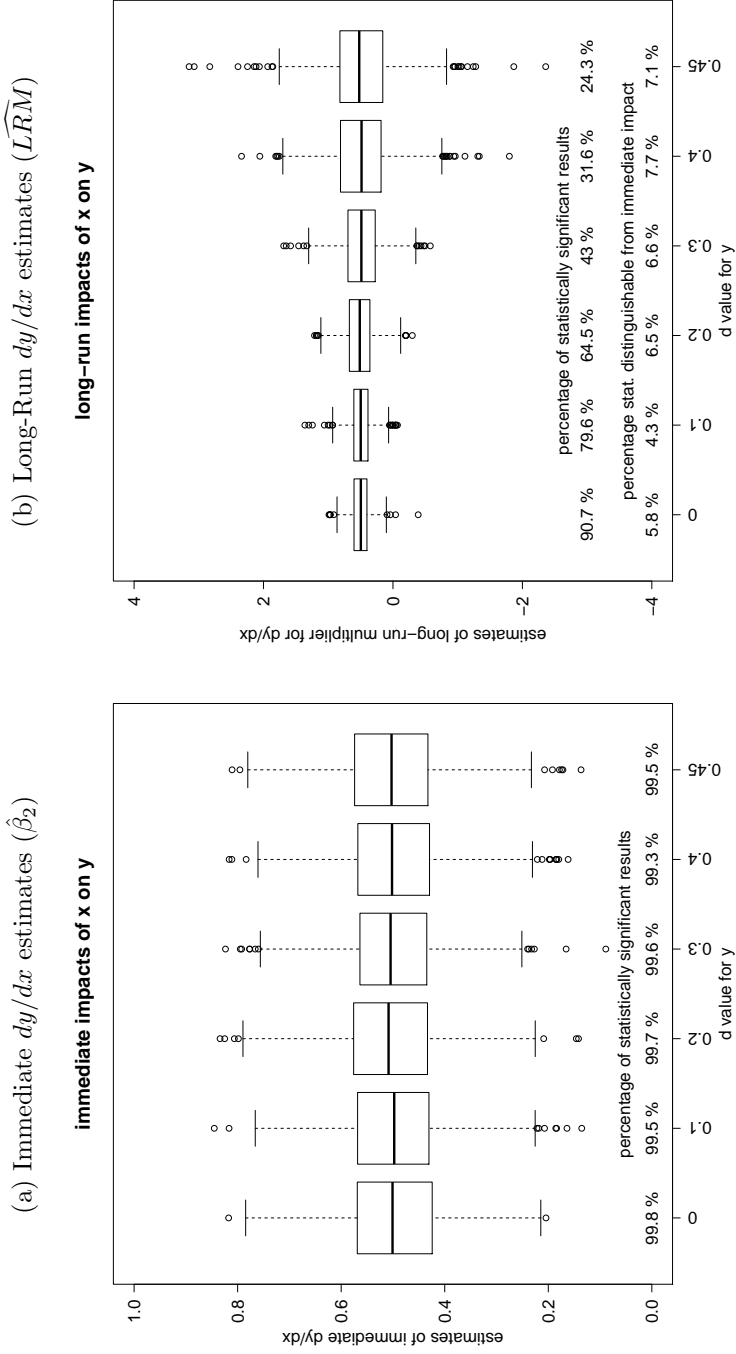
When $\gamma_x = 0.5$ in equation (3), the immediate impact of y on x is the same as the long-run impact: 0.5. If the ADL model can accurately approximate the relationships in this data set, it should show an accurate and identical immediate and long-term impact ($\hat{\beta}_2 = \widehat{LRM} = 0.5$). I designate a true positive immediate impact as a statistically significant value of $\hat{\beta}_2$ using a two-tailed t -test, $\alpha = 0.05$; I use the same test for detecting true positive \widehat{LRM} values.

Figure 2 shows the estimated values of $\hat{\beta}_2$ and \widehat{LRM} for 1000 data sets simulated under these conditions. Both immediate and long-run impact estimates are properly centered on the correct value of 0.5. However, the degree of noise in the estimate of long-run impacts increases as d gets closer to 0.5 and y gets closer to being non-stationary. This additional noise hurts the ADL model's ability to distinguish the estimated LRM from zero; when $d = 0.45$, only about a quarter of true positive long-run relationships are detected.

However, I believe that in the presence of an immediate impact of x on y , it may sometimes be reasonable to assert that the null expectation for the long-run relationship should be equal to the immediate impact. If changes in x are not theoretically expected to “wear off” over time, then the effect of y caused by a change in x at time t^* should persist by inertia beyond that time point. Consequently, I also test the hypothesis that $\beta_2 \neq LRM$ against the null that $\beta_2 = LRM$, showing the results in a second line of numbers in Figure 2b.⁷ For this test, the ADL (falsely) rejects this null only slightly more often than the expected $\alpha = 0.05$ rate.

⁷To test the hypothesis that $\beta_2 \neq LRM$, I draw 1000 simulated values from asymptotic distribution of the ADL model and subtract the draws for $\hat{\beta}_2$ from those for the $LRM = \hat{\beta}_3 / (1 - \hat{\beta}_1)$, subtract the 1000 draws, then use the difference as a parametric bootstrap approximation of the difference between the immediate and long run impact. I then examine the 2.5th and 97.5th quantiles of these differences to test whether any difference is statistically significant.

Figure 2: ADL approximation of ARFIMA models with identical immediate and long-run impacts



These figures depict the result of analyzing 1000 simulated data sets from the ARFIMA process in (equations 2-4) with $\gamma_x = 0.5$ using the ADL model in equation (5). The left figure depicts the estimates of $\hat{\beta}_2$ from the model; these should be centered on 0.5 if the estimates are accurate. The line of numbers at the bottom of the left figure depicts the proportion of $\hat{\beta}_2$ estimates that are statistically significant ($\alpha = 0.05$, two-tailed) for each value of d . The right figure depicts the estimates of \widehat{LRM} from the Bewley transformation of the model; these should also be centered on 0.5 if the estimates are accurate. The top line of numbers at the bottom of the right figure depicts the proportion of \widehat{LRM} estimates that are statistically significant ($\alpha = 0.05$, two-tailed) for each value of d . The bottom line of numbers at the bottom of the right figure depicts the proportion of \widehat{LRM} estimates that are statistically distinguishable from $\hat{\beta}_2$ using the same test criterion.

Fractionally-integrated DGPs that include long-term adjustment

To simulate fractionally integrated data for y and x that includes a long-term adjustment, I slightly modify the ARFIMA process in equations (2-4):

$$(1 - 0.5L)(1 - L)^d(y_t) = \kappa_t \quad (6)$$

$$\kappa_t = \gamma_x x_t + \varepsilon_t \quad (7)$$

$$(1 - 0.5L)(1 - L)^{0.3}x_t = \psi_t \quad (8)$$

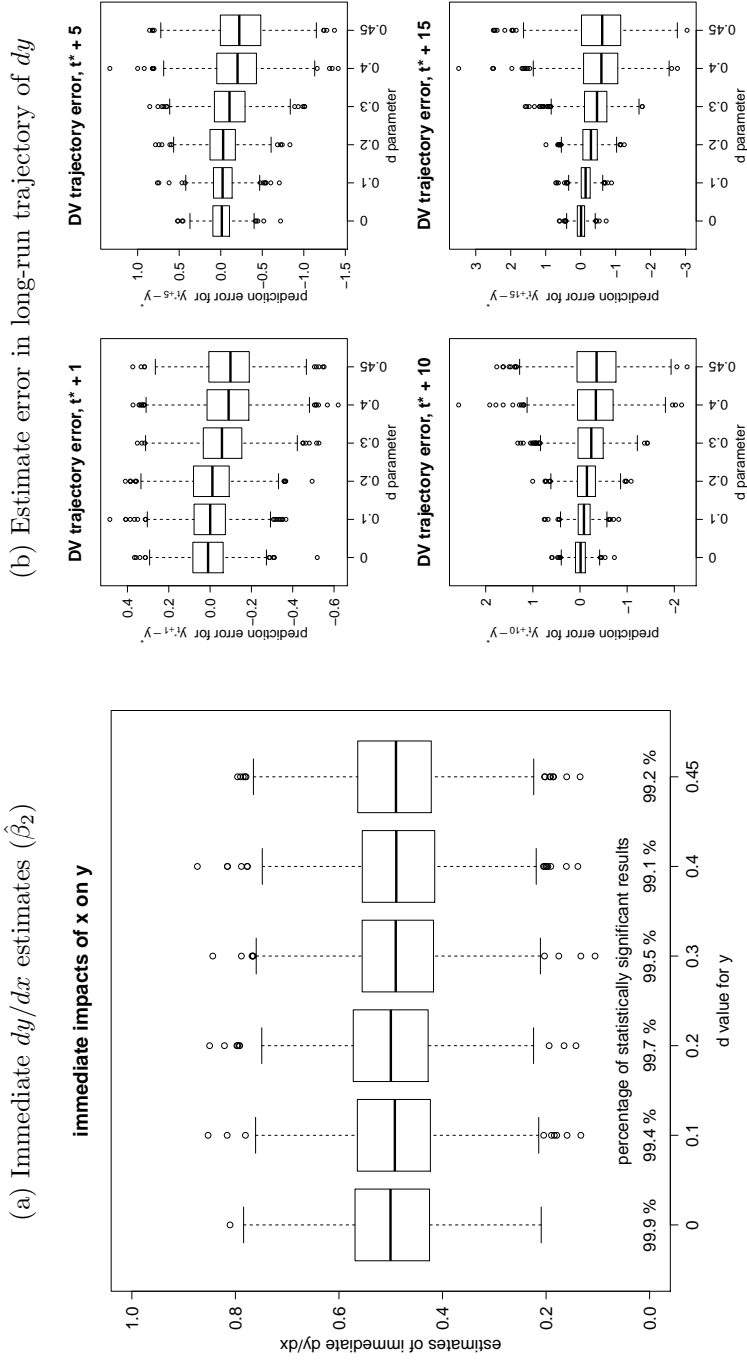
This allows changes in x to propagate into the ARFIMA process by directly entering the noise term; a change in x at time t^* will have continuing impacts on y after that time as the initial impacts echo through the lag structure of y .

I assess the accuracy of the immediate impact of x on y as before: $\hat{\beta}_2$ should be statistically significant and $= 0.5$ if the ADL yields accurate results. Given the existence of a gradual, long-term adjustment in y initiated by a change in x , I also assess how well the ADL can match the trajectory of change in y following an single permanent change in x . Specifically, once an ADL model is fitted, I set $x = 0$, set the lagged value of y_{t-1} in order to put the system into an equilibrium y^* , simulate a change in x of 1 at time t^* , then calculate \hat{y} from this model for $t^* + c$ from $c = 1 \dots 15$. I then compare the difference between y^* and \hat{y}_{t^*+c} calculated from the ADL to the *true* difference as simulated using the true parameters and `arfima.sim`. This process gives a sense of how well the ADL is able to approximate the unfolding of the data generating process over time after a change in x .

The results are shown in Figure 3. As in the case with identical immediate and long-run impacts, the ADL does an excellent job of accurately measuring the immediate dy/dx and rejecting the null hypothesis. The estimated \widehat{LRM} is statistically significant over 93% of the time (and over 98% when $d < 0.45$); additionally, it is statistically distinguishable from the immediate impact $\hat{\beta}_2$ over 92% of the time.⁸

⁸I use the same parametric bootstrapping technique for this test as laid out in footnote 7.

Figure 3: ADL approximation of ARFIMA models with long-run adjustment



These figures depict the result of analyzing 1000 simulated data sets from the ARFIMA process in (equations 6-8) with $\gamma_x = 0.5$ using the ADL model in equation (5). The left figure depicts the estimates of $\hat{\beta}_2$ from the model; these should be centered on 0.5 if the estimates are accurate. The line of numbers at the bottom of the left figure depicts the proportion of $\hat{\beta}_2$ estimates that are statistically significant ($\alpha = 0.05$, two-tailed) for each value of d . The right figure depicts the accuracy of ADL estimates of the change in y at time $t^* + c$ (compared its initial equilibrium) caused by a one-time permanent change in x at time t^* ; these should be centered on 0 if the ADL estimates are accurate (as the figure shows the difference between the true trajectory and the ADL estimate).

As shown in Figure 3b, the trajectory of long-run changes in y over time is estimated with progressively greater noise as the temporal distance from the intervention gets larger; this noise also gets larger as d grows. Additionally, there is a tendency of the ADL model to underestimate the magnitude of long-run changes, especially when d is close to 0.5. The underestimation of long-run impacts makes sense: fractionally integrated time series are equivalent to very long autodistributed lag models (Shumway and Stoffer, 2010, pp. 268-269) while the ADL includes just one lag. Consequently, a change in x has effects that propagate cumulatively over a long period of time; the ADL must approximate this process with a much smaller number of lags of y .

Conclusion

My simulation study leads to four conclusions about the use of ADL/ECM models to study the immediate and long-run effects of a fractionally integrated (but stationary) and weakly exogenous variable x on a fractionally integrated y . Under the conditions of the simulation:

1. ADL/ECM models **did not find immediate or instantaneous effects of x on y in fractionally integrated data where they do not exist**; standard t -tests for the coefficient on Δx_t in the ADL produced false positives at close to the expected rate.
2. ADL/ECM models **did not find long-run effects of x on y in fractionally integrated data where they do not exist**. Again, standard t -tests for the long-run multiplier (LRM) estimated by the Bewley transformation of the ADL produced false positives at close to the expected rate.
3. ADL/ECM models **accurately identified and recovered the magnitude and direction of immediate impacts of x on y in fractionally integrated data where they existed**.
4. ADL/ECM models **detected the presence of long-run impacts greater than**

the immediate effect of x on y in fractionally integrated data, but the **magnitudes were underestimated** and tests for distinguishability from the immediate impact were more useful than tests against a long-run impact of zero.

Based on these conclusions, it appears that ADL/ECM models are very useful for recovering the immediate impact of x on y , despite fractional integration. The results for long-run impacts are not quite as robust: these impacts are likely to be underestimated by an ADL/ECM run on fractionally integrated data, possibly because the very long memory of such series allows for an especially extended impact on y of a one-time permanent change in x . Nevertheless, hypothesis tests on the *LRM* do allow it to be distinguished from immediate impacts where appropriate, and do *not* allow it to be distinguished when there is no long-term cumulative effect.

My overall recommendation is to slightly refine the advice of Grant and Lebo (2015) and Keele, Linn and Webb (2015). In non-fractionally-cointegrated data sets with many temporal observations T , it seems appropriate to estimate d with an ARFIMA model and fractionally difference a variable prior to estimation as indicated in Keele, Linn, and Webb's Table 5. But an ADL/ECM provides a serviceable approximation in a short T data set, where d is inaccurately estimated and overfitting is a concern. This recommendation does not absolve a researcher of the responsibility to establish that the studied data are stationary (and that the independent variables are weakly exogenous) before applying the ADL/ECM; based on prior research, I would still expect the ADL/ECM to be a very problematic choice for non-stationary (and non-cointegrated) or endogenously related variables.

References

- Beck, Nathaniel and Jonathan N. Katz. 1996. "Nuisance vs. Substance: Specifying and Estimating Time-Series-Cross-Section Models." *Political Analysis* 6(1):1–36.
- De Boef, Suzanna and Luke Keele. 2008. "Taking Time Seriously." *American Journal of Political Science* 52:184–200.
- Enders, Walter. 2015. *Applied Econometric Time Series, 4th edition*. Wiley.

Grant, Taylor and Matthew Lebo. 2015. "Error Correction Methods with Political Time Series." *Political Analysis* forthcoming.

Keele, Luke, Suzanna Linn and Clayton McLaughlin Webb. 2015. "Treating Time with All Due Seriousness." *Political Analysis* forthcoming.

Shumway, Robert H. and David S. Stoffer. 2010. *Time Series Analysis and Its Applications*. Springer.