Interpreting Lagged Effects of the Independent Variable: How does the Local Economy Affect Welfare Caseloads?

by

Terra G. McKinnish
Department of Economics
Campus Box 256
University of Colorado
Boulder, CO 80309-0256
(303) 492-6770
fax (303) 492-8960
terra.mckinnish@colorado.edu

July 2002

I thank participants in the Measurement Error session at the November 2000 meetings of the Southern Economic Association and the Microeconometric Methods Session at the January 2002 meetings of the Econometrics Society for helpful comments and suggestions. Comments from Dan Black, Hillary Hoynes and Jeffrey Zax were very helpful. Rebecca Blank very generously and quickly shared her data. Financial support from the National Science Foundation under grant #SES-9819142 is gratefully acknowledged.

Abstract

Coefficient estimates on lagged independent variables, such as those recently estimated in the AFDC caseload literature, potentially reflect omitted variable or measurement error bias. This paper demonstrates that specification tests will generally not distinguish between a true lagged effect, measurement error and omitted variable bias. I argue that lagged effects often reflect the differential effect of short and long-term changes in conditions. A comparison of specification test results analyzing AFDC caseload to those obtained analyzing UI caseloads provides some support for this claim, suggesting that the lagged effects estimated in the AFDC caseload literature are not mere artifacts of omitted variable or measurement error bias.

1. Introduction

Panel data models have become very popular because they can eliminate bias due to time-constant omitted variables. There is often the concern, however, that other forms of misspecification remain, such as omitted lagged effects of the independent variable, measurement error, and omitted time-varying characteristics. In particular, it has become increasingly the case that researchers use panel data to investigate the presence of lagged effects of economic, demographic or government policy variables. The coefficients on the contemporaneous and lagged variables are combined to obtain a long-run effect.

This paper investigates misspecification in panel data models to make the following points. First, coefficient estimates on lagged independent variables may merely reflect the presence of omitted variable or measurement error bias. Therefore, adding up the contemporaneous and lagged coefficients can actually increase, rather than reduce, bias. Second, specification tests based on the data at hand (as opposed to external data, such as external instruments), will generally not allow one to distinguish between a true lagged effect, measurement error and omitted variable bias. Finally, this paper argues that lagged effects that do not reflect omitted variable bias or bias due to a mismeasured independent variable often reflect the differential effect of short and long-term changes in conditions.

The first two points are demonstrated by analytical results that compare the outcomes of three different specification tests in the presence of each of the three forms of misspecification: omitted lag, measurement error and omitted time-varying characteristic. The three specification tests are the long-differences specification recommended by Griliches and Hausman (1986) to test for measurement error, the Heckman-Hotz (1989) test designed to test the fixed-effects assumption, and the addition of the lagged value of the independent variable to check for omitted lags. The analytical results demonstrate that under very reasonable conditions, all three forms of misspecification can produce indistinguishable failures of each specification test. The specification tests generally cannot be used to distinguish between the three forms of misspecification discussed.

This paper illustrates these points with an analysis of the effect of local economic conditions on participation in Aid to Families with Dependent Children (AFDC).[1] The literature on AFDC caseloads has recently given rise to a number of papers that have estimated sizeable lagged effects of economic conditions. Specification test results are consistent with all three misspecifications, which leaves unsettled the question of whether the lagged effects can be interpreted as "real." This paper argues that these lagged effects reflect the fact that AFDC participation responds more to long-term changes in economic conditions than transitory fluctuations. Given contemporaneous changes in economic conditions, lagged changes in economic conditions provide a measure of the duration of the change. Therefore, the estimated lagged effects in the AFDC literature may reflect this differential response to short and long-term changes in economic conditions.

In order to test this hypothesis, I assume that Unemployment Insurance (UI) participation should be substantially more sensitive to short-term variation in economic conditions than AFDC participation. If true, the analytical results predict a number of differences in the specification test results when the dependent variable is changed to measure the size of the UI program. Some of these predicted changes are observed using state-level data and all are observed using county-level data.

[1] AFDC was a welfare program that provided monthly benefits to low-income single mothers, which was replaced with Temporary Aid to Needy Families (TANF) by the welfare reform bill passed in 1996. Like AFDC, TANF provides monthly benefits to low-income families with children, but TANF contains provisions such as time limits and work requirements that were not present in AFDC.

2. Models and Specification Tests

If the true model is:

$$Y_{it} = \alpha_{0i} + \alpha_1 X_{it} + \varepsilon_{it},$$

then estimating the model in differences,

$$(Y_{it} - Y_{it-j}) = \alpha_1 (X_{it} - X_{it-j}) + (\varepsilon_{it} - \varepsilon_{it-j}), \tag{1}$$

yields a consistent estimate of α_1. Three models below provide simple representations of the most common specification problems that produce inconsistent estimates from the estimation of equation (1).

The Omitted Lag (OL) model contains a lagged value of the independent variable:

$$Y_{it} = \alpha_{0i} + \alpha_1 X_{it} + \alpha_2 X_{it-1} + \varepsilon_{it}, \tag{2}$$

$$X_{it} = \rho X_{it-1} + \theta_{it}, \tag{3}$$

where ε and θ are i.i.d disturbances with means zero and variances σ_ε^2 and σ_θ^2, respectively.

For the measurement error (ME) model, the true model is:

$$Y_{it} = \alpha_{0i} + \alpha_1 Z_{it} + \varepsilon_{it}, \tag{4}$$

but *Z* is not observed directly, only a noisy realization:

$$X_{it} = Z_{it} + v_{it}. \tag{5}$$

Where:

$$Z_{it} = \rho Z_{it-1} + \theta_{it}, \tag{6}$$

and ε, v, and θ are all i.i.d. with means zero and variances $\sigma_\varepsilon^2, \sigma_v^2$ and σ_θ^2 respectively.

In the Omitted Variable (OV) model, there exists an unobserved characteristic that is correlated with both the dependent and independent variables. This variable is not time-constant;

therefore differencing the data does not eliminate the confounding variable. One very simple model with this property is:

$$Y_{it} = \alpha_{0i} + \alpha_1 X_{it} + \alpha_2 W_{it} + \varepsilon_{it}, \tag{7}$$

$$X_{it} = \rho X_{it-1} + \theta_{it}, \tag{8}$$

$$W_{it} = \delta W_{it-1} + \theta_{it}, \tag{9}$$

where W is unobserved and ε and θ are both i.i.d disturbances with means zero and variances σ_ε^2 and σ_θ^2, respectively.

Specification tests can be estimated for each of the above forms of misspecification. To check for omitted lagged effects of the independent variable, a lag is simply included in the model:

$$Y_{it} - Y_{it-j} = \beta_j (X_{it} - X_{it-j}) + \gamma_j (X_{it-1} - X_{it-j-1}) + (e_{it} - e_{it-j}). \tag{10}$$

Specifications of this form are used in a wide variety of studies.[2] A good example of a literature in which lagged values of the independent variable are included in the model is the recent literature on AFDC caseloads. Models with lagged values of economic conditions have been estimated in work by Bartik and Eberts (1999), CEA (1999), Figlio and Ziliak (1999), Wallace and Blank (1999), Ziliak, Figlio, Davis and Connolly (2000), Mueser et al. (2000), and Blank (2001).[3]

For the measurement error model, Griliches and Hausman (1986) recommend estimating:

$$Y_{it} - Y_{it-j} = \beta_j (X_{it} - X_{it-j}) + (e_{it} - e_{it-j}), \tag{11}$$

---

[2] See, for example, Krueger and Rouse (1998), Hildreth and Oswald (1997) and Van Reened (1997).

[3] Some of these papers use fixed-effects, rather than differences specifications. This paper does not consider a fixed-effects specifications for several reasons. First, the fixed-effects specification cannot be used for the Heckman-Hotz test. Second, the analytical predictions for fixed-effects specifications rely on very specific characteristics of the underlying correlation structures generating the right-hand side variables. Finally, because the ability to make strong predictions about the outcome of the specification tests estimated in a fixed-effects format is limited, adding the fixed-effects form of the specification tests is unlikely to improve our ability to distinguish between the different forms of misspecification. If the more restrictive differences forms of the specification tests cannot be used to

for different values of j. Griliches and Hausman show that under common assumptions $\left|\text{plim}(\hat{\beta}_{j+1})\right| > \left|\text{plim}(\hat{\beta}_j)\right|$ if the independent variable is measured with error.

Heckman and Hotz (1989) suggest a simple method of testing the fixed-effects assumption. They propose including a lagged value of the dependent variable in the regression:

$$Y_{it} - Y_{it-j} = \beta_j (X_{it} - X_{it-j}) + \lambda_j Y_{it-j-1} + (e_{it} - e_{it-j}) . \tag{12}$$

If the differences model in equation (1) is correctly specified, then the coefficient on a lagged value of the dependent variable should be zero. If, however, there are unobserved time-varying confounders, then the lagged value of the dependent variable will be correlated with these omitted characteristics, and the coefficient on the lagged value of Y should be non-zero.[4]

3. Analytical Results

In this section, the results of each specification test, under each of the three models described in equations (2)-(9), are predicted. Specifically, what is the predicted sign of the relevant coefficient? Second, does the magnitude of the coefficient increase or decrease as longer differences are used (as j increases)?

In all cases, it is assumed that all error terms ($\varepsilon, v,$ and θ) are "best case" i.i.d. disturbances and that the right-hand side variables are generated by an AR(1) process so that they are positively correlated, but changing, over time ($0 < \rho < 1$ and $0 < \delta < 1$). It is further assumed that the coefficient on the omitted variable or omitted lag, α_2, is of the same sign as α_1.[5] Proofs of all propositions appear in Appendix A.

---

distinguish between the models, the fixed-effects forms certainly will not.

[4] The Heckman-Hotz test has been implemented diverse literatures. See, for example: Raaum and Torp, 1999; Regner, 1999; Neumark and Wascher, 1992; Baker, Benjamin and Stanger, 1999; Black and Nagin, 1998.

[5] Without these assumptions, it is very difficult to make analytical predictions about the results of these specification tests. Generalizing the three models would merely reinforce the results of this paper. Under more general models, the specification tests could produce coefficient estimates of either sign and with a wide range of patterns in j. If these three specification tests cannot distinguish between the three forms of misspecification when they are

Proposition 1: Estimating the long-differences model in equation (11) in the case of the OL model produces a coefficient estimate with the properties:

a) The sign of $\text{plim}(\hat{\beta}_{\text{j}})$ is the sign of α_1, $\forall j > 1$

b) For j=1, the sign of $\text{plim}(\hat{\beta}_{\text{j}})$ is the sign of α_1 iff $2\alpha_1 > \alpha_2(1-\rho)$.

c) $\left|\text{plim}(\hat{\beta}_{j+1})\right| > \left|\text{plim}(\hat{\beta}_j)\right| \;\; \forall j$

Proposition 2: Estimating the long-differences model described in equation (11) in the case of the ME model produces a coefficient estimate with the properties:

a) The sign of $\text{plim}(\hat{\beta}_{\text{j}})$ is the sign of α_1, $\forall j$

b) $\left|\text{plim}(\hat{\beta}_{j+1})\right| > \left|\text{plim}(\hat{\beta}_j)\right| \;\; \forall j$

Proposition 3: Estimating the long-differences model described in equation (11) in the case of the OV model produces a coefficient estimate with the properties:

a) The sign of $\text{plim}(\hat{\beta}_{\text{j}})$ is the sign of α_1, $\forall j$

b) $\left|\text{plim}(\hat{\beta}_{j+1})\right| > \left|\text{plim}(\hat{\beta}_j)\right| \;\; \forall j$ iff $\delta > \rho$

Propositions 1-3 therefore show that all three forms of misspecification can generate coefficient estimates from the long-differences model that are the same sign as the true coefficient and increase in magnitude as longer differences are used

Proposition 4: Estimating the lagged-effect model described in equation (10) in the case of the OL model produces a consistent coefficient estimate.

Proposition 5: Estimating the lagged-effect model described in equation (10) in the case of an OL model with an additional lagged effect:

represented by such specific models, then they certainly could not distinguish between the three forms of misspecification if more general representations were allowed.

$$Y_{it} = \alpha_{0i} + \alpha_1 X_{it} + \alpha_2 X_{it-1} + \alpha_3 X_{it-2} + \varepsilon_{it} .$$

produces a coefficient estimate with the properties:

a) The sign of $\text{plim}(\hat{\gamma}_{\text{j}})$ is the sign of α_1, $\forall j > 1$

b) For j=1, the sign of $\text{plim}(\hat{\gamma}_{\text{j}})$ is the sign of α_1 iff $5\alpha_2 > \alpha_3(1-\rho)(2-\rho)(1+\rho)$.

c) $\left|\text{plim}(\hat{\gamma}_{j+1})\right| > \left|\text{plim}(\hat{\gamma}_j)\right| \ \forall j$

Proposition 6: Estimating the lagged-effects model described in equation (10) in the case of the ME model produces a coefficient estimate with the properties:

a) The sign of $\text{plim}(\hat{\gamma}_{\text{j}})$ is the sign of α_1, $\forall j$

b) $\left|\text{plim}(\hat{\gamma}_{j+1})\right| > \left|\text{plim}(\hat{\gamma}_j)\right| \ \forall j > 1$

c) For j=1, $\left|\text{plim}(\hat{\gamma}_{j+1})\right| > \left|\text{plim}(\hat{\gamma}_j)\right|$ iff

$$\sigma_v^4(3\rho - 1) > (1-\rho)^2(1+\rho)(2\sigma_Z^2\sigma_v^2 - \sigma_Z^4(\rho^2 - \rho - 1)(1+\rho))$$

Propositions 4-6 therefore show that all three forms of misspecification can generate coefficients estimates from the lagged-effects model that are the same sign as the true coefficient and increase in magnitude as longer differences are used.

Proposition 7: Estimating the lagged-effect model described in equation (10) in the case of the OV model produces a coefficient estimate with the properties:

a) The sign of $\text{plim}(\hat{\gamma}_{\text{j}})$ is the sign of α_1, $\forall j$

b) $\left|\text{plim}(\hat{\gamma}_{j+1})\right| > \left|\text{plim}(\hat{\gamma}_j)\right| \ \forall j$ iff $\delta > \rho$

Proposition 8: Estimating the Heckman-Hotz test described in equation (12) in the case of the OL model produces a coefficient estimate with the properties:

a) $\text{plim}(\hat{\lambda}_{\text{j}}) < 0, \ \forall j$

b) $\left|\text{plim}(\hat{\lambda}_{j+1})\right| > \left|\text{plim}(\hat{\lambda}_{j})\right| \ \ \forall j$

Proposition 9: Estimating the Heckman-Hotz test described in equation (12) in the case of the ME model produces a coefficient estimate with the properties:

a) $\text{plim}(\hat{\lambda}_{\text{j}}) < 0, \ \forall j$

b) $\left|\text{plim}(\hat{\lambda}_{j+1})\right| > \left|\text{plim}(\hat{\lambda}_{j})\right| \ \ \forall j$

Proposition 10: Estimating the Heckman-Hotz test described in equation (12) in the case of the OV model produces a coefficient estimate with the properties:

a) It is possible that $\text{plim}(\hat{\lambda}_{\text{j}}) < 0, \ \forall j$

b) It is possible that $\left|\text{plim}(\hat{\lambda}_{j+1})\right| > \left|\text{plim}(\hat{\lambda}_{j})\right|$

Proposition 10 is actually true for a large set of parameter values. A large grid search of the parameter space indicated that the coefficient frequently has both of these properties, particularly, but not always, if $\delta > \rho$. Both the expression for the probability limit and the relationships between the underlying parameters and properties of the estimate were sufficiently complex that it was not possible to produce a general set of conditions for which Proposition 10 would hold.

Propositions 1-10 are summarized in Table 1. From the table, it is clear that under very broad conditions, each of the three forms of misspecification generate very similar patterns in the coefficient estimates from the three different specification tests. All three forms of misspecification can generate coefficients in the long-differences and lagged effects models that are the same sign as the true coefficient and increase in magnitude as longer differences are used.

All three forms of misspecification can generate coefficients in the Heckman-Hotz model that are negative and increasing in magnitude as longer differences are used.

4. Application to AFDC Participation

There is a relatively large literature that has examined the relationship between local economic conditions and welfare participation. The earlier research on this topic (Fitzgerald, 1995; Miller and Sanders,1997; Berger and Black, 1998; and Hoynes, 2000) estimated the relationship between contemporaneous economic conditions and AFDC participation using micro-level data.[6] The more recent studies (Bartik and Eberts, 1999; CEA,1999; Figlio and Ziliak, 1999; Wallace and Blank, 1999; Ziliak, Figlio, Davis and Connolly, 2000; Mueser et al., 2000; and Blank, 2001) have used aggregate caseload data and include lagged values of economic conditions in their specification.[7] These studies have found that the resulting long-run elasticities can be substantial in magnitude.

While these previous studies have found that the lagged effects of economic conditions on AFDC participation are important, it is not immediately apparent whether these estimates reflect true lagged effects, or merely measurement error or omitted time-varying characteristics. In this section, we investigate this issue by first estimating the three different specification tests described in equations (10)-(12).

This analysis uses two different data sets. The first data set is that assembled by Blank (2001). Her regression of state-level per capita AFDC caseloads on contemporaneous and

[6]Bollinger and David (1997 and 2001) show that self-reports of program participation, used in much of the micro-data research, contain response bias that is correlated with both the true status and demographic characteristics. Hoynes (2000) using micro-level administrative data and therefore avoids this source of bias.

[7] Ziliak et al (2000) use monthly data and include a lagged value of AFDC caseload in the model to control for what they describe as sluggish adjustment in caseloads. In their specification, the lagged dependent variable belongs in the model and therefore cannot be used as a test of the fixed-effects assumption. Klerman and Haider (2001), discussed in more detail in footnote 8, also point out that the coefficient on the lagged depend variable in the Ziliak et al specification is likely biased due to omitted variables and this bias can cause one to substantially overstate the effect of the independent variables. If the coefficient on the lagged dependent variable does not reflect omitted variable bias, the coefficient should not increase in magnitude as longer-differences are used, as it does in our empirical results.

lagged state unemployment rates is very typical of the regressions estimated in this literature. While Blank estimates fixed-effects models, I re-specify the model in differences, using as a base model:

$$\Delta_j \log(AFDC\_Case_{st} / Pop_{st}) = \beta_0 + \beta_j \Delta_j Unempl\_Rate_{st} + \Delta_j X_{st}\phi + Year_t\theta + \varepsilon_{st} \qquad (13)$$

where $\Delta_j X_t = X_t - X_{t-j}$, *AFDC_Case* is the AFDC caseload, Pop is state population, *Unempl_Rate* is the state unemployment rate, and *Year* is a vector of year indicators. One of Blank's contributions to the literature is a much more detailed set of control variables than had previously been included in caseload regressions. *X* is therefore a vector of state characteristics including data on wages, nonmarital births, female-headed households, immigration, state politics, AFDC and Medicaid benefits, and state demographics.[8]

The second data set contains annual data on county-level employment, AFDC expenditures and population for all counties in the US from 1969-98 obtained from the Bureau of Economic Analysis (BEA) Regional Economic Information System (REIS). Because a full time series of caseload and unemployment data are not available at the county level, per capita AFDC expenditures and the employment-to-population ratio are used instead. The base model is:

$$\Delta_j \log(AFDC\_Exp_{cst} / Pop_{cst}) = \beta_0 + \beta_j \Delta_j (Emp_{cst} / Pop_{cst}) + State_s * Year_t\theta + \varepsilon_{cst} \qquad (14)$$

where $AFDC\_Exp_{cst}$ is AFDC expenditures per capita for county *c* in state *s* in year *t*.[9] *Emp* is county-level employment, and *State* is a vector of state indicators. One advantage of using county-level data is that the state-year effects purge out any characteristics, such as AFDC

[8] Specifically, X includes: log(median wage), log(20th wage percentile), share nonmarital births, share single female heads, avg years of education, share black, share elderly, share immigrants (lagged 1 and 2 years), party of governor, indicator for both state house and senate democrat, indicator for both state house and senate republication, log(maximum AFDC benefit), indicator for AFDC-UP established, log(average medicaid expenditures), indicator for any major welfare reform waiver. I continue Blank's practice of weighting the regression by state population.

[9] Because AFDC benefit levels are set at the state level, however, most of the variation in expenditures at the county level should largely reflect changes in caseload.

benefit levels or other differences in welfare policy, that tend to vary only at the state level. This reduces the potential for omitted-variable bias due to unobserved time-varying characteristics.

The results obtained using Blank's data are reported in Panel A of Table 2. The long-differences results reported in the first column are obtained by estimating equation (13), using j=1 to 5. Lagged effects and Heckman-Hotz results are obtained by separately adding lagged changes in the unemployment rate and lagged values of per capita caseload to the right-hand side of the model.[10] We expect the coefficient on the unemployment rate to be positive. The analytical predictions in Table1 therefore tell us that coefficients on the unemployment rate in the long-differences model that are positive and increasing in magnitude, coefficients on the lagged unemployment rate in the lagged-effects model that are positive and increasing in magnitude and coefficients on lagged per capita caseloads that are negative and increasing in magnitude are consistent with all three models of misspecification. The results in Table 2 are consistent with all three of these predictions.[11]

The results obtained using county-level BEA data are reported in Panel A of Table 3. The long-differences results reported in the first column are obtained by estimating equation (14) using j=1 to 5. Lagged effects and Heckman-Hotz results are obtained by separately adding lagged changes in the employment-to-population ratio and lagged values of per capita AFDC

[10] Klerman and Haider (2001) demonstrate that coefficient estimates from models of the form in equations (13) and (14), even those including multiple lags of economic conditions, are also likely biased because they fail to take into account the fact that caseloads are the net result of flows on and off the program. They show that a simple stock-flow model of welfare caseloads indicates that any regression analysis of the stock requires a full lag structure, and interactions of those lags, that is equal in length to the longest time people spend on the welfare program. Dealing with such a complicated lag structure is beyond the scope of this paper. One way to interpret their analysis of welfare case flows is that lagged values of economic conditions matter because they affect entry and exit to and from welfare in previous periods. Lagged economic conditions therefore affect the composition of the current welfare caseload and the extent to which the current caseload will respond to changes in economic conditions. This explanation suggests a link between the lagged effects model and the omitted time-varying characteristic model, since detailed information on caseload composition could potentially substitute for the lagged effects.

[11] Blank uses two lags of the unemployment rate in her model. While I do not provide analytical results for this case, I re-estimated results using her data with a lag in the baseline model in equation (13) and a second lag in the lagged effects model. The results still displayed the same pattern, except that the coefficient on the lagged effect decreased in magnitude between j=2 and 3. When I performed a similar sensitivity analysis for the UI results reported below, the lagged effect and Heckman-Hotz results displayed the same patterns reported in Panel B of Table 2, but several of the long-differences coefficients were inexplicably negative.

expenditures to the right-hand side of the model. We expect the coefficient on the employment rate to be negative. The analytical predictions in Table1 therefore tell us that coefficients on the employment rate in the long-differences model that are negative and increasing in magnitude, coefficients on the lagged employment rate in the lagged-effects model that are negative and increasing in magnitude and coefficients on lagged per capita AFDC expenditures that are negative and increasing in magnitude are consistent with all three models of misspecification. The results in Table 3 are largely consistent with these predictions. The one violation is that the coefficient on the lagged effect decreases in magnitude between j=3 and j=4. [12]

In Tables 2 and 3, the results indicate that the coefficients on the lagged unemployment or employment rates could reflect true lagged effects, or they could merely reflect the presence of measurement error or omitted time-varying characteristics. As a result, we do not know whether or not the larger long-run effect obtained by adding up the contemporaneous and lagged effects reflects bias due to misspectification.

5. Short-Term Changes, Long-Term Changes and Unemployment Insurance

McKinnish (2000) and Black, McKinnish and Sanders (2002) argue that the coefficient estimates obtained by estimating regressions of the form in equations (13) and (14) are biased due to the fact that some of the time-series variation in economic conditions is highly transitory in nature and unlikely to affect AFDC participation. [13] This suggests an alternative interpretation for the lagged effects estimated in this literature. Given contemporaneous changes in economic conditions, lagged changes in economic conditions provide a measure of the duration of the

[12] One concern might be that the sample shrinks as longer-differences are used. The analysis in Tables 2 and 3 was replicated using only the years after 1982 and 1975, respectively, in order to maintain a relatively constant sample. This has little effect on the results.

[13] Keane and Wolpin (2002) and Baker, Benjamin and Stanger (1999) make similar points in other literatures.

change. Therefore, the estimated lagged effects reflects in the AFDC literature may reflect in part the larger response of AFDC participation to long-term changes in economic conditions.

One way to model this differential effect of short and long-term changes in economic conditions is as a measurement error model in which Z represents sustained changes in economic conditions that do affect AFDC participation and v represents transitory fluctuations in economic conditions that do not affect AFDC participation. If AFDC participation is less responsive to v than Z, the coefficient on contemporaneous economic conditions will suffer from classical measurement error attenuation bias. If v is uncorrelated over time, this model is identical to the ME model in equations (4)-(6).

It would be nice to distinguish between two cases. In the first case, lagged effects, and the larger long-run elasticities resulting from them, are mere artifacts of omitted variable bias or actual mismeasurement of the annual unemployment or employment rate. In this case the larger long-run effects should be disregarded. In the second case, the lagged effects and larger long-run elasticities associated with them reflect the fact that the full effect of economic conditions is in fact larger than what is estimated using only contemporaneous economic conditions, perhaps because of the differential impact of short-term and long-term changes. One way to at least provide empirical evidence that is consistent with the second case is to find a dependent variable similar to AFDC expenditures, but one that is more responsive to transitory changes in economic conditions than AFDC expenditures. If such a dependent variable were used, then there should be a very different pattern predicted for the specification test results.

Like AFDC, Unemployment Insurance (UI) is an income maintenance program, but one that is specifically designed to act as a buffer to business cycle fluctuations. The program typically provides qualified recipients 50-70% of their previous wages for up to 26 weeks. As the UI program is designed to sustain workers through temporary job losses, UI expenditures should be quite sensitive to transitory fluctuations in the unemployment or employment rates. Because of limits on the duration of benefits, it is unlikely that UI expenditures are as responsive to long-term changes in economic conditions as AFDC expenditures.[14]

Consider an slightly more general form of the measurement error model described in equations (5) and (6):

$$X_{it} = Z_{it} + v_{it}, \tag{15}$$

$$Z_{it} = \rho Z_{it-1} + \theta_{it}, \tag{16}$$

$$v_t = \delta v_{t-1} + \mu_t \quad . \tag{17}$$

where $0 < \delta < \rho < 1$. So that *Z* is more highly correlated over time than *v*. Z would be the appropriate right-hand side variable if the dependent variable is AFDC expenditures.[15] When the dependent variable is UI expenditures, however, we would model the relationship as:

$$Y_{it} = \alpha_{oi} + \alpha_1 v_{it} + \varepsilon_{it}. \tag{18}$$

For the case described in equation (18), in which the dependent variable is a function of the part of *X* that is less correlated over time, *v*, we can derive new analytical results for the three different specification tests.

Proposition 11: Estimating the long-differences model described in equation (11) in the case of the ME model described in equations (15)-(18) produces a coefficient estimate with the properties:

a) The sign of $\text{plim}(\hat{\beta}_{\text{j}})$ is the sign of α_1, $\forall j$

[14] Making a similar argument, Black, Daniels and Sanders (2002) show that UI expenditures were relative non-responsive to the large shocks to the coal economy during the 1970's and 1980's.

[15] Allowing *v* to be positively correlated over time, so that $0 < \delta < \rho$ does not change the patterns predicted in Table 1 for the measurement error model in the case that *Y* is a function of *Z*.

b) $\left|\text{plim}(\hat{\beta}_{j+1})\right| < \left|\text{plim}(\hat{\beta}_j)\right| \quad \forall j$

Proposition 12: Estimating the lagged-effects model described in equation (10) in the case of the ME model described in equations (15)-(18) produces a coefficient estimate with the properties:

a) The sign of $\text{plim}(\hat{\gamma}_{\text{j}})$ is the sign of $-\alpha_1$, $\forall j$

b) It is possible that $\left|\text{plim}(\hat{\gamma}_{j+1})\right| > \left|\text{plim}(\hat{\gamma}_j)\right|$

Property b of Proposition 12 is actually true for a large set of parameter values. A large grid search failed to turn up any violations of the property for cases in which $j \geq 2$. It tended to be the case that $\text{plim}(\hat{\gamma}_1) < \text{plim}(\hat{\gamma}_2)$ only when either ρ or $\sigma_\mu^2 / \sigma_\theta^2$ was high.

Proposition 13: Estimating the Heckman-Hotz model described in equation (12) in the case of the ME model described in equations (15)-(18) produces a coefficient estimate with the properties:

a) $\text{plim}(\hat{\beta}_{\text{j}})$ <0 $\forall j$

b) $\left|\text{plim}(\hat{\beta}_{j+1})\right| > \left|\text{plim}(\hat{\beta}_j)\right| \quad \forall j$ if $2\sigma_y^2 > \alpha_1{}^2\delta^2$

If UI expenditures are largely responsive to short-term variation in economic conditions, as described in equations (15)-(18), then specification test results using UI caseloads or expenditures as a dependent variable should display some important differences from the results obtained using AFDC caseloads and expenditures. First, according to Proposition 11 and in contrast to the AFDC results, the coefficient estimate from the long-differences analysis should *decrease* in magnitude as *j* increases. Second, according to proposition 12, the sign of the coefficient on the lagged effect of economic conditions in the UI regressions should be opposite that estimated in the AFDC regressions. In the Heckman-Hotz test, however, the coefficients on lagged UI caseloads or expenditures should remain negative and increasing in magnitude, as they were in the AFDC regressions. Suppose instead the specification test results reported in Panel A of Table 2 and Panel A of Table 3 are generated by omitted variable bias or by economic indicators that are measured with error. In this case the UI results should display the same patterns as the AFDC results, because they should be subject to the same biases.

For the results reported in Panel B of Table 2, we simply substitute UI caseloads for AFDC caseloads in the base model described in equation (13).[16] Only some of the predictions from Propositions 11-13 are observed in the UI results reported in Table 2. In the long-differences results reported in the first column, contrary to the prediction in Proposition 11, the magnitude of the coefficient estimates still increase in magnitude. It should be noted, however that the long-differences coefficient tripled in magnitude in the AFDC results and only increases roughly 30% in the UI results. This is consistent with the argument that UI participation is more responsive to short-term economic variation than AFDC participation. We do observe the sign flip on the lagged effect predicted in Proposition 12. Where the coefficients on the lagged effect in the AFDC results were positive, four of the five coefficients on the lagged effect in the UI results are negative. It is also the case, however, that the magnitude of the coefficient on the lagged unemployment rate decreases in magnitude after j=3, which is not predicted in Proposition 12. Finally, the Heckman-Hotz results are as predicted. As with the AFDC results, the coefficients on lagged per capita UI caseloads are negative and increasing in magnitude.

For the results reported in Panel B of Table 3, we substitute UI expenditures for AFDC expenditures in the base model described in equation (14). All of the predictions from Propositions (11)-(13) are observed in the results. As predicted by Proposition 11, the

[16] UI caseload is measured as total weeks of UI benefits paid by the state, available on-line at http://ows.doleta.gov/lpbin20/lpext.dll/HB/HB%20394/tables.htm#state

coefficient estimates from the long-differences model remain negative, but *decrease* in magnitude as longer differences are used. In contrast, the coefficient estimates obtained using AFDC expenditures were increasing in magnitude. As predicted by Proposition 12, we observe a sign flip on the lagged value of the employment rate. While the coefficients on the lagged employment rate were *negative* and increasing in magnitude in the AFDC results, the coefficients on the lagged employment rate are *positive* and increasing in magnitude in the UI results. Finally, as predicted by Proposition 13 the coefficients on lagged per capita UI expenditures are negative and increasing in magnitude, just as they were in the AFDC results.[17]

These Unemployment Insurance results are moderately consistent, in the case of the Blank data, and entirely consistent, in the case of the BEA data, with the argument that social program participation responds differentially to long-term and short-term changes in local economic conditions. One potential explanation for the difference between the results using the Blank data and those obtained using the BEA data is that the county-level data allows the estimation of state-year effects which purge out omitted state-level variables. Despite Blank's extensive attempts to control for all relevant changes in state characteristics, the state-level results might still be biased in part due to omitted variables, and this bias muddies the predictions from Propositions 11 and 12.

The UI results suggest that the lagged effects of economic conditions estimated in this literature are not purely artifacts of omitted variable or measurement error bias, but reflect true larger impacts of economic conditions on welfare participation. These results also suggest a reasonable interpretation for these lagged effects. Given contemporaneous changes in economic conditions, lagged changes in economic conditions provide a measure of the duration of the

[17] As was the case with the AFDC results, I re-estimated all three models on both data sets, truncating the data sets so that the sample size is not affected by the size of the long differences. There was very little effect on the results.

change. Therefore, the estimated lagged effects reflects in the AFDC literature may reflect in part the larger response of AFDC participation to long-term changes in economic conditions.

6. Conclusions

The analytical and empirical results presented in this paper are intended to make two points. The first is that estimated lagged effects must be interpreted with caution. Like any other variable, they are subject to biases due to omitted variables or measurement error. Because the contemporaneous and lagged effects are often added up to claim that the total effect is larger than previously suspected, it is particularly important to be aware that estimating lagged effects has the potential to exacerbate bias. The second point is that there are many behaviors that are potentially more responsive to long-term changes in economic conditions or government policies than transitory fluctuations in conditions. In this case the lagged values of conditions control for the duration of the change and the presence of lagged effects reflects this differential response to short and long-term variation in conditions.

References

Baker, Michael, Dwayne Benjamin, and Shuchita Stanger. 1999. "The Highs and Lows of the Minimum Wage Effect: A Time-Series Cross-Section Study of the Canadian Law," *Journal of Labor Economics*, 17(2):318-50.

Bartik, Timothy J. and Randall W. Eberts. 1999. "Examining the Effect of Industry Trends and Structure on Welfare Caseloads." In Sheldon Danziger (ed.), *Welfare Reform and the Economy: What Will Happen When a Recession Comes?* Kalamazoo, MI: Upjohn Institute for Employment Research, pp. 119-57.

Berger, Mark C. and Dan A. Black. 1998. "The Duration of Medicaid Spells: An Analysis Using Flow and Stock Samples," *Review of Economics and Statistics,* 80(4): 667-75.

Black, Dan A. and Daniel Nagin. 1998. "Do Right-to-Carry Laws Deter Violent Crime?" *Journal of Legal Studies*, 27(1): 209-219.

Black, Dan A., Kermit Daniel, and Seth G. Sanders. 2002. "The Impact of Economic Conditions On Participation in Disability Programs: Evidence from the Coal Boom and Bust" *American Economic Review* 92(1):27-50.

Black, Dan A., Terra McKinnish, and Seth G. Sanders. 2002. "How the Availability of High-Wage Jobs to Low-skilled Men Affects AFDC Expenditures: Evidence from Shocks to the Steel and Coal Economies." *Journal of Public Economics*, forthcoming.

Blank, Rebecca M. 2001. "What Causes Public Assistance Caseloads to Grow?" *Journal of Human Resources*, 36(1): 85-118.

Bollinger, Chris and Martin David. 1997. "Modeling Food Stamp Program Participation in the Presence of Reporting Errors." *Journal of the American Statistical Association*. 92:827-935.

Bollinger, Chris and Martin David. 2001. "Estimation with Response Error and Non-Response: Food Stamp Participation in the SIPP." *Journal of Business and Economic Statistics*. 19(2): 129-41.

Council of Economic Advisers. 1999. "Economic Expansion, Welfare Reform, and the Decline in Welfare Caseloads: An Update. Technical Report," Washington, D.C.: Executive Office of the President.

Figlio, David and James Ziliak. 1999. "Welfare Reform, the Business Cycle and the Decline in AFDC Caseloads." In Sheldon Danziger (ed.), *Welfare Reform and the Economy: What Will Happen When a Recession Comes?* Kalamazoo, MI: Upjohn Institute for Employment Research, pp. 19-48.

Fitzgerald, John. 1995. "Local Labor Markets and Local Area Effects on Welfare Duration," *Journal of Applied Policy and Management,* 14: 43-67.

Griliches, Zvi and Jerry A. Hausman. 1986. "Errors in Variables in Panel Data." *Journal of Econometrics*. 31:93-118.

Heckman, James and V. Joseph Hotz. 1989. "Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs." *Journal of the American Statistical Association*. 84:862-74.

Hildreth, Andrew and Andrew Oswald. 1997. "Rent-Sharing and Wages: Evidence from Company and Establishment Panels." *Journal of Labor Economics*. 15(2):318-37.

Hoynes, Hillary. 2000. "Local Labor Markets and Welfare Spells: Do Demand Conditions Matter?" *Review of Economics and Statistics* 82(3) 351-68.

Keane, Michael and Kenneth Wolpin. 2002. "Estimating Welfare Effects Consistent with Forward-Looking Behavior." *Journal of Human Resources* 37(3):570-622.

Klerman, Jacob and Steven Haider. 2001. "A Stock-Flow Analysis of the Welfare Caseload: Insights from California Economic Conditions," Labor and Population Program Working Paper #01-02. Santa Monica, CA: RAND.

Krueger, Alan and Cecilia Rouse. 1998. "The Effect of Workplace Education on Earnings, Turnover, and Job Performance." *Journal of Labor Economics*. 16(1):62-93.

McKinnish, Terra. 2000. "Model Sensitivity in Panel Data Analysis: Some Caveats About the Interpretation of Fixed-Effects and Differences Estimates." mimeo, University of Colorado.

Miller, Robert and Seth Sanders. 1997. "Human Capital Development and Welfare Participation." *Carnegie-Rochester Conference Series on Public Policy,* 46: 1-47.

Mroz, T. 1987. "The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions." *Econometrica*. 55(4):765-99.

Mueser, Peter, Julie Hotchkiss, Christopher King, Phillip Rokicki, and David Stevens. 1999. "The Welfare Caseload, Economic Growth and Welfare-to-Work Policies," mimeo. Columbia, MO: University of Missouri-Columbia.

Neumark, David and William Wascher. 1992. "Employment Effects of Minimum and Subminimum Wages: Panel Data on State Minimum Wage Laws." *Industrial and Labor Relations Review*. 46:55-81.

Raaum, Oddbjorn. 1999. "Labour Market Training in Norway- Effect on Earnings." mimeo, University of Oslo.

Regner, Hakan. 1999. "A Nonexperimental Evaluation of Training Program for the Unemployed in Sweden." mimeo, Stockholm University.

Van Reenen, John. 1997. "Employment and Technological Innovation: Evidence from UK Manufacturing Firms*." Journal of Labor Economics*. 15(2):255-84.

Wallace, Geoffrey and Rebecca Blank. 1999. "What Goes Up Must Come Down? Explaining Recent Changes in Public Assistance Caseloads." In Sheldon Danziger (ed.), *Welfare Reform and the Economy: What Will Happen When a Recession Comes?* Kalamazoo, MI: Upjohn Institute for Employment Research, pp. 49-90.

Ziliak, James, David Figlio, Elizabeth Davis and Laura Connolly. 2000. "Accounting for the Decline in AFDC Caseloads." *Journal of Human Resources*. 35(3):570-86.

Table 1: Analytical Predictions

	Omitted Lag	Measurement Error	Omitted Variable
Long-Differences			
Sign	sign of α_1 [a]	sign of α_1	sign of α_1
Magnitude	$\uparrow$ in j	$\uparrow$ in j	$\uparrow$ in j iff $\delta > \rho$
Lagged Effects			
Sign	sign of α_1 [b]	sign of α_1	sign of α_1
Magnitude	$\uparrow$ in j	$\uparrow$ in j [c]	$\uparrow$ in j iff $\delta > \rho$
Heckman-Hotz			
Sign	negative	negative	negative[d]
Magnitude	$\uparrow$ in j	$\uparrow$ in j	$\uparrow$ in j[d]

[a] For j=1, iff $2\alpha_1 > \alpha_2(1-\rho)$

[b] For j=1, iff $5\alpha_2 > \alpha_3(1-\rho)(2-\rho)(1+\rho)$

[c] For j=1, iff $\sigma_v^4(3\rho-1) > (1-\rho)^2(1+\rho)(2\sigma_Z^2\sigma_v^2 - \sigma_Z^4(\rho^2-\rho-1)(1+\rho))$

[d] Grid search indicates often, but not generally, the case. More likely if $\delta > \rho$.

Table 2: Analysis Using Data from Blank (2000): US States, 1977-96

Panel A: Regressions of Per Capita AFDC Caseloads on Unemployment Rates

	Long Differences: $\beta(X_{it}-X_{it-j})$ Estimate of β	N	Lagged Effects: $\beta(X_{it}-X_{it-j})+\gamma(X_{it-1}-X_{it-j-1})$ Estimate of γ	N	Heckman-Hotz: $\beta(X_{it}-X_{it-j})+\lambda Y_{it-j-1}$ Estimate of λ	N
j=1	.0097 (.0021)	969	.0163 (.0020)	969	-.0212 (.0043)	918
j=2	.0188 (.0025)	918	.0217 (.0030)	918	-.0448 (.0074)	867
j=3	.0223 (.0027)	867	.0306 (.0036)	867	-.0636 (.0100)	816
j=4	.0260 (.0028)	816	.0316 (.0041)	816	-.0820 (.0122)	765
j=5	.0267 (.0030)	765	.0370 (.0047)	765	-.1096 (.0143)	714

Panel B: Regressions of Per Capita UI Caseloads on Unemployment Rates

	Long Differences: $\beta(X_{it}-X_{it-j})$ Estimate of β	N	Lagged Effects: $\beta(X_{it}-X_{it-j})+\gamma(X_{it-1}-X_{it-j-1})$ Estimate of γ	N	Heckman-Hotz: $\beta(X_{it}-X_{it-j})+\lambda Y_{it-j-1}$ Estimate of λ	N
j=1	.0724 (.0045)	969	.0012 (.0045)	969	-.0519 (.0086)	969
j=2	.0877 (.0047)	918	-.0158 (.0056)	918	-.0785 (.0134)	918
j=3	.0928 (.0044)	867	-.0187 (.0061)	867	-.0850 (.0160)	867
j=4	.0931 (.0042)	816	-.0165 (.0063)	816	-.0963 (.0179)	816
j=5	.0964 (.0033)	765	-.0096 (.0070)	765	-.1153 (.0197)	765

Table 3: Analysis Using BEA Data: US Counties, 1969-98

Panel A: Regressions of Per Capita AFDC Expenditures on Employment Rates

	Long Differences: $\beta(X_{it} - X_{it-j})$		**Lagged Effects:** $\beta(X_{it} - X_{it-j}) + \gamma(X_{it-1} - X_{it-j-1})$		**Heckman-Hotz:** $\beta(X_{it} - X_{it-j}) + \lambda Y_{it-j-1}$	
	Estimate of β	**N**	**Estimate of** γ	**N**	**Estimate of** λ	**N**
j=1	-.2327 (.0359)	81,454	-.4859 (.0487)	79,094	-.0179 (.0010)	78,309
j=2	-.4496 (.0555)	78,312	-.5797 (.0625)	75,952	-.0364 (.0018)	75,217
j=3	-.5855 (.0671)	75,310	-.6547 (.0669)	72,951	-.0550 (.0026)	72.232
j=4	-.6807 (.0738)	72,349	-.5896 (.0707)	69,990	-.0730 (.0033)	69,269
j=5	-.7326 (.0782)	69,392	-.6101 (.0762)	67,032	-.0906 (.0039)	66,318

Panel B: Regressions of Per Capita UI Expenditures on Employment Rates

	Long Differences: $\beta(X_{it} - X_{it-j})$		**Lagged Effects:** $\beta(X_{it} - X_{it-j}) + \gamma(X_{it-1} - X_{it-j-1})$		**Heckman-Hotz:** $\beta(X_{it} - X_{it-j}) + \lambda Y_{it-j-1}$	
	Estimate of β	**N**	**Estimate of** γ	**N**	**Estimate of** λ	**N**
j=1	-1.783 (.1891)	84,204	1.123 (.1116)	81,807	-.1037 (.0025)	53,952
j=2	-1.508 (.1791)	80,953	1.642 (.1576)	78,561	-.1493 (.0035)	53,600
j=3	-1.111 (.1555)	77,907	1.736 (.1817)	75,528	-.1768 (.0043)	53,277
j=4	-0.893 (.1299)	74,939	1.849 (.1630)	72,563	-.2074 (.0051)	52,869
j=5	-0.757 (.1132)	71,951	1.859 (.1714)	69,567	-.2369 (.0058)	52,255

Appendix A

All proofs assume that random variables have been demeaned, all error terms ($\varepsilon, v,$ and θ) are "best case" i.i.d. disturbances and that the right-hand side variables are generated by an AR(1) process. Three other assumptions are:

A1. $0 < \rho < 1$

A2. $0<\delta<1$ in the OV and modified ME models

A3. α_2 is of the same sign as α_1 in the OL and OV models

Proposition 1:

$$\text{plim}(\hat{\beta}_j) = \alpha_1 + \alpha_2 \frac{Cov(X_{it-1} - X_{it-j-1}, X_{it} - X_{it-j})}{Var(X_{it} - X_{it-j})} = \alpha_1 + \alpha_2 \left(\frac{2\rho - \rho^{j+1} - \rho^{j-1}}{2(1-\rho^j)} \right)$$

The sign of $\text{plim}(\hat{\beta}_{\text{j}})$ is the sign of α_1 given A1 and A3. $\left|\text{plim}(\hat{\beta}_{\text{j+1}})\right| > \left|\text{plim}(\hat{\beta}_{\text{j}})\right| \forall j$ iff $\frac{2\rho - \rho^{j+2} - \rho^j}{2(1-\rho^{j+1})} > \frac{2\rho - \rho^{j+1} - \rho^{j-1}}{2(1-\rho^j)}$, which is true iff $(1-\rho^2)(1-\rho) > 0$, which true is under A1.

Proposition 2:

$$\text{plim}(\hat{\beta}_{\text{j}}) = \frac{Cov(Y_{it} - Y_{it-j}, X_{it} - X_{it-j})}{Var(X_{it} - X_{it-j})} = \alpha_1 \frac{Var(Z_{it} - Z_{it-j})}{Var(Z_{it} - Z_{it-j}) + Var(v_{it} - v_{it-j})} = \alpha_1 \left(\frac{\sigma_Z^2(1-\rho^j)}{\sigma_Z^2(1-\rho^j) + \sigma_v^2} \right)$$

Clearly, $\text{plim}(\hat{\beta}_{\text{j}})$ is the sign of α_1 and increasing in j under A1.

Proposition 3:

$$\text{plim}(\hat{\beta}_j) = \alpha_1 + \alpha_2 \frac{Cov(X_{it} - X_{it-j}, W_{it} - W_{it-j})}{Var(X_{it} - X_{it-j})} = \alpha_1 + \alpha_2 \left(\frac{\sigma_{XW}(2-\delta^j - \rho^j)}{2\sigma_X^2(1-\rho^j)} \right)$$

The sign of $\text{plim}(\hat{\beta}_{\text{j}})$ is the sign of α_1 under A1 and A3. $\left|\text{plim}(\hat{\beta}_{\text{j+1}})\right| > \left|\text{plim}(\hat{\beta}_{\text{j}})\right| \forall j$ iff: $\frac{2-\delta^{j+1} - \rho^{j+1}}{2\sigma_X^2(1-\rho^{j+1})} > \frac{2-\delta^j - \rho^j}{2\sigma_X^2(1-\rho^j)}$, which is true iff: $\delta^j(1-\delta) - \rho^j(1-\rho) + \rho^j\delta^j(\delta - \rho) > 0$, which under A1 and A2 is true iff $\delta > \rho$.

Proposition 4: Self-evident.

Propositions 5-7:

$$\text{plim}(\hat{\gamma}_{\text{j}}) = \frac{Var(X_{it} - X_{it-j})Cov(Y_{it} - Y_{it-j}, X_{it-1} - X_{it-j-1}) - Cov(Y_{it} - Y_{it-j}, X_{it} - X_{it-j})Cov(X_{it} - X_{it-j}, X_{it-1} - X_{it-j-1})}{Var(X_{it} - X_{it-j})Var(X_{it-1} - X_{it-j-1}) - Cov(X_{it} - X_{it-j}, X_{it-1} - X_{it-j-1})^2}$$

Proposition 5:

$$\text{plim}(\hat{\gamma}_1) = \alpha_2 - \alpha_3(1-\rho)(2-\rho)(1+\rho)/5,$$

$$\text{plim}(\hat{\gamma}_j) = \alpha_2 + \alpha_3 \frac{(2\rho - \rho^{j+1} - \rho^{j-1})[2(1-\rho) - \rho(2\rho^2 - \rho^{j+2} - \rho^{j-2})]}{4(1-\rho^j)^2 - (2\rho - \rho^{j+1} - \rho^{j-1})^2} \quad \forall j > 1.$$

Denominator is greater than zero by Cauchy-Schwarz Inequality.
$2\rho - \rho^{j+1} - \rho^{j-1}$ if j>1, and $2(1-\rho^j) - (2\rho^2 - \rho^{j+2} - \rho^{j-2}) = 2(1-\rho^2) + \rho^{j-2}(\rho^2 - 1)^2$ is positive under A1. Therefore $\text{plim}(\hat{\gamma}_{\text{j}})$ is the sign of α_1 for *j>1*.
$\left|\text{plim}(\hat{\gamma}_{\text{j+1}})\right| > \left|\text{plim}(\hat{\gamma}_{\text{j}})\right| \forall j > 1$ iff:

$$\frac{(2\rho - \rho^{j+2} - \rho^j)[2(1-\rho) - \rho(2\rho^2 - \rho^{j+3} - \rho^{j-1})]}{4(1-\rho^{j+1})^2 - (2\rho - \rho^{j+2} - \rho^j)^2} > \frac{(2\rho - \rho^{j+1} - \rho^{j-1})[2(1-\rho) - \rho(2\rho^2 - \rho^{j+2} - \rho^{j-2})]}{4(1-\rho^j)^2 - (2\rho - \rho^{j+1} - \rho^{j-1})^2}$$

which is true iff: $4\rho^{2j-3}(1+\rho)^3(1-\rho)^4(\rho - \rho^{j+1} + 1) > 0$, which is true under A1.

Proposition 6:

$$\text{plim}(\hat{\gamma}_1) = \frac{2\alpha_1\sigma_v^2\sigma_z^2\rho(1-p)}{4(\sigma_z^2(1-\rho) + \sigma_v^2)^2 - (\sigma_z^2(1-\rho)^2 + \sigma_v^2)^2},$$

and, $$\text{plim}(\hat{\gamma}_j) = \frac{2\alpha_1\sigma_v^2\sigma_z^2(2\rho - \rho^{j+1} - p^{j-1})}{4(\sigma_z^2(1-\rho^j) + \sigma_v^2)^2 - \sigma_z^4(2\rho - \rho^{j+1} - p^{j-1})^2} \quad \forall j > 1.$$

For both results, denominator is positive by Cauchy-Schwarz Ineqaulity, and numerator is the sign of α_1 under A1. $\left|\text{plim}(\hat{\gamma}_{\text{j+1}})\right| > \left|\text{plim}(\hat{\gamma}_{\text{j}})\right| \forall j > 1$ iff:

$$\frac{2\alpha_1\sigma_v^2\sigma_z^2(2\rho - \rho^{j+2} - p^j)}{4(\sigma_z^2(1-\rho^{j+1}) + \sigma_v^2)^2 - \sigma_z^4(2\rho - \rho^{j+2} - p^j)^2} > \frac{2\alpha_1\sigma_v^2\sigma_z^2(2\rho - \rho^{j+1} - p^{j-1})}{4(\sigma_z^2(1-\rho^j) + \sigma_v^2)^2 - \sigma_z^4(2\rho - \rho^{j+1} - p^{j-1})^2},$$

which is true iff:
$4\sigma_v^2[2\sigma_Z^2(1-\rho)(1+\rho) + \sigma_v^2(1+\rho^2)] + \sigma_Z^4(1-\rho)^2(1+\rho)^2[4 + \rho^{2j-1}(1+\rho^2) - 2\rho^j(1+\rho)] > 0$ which is true since $2\rho^j(1+\rho) \le 4 \le 4 + \rho^{2j-1}(1+\rho^2)$ under A1.
Furthermore: $\left|\text{plim}(\hat{\gamma}_2)\right| > \left|\text{plim}(\hat{\gamma}_1)\right|$ iff

$$\frac{2\alpha_1\sigma_Z^2\sigma_v^2\rho(1-\rho^2)}{4[\sigma_z^2(1-\rho^2) + \sigma_v^2]^2 - \sigma_z^4\rho^2(1-\rho^2)^2} > \frac{2\alpha_1\sigma_Z^2\sigma_v^2\rho(1-\rho)}{4[\sigma_z^2(1-\rho) + \sigma_v^2]^2 - [\sigma_z^2(1-\rho^2) + \sigma_v^2]^2}$$, which is true iff:

$\sigma_v^4(3\rho-1)-(1-\rho)^2(1+\rho)[2\sigma_Z^2\sigma_v^2-\sigma_Z^4(\rho^2-\rho-1)(1+\rho)]>0$

Proposition 7:

$$\text{plim}(\hat{\gamma}_j)=\frac{\alpha_2\sigma_{XW}\left[2(1-\rho^j)(2\delta-\delta^{j+1}-\rho^{j-1})-(2-\delta^j-\rho^j)(2\rho-\rho^{j+1}-\rho^{j-1})\right]}{4\sigma_X^2(1-\rho^j)^2-\sigma_X^2(2\rho-\rho^{j+1}-\rho^{j-1})^2}$$

The sign of $\text{plim}(\hat{\beta}_{\text{j}})$ is the sign of α_1 iff
$2(1-\rho)(2\delta-\delta^2-1)-(2-\delta-\rho)(2\rho-\rho^2-1)>0$ which is true iff:
$(2-\delta-\rho)(1-\rho)-(2-\delta-\delta)(1-\delta)>0$, which is true under A1-A2 iffδ>ρ

For Propositions 8-10:

$$\text{plim}(\hat{\lambda}_{\text{j}})=\frac{Var(X_{it}-X_{it-j})Cov(Y_{it}-Y_{it-j},Y_{it-j-1})-Cov(Y_{it}-Y_{it-j},X_{it}-X_{it-j})Cov(X_{it}-X_{it-j},Y_{it-j-1})}{Var(X_{it}-X_{it-j})Var(Y_{it-j-1})-Cov(X_{it}-X_{it-j},Y_{it-j-1})^2}$$

Proposition 8:

$$\text{plim}(\hat{\lambda}_j)=\frac{-\alpha_2(\alpha_1+\alpha_2\rho)\sigma_X^2(1-\rho^2)(2-\rho^j)}{2\sigma_Y^2-\sigma_X^2(1-\rho^j)\rho^2(\alpha_1+\alpha_2\rho)^2}$$

Denominator is positive by Cauchy-Schwarz Inequality, and numerator is negative under A1 and A3, so $\text{plim}(\hat{\lambda}_{\text{j}})<0$. Verify $\text{plim}(\hat{\lambda}_{\text{j+1}})<\text{plim}(\hat{\lambda}_j)\ \forall j$: This is true iff:

$$\frac{-\alpha_2(\alpha_1+\alpha_2\rho)\sigma_X^2(1-\rho^2)(2-\rho^{j+1})}{2\sigma_Y^2-\sigma_X^2(1-\rho^{j+1})\rho^2(\alpha_1+\alpha_2\rho)^2}<\frac{-\alpha_2(\alpha_1+\alpha_2\rho)\sigma_X^2(1-\rho^2)(2-\rho^j)}{2\sigma_Y^2-\sigma_X^2(1-\rho^j)\rho^2(\alpha_1+\alpha_2\rho)^2}$$ which is true iff:

$[2\sigma_X^4+\sigma_X^2\rho^2(\alpha_1+\alpha_2\rho)^2]\rho^j(1-\rho)>0$ which is true under A1.

Proposition 9:

$$\text{plim}(\hat{\lambda}_{\text{j}})=\frac{-2\alpha_1^{\ 2}\sigma_z^2\sigma_v^2\rho(1-\rho^j)}{\sigma_y^2(2\sigma_z^2(1-\rho^j)+2\sigma_v^2)-\alpha_1^{\ 2}\sigma_z^4\rho^2(1-\rho^j)^2}$$

Denominator is positive by Cauchy-Schwarz Inequality, and numerator is negative given A1, so $\text{plim}(\hat{\lambda}_{\text{j}})<0$. $\text{plim}(\hat{\lambda}_{\text{j+1}})<\text{plim}(\hat{\lambda}_j)\ \forall j$ is true iff:

$$\frac{-2\alpha_1^{\ 2}\sigma_z^2\sigma_v^2\rho(1-\rho^{j+1})}{\sigma_y^2(2\sigma_z^2(1-\rho^{j+1})+2\sigma_v^2)-\alpha_1^{\ 2}\sigma_z^4\rho^2(1-\rho^{j+1})^2}<\frac{-2\alpha_1^{\ 2}\sigma_z^2\sigma_v^2\rho(1-\rho^j)}{\sigma_y^2(2\sigma_z^2(1-\rho^j)+2\sigma_v^2)-\alpha_1^{\ 2}\sigma_z^4\rho^2(1-\rho^j)^2}$$

which is true iff: $2\sigma_V^2\sigma_Y^2\rho^j(1-\rho)+\alpha_1^{\ 2}\sigma_Z^4\rho^2\rho^j(1-\rho^j)(1-\rho^{j+1})(1-\rho)>0$, which is true under A1.

Proposition 10:

$$\text{plim}(\hat{\lambda}_{\text{j}}) = \frac{\alpha_2\sigma_{XW}\,\rho(\alpha_1\sigma_X^2 + \alpha_2\sigma_{XW})(2-\delta^j-\rho^j) - 2\alpha_2\sigma_X^2\,\delta(\alpha_1\sigma_{XW} + \alpha_2\sigma_W^2)(1-\delta^j)}{2\sigma_X^2\sigma_Y^2 - \rho^2(1-\rho^j)(\alpha_1\sigma_X^2 + \alpha_2\sigma_{XW})^2}$$

There are many values for which $\text{plim}(\hat{\lambda}_{\text{j}}) < 0$ and $\text{plim}(\hat{\lambda}_{\text{j+1}}) < \text{plim}(\hat{\lambda}_j)$. For example: $\alpha_1 = \alpha_2 = 1$, $\sigma_Y^2 = 5\sigma_\theta^2$ or $\sigma_Y^2 = 100\sigma_\theta^2$, ρ is equal to .1, .2, .3 or .4, and $\delta > \rho$.

Proposition 11:

$$\text{plim}(\hat{\beta}_{\text{j}}) = \frac{Cov(Y_{it} - Y_{it-j}, X_{it} - X_{it-j})}{Var(X_{it} - X_{it-j})} = \alpha_1 \frac{Var(v_{it} - v_{it-j})}{Var(Z_{it} - Z_{it-j}) + Var(v_{it} - v_{it-j})} = \alpha_1\left(\frac{\sigma_v^2(1-\delta^j)}{\sigma_v^2(1-\delta^j) + \sigma_z^2(1-\rho^j)}\right)$$

The sign of $\text{plim}(\hat{\beta}_{\text{j}})$ is the sign of α_1 under A1-A3. $\left|\text{plim}(\hat{\beta}_{\text{j+1}})\right| > \left|\text{plim}(\hat{\beta}_{\text{j}})\right| \;\forall j$ iff:

$\frac{1-\rho^j}{1-\delta^j} < \frac{1-\rho^{j+1}}{1-\delta^{j+1}}$ which is true iff $\delta < \rho$.

Proposition 12:

$$\text{plim}(\hat{\gamma}_{\text{j}}) = \frac{Var(X_{it} - X_{it-j})Cov(Y_{it} - Y_{it-j}, X_{it-1} - X_{it-j-1}) - Cov(Y_{it} - Y_{it-j}, X_{it} - X_{it-j})Cov(X_{it} - X_{it-j}, X_{it-1} - X_{it-j-1})}{Var(X_{it} - X_{it-j})Var(X_{it-1} - X_{it-j-1}) - Cov(X_{it} - X_{it-j}, X_{it-1} - X_{it-j-1})^2}$$

$$= \frac{2\alpha_1\sigma_v^2\sigma_z^2[(1-\rho^j)(2\delta - \delta^{j+1} - \delta^{j-1}) - (1-\delta^j)(2\rho - \rho^{j+1} - p^{j-1})]}{4[\sigma_z^2(1-\rho^j) + \sigma_v^2(1-\delta^j)]^2 - [\sigma_z^2(2\rho - \rho^{j+1} - p^{j-1}) + \sigma_v^2(2\delta - \delta^{j+1} - \delta^{j-1})]^2}.$$

$\text{plim}(\hat{\gamma}_{\text{j}})$ is the sign of α_1 iff $(1-\rho)(2\delta - \delta^2 - 1) - (1-\delta)(2\rho - \rho^2 - 1) < 0$, which is true iff: $(\delta - \rho)(1-\delta)(1-\rho) > 0$ which under A1, A2 and $\delta < \rho$.

Proposition 13:

$$\text{plim}(\hat{\lambda}_{\text{j}}) = \frac{Var(X_{it} - X_{it-j})Cov(Y_{it} - Y_{it-j}, Y_{it-j-1}) - Cov(Y_{it} - Y_{it-j}, X_{it} - X_{it-j})Cov(X_{it} - X_{it-j}, Y_{it-j-1})}{Var(X_{it} - X_{it-j})Var(Y_{it-j-1}) - Cov(X_{it} - X_{it-j}, Y_{it-j-1})^2}$$

$$= \frac{-2\alpha_1{}^2\sigma_z^2\sigma_v^2\delta(1-\rho^j)(1-\delta^j)}{\sigma_y^2[2\sigma_z^2(1-\rho^j) + 2\sigma_v^2(1-\delta^j)] - \alpha_1{}^2\sigma_v^4\delta^2(1-\delta^j)^2}$$

Denominator is positive by Cauchy-Schwarz Inequality, and numerator is negative under A1 and A2, so $\text{plim}(\hat{\lambda}_{\text{j}}) < 0$. $\text{plim}(\hat{\lambda}_{\text{j+1}}) < \text{plim}(\hat{\lambda}_j) \;\forall j$ iff:

$$\frac{(1-\rho^{j+1})(1-\delta^{j+1})}{\sigma_y^2[2\sigma_z^2(1-\rho^{j+1}) + 2\sigma_v^2(1-\delta^{j+1})] - \alpha_1{}^2\sigma_v^4\delta^2(1-\delta^{j+1})^2} > \frac{(1-\rho^j)(1-\delta^j)}{\sigma_y^2[2\sigma_z^2(1-\rho^j) + 2\sigma_v^2(1-\delta^j)] - \alpha_1{}^2\sigma_v^4\delta^2(1-\delta^j)^2}$$

which is true iff:

$$2\sigma_y^2\sigma_z^2(1-\rho^j)(1-\rho^{j+1})\delta^j(1-\delta)+2\sigma_y^2\sigma_v^2(1-\delta^j)(1-\delta^{j+1})\rho^j(1-\rho)$$
$$+\alpha_1{}^2\sigma_v^2\delta^2(1-\delta^j)(1-\delta^{j+1})[\delta^j(1-\delta)-\rho^j(1-\rho)+\rho^j\delta^j(\delta-\rho)]>0$$

which is clearly always true if $\delta \geq \rho$. But, for $\delta < \rho$ it s helpful to rearrange:

$$2\sigma_y^2\sigma_z^2(1-\rho^j)(1-\rho^{j+1})\delta^j(1-\delta)+\sigma_v^2(1-\delta^j)(1-\delta^{j+1})[\alpha_1{}^2\delta^{j+2}(1-\delta-\rho^j(\rho-\delta))$$
$$+\rho^j(1-\rho)(2\sigma_y^2-\alpha_1{}^2\delta^2)]>0$$

Since $(1-\delta)-\rho^j(\rho-\delta)>0$ under A1 and A2, the case where $2\sigma_y^2 > \alpha_1{}^2\delta^2$ is a sufficient condition.