IS THE TIME-SERIES EVIDENCE ON MINIMUM WAGE EFFECTS CONTAMINATED BY PUBLICATION BIAS?

DAVID NEUMARK and WILLIAM WASCHER*

Existing meta-analysis approaches to testing for publication bias are problematic when applied to time-series studies in economics, because changes in parameters can generate spurious evidence of publication bias. We suggest an alternative test in such contexts, and apply it to time-series studies of the effects of minimum wages on employment. In contrast to recent research by Card and Krueger [1995a], we find that the results of published time-series studies of minimum wage effects are consistent with structural change, and that the null hypothesis of no publication bias is not rejected by the evidence. (*JEL* B41, C22, C40, J23)

I. INTRODUCTION

Economists often express concern that the journal review process has the unintended effect of biasing published empirical results towards the rejection of economic null hypotheses, in favor of statistically significant results that are consistent with economists' priors or predictions. In other disciplines, meta-analyses of published papers—generally focusing on clinical trials—provide some evidence consistent with a bias towards publication of significant results (see, for example, Berlin et al. [1989]), under the (reasonable) presumption that the parameter of interest, such as the effect of a drug, should be constant across studies.[1] In economics, however, the parameters of statistical models are more likely to vary across samples—especially when the samples cover different time periods—because of changed behavior (e.g., Lucas [1976]), model misspecification, or measurement changes. The possibility of parameter instability makes it difficult to apply standard meta-analysis techniques to the economics literature, because these techniques draw inferences regarding publication bias from changes in estimates across samples.[2]

In this paper, we suggest a method of testing the null hypothesis of no publication bias when parameters are changing. In particular, if there is a specification that is plausibly free of selective specification search, then that specification can be used as a benchmark with which to compare changes in estimates across the published studies. A statistical test for whether the changes in estimates across the published studies are significantly different from those for the benchmark specification es-

* The views expressed in this paper do not necessarily represent those of the Federal Reserve Board or its staff. We thank Jeff Biddle, Harry Holzer, Alan Krueger, Robert Lalonde, Kevin Lang, Peter Schmidt, John Strauss, Jeff Wooldridge, and seminar participants at Michigan State, the Milken Institute, MIT, UCSD, and the University of Washington for helpful comments. David Card and Alan Krueger for making their data available, and Daniel Hansen for research assistance.

Neumark: Professor, Department of Economics, Michigan State University, East Lansing and Research Associate National Bureau of Economic Research, Cambridge Mass., Phone 1-517-353-7275, Fax 1-517-432-1068 E-mail neumarkd@pilot.msu.edu
Wascher: Senior Economist, Board of Governors of the Federal Reserve System, Washington, D.C. Phone 1-202-452-2812, Fax 1-202-452-5296 E-mail wwascher@frb.gov

1. In the clinical trial literature—in which there is little room for manipulating specifications—the typical assertion is that it is difficult to publish insignificant results (the "file-drawer" problem). If a similar problem exists in empirical economics, one response by researchers might be to search across specifications to look for significant or "expected" results (e.g., Leamer [1978]). For ease of exposition, in this paper we use the term publication bias to refer to both the file-drawer problem and the selective specification search that publication bias may induce, although the latter is the issue with which we are most concerned.

2. We do not distinguish between changes in underlying behavioral parameters and changes in the parameters of a statistical model conditional on the model specification and measurement of variables. Below, we use the term "structural change" as a short-hand for such changes in the parameters of the statistical model—whether these stem from changes in behavior (i.e., behavioral parameters), model misspecification, or measurement changes—which together constitute the alternative hypothesis to publication bias.

Economic Inquiry
(ISSN 0095-2583)
Vol. XXXVI, July 1998, 458–470

timated over different sample periods can then be used to assess the evidence of publication bias in the published studies under consideration.

We apply our approach to the question of whether there is evidence of publication bias in the time-series literature on the employment effects of the minimum wage. We choose this application for three reasons. First, the economic impact of the minimum wage is an important issue for policy. Second, widespread interest in this question has generated sufficient research to make the application of meta-analysis techniques feasible. Third, it is the topic of one of the rare applications of meta-analysis in economics (Card and Krueger [1995a,b]). We show that when we recognize the special complexities associated with meta-analysis of studies using time-series data, the evidence that the minimum wage literature is contaminated by publication bias is much weaker.

II. EVIDENCE OF PUBLICATION BIAS

The approach used by Card and Krueger to test for publication bias in the minimum wage literature is relatively straightforward. First, in the absence of publication bias and if parameters are stable, then estimated t-statistics for employment effects of minimum wages should rise as the sample period is extended; more precisely, the log of the absolute t-ratios should rise one-for-one with the log of the square root of the degrees of freedom. This is closely related to the approach of Berlin, et al. [1989], who test for publication bias in research from clinical trials by asking whether the estimated magnitudes of the effects in published studies are negatively related to sample sizes. The intuition is that if only significant results get published, the estimated effect will be bigger when it is based on a smaller sample. Second, a tendency for published t-statistics for employment effects to be clustered around two can be interpreted as an indication of publication bias.

Using these techniques, Card and Krueger report evidence that is consistent with publication bias in the minimum wage literature. Focusing on the existing time-series studies that use quarterly data, they first report that a regression of the log of the absolute t-statistics for the estimated minimum wage effects from these studies on the log of the square root of the degrees of freedom yields a negative coefficient estimate, rather than the positive coefficient (of one) implied by the null hypothesis of no publication bias. Second, they provide a graph suggesting that the published absolute t-statistics are clustered around two; more precisely, the graph shows that a scatterplot of estimated elasticities and standard errors of these estimates fits a line through the origin with slope two quite well.

However, such evidence might instead reflect structural change over time, due to changes in behavioral parameters, model misspecification, or measurement changes. If, for example, the absolute size of the minimum wage coefficient in the time-series employment equations has declined over time, the absolute t-statistic may have fallen as more years were added to the time-series regressions, even though the standard errors were also falling; a declining coefficient would also explain why, as shown below, the Berlin et al. [1989] test generates evidence consistent with publication bias. Card and Krueger clearly acknowledge the possibility of structural change as an alternative explanation of their findings, but they interpret their evidence as indicating that publication bias is the more "likely" or "plausible" explanation [1995a, 242; 1995b, 192].[3] In particular, they conclude that "the studies in the literature have been affected by specification searching and publication biases, induced by editors' and authors' tendencies to look for negative and statistically significant estimates of the employment effect of the minimum wage" [1995a, 242].[4]

3. Their conclusion may be based, in part, on their interpretation of the exchange between Welch [1974; 1976; 1977] and Siskind [1977]—discussed in Chapter 6 of *Myth and Measurement*—as an example of publication bias. We leave it to the reader to examine the original articles and assess this interpretation. However, even if one accepts Card and Krueger's view of this exchange, this particular incident provides no basis for interpreting other studies similarly.

4. As Card and Krueger note, De Long and Lang [1992] have extended the meta-analysis approach to study publications in economics more generally. However, De Long and Lang find evidence of a "reverse" form of publication bias, namely that "journals tend to publish papers that fail to reject their null hypotheses only when the null hypotheses are likely to be false" [1992, 1,257]. In contrast to Card and Krueger's conclusion, this finding would cast doubt on studies reporting *no* statistically significant effects of minimum wages on employment.

Before turning to an assessment of the evidence, using our approach to testing for publication bias accounting for the possibility of structural change, we want to address the question of whether this particular application of meta-analysis is of substantive interest in its own right. Some researchers have argued that the more recent literature on minimum wage effects—which has used panel data across states, firms, or establishments—has generated more reliable inferences regarding the employment effects of minimum wages, and that the time-series literature should therefore be regarded as essentially obsolete. As a general proposition, we agree that the use of panel data represents a significant step forward in this literature. However, it is important to distinguish between tests of the theory of labor demand exploiting exogenous variation in wages—for which these panel data are especially useful—and policy evaluations of increases in the federal minimum wage. For the latter, analyses performed at the sub-national level—using state-level data, for example—may provide misleading estimates, because the capacity for labor and other inputs to flow across sub-national borders may be greater than the ability of these inputs to flow across national borders. That is, because minimum wage effects may vary at different levels of aggregation, time-series estimates using aggregate data still provide useful information about the overall effects of changes in the federal minimum wage.

In addition, given that recent evidence using the longest available time series indicates no significant disemployment effects of minimum wages (Card and Krueger [1995b]), some have argued that it is irrelevant whether earlier studies obtained estimates that were influenced by publication bias. We disagree with this logic. If the earlier studies were influenced by publication bias, then the "true" time-series evidence would have pointed to no significant disemployment effects for a long time, and only because of publication bias did earlier researchers using time-series data reach opposite conclusions. If, instead, the decline in minimum wage effects in published studies reflects structural change, then variation in estimated minimum wage effects over different sample periods would suggest that the aggregate effects of minimum wages are negative and significant at some times, but not at other times. In this case, time-series regressions some years hence might well reveal a stronger negative effect of minimum wages on employment. More generally, evidence pointing to structural change as the source of differing parameter estimates would suggest that investigating the conditions under which minimum wage effects are important might be a fruitful avenue for future research.

III. ASSESSING THE EVIDENCE

Is publication bias the most likely explanation of the pattern of declining minimum wage effects (or t-statistics) in published studies? Our approach to answering this question can be interpreted as attempting to mimic the following historical "experiment." Suppose researchers at different times in the past had approached the time-series data with a regression specification that was not influenced by selective specification searches aimed at producing significant negative effects of minimum wages. Would the coefficients from this "benchmark" specification, estimated over varying numbers of years, have been different from those reported in the literature? A positive conclusion—in particular, if the findings of these hypothetical researchers would have generated stable estimates of minimum wage effects (and rising absolute t-statistics)—would support the publication bias explanation. A negative answer, in contrast, would point to changing parameters rather than publication bias as the more likely explanation for the observed pattern of t-statistics and coefficient estimates in the literature; more formally, such evidence would fail to reject the null hypothesis of no publication bias.[5]

This approach is related to the "gold standard" approach in the medical literature: using a high-quality, large study with which to compare other published studies (Chalmers et al. [1987]). Of course, the existence of such a study—or a "benchmark" specification in our case—is not always a given, and the selection of such a study or specification from a set of

5. Of course, the benchmark specification could generate rising (absolute) minimum wage effects. In this case, structural change would obscure any evidence of publication bias using the meta-analysis approach employed by Card and Krueger. In contrast, such bias would still be detected by comparing the published estimates to the time series of coefficient estimates generated by the benchmark specification.

alternatives is a subjective matter. Consequently, the approach we pursue in this particular context may not always be applicable and may be influenced by the subjective decisions of the researcher. In the present case, however, we believe that a good set of plausible benchmark specifications is available.

The first benchmark specification we consider is one developed by Solon [1985] and later used by Wellington [1991]. This appears to be a consensus specification, based on earlier research by Brown et al. [1983], supplemented with a more flexible serial correlation structure. In this specification—which is also used by Card and Krueger [1995b] to update the existing time-series evidence—the log of the teen employment rate is regressed on the log of the Kaitz index (an industry-coverage-weighted average of the minimum wage relative to the average industry wage), the log of the unemployment rate of adult males, the fraction of 16–19 year-olds who are 16–17, the fraction of 16–19 year-olds in the armed forces, the log of the fraction of 16–19 year-olds in the population, quarterly seasonal dummies (since the data used are not seasonally adjusted), a time trend and its square, and interactions of the time trend and its square with each of the quarter dummies. We use the same data as Card and Krueger, which they have made available via the Internet.

We begin our analysis with this particular benchmark specification because it reflects some features that are relatively common across the entire array of time-series minimum wage studies (such as the Kaitz index specification, the inclusion of an aggregate demand indicator, and the inclusion of some supply indicators), and because it excludes some variables over which there has been some disagreement (such as the school enrollment rate and lagged values of the Kaitz index). In addition, this specification includes some features that were only adopted in work that followed most of the earlier time-series studies. For example, although Brown, et al. [1983] note that correcting for serial correlation was rare in the work they review, later research corrected for first-order serial correlation with the Cochrane-Orcutt method. Similarly, Solon [1985] finds that seasonal effects are not constant over time, necessitating the inclusion of the time-quarter interactions (in part, to get to a model consistent with an AR(1) specification). Because this specification builds on much earlier work, uses statistical tests to identify the appropriate error structure, and excludes variables that are contentious, it seems a potentially useful one with which to mimic our historical experiment. Below we explore other, perhaps more compelling, benchmark specifications to assess the robustness of our results.

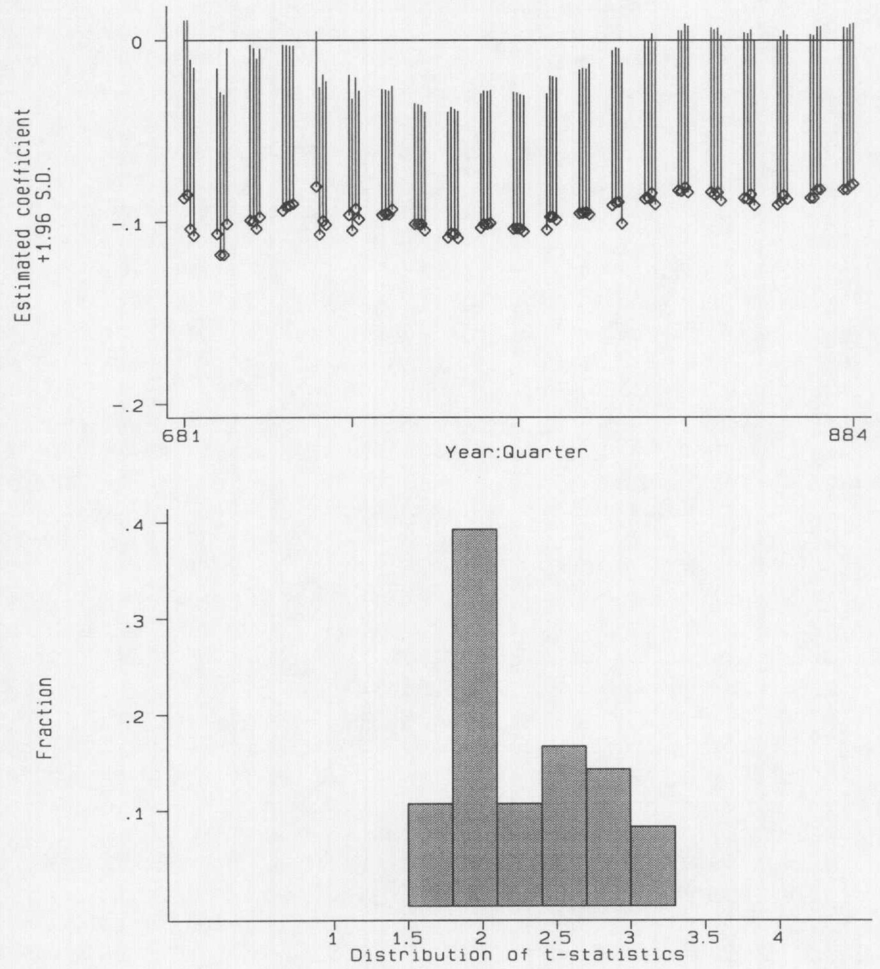
The first panel of Figure 1 reports the estimated elasticities of teen employment with respect to the minimum wage that we obtain when we repeatedly reestimate this specification after adding an additional quarter of data ("rolling regressions"). We begin with a sample period beginning in the first quarter of 1954 and ending in 1968, the earliest ending year of the sample periods in the time-series studies that Card and Krueger consider. We continue through 1988—the last year of data in the published studies that they cover. We also show the upper end of the 95% confidence interval, so that the reader can see which estimates are significantly different from zero. The graph shows that the estimated elasticity has generally fallen in absolute value over time.

Next, we regressed the log of the absolute t-statistics generated by this procedure against the log of the square root of the degrees of freedom; OLS estimates of this meta-analysis regression are reported in column (1) in panel A of Table I. The estimated coefficient is negative (–.41) and significant.[6] Also reported in Table I is the regression of the absolute value of the estimated coefficient on the square root of the degrees of freedom (the test proposed by Berlin et al. [1989]). This estimated slope is also negative (–.004) and significant (see column (1), panel B). Finally, the second panel of Figure 1 shows that the absolute t-statistics from the rolling regressions are clustered around two.

What are the implications of these results? Card and Krueger interpreted qualitatively similar results for published studies as evidence of publication bias. However, the results in Figure 1 and in column (1) of Table I suggest that studies done by researchers using the same specification at different points in

6. As Card and Krueger [1995b] note, the standard errors in Table I cannot be taken literally, because the samples on which the estimates are based overlap substantially.

FIGURE 1
Solon Specification, 1968–1988



time would also have generated evidence consistent with publication bias. That is, the findings from the published studies may reflect nothing but objective researchers estimating a relationship that was, in fact, changing over time.[7]

To differentiate between publication bias and structural change, we need a method of assessing whether the estimated coefficients from the meta-analysis regressions based on the published studies are different from those based on the estimates from the rolling regressions. A statistical approach to assessing this evidence can be developed using the Hausman test framework. In particular, under the null hypothesis of no publication bias, there are two consistent estimates of changes over time in the absolute value of the minimum wage effect (or the associated t-statistic): the estimate in Table I, based on a benchmark specification (b_B); and the estimate based on the published studies (b_P). That is, b_B and b_P can be viewed as alternative estimates of the parameter β or β' in the two meta-analysis regressions that we consider:

7. There is no compelling reason why the t-statistics from the rolling regressions should be clustered around two. However, if the t-statistics for the shorter samples tended to be near two and the estimated coefficients fell over time, then such clustering could occur.

TABLE I
Meta-Analysis OLS Regressions for Time-Series Estimates of Employment Effects of Minimum Wages, Quarterly Data, 1968–1988

	Solon Specification (1)	Incl. Adult Male Empl./Pop. (2)	Also Incl. Log Real Manu. Wage (3)
A. *Regressions of Log Absolute Values of t-Statistics on Log Square Root of Degrees of Freedom*			
Constant	1.69	2.18	2.31
	(.25)	(.27)	(.23)
Log square root of degrees of freedom	–.41	–.66	–.73
	(.12)	(.12)	(.11)
R^2	.13	.26	.36
B. *Regressions of Absolute Coefficient Estimates on Square Root of Degrees of Freedom*			
Constant	.13	.16	.16
	(.005)	(.005)	(.005)
Square root of degrees of freedom	–.004	–.008	–.008
	(.0006)	(.0006)	(.0005)
R^2	.38	.67	.75

Standard errors are reported in parentheses. There are 84 observations used in each column. Minimum wage employment equations from which the coefficients and t-statistics are obtained are described in the text.

$$\ln(t) = \alpha + \beta \ln(N^{1/2}) + \varepsilon$$

and

$$g = \alpha' + \beta' N^{1/2} + \varepsilon,$$

where t and g are the time series of absolute values of t-statistics or estimated minimum wage effects from the published studies (in which case the estimates of β or β' are denoted b_P) or the rolling benchmark regressions (in which case they are denoted by b_B), and N is the degrees of freedom.[8]

Although b_B and b_P are consistent estimates of β or β' under the null hypothesis, under the alternative hypothesis of publication bias, only the estimate b_B (of either β or β') provides a consistent estimate of how the minimum wage effect changes over time. Thus, a test of the statistical significance of the difference between b_P and b_B provides a test for publication bias. More specifically, as long as the estimate b_B from the benchmark specification is efficient, then the results in Hausman [1978] imply that under the null hypothesis of no publication bias, the statistic

$$(b_P - b_B)^2 / [\mathrm{Var}(b_P) - \mathrm{Var}(b_B)]$$

is distributed as $\chi^2(1)$, and the null hypothesis is rejected at the α-percent level if this statistic exceeds the α-percent critical value; the efficiency of b_B eliminates the covariance between the two estimates from the test statistic.[9]

8. If there is structural change, then the employment equation we use to obtain the series t and g is misspecified, because it assumes a constant minimum wage effect. However, the equations estimated in the published literature are also specified with constant minimum wage effects, and we are attempting to replicate the change in estimated minimum wage effects that these published studies would have reported had they used the benchmark specification; the rolling regression estimates of a valid benchmark specification provide a consistent estimate of this change.

9. We are grateful to Kevin Lang for suggesting this framework. Note that this case differs slightly from the application of Hausman tests to model specification. In specification tests, the estimate that is efficient under the null is inconsistent under the alternative (this is the reason the specification test is of interest, as it tells us whether or not we can use the efficient estimator). However, the distributional results in Hausman [1978] for the statistic in the text do not require this restriction, and as we know that b_P is inefficient under the null hypothesis (both because it is based on a limited number of possible sample periods and because it is subject to random variation in specifications across published studies), the application of the test to the question of publication bias instead requires the estimate from the benchmark specification (b_B) to be efficient under the null.

Card and Krueger report the estimated coefficient from a regression of the log absolute values of the t-statistics from the published studies on the log square root of the degrees of freedom. This estimate, which corresponds to the regression in panel A of Table I, is –.81, with a standard error of .70. As noted above, implementation of the above test statistic also requires an asymptotically efficient estimate (under the null) of the corresponding coefficient from the time-series of estimated absolute t-statistics from the benchmark specification (b_B). LM tests indicated significant evidence of first-order serial correlation (but no higher-order serial correlation), as well as heteroscedasticity corresponding to increased precision of the estimates with more degrees of freedom. We therefore reestimated the meta-analysis regression correcting for first-order serial correlation and for heteroscedasticity of this form.[10]

The corresponding estimate is reported in column (1) of panel A of Table II. The efficient estimate of the coefficient from the t-statistic regression is –.39, with a standard error of .44. As the second-to-last row of panel A reports, the *p*-value from the Hausman test of the null hypothesis of no publication bias is .44, indicating that we fail to reject this null hypothesis.[11] This implies that the deviations of the elasticities in the published studies from the regression line produced by the rolling regression, although in the direction of publication bias, could easily have been produced by chance.[12] A similar analysis in panel B, based on our estimates of the regression of the absolute coefficient estimates on the square root of the degrees of freedom, leads to the same conclusion. Indeed, in this case the *p*-value is .99, as the estimated change in the minimum wage effect from the rolling regressions is almost identical to that from the published studies.

Our analysis to this point hinges on accepting Solon's specification as free of publication bias, and hence as a valid benchmark specification. If the Solon specification itself reflects publication bias, then a comparison of estimates from the Solon specification with those from published studies does not provide a valid test for publication bias. To see this, suppose that researchers tended to select specifications by looking for negative effects of minimum wages. They would tend to report those specifications with relatively large negative effects, for the sample period they were using. However, because the estimated effects are partly random, we would expect the estimated minimum wage effect to be attenuated as the sample period is extended, reflecting regression to the mean in the estimated minimum wage effect.[13] In other words, rolling regression estimates of a specification that itself reflects publication bias might generate the same qualitative patterns in the estimated coefficients or t-statistics that publication bias generates in a sequence of published studies, hence biasing our test towards finding no evidence of publication bias.

We address this concern in two ways. First, we repeat the previous analysis for two alternative benchmarks based on specifications used by Card and Krueger [1995b] in their assessment of the evidence from time-series estimates of minimum wage effects. The first specification adds the log of the adult male employment-to-population ratio as an inde-

10. Because of the overlapping samples, one could, with some assumptions, solve for the error structure in the meta-analysis regressions using the rolling regression estimates. However, the first-order serial correlation structure fits the residuals very well, and therefore seems a good approximation to the true process.

11. The standard error of the Card and Krueger estimate is undoubtedly biased downward because of overlapping samples, but a larger standard error would only strengthen our failure to reject the null, since the denominator of the test statistic would be larger.

12. The estimates of ρ near unity in panel A suggest that the residual may be nearly I(1), which might argue for using a first-difference estimator instead of the AR correction used in Table II. However, in a closely-related estimation problem (with unknown ρ), Canjels and Watson [1997] show that the Prais-Winsten estimator is the preferred estimator when ρ is likely to be near one, but may not be exactly one. (In particular, this estimator is more robust to variation in ρ than is the first-difference estimator.) Using our data, the Prais-Winsten estimates were very similar to the maximum likelihood estimates in Table II. For example, the estimated coefficient (standard error) of the log square root of degrees of freedom corresponding to column (1) of Table II was –.47 (.32). The Prais-Winsten estimates were also similar to the maximum likelihood estimates in the other columns of this panel, discussed below. Straight first-difference estimates (correcting the differenced regression for heteroscedasticity) also were negative, with larger absolute values than the estimates from the published studies reported by Card and Krueger (which is inconsistent with publication bias); however, the first-difference estimates were quite imprecise.

13. This is a potential issue because Solon's paper used data only through 1979.

TABLE II
Meta-Analysis Regressions for Time-Series Estimates of Employment Effects of Minimum Wages, Quarterly Data, 1968–1988: AR(1), Heteroscedasticity-Corrected Estimates

	Solon Specification (1)	Incl. Adult Male Empl./Pop. (2)	Also Incl. Log Real Manu. Wage (3)
A. *Regressions of Log Absolute Values of t-Statistics on Log Square Root of Degrees of Freedom*			
Constant	1.57	2.02	2.06
	(1.00)	(1.12)	(.94)
Log square root of degrees of freedom	–.39	–.64	–.65
	(.44)	(.49)	(.41)
ρ	.94	.95	.94
	(.04)	(.03)	(.03)
P-value, H_0: no publication bias	.44	.74	.77
Data through 1993	.24	.56	.61
B. *Regressions of Absolute Coefficient Estimates on Square Root of Degrees of Freedom*			
Constant	.15	.17	.17
	(.01)	(.01)	(.01)
Square root of degrees of freedom	–.006	–.009	–.009
	(.001)	(.001)	(.001)
ρ	.74	.65	.69
	(.08)	(.09)	(.08)
P-value, H_0: no publication bias	.99	.80	.78
Data through 1993	.79	.98	.96

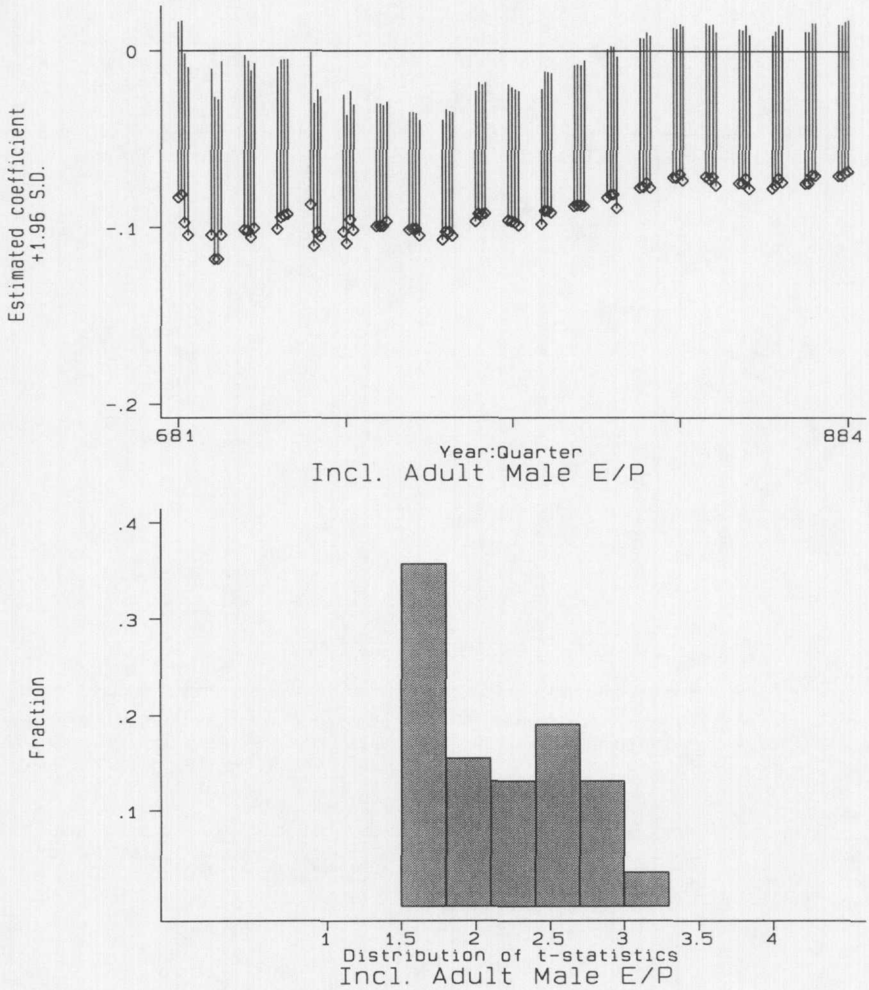
Standard errors are reported in parentheses. There are 84 observations used in each column. Minimum wage employment equations from which the coefficients and t-statistics are obtained are described in the text. The heteroscedasticity correction is based on a regression of the squared residuals on the log of the degrees of freedom in Panel A, and the degrees of freedom in Panel B. After transforming the data to obtain homoscedastic errors, maximum likelihood was used to estimate the AR(1) model. The test for publication bias is based on a Hausman test of the significance of the difference between the estimated coefficient from these regressions and the estimate from the published studies. For Panel A, the estimate from the published studies, reported by Card and Krueger, is –.81 (with a standard error of .70). For Panel B, the estimate (computed from the estimates in Table 6.1 of *Myth and Measurement*, and degrees of freedom supplied by Alan Krueger) is –.006 (with a standard error of .012).

pendent variable. The second specification, for which they offer a priori theoretical arguments, also adds the log of the average real manufacturing wage.[14] If these specifications were not considered by earlier researchers, or were discarded as part of their selective specification search, then the estimated minimum wage effects from these specifications should not show the same attenuation that earlier published specifications would show had the latter been chosen with the goal of generating negative minimum wage effects for the sample periods under study.

14. We believe it is fair to assume that, in suggesting these alternative specifications, Card and Krueger did not engage in selective specification search to find negative minimum wage effects.

The results are reported in columns (2) and (3) of Tables I and II, and in Figures 2 and 3. For both specifications, the estimated minimum wage effects decline over time and the absolute values of the t-statistics are clustered around two. This is evidence that—were it to come from published studies—would point to publication bias; indeed, the absolute values of the estimated slope coefficients in both tables are larger for these specifications than for the Solon specification. Repeating the Hausman tests for these specifications, we see that, again, the null hypothesis of no publication bias cannot be rejected. In fact, the evidence in panel B is not even in the direction of publication bias, as the estimated coefficients from the rolling regression estimates are more negative than that from the published studies

FIGURE 2
Card/Krueger Specifications, 1968–1988



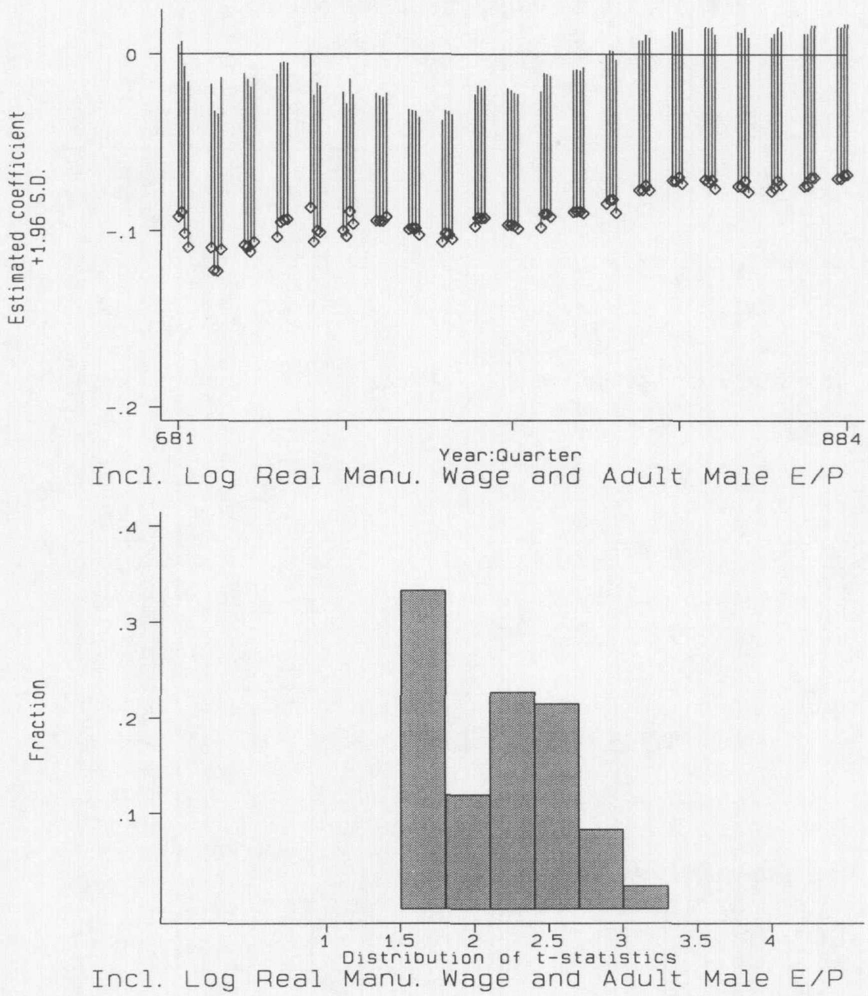
(reported in the note to Table II). This latter evidence is quite important, because one might question the power of the Hausman test we propose. Obviously, though, if b_B is more negative than b_P, then this is not an issue.

We also re-computed the tests for publication bias using data through 1993, although it seems to us that these additional observations are less relevant to the question of publication bias. Including the more recent data may provide a better estimate of the change in the minimum wage effect over time, but it seems more appropriate to base our conclusions on a comparison between the published studies (which use data only through 1988) and what these studies would have obtained had they used the benchmark specification. Regardless, the qualitative conclusions, as reported in the last row in each panel of Table II, are unchanged; there is still no significant evidence against the null hypothesis of no publication bias.[15]

The second way in which we address our concern that the Solon specification might not be a valid benchmark is to estimate the meta-

15. The differences in p-values indicate that the evidence is a bit more in the direction of publication bias in the t-statistic regressions when data through 1993 are included, and a bit less so in the regressions for the absolute value of the estimated coefficient.

FIGURE 3
Card/Krueger Specifications, 1968–1988



analysis regressions using only the estimates for samples that included observations for the years after 1979, the last year included in Solon's study. In particular, if Solon's specification is contaminated by publication bias, we might expect the attenuation of the estimated minimum wage effect to emerge only in subsequent years. As shown in column (1) of Table III, the estimated minimum wage effects using this more limited sample of estimates do decline faster than in the longer sample. However, the estimated changes in the absolute minimum wage effect are even more negative in the two alternative specifications proposed by Card and Krueger, suggesting that the sharper decline in estimated minimum wage effects in this latter period using the Solon specification reflects parameter changes, rather than publication bias in that specification.

Finally, although our formal test for publication bias hinges on a number of statistical assumptions, we regard this test as a useful heuristic device for assessing the statistical strength of the evidence of publication bias. However, the key result in this paper is that we are able to replicate the results in Card and Krueger's meta-analysis of published studies, which they interpret as suggestive of publication bias, using benchmark specifications—

TABLE III
Meta-Analysis Regressions for Time-Series Estimates of Employment Effects of Minimum Wages, Quarterly Data, 1980–1988: AR(1), Heteroscedasticity-Corrected Estimates

	Solon Specification (1)	Incl. Adult Male Empl./Pop. (2)	Also Incl. Log Real Manu. Wage (3)
A. *Regressions of Log Absolute Values of t-Statistics on Log Square Root of Degrees of Freedom*			
Constant	3.63	3.89	3.96
	(.67)	(.75)	(.77)
Log square root of degrees of freedom	–1.28	–1.44	–1.47
	(.28)	(.32)	(.33)
ρ	.69	.70	.72
	(.12)	(.12)	(.12)
B. *Regressions of Absolute Coefficient Estimates on Square Root of Degrees of Freedom*			
Constant	.16	.17	.17
	(.02)	(.02)	(.02)
Square root of degrees of freedom	–.007	–.009	–.009
	(.002)	(.002)	(.002)
ρ	.47	.59	.61
	(.15)	(.14)	(.13)

Standard errors are reported in parentheses. There are 36 observations used in each column. See notes to Table II for details.

that are arguably free of publication bias—estimated over successively longer periods.

IV. WHY DID ESTIMATED MINIMUM WAGE EFFECTS DECLINE?

A natural question that arises from our analysis is why "objective" estimates of minimum wage effects apparently declined as later years were added to the samples. Although this question is not the focus of this paper, our conclusion that we cannot reject the null of no publication bias, and that the data are consistent with declining minimum wage effects over time, would be bolstered if these parameter changes could be explained. We can think of three possible sources of structural change.[16]

First, the Kaitz index fell over much of the 1980s, implying that the percentage of teens bound by the minimum wage was declining during this period. Because the time-series estimate of the minimum wage effect in the simple linear employment equation can be viewed roughly as an average of the minimum wage effect prevailing in each year, a nonlinearity in the true relationship between employment and the Kaitz index could produce changes over time in the estimated coefficient in the linear model.[17] In particular, a positive relationship between the absolute size of the minimum wage effect and the level of the Kaitz index could explain the declining minimum wage effect found when data from the 1980s are added to the sample. As it turns out, however, the data are inconsistent with this explanation. Because the values of the Kaitz index in the 1980s were still well above those in the 1950s or 1960s, the sample average of the Kaitz index continues to rise as observations from the 1980s are added. This suggests that a nonlinearity of the type that could have produced an increase in the estimated minimum wage coefficient in the 1970s also would have produced a further increase as the 1980s data were included, rather than the decline we observe.

16. Williams and Mills [1995] offer an alternative explanation associated with the time-series properties of teenage employment rates and the Kaitz index. In particular, they argue that the decline in the minimum wage coefficient in standard regressions occurs because these two variables are integrated of different orders.

17. Such a nonlinearity could arise for reasons discussed in Neumark and Wascher [1994].

Second, early increases in the Kaitz index partly reflected changes in coverage by minimum wage laws, whereas later changes were generated mainly from changes in minimum wage levels or average wages. If the effects of coverage changes are larger than are the effects of changes in the relative minimum wage, the effective absence of further increases in coverage after the early 1970s could generate a declining employment effect.

Finally, if the workers for whom minimum wage workers are most substitutable are not those earning the average wage, but those earning a wage lower in the distribution (such as the 25th centile), then the well-documented rise in wage dispersion over the 1980s suggests that the standard Kaitz index, which is measured using the mean wage, may not accurately measure the decline in the effective minimum wage during that period. In particular, the Kaitz index measured using the average wage would decline more over the 1980s than a Kaitz index measured using the 25th-centile wage. The use of the "incorrect" Kaitz index based on the average wage would then produce an underestimate of the absolute minimum wage elasticity. A complete assessment of this hypothesis entails a major undertaking, because it requires micro-data on wage distributions over a long period. Nonetheless, given evidence that estimated minimum wage effects declined more sharply in the 1980s, we think this is the most plausible of the three explanations that we have offered.[18]

V. CONCLUSION

This paper presents an approach to testing for publication bias in the economics literature, where it may be difficult to distinguish the effects of publication bias from changes in parameters. The basic approach is to identify a benchmark specification that is unlikely to have been influenced by selective specification search, and to compare changes in the estimate of interest across published studies using different sample periods to changes in estimates from the benchmark specification estimated over similar periods. This approach contrasts with the clinical trial literature, in which the null hypothesis of no publication bias is based on a comparison of estimates across published studies using different sample sizes with a parameter assumed to be constant.[19] The key element in our approach, of course, is the identification of a benchmark specification that does not reflect publication bias. While this injects an element of subjectivity, and one researcher's choice may not fully convince others, we believe we have demonstrated that choosing a benchmark specification is necessary in testing for publication bias in time-series studies.

We apply this approach to evaluate the suggestions by Card and Krueger [1995a; 1995b] that the time-series evidence on minimum wage effects reflects selective specification search motivated by authors' and editors' predisposition towards finding and publishing results that conform to conventional economic theory. As Card and Krueger report, the absolute magnitudes of estimated minimum wage effects in published studies fell as the time series were extended, consistent with publication bias. However, we find that the absolute magnitudes of the estimated effects also fall over time for specifications that emerged in the literature subsequent to all or most of the published work that Card and Krueger reevaluate, including specifications that they advocate in part on a priori theoretical grounds. In addition, statistical tests indicate that the reduction over time in the estimated minimum wage effect in the published studies is not significantly larger than the reduction in the estimates from the benchmark specifications using successively longer sample periods. Thus, conditional on having identified a benchmark specification that is free of publication bias, we cannot reject the hypothesis of no publication bias in time-series studies of minimum wage effects on employment.

18. This explanation was suggested by Lawrence Katz at an AEI conference on minimum wages, and is supported in recent unpublished empirical evidence reported by Donald Deere, Kevin Murphy, and Finis Welch.

19. In this paper, we discuss the approach as applied to time-series studies. In principle, a similar approach could be applied to cross-sectional studies, where true parameters may vary by location, demographic group, etc., perhaps leading to corresponding variation in parameter estimates across published studies. However, because we would not necessarily expect variation in sample size to be correlated with variation in the representation of different locations or demographic groups in the data, such parameter variation is less likely to bias tests for publication bias when meta-analysis methods from the clinical trial literature are used.

While we do not make any claim that these findings generalize to other areas of empirical research in economics, provisionally, at least, our evidence should make researchers less skeptical of published evidence that is consistent with widely-held priors.

REFERENCES

Berlin, Jesse A., Colin B. Begg and Thomas A. Louis. "An Assessment of Publication Bias Using a Sample of Published Clinical Trials." *Journal of the American Statistical Association*, 1989, 381–92.

Brown, Charles, Curtis Gilroy and Andrew Kohen. "Time-Series Evidence on the Effect of the Minimum Wage on Youth Employment and Unemployment." *Journal of Human Resources*, Winter 1983, 3–31.

Canjels, Eugene and Mark W. Watson. "Estimating Deterministic Trends in the Presence of Serially Correlated Errors." *Review of Economics and Statistics*, May 1997, 184–200.

Card, David and Alan B. Krueger. "Time-Series Minimum-Wage Studies: A Meta-Analysis." *American Economic Review*, May 1995a, 238–43.

______. *Myth and Measurement: The New Economics of the Minimum Wage*. Princeton, N.J.: Princeton University Press, 1995b.

Chalmers, T. C., H. Levin, H. S. Sacks, D. Reitman, J. Berrier and R. Nagalingam. "Meta-analysis of Clinical Trials as a Scientific Discipline, I: Control of Bias and Comparison With Large Cooperative Trials." *Statistics in Medicine*, 1987, 315–25.

De Long, J. Bradford and Kevin Lang. "Are All Economic Hypotheses False?" *Journal of Political Economy*, December 1992, 1257–72.

Hausman, Jerry. "Specification Tests in Econometrics." *Econometrica*, 1978, 1251–71.

Leamer, Edward P. *Specification Searches: Ad Hoc Inference with Nonexperimental Data*. New York: John Wiley and Sons, Inc., 1978.

Lucas, Robert E. Jr. "Econometric Policy Evaluation: A Critique," in *The Phillips Curve and Labor Markets*, Karl Brunner and Allan Meltzer, eds. Amsterdam: North Holland, 1976.

Neumark, David and William Wascher. "Minimum Wage Effects and Low-Wage Labor Markets: A Disequilibrium Approach." NBER Working Paper No. 4617, 1994.

Siskind, Frederic. "Minimum Wage Legislation in the United States: Comment." *Economic Inquiry*, 1977, 135–38.

Solon, Gary. "The Minimum Wage and Teenage Employment: A Reanalysis with Attention to Serial Correlation and Seasonality." *Journal of Human Resources*, Spring 1985, 292–7.

Welch, Finis. "Minimum Wage Legislation in the United States." *Economic Inquiry*, 1974, 285–318.

______. "Minimum Wage Legislation in the United States," in *Evaluating the Labor Market Effects of Social Programs*, Orley Ashenfelter and James Blum, eds. Princeton, N.J.: Princeton University Press, 1976.

______. "Minimum Wage Legislation in the United States: Reply." *Economic Inquiry*, 1977, 139–42.

Wellington, Alison J. "Effects of the Minimum Wage on the Employment Status of Youths: An Update." *Journal of Human Resources*, Winter 1991, 27–46.

Williams, Nicolas and Jeffrey A. Mills. "The Minimum Wage and Teenage Employment: Is There a Relationship?" Manuscript, University of Cincinnati, 1995.