Reply

Russell Davidson and James G. MacKinnon

We welcome the comments on our paper by Ackerberg and Devereux (AD) and Blomquist and Dahlberg (BD), and we are happy to take this opportunity to respond briefly to them.

We agree wholeheartedly with both AD and BD that is it unprofitable to try to obtain meaningful empirical results when all available instruments are very weak. At least in our experiments, however, it is only in this case that JIVE1 ever outperforms LIML in terms of dispersion; see Figure 5 and the top left panel of Figure 6. Thus we disagree with BD that our "categorical rejection of JIVE" is not in accord with our simulation results. BD are, however, absolutely correct to point out that our experiments deal with a very simple case and do not tell the whole story.

AD discuss two quite new estimators, IJIVE and UIJIVE, which apparently improve upon the JIVE1 estimator. Their comment suggests that our conclusion that JIVE1 is inferior to LIML does not necessarily apply to these estimators. In order to investigate the properties of these two estimators, we performed a new set of simulation experiments, with the same design as those we describe in the paper. The results of the new experiments are presented graphically in Figures 1a–6a, which, like Figures 1–13 from the paper itself, are available from the JAE Data Archive website at `www.econ.queensu.ca/jae/`. The new figures deal with the same cases as Figures 1–6 of the paper. Results for IJIVE and UIJIVE are added, and, for readability, results for 2SLS are removed.

The superiority of IJIVE and, especially, UIJIVE relative to JIVE1 emerges quite clearly from the new figures. What is particularly interesting is that UIJIVE tends to be substantially less dispersed than the other JIVE estimators. When the instruments are weak, it is often much less dispersed than LIML. Our tentative conclusion is that UIJIVE is the best JIVE estimator to date and may well be worth using in practice.

The simulation results suggest that IJIVE and UIJIVE, like JIVE1 and LIML, have no moments, and we have confirmed this analytically. This leads us to question the interpretation of the results in Phillips and Hale (1977) and in Ackerberg and Devereux (2003) that purport to yield approximate biases for these estimators. These results are based on stochastic expansions which can be represented schematically as

$$n^{1/2}(\hat{\beta}_2 - \beta_{20}) = t_0 + n^{-1/2}t_1 + o_p(n^{-1/2}), \tag{1}$$

where $\hat{\beta}_2$, as in the paper, is an estimator of the coefficient of the endogenous regressor in the second-stage regression, β_{20} is the true value of the parameter, t_0 is a random variable that has a normal distribution with zero expectation, and t_1 is a random variable with a nonzero expectation, which provides the approximate bias of the estimator.

The catch is that the $o_p(n^{-1/2})$ remainder in (1) has no moments. However, the fact that it does tend to zero in probability implies that the distribution of the first two terms of the truncated expansion converges to that of the estimator itself for large sample sizes n. We have expressed the stochastic expansion as an expansion in powers of $n^{-1/2}$, but it is just as possible to express it in the form of small-sigma asymptotics, as is done by Phillips and Hale. The two expansions appear to be equivalent, at least to order $n^{-1/2}$.

The techniques proposed in Appendix 1 of Phillips and Hale for simplifying the calculation of JIVE-type estimators can equally well be applied to IJIVE and UIJIVE. In all cases, the instruments $\tilde{\boldsymbol{Y}}_t$ can be expressed in terms of the fitted values $\boldsymbol{P_W Y}$ of the first-stage regressions of 2SLS and the diagonal elements h_t of $\boldsymbol{P_W}$, as we show for JIVE1 in the paper. Since IJIVE is just JIVE1 using projected variables, the same formulas can be used for it. For UIJIVE, it is easy to see that the instruments are given by a very similar formula, namely,

$$\tilde{\boldsymbol{Y}}_t = \frac{1}{1 - h_t + \omega}\big((\boldsymbol{P_W Y})_t - (h_t - \omega)\boldsymbol{Y}_t\big), \tag{2}$$

in which h_t is replaced by $h_t - \omega$, with $\omega = 2/n$ in the case with just one included endogenous variable, or $\omega = (g+1)/n$ more generally when there are g included endogenous variables. Of course, here $\boldsymbol{Y}$ and $\boldsymbol{W}$ should be interpreted as matrices of endogenous variables and instruments that have been projected off the included exogenous variables.

The formula (2) suggests that, whereas JIVE1 and IJIVE involve "omit-1" fitted values, UIJIVE involves "omit-less-than-1" fitted values. Using this formula together with equation (3) of our paper to calculate the UIJIVE estimator is orders of magnitude faster, for large n, than using the elegant formula in Ackerberg and Devereux (2003), which involves manipulating $n \times n$ matrices.

A point that emerges clearly from our simulations and those presented by AD in their comment is that, although UIJIVE is constructed so as to reduce mean bias, it does not do nearly so well as regards median bias. This implies that the estimator is significantly skewed, a point that would be worth subsequent investigation. It would also be interesting to look more closely at the modified LIML estimator proposed by Fuller (1977), for which moments exist, to see how well it performs relative to UIJIVE.