Introduction: Essays in Honor of Robert J. LaLonde

Orley Ashenfelter, *Princeton University*
Guest Editor

David Card, *University of California, Berkeley*
Guest Editor

It is our pleasure to introduce this volume of papers, all of which were presented at a conference in honor of Robert LaLonde held in Chicago in the spring of 2015.[1]

It is conventional for a Festschrift conference ("a gathering of scholars to present papers in honor of a colleague") to be held on the honoree's sixtieth birthday or thereabouts. In Bob's case, it was decided to move things up a little. Diagnosed with what the National Institutes of Health have concluded is a unique neurological disease, it was hoped that Bob could both enjoy his colleague's admiration and participate in the conference to honor him. On both scores, the conference and this set of papers were a remarkable success.

Bob earned his undergraduate degree at the University of Chicago, and, apart from a few years in East Lansing, he has spent almost all his academic career there. But there were two important exceptions. In 1987–88, Bob served as a senior staff economist at the president's Council of Economic Advisors (CEA). Fresh from finishing a pioneering dissertation focused on the quantitative evaluation of employment and training programs, Bob stepped into the Washington debates with just what everyone wanted: a strong case for the value of evidence-based public policy.

[1] The conference was held at the Federal Reserve Bank of Chicago. We are deeply grateful to Daniel Sullivan and the staff of the Bank for their hospitality and generosity.

[*Journal of Labor Economics*, 2017, vol. 35, no. S1]

The other major break was the 5 years he spent at Princeton. Bob completed his dissertation at a moment when the techniques of program evaluation were being extended to settings not previously imagined. There were two important forces behind this development. One was the focus on econometric methods and their reliability. Causal inference, as opposed to correlation, was a novel idea, and the search for credible methods for establishing causality was a driving force behind many dissertations and research projects. Bob's time at the CEA demonstrated how important the credibility of causal inference had become. Those were the Reagan years, and if a program did not pay off, it was in serious jeopardy. The underlying justification really did need to be based on evidence.

A second factor in the program evaluation revolution was the increasing availability of microeconomic data and the advent of field experimentation for the study of economic behavior. Bob's dissertation, and his landmark 1986 paper in the *American Economic Review*, really was convincing. What is more, it demonstrated, by comparing standard econometric estimators with the results of a field experiment, using real data, that credible program evaluation is hard to do. Many nonexperimental estimators gave misleading results. Even worse, it was difficult to predict which would do well.

In retrospect, Bob's work cast a breathtakingly long shadow. Fortunately, the researcher behind this work, careful and hardworking but completely without guile, was the perfect ambassador for the message. Today we see many new papers each week reporting the results of field experiments, routinely summarizing tests for the balance of covariates, and paying careful attention to sources of data and issues of follow-up. All of this and more was in Bob's early work, published some 30 years ago.

A glance through Bob's curriculum vitae shows a long list of studies of topics that are sometimes far removed from the evaluation of employment and training programs. These cover everything from collective bargaining to the analysis of wine prices (full disclosure: one of us prompted that work).

The economics of immigration and the analysis of the causes and consequences of job displacement are two especially large aspects of Bob's scholarly work. Both of these areas remain highly topical, and they show just how much Bob's work has been inspired by deep-seated social issues that raise difficult empirical questions for economists.

A prominent feature of Bob's scholarship, including his many papers on worker displacement and the economics of immigration, is the close connection to his early work on program evaluation. There is always a careful explication of the problem at hand, a clearly elucidated measurement method to study the problem, and then an ambitious, but careful, data analysis. And, in the framing of the problem, there is always the goal of evaluation and its implications for public policy.

The papers in this volume cover a wide array of problems, but we think they meet the standards that Bob would set for himself and his students for credible and useful research. And we are especially pleased that we could actively participate in the discussions that followed many of the papers and led to substantive improvements in the final drafts.

Three of the papers follow up on Bob's evaluation of nonexperimental program estimators. Griffen and Todd compare alternative estimates of the effect of the Head Start program to experimentally derived estimates from the Head Start Impact Study. Departing slightly from Bob's original design, the authors use data from a single source (the Early Childhood Longitudinal Study–Birth Cohort) for both the program participants and nonparticipants. As they note, this resolves one of the major concerns in LaLonde's analysis, which is that the set of control variables and outcomes available for the treatment and comparison groups can differ when the two samples are drawn from different sources. They also implement matching-style estimators (which were relatively unknown to economists in the early 1980s) and conduct some of the specification tests later proposed by Heckman and Hotz (1989) to help distinguish between alternative estimators. Interestingly, despite the closer alignment of the data for the treatment and comparison groups and the availability of new methods and specification tests, they reach a similar conclusion to Bob's, noting that while "some of the methods for some of the outcomes reproduce the experimental results fairly closely . . . a priori it would be difficult to know which estimator would work well for any particular outcome."

Calónico and Smith hew even closer to LaLonde's original paper, providing a replication and update of his analysis of female participants in the National Supported Work (NSW) Demonstration program. Their findings offer broad support for the lessons that have been drawn over the past three decades from LaLonde's original paper and subsequent re-analyses of his results for males. Specifically, they confirm Smith and Todd's (2005) conclusion that traditional econometric methods perform about as well as newer matching style estimators. They also affirm LaLonde's (1995) conclusion that selection biases for female participants in compensatory training programs like NSW are less severe than for male participants.

Bisbee, Dehejia, Pop-Eleches, and Samii build on LaLonde (1986) in a different way, asking whether quasi-experimental estimates derived from one setting can be statistically extrapolated to other settings. Their approach is an interesting alternative to the theoretical model–based extrapolation sometimes suggested by economists (e.g., Deaton and Cartwright 2016). They focus on instrumental variables (IV) estimates of the effect of having a third child on a mother's labor supply, identified using the gender mix of the first two children as an instrument for additional fertility (Angrist and Evans 1998). They present and analyze a series of IV estimates for 52 differ-

ent countries and up to five different time periods per country. Their findings suggest that relatively simple statistical models of the heterogeneity in underlying causal effects can be useful in transporting findings across highly diverse settings.

In an important set of papers coauthored with John Ham (Ham and LaLonde 1996; Eberwein, Ham, and LaLonde 1997), Bob showed how researchers could estimate *dynamic treatment effects* on the transition rates between labor market states using a combination of experimental data and econometric techniques. Ba, Ham, LaLonde, and Li apply these methods to re-examine the dynamic treatment effects of the Job Training Partnership Act (JTPA) classroom training programs on female participants. JTPA participants, like those in many other training programs, often have a delay between assignment to a program and the commencement of training. Moreover, trainees can work and participate in training at the same time. Despite these complexities, the authors confirm the earlier conclusion of Ham and LaLonde (1996) and Eberwein et al. (1997) that the JTPA program improved participants' ability to successfully move from unemployment to employment, leading to shorter average spells of nonemployment.

Reflecting his long-standing concern with the plight of disadvantaged families, Bob has written a number of papers on the effects of incarceration on female offenders. The paper by Butcher, Park, and Piehl takes a step back from LaLonde's work and asks how women and men who have been convicted of a felony are treated by the criminal justice system. Using data on the universe of convicted felons in Kansas from the past decade, they reach two main conclusions. First, conditional on the observed facts of the case, women are substantially less likely to serve time in prison than men. This suggests that women who do serve time are likely to be even more negatively selected than their male counterparts. Second, there is substantial variation across judges in the harshness of the sentences they give to female offenders that is highly correlated with the relative severity of sentences they give to male offenders. Like their male counterparts, the amount of time that female offenders spend in prison is therefore highly dependent on the identity of the judge who hears their case.

As noted earlier, a major theme in Bob's work is the analysis of job displacement. The paper by Farber uses data from the Displaced Worker Surveys of the Current Population Survey to shed new light on job losses during the Great Recession of 2007–9. The job loss rate in these years surpassed even the rate in the 1982–83 recession; levels of unemployment and underemployment remained elevated for many years after. Farber shows that the average earnings losses of displaced workers in the United States are largely driven by the loss of full-time work: many job losers move at least temporarily to part-time work, while others remain unemployed up to 4 years after their displacement. Even among those who return to full-time work, however, there is wide variation in their outcomes, with older

and long-tenure workers tending to experience larger losses and a surprising fraction of younger, short-tenure workers actually experiencing a gain in their average pay.

The final three papers in this issue present evaluations of ongoing or experimental programs. The paper by Jepsen, Mueser, and Troske examines the effect of passing the GED exam on postsecondary schooling enrollment and course completion. Using regression discontinuity methods, the authors conclude that successfully passing the GED is associated with a modest positive effect on college enrollment rates, particularly for female test-takers. They find a comparable effect on course completion rates, suggesting that GED qualifications are indeed adequate to allow students to pass the courses they attempt. Nevertheless, the impact on the total number of courses completed is small—around one college-level class for male students and two for female students—potentially explaining the relatively small earnings impacts for GED completion that have been estimated in earlier literature.

The paper by Brinkman, Hasan, Jung, Kinnell, and Pradhan examines the impacts of a pre-school playgroup program for children in rural Indonesia. The evaluation includes two components: one based on randomized order for program implementation across sites and the other based on quasi-experimental comparisons between children in program sites and comparison sites with no program. Both designs show a comparable but modest effect on enrollment in playgroups, but the randomized order of treatment design implies relatively small treatment effects on overall months of exposure to playgroups, which the authors argue were too small to significantly affect child outcomes. The nonexperimental design created a larger gap in exposure to playgroups, which was associated with a modest positive impact on one outcome domain. A lesson the authors draw is that the specific package of playgroup services offered in this program led to an unintended fall in kindergarten enrollment that may have offset the gains to the program itself—the kind of "substitution problem" that has been identified in Bob's work and many other studies in the United States.

The final paper, by Jacobson and Davis, addresses a long-standing question in the job training literature, one raised in LaLonde's work and many other studies: Why do female participants seem to gain more from training then males? The analysis makes use of unusually detailed data for participants in the Workforce Investment Act program in Florida, including many years of pre- and post-program earnings data and precise information on the field of training for each participant. Consistent with the existing literature, these authors find that female participants have higher average earnings gains from 5 years before to 6 years after training than males. Surprisingly, this difference is entirely attributable to the fact that females tend to select relatively "high-gain" fields for their training. Within field, males actually have slightly greater average gains than females. The authors conclude that incorporating information on the field of training could be im-

portant both for improving our evaluations of training programs and for the ultimate operation of the programs.

References

Angrist, Joshua D., and William Evans. 1998. Children and their parents' labor supply: Evidence from exogenous variation in family size. *American Economic Review* 99, no. 3:450–77.

Deaton, Angus, and Nancy Cartwright. 2016. Understanding and misunderstanding randomized controlled trials. NBER Working Paper no. 22594, National Bureau of Economic Research, Cambridge, MA.

Eberwein, Curtis J., John C. Ham, and Robert J. LaLonde. 1997. The impact of being offered and receiving classroom training on the employment histories of disadvantaged women: Evidence from experimental data. *Review of Economic Studies* 64, no. 4:655–82.

Ham, John C., and Robert J. LaLonde. 1996. The effect of sample selection and initial conditions in duration models: Evidence from experimental data on training. *Econometrica* 64, no. 1:175–205.

Heckman, James J., and V. Joseph Hotz. 1989. Choosing among alternative nonexperimental methods for estimating the impact of social programs: The case of Manpower training. *Journal of the American Statistical Association* 84, no. 408:862–80.

LaLonde, Robert J. 1986. Evaluating the econometric evaluations of training programs using experimental data. *American Economic Review* 76, no. 4:604–20.

———. 1995. The promise of public sector–sponsored training programs. *Journal of Economic Perspectives* 9, no. 2:149–68.

Smith, Jeffrey A., and Petra E. Todd. 2005. Does matching overcome LaLonde's critique of nonexperimental methods? *Journal of Econometrics* 125, nos. 1–2:305–53.