

Comments

María Laura Alzúa: Lucas Ronconi evaluates the income effect of a workfare program in Argentina, known as *Trabajar*, focusing mainly on income effects. Given the lack of experimental data with which to conduct the evaluation, the author relies on nonexperimental methods. The topic of the paper is highly relevant, since workfare has become one of the most important policies not only in Argentina, but in other Latin American countries as well. Despite the high cost of such workfare programs, most impact evaluation is conducted with poor-quality data, if they are conducted at all. Moreover, the author suggests that both the lack of good targeting and low enforcement causes the program to be ineffective. One of the main policy implications of the paper, besides the self-targeting and enforcement implications, is that serious impact evaluations need to be conducted and should be mandated in the project design.

However, a number of issues need to be addressed in further extensions of the paper. Data availability allows the author to use outcomes like hours worked to measure whether the working conditionality serves its purpose, and the findings support the paper's premise of low enforcement. At the same time, a few points on the methodology used to assess the impact deserve some attention. One is the credibility of the identification assumptions needed when one uses propensity score matching and nonexperimental data, in particular the assumption that unobserved heterogeneity does not affect the participation decision. This assumption can be very strong in programs like *Trabajar*. A useful discussion about why the paper makes this assumption would help the reader trust the chosen methodology.

With regard to pursuing the difference-in-differences approach to study workfare programs, Ashenfelter's dip may be a concern, since program beneficiaries may be experiencing a temporary drop in income.¹ If this is the

1. Ashenfelter (1978).

case, the difference-in-differences methodology may be overestimating program effects. Another issue is that the methodology is not well specified. Matching estimators are not invariant to the window specified in the matching. Consequently, robustness could be improved by adding estimations with different windows. Finally, there is an important debate about the validity of bootstrapping standard errors when one uses propensity score matching estimators. The paper would greatly benefit from an explanation of the assumptions under which such calculations are valid.

G. David Rosas Shady: Lucas Ronconi's paper presents evidence on the targeting and the impact on revenues of Argentina's workfare program *Trabajar*.¹ Workfare or public works programs, which can be considered safety nets or active labor market interventions, have been implemented in many countries around the world to counter natural disasters and economic crisis. These programs usually provide short-term employment opportunities to unskilled workers through different national or regional projects, such as road construction. One of their main goals is to provide income transfers to the poor and to smooth their consumption during the crisis.

In Latin America, workfare programs were first implemented in the 1980s as a short-term response to the economic crisis and the increase in unemployment rates that affected many countries, and in many cases they were associated with the structural reforms and stabilization measures that were common to the region in that period. The persistence of high poverty rates and the inability of governments to improve labor market conditions and efficiently protect poor people drove some countries to better focalize their efforts and to convert these programs into permanent interventions.² These changes were consistent with the experience of the countries in the Organization for Economic Cooperation and Development (OECD), where similar programs were reoriented toward the long-term unemployed and vulnerable people.³

1. My comments are based on a previous version of Ronconi's paper, and the author has incorporated many of these comments in the current version.

2. For example, in response to the economic crisis of 2001, the Argentine government replaced *Trabajar* with *Jefes the Hogar*, which expanded the number of beneficiaries. Almost at the same time, Colombia introduced *Empleo en Acción*, a program oriented to the poorest unemployed workers. The Peruvian government designed a program similar to *Trabajar*, called *A Trabajar Urbano*, to help the poorest urban workers during the economic recession of 1999–2001. The implementation of this program began in 2002 when the economic recession was over, and it continues today under the name of *Construyendo Peru*.

3. Samaniego (2004).

Unfortunately, few well-designed impact evaluations of the Latin American workfare programs have been undertaken. The available evidence shows that while these programs can be successful in targeting poor workers, there is room for improvement. These evaluations also show that workfare programs can be effective in protecting the income of poor people during an economic crisis.⁴ Evidence from a large sample of developing countries confirms these results, but also emphasizes that the success of each program depends on its design features—for example, on the levels of the benefit (wage rate) that is offered to the beneficiaries and of the labor intensity on the specific projects that are financed.⁵ The evidence is less conclusive in the case of the European countries.⁶ The workfare programs seem to have positive benefits in the short term, but their medium-term results are less clear.⁷

The Argentine government initiated *Trabajar* in 1996 in response to growing unemployment, with financial and technical support from the World Bank. The program was designed to be a social safety net: the main objective was to reduce poverty through a temporary income support (namely, a low wage subsidy conditioned on participation in community work programs) to unemployed poor people. The targeting mechanism was based on the assumption that only poor people, with few prospects for employment, would participate in the program because the wage rate offered was relatively low. Following the strong economic and political crisis that affected Argentina at the end of the 1990s, in 2002 the number of program beneficiaries was expanded and the name changed to *Jefes de Hogar*.

The first impact evaluation of *Trabajar* was carried out by Jalan and Ravallion in 1999 in the context of the World Bank support.⁸ This evaluation was based on a quasi-experimental design and used the 1997 national household survey (*Encuesta de Desarrollo Social*, EDS) to construct the control group. The survey of beneficiaries was conducted almost at the same time as the EDS, and it used the same questionnaire and interview team. The authors found that the beneficiaries were more likely to be poor than were nonparticipants, and they found positive impacts on incomes. These results have been used to present *Trabajar* as an example of a successful public works program.⁹

4. Bouillon and Tejerina (2007).

5. Subbarao (2003).

6. Martin (1998).

7. Mato Díaz (2004).

8. Jalan and Ravallion (1999).

9. See, for example, Subbarao (2003).

Ronconi uses different data than Jalan and Ravallion and finds less positive results. He stresses the problems of targeting and enforcement that characterized *Trabajar*. In that sense, this paper represents an interesting contribution to the literature on the evaluation of public works programs. Nevertheless, the author could have placed more emphasis on the new evidence obtained using different data; included a discussion about the dynamics of the program; considered a more comprehensive literature review; and developed an analysis of the policy implications of its main conclusions. I address each of these issues in turn.

First, with regard to the new evidence, the author argues that the program was less pro-poor than the previous estimates suggested, and he considers that Jalan and Ravallion overestimate the impacts of *Trabajar*. Ronconi advances three main reasons, which are well supported by evidence: the self-targeting of the beneficiaries was unlikely; a negative incidence could exist, especially if the enforcement of the program was weak; and the unobserved heterogeneity probably biases the results.¹⁰ The author allows for negative net incidence of the program and controls for unobserved time-invariant heterogeneity. His paper thus seems to use a more representative sample of beneficiaries than Jalan and Ravallion, and the results are probably more credible. Nevertheless, the author could have better examined whether the level of the benefit that was offered to the recipients in 1997 and in subsequent years was low enough to support the self-targeting hypothesis. Also, he evaluates the impact of *Trabajar* in a different period than Jalan and Ravallion (he uses the EPH longitudinal data from 2000–02). If his purpose was to reassess the impacts found by Jalan and Ravallion, he could have used data from 1997 to replicate the analysis and then control for some of the potential bias mentioned before.¹¹ This would clarify the differences under the same circumstances.

It is not clear in the paper why a small sample size is more representative than in previous studies. The EPH survey was not designed to measure workfare recipients in a way that is statistically representative. Moreover, it is impossible to distinguish workfare recipients from other recipients of social aid, which in this case could be as high as 20 percent of the observed sample. Also, the author could have developed a more detailed discussion about the propensity score limitations, especially with a small sample of beneficiaries. Ronconi could also have recognized the importance of Jalan and Ravallion's

10. Ronconi recognizes this last point as a limitation of the results and presents a test for selection bias in the matching estimator (namely, the Sargan-Wu-Hausman test).

11. One possibility would be to use the 1997 EDS to build the control and treatment groups.

paper in the literature (it is only mentioned in a footnote). This paper was innovative in providing one of the first applications of the propensity-score matching technique for welfare analysis. It became a reference point for the use of national surveys (or censuses) to draw a comparison group when a proper evaluation design is not available.

Second, Ronconi could have considered a detailed discussion about the dynamics of the program during the 1997–2008 period. The author only mentions that Argentina had different public works programs at that time, with *Trabajar* being the largest, and that these programs share some basic features. He does not explain the main modifications that *Trabajar* underwent during the period and does not analyze whether those changes could affect the functioning and potential impacts of the program. More importantly, the operation of *Trabajar* could have changed over time. For example, it would be interesting to identify whether the enforcement mechanism was improved in *Jefes de Hogar*. Also, he does not consider the implications of the macroeconomic performance of the country in those quite different periods.

Third, the paper could have included a detailed discussion of the policy implications of its main results. This discussion would be facilitated by the development of a more detailed literature review. For example, it would be interesting to know whether Ronconi's results are consistent with the literature, especially in terms of the magnitude of workfare program impacts. Also, further analysis on the magnitude of the impact on incomes would help determine whether the program accomplished its expected objectives. In other words, were 60 pesos enough to alleviate short-term poverty? In that sense, a cost-benefit analysis would also be relevant. Moreover, the author should discuss the results in the framework of the implications of safety net programs and the implications for the subsequent *Jefes de Hogar* program. All this could allow Ronconi to support his conclusions and recommendations on the importance of enforcement and the ineffectiveness of *Trabajar* as an instrument for reducing short-term poverty.

Finally, Ronconi could have also emphasized the problems that arise from using ex post evaluation data and the importance of considering a proper impact evaluation design for this type of program, even during a crisis situation.

References

- Abadie, Alberto, and Guido Imbens. 2006. "On the Failure of the Bootstrap for Matching Estimators." Harvard University, John F. Kennedy School of Government.
- Ashenfelter, Orley. 1978. "Estimating the Effect of Training Programs on Earnings." *Review of Economics and Statistics* 60(1): 47–57.
- Ashenfelter, Orley, and David Card. 1985. "Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs." *Review of Economics and Statistics* 67(4): 648–60.
- Betcherman, Gordon, Karina Olivas, and Amit Dar. 2004. "Impacts of Active Labor Market Programs: New Evidence from Evaluations with Particular Attention to Developing and Transition Countries." Social Protection Discussion Paper 402. Washington: World Bank.
- Bouillon, Cesar P., and Luis Tejerina. 2007. "Do We Know What Works? A Systematic Review of Impact Evaluations of Social Programs in Latin America and the Caribbean." Washington: Inter-American Development Bank.
- Dar, Amit, and Zafiris Tzannatos. 1999. "Active Labor Market Programs: A Review of the Evidence from Evaluations." Social Protection Discussion Paper 9901. Washington: World Bank.
- Dehejia, Rajeev, and Sadek Wahba. 1999. "Causal Effects in Nonexperimental Studies: Re-evaluating the Evaluation of Training Programs." *Journal of the American Statistical Association* 94(448): 1053–62.
- Gaiha, Raghav, and Katsushi Imai. 2002. "Rural Public Works and Poverty Alleviation: The Case of the Employment Guarantee Scheme in Maharashtra." *International Review of Applied Economics* 16(2): 131–51.
- Haddad, Lawrence, and Michelle Adato. 2001. "How Efficiently Do Public Works Programs Transfer Benefits to the Poor? Evidence from South Africa." FCND Discussion Paper 108. Washington: International Food Policy Research Institute.
- Heckman, James, Hidehiko Ichimura, and Petra Todd. 1998. "Matching as an Econometric Evaluation Estimator." *Review of Economic Studies* 65(2): 261–94.
- IDB (Inter-American Development Bank). 2005. *The Politics of Policies*. Washington.
- Jalan, Jyotsna, and Martin Ravallion. 1999. "Income Gains to the Poor from Workfare: Estimates for Argentina's *Trabajar* Program." Policy Research Working Paper 2149. Washington: World Bank.
- . 2003. "Estimating the Benefit Incidence of an Antipoverty Program by Propensity-Score Matching." *Journal of Business and Economic Statistics* 21(1): 19–30.
- LaLonde, Robert. 1986. "Evaluating the Econometric Evaluations of Training Programs Using Experimental Data." *American Economic Review* 76(4): 602–20.
- Lanjouw, Peter, and Martin Ravallion. 1999. "Benefit Incidence, Public Spending Reforms, and the Timing of Program Capture." *World Bank Economic Review* 13(2): 257–73.

- Lee, Joohee. 2000. "Income Assistance and Employment Creation through Public Works in Korea." Paper presented at the International Conference on Economic Crisis and Labor Market Reform: The Case of Korea. Korean Confederation of Trade Unions (KCTU), Seoul, May 18–20.
- Martin, John P. 1998. "What Works among Active Labour Market Policies: Evidence from OECD Countries' Experiences." Labour Market and Social Policy Occasional Paper 35. Washington: Organization for Economic Cooperation and Development.
- Mato Díaz, Francisco J. 2004. "Recursos para las políticas del mercado de trabajo: su impacto según la experiencia europea." In "En búsqueda de efectividad, eficiencia y equidad: las políticas del mercado de trabajo y los instrumentos de su evaluación," edited by Jürgen Weller. Santiago: Economic Commission for Latin America and the Caribbean.
- Newman, John, Steen Jorgensen, and Menno Pradhan. 1991. "How Did Workers Benefit from Bolivia's Emergency Social Fund?" *World Bank Economic Review* 5(2): 367–93.
- Ravallion, Martin, and others. 2005. "What Can Ex-Participants Reveal about a Program's Impact?" *Journal of Human Resources* 40(1): 208–30.
- Ronconi, Lucas. 2002. "El Programa Trabajar." Working Paper 63. Buenos Aires: Centro de Estudios para el Desarrollo Institucional.
- Ronconi, Lucas, and Ignacio Franceschelli. 2007. "Clientelism, Public Workfare, and the Emergence of the Piqueteros." In *Political Institutions and Development: Failed Expectations and Renewed Hopes*, edited by Natalia Dinello and Vladimir Popov. Cheltenham, U.K.: Edward Elgar.
- Rosenbaum, Paul R., and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70(1): 41–55.
- Samaniego, Norma. 2004. "Las primeras experiencias de evaluación de las políticas del mercado de trabajo en América Latina: ¿Qué hemos aprendido?" In "En búsqueda de efectividad, eficiencia y equidad: las políticas del mercado de trabajo y los instrumentos de su evaluación," edited by Jürgen Weller. Santiago: Economic Commission for Latin America and the Caribbean.
- Subbarao, Kalanidhi. 1997. "Public Works as an Anti-Poverty Program: An Overview of Cross-Country Experience." *American Journal of Agricultural Economics* 79(2): 678–83.
- . 2003. "Systemic Shocks and Social Protection: Role and Effectiveness of Public Works Programs." Social Protection Discussion Paper 302. Washington: World Bank.