Policy Research Working Papers, vol. 57, núm. 5, 2004, pp. 128-142.

Assisting the Transition from Workfare to Work: A Randomized Experiment.

GALASSO, Emanuella, RAVALLION, Martin, SALVIA, Agustín y HECTOR AGUSTIN SALVIA.

ASSISTING THE TRANSITION FROM WORKFARE TO WORK: A RANDOMIZED EXPERIMENT

EMANUELA GALASSO, MARTIN RAVALLION, and AGUSTIN SALVIA*

Argentina's Proempleo Experiment, conducted in 1998–2000, was designed to assess whether a wage subsidy and specialized training could assist the transition from workfare to regular work. Randomly sampled workfare participants in a welfare-dependent urban area were given a voucher that entitled an employer to a sizable wage subsidy; a second sample also received the option of skill training; and a third sample formed the control group. Voucher recipients, the authors find, had a higher probability of employment than did the control group, even though the rate of actual take-up of vouchers by the hiring employers was very low. The employment gains were in the informal sector and largely confined to female workers, younger workers, and more educated workers. Skill training had no statistically significant impact overall, though once the analysis corrects for selective compliance, an impact for those with sufficient prior education is found.

Wage subsidies and training programs are often used by governments to help get able-bodied adults off the unemployment or workfare rolls into regular jobs. While there is some evidence that both interventions can help in the transition to regular work, results have varied greatly according to the setting and the method used to assess the impact.[1] It has

*Emanuela Galasso and Martin Ravallion are economist and research manager, respectively, with the Development Research Group of the World Bank, and Agustin Salvia is a consultant to the Ministry of Labor, Government of Argentina. The work reported in this paper is part of the ex-post evaluation of the World Bank's Social Protection III Project in Argentina. The authors thank the staff of the Trabajar project office in the Ministry of Labor, Government of Argentina, who helped in countless ways, and the Bank's Manager for the project, Polly Jones, for her continuing support of the evaluation effort. They also thank Julio Zelarayan, who managed the operations of the experiment and the data processing. For comments, the authors are grateful to Robert Boruch, Polly Jones, Michael Lokshin, Biju Rao, Dominique van de Walle, and participants at presentations given at the Ministry of Labor Argentina, the Latin American and Caribbean Economics Association, the World Bank, and Yale University. These are the views of the authors, and need not reflect those of the Government of Argentina or the World Bank.

A data appendix with additional results, copies of the computer programs used to generate the results, and the data set used here are available from Emanuela Galasso at the World Bank (MSN MC3–306), 1818 H Street NW, Washington, D.C. 20433; egalasso@worldbank.org.

Correspondence: Martin Ravallion, World Bank (MSN MC3–306), 1818 H Street NW, Washington, D.C. 20433; mravallion@worldbank.org.

[1]An overview of the arguments for and against wage subsidies can be found in Katz (1996), Bell et al. (1999), and Blundell (2001). Impact assessments can be found in Burtless (1985), Woodbury and Spiegelman (1987) and Dubin and Rivers (1993). The theory and evidence on training programs are reviewed by Heckman et al. (1999), and empirical studies include LaLonde (1986), Heckman et al. (1997), Dehejia and Wahba (1999), and Smith and Todd (2001).

Industrial and Labor Relations Review, Vol. 57, No. 5 (October 2004). © by Cornell University.
0019-7939/00/5705 $01.00

often proved difficult to get robust estimates of the impact using non-experimental methods.[2]

This paper reports on the "Proempleo Experiment" conducted by Argentina's Ministry of Labor in 1998–2000. The experiment was motivated by concerns about welfare dependency in "company towns" that had seen heavy retrenchments by their principal employer. The main form of welfare assistance provided to such towns in Argentina, as in most developing countries, is temporary work, at a relatively low wage, oriented to social infrastructure or community services. In the study towns, the heavy dependence on such workfare programs emerged in the wake of the privatization of the public oil refinery and subsequent sharp contraction in employment. Workfare participants in these towns may well need assistance in getting regular employment in the private sector. Wage subsidies and training programs seem obvious responses. But will they work, and at what cost?

The Proempleo Experiment was a randomized trial designed to assess the efficacy of providing a wage subsidy and specialized training in assisting the transition from workfare to regular work. The wage subsidy and training were provided to a random sample of workfare participants. At the time the sample frame was formed, all participants in the experiment were registered in workfare programs, mainly Argentina's "Trabajar" program. The design features of this program assured that it was well targeted. The wage rate in the program was deliberately set at a low level.[3] Taking account of the cost to participants of the work requirement through foregone earnings, 80% of participants came from the poorest quintile nationally on the basis of estimated pre-intervention household income per person (Jalan and Ravallion 2003). Thus the Proempleo Experiment was implicitly targeted to low-wage workers who tended to come from poor families.

In the experiment, one randomly chosen sample of Trabajar workers received a voucher that entitled a private-sector employer to a wage subsidy covering part of the total wages paid to the employee. A second sample was offered limited training as well. A third random sample formed the control group. After a baseline survey, follow-up surveys were conducted at six-month intervals for a total of 18 months. Using data from these surveys, we compare employment and incomes over time across these three samples. To our knowledge, the randomized experiment we study is the first ever to assess options for promoting the private-sector employment of workfare participants (though there have been randomized evaluations of other labor market interventions; we refer to examples later). Our study is also unusual in that we consider two interventions simultaneously.

Evaluation Methods for Active Labor Market Programs

There have been a number of attempts to assess how much active labor market programs help the transition to regular employment and raise the incomes of poor or otherwise disadvantaged groups. Most evaluations have been plagued by concerns over non-random assignment. Selective placement (through individual choice or purposive targeting) means that data on non-participants among those eligible do not reveal well the likely circumstances of

[2]A classic study by LaLonde (1986) found large biases in non-experimental methods when compared to a randomized evaluation of a U.S. training program. Using the same data set, Dehejia and Wahba (1999) found that propensity-score matching achieved a good approximation—much better than the non-experimental methods studied by LaLonde. However, Smith and Todd (2001), again using the same data set, questioned this finding, arguing that Dehejia and Wahba's results were sensitive to choices made in sample selection and model specification.

[3]Earnings data from the October 2000 Permanent Household Survey indicate that 95% of workers in full-time jobs (35 hours or more per week) earned more than the prevailing Trabajar wage rate.

participants in the absence of the program.

Various methods of dealing with this problem can be found in the literature. One possibility is to assess the counterfactual using a control group of non-participants matched on observable characteristics or some scalar aggregate of those characteristics, such as the propensity score (following Rosenbaum and Rubin 1983). An alternative approach is to use an instrumental-variables estimator, in which the instrumental variable (IV) identifies the exogenous variation in participation. Naturally these non-experimental methods require assumptions to make up for the missing data on outcomes in the absence of the intervention. Matching on the basis of propensity scores requires the conditional independence (sometimes called "strong ignorability") assumption, namely the assumption that pre-intervention outcomes are independent of participation given the observable covariates. Instrumental-variables methods require an alternative conditional independence assumption—the exclusion restriction that the IV is uncorrelated with outcomes conditional on participation and the values taken by the control variables.

In a few cases, evaluations of ALMP's have used randomized assignment. In the case of training programs, two examples are the U.S. Job Training Partnership Act (see, for example, Heckman et al. 1997) and the U.S. National Supported Work Demonstration (studied by Lalonde [1986] and Dehejia and Wahba [1999], among others). Randomized evaluations of wage subsidy programs have been done by Burtless (1985), Woodbury and Spiegelman (1987), and Dubin and Rivers (1993)—all for targeted wage subsidy schemes in the United States. (Woodbury and Spiegelman [1987] and Dubin and Rivers [1993] studied the same experiment by the Illinois Department of Employment Security in the mid-1980s.)

Such randomized evaluations attempt to approximate the theoretical ideal in which the distributions of all (observed or unobserved) covariates are balanced between the treatment and control groups. If everyone who is given access to the training automatically takes it up and access is assigned randomly, then an unbiased estimate of mean impact for those treated can be obtained by taking the mean difference in the outcome measure (employment, say) between the treatment and control groups. This is equivalent to the regression coefficient of the outcome measure on a dummy variable for which group one belongs to (treatment/control). This provides an unbiased estimate of impact, given that the dummy variable is exogenous under randomization with full compliance.

However, it is often the case in randomized evaluations that some of those selected for the program do not want to participate. When that is the case, actual treatment ceases to be exogenous, even when assignment to treatment is exogenous. In one common method of dealing with this problem, the first step is to calculate what is referred to as the "intention to treat" (ITT) given by the difference in mean outcomes between those assigned to the program (whether they take it up or not) and those not assigned. Next, the ITT is divided by the "compliance rate," given by the proportion of those assigned to the program who in fact take it up.[4] It is readily demonstrated that this method is equivalent to estimating impact using the Two-Stage Least Squares (2SLS) regression coefficient of the outcome measure on a treatment dummy variable, with a dummy variable for assignment as the IV.[5] As with all IV estimators, this requires an exclusion restriction, namely that being randomly assigned to the program only affects potential outcomes via actual participation.[6]

[4]Bloom (1984) appears to have originated this method of calculation. An example in the context of assessing the Illinois experimental wage subsidy program can be found in Dubin and Rivers (1993).

[5]The working paper version includes a formal demonstration of this equivalence and further discussion (Galasso et al. 2002).

[6]Angrist, Imbens, and Rubin (1996) provided a precise statement of the exclusion restriction and other conditions under which this 2SLS estimator gives a consistent estimate of the impact of treatment on the treated.

The Proempleo Experiment

Workfare programs impose work requirements on welfare recipients, typically at benefit levels at or below prevailing market wages for relatively unskilled labor. Two incentive arguments are made in favor of such programs. The first is that by setting a low benefit level, a workfare program will be self-targeted to those most in need; few of the non-poor, in particular, will want to participate. The second argument is that a low benefit level assures that participants do not become dependent on the program, since they will likely turn to regular work when it becomes available.[7]

Workfare wages are typically fixed across participants and across geographic areas. This is a defensible design feature, given fairness considerations and constraints on the information available to policy-makers.[8] However, with a fixed benefit level and heterogeneity in local labor market conditions, there can be a subset of workfare participants who become dependent on the scheme, even though this is not generally so.

The Proempleo Experiment was conducted in two adjacent towns, Cutral Co and Plaza Huincul, which together make up the bulk of the department of Confluencia in the province of Neuquen. Though officially distinct, these towns form a relatively homogeneous urban conglomerate with a population of about 50,000. Both had been affected by the severe contraction in employment following the downsizing in 1993 and then privatization of the largest employer, the state-owned oil company.[9] Even five years later, the Trabajar participation rate appeared to be unusually high in Confluencia, relative to expected demand for the program. In late 1998, the average number of Trabajar participants per month represented 28% of the total number of people living in households that were poor (with income below Argentina's official poverty line) and also included an unemployed worker; the corresponding national figure was 5%.[10] However, the incidence of poverty was not unusually high; the proportion of the population living in households that had income per person below the poverty line and at least one unemployed person was 3.5% in Confluencia versus 4.2% nationally. In the light of these figures, the Ministry of Labor wanted to explore policy options for assisting workfare participants in Confluencia to find regular private-sector jobs.

The target population for the experiment was the set of beneficiaries of temporary employment programs, managed by the Ministry of Labor. The main program was the aforementioned Trabajar program. (There was a much smaller provincial workfare program, the beneficiaries of which were not included in the study.) The baseline survey aimed to cover everyone on the official list of participants in temporary programs drawn in October 1998, though some had to be dropped because they could not be located. In all, 953 workfare participants and their households answered the baseline questionnaire in full.

---

[7]Besley and Coate (1992) provided a theoretical model incorporating both arguments.

[8]Indeed, good targeting with modest information requirements is one of the well-recognized advantages of workfare programs over alternatives, notably (though not only) in developing countries where there are severe constraints on the information available for targeting. For overviews of these arguments and the literature, see Besley and Kanbur (1993) and Lipton and Ravallion (1995, Sec. 6).

[9]The oil company had facilities for scouting, extracting, and refining oil in Plaza Huincul and Cutral Co, which made these towns the logistic and demographic center of a large area with oil-based activity. For historical and sociological studies on the formation and development of these towns and on the social and employment effects of downsizing the oil company, see Favaro (1999) and Salvia (2001).

[10]For the purpose of this calculation, participants in temporary employment programs were counted as unemployed. The calculation of the number of people living in poor households with an unemployed member was done using the Permanent Household Survey for October 1998.

Three random samples were then drawn by lottery. One sample got just the voucher, one got the voucher and training, and one was reserved as the control group. All three were given the same questionnaire as for the Permanent Household Survey, which is administered twice a year by the statistical office of the Government of Argentina. Some questions were added that were specific to temporary employment programs. All interviews were conducted at participants' homes.

The experiment was not announced publicly by the Ministry of Labor. Nor were any of the beneficiaries (in either the treatment group or the control group) told that they were part of an experiment; they believed that they were selected for a special study, but had no knowledge of how they were assigned to it. There is no official channel by which they could have found out that others had been chosen to receive benefits that they did not have access to. The local Ministry of Labor staff were not told how beneficiaries had been selected. So there does not appear to be any way in which program participants or potential employers could have known that the assignment was random. Efforts were made to prevent members of one treatment group from meeting those in another, notably by scheduling their visits to the Ministry's local offices on different days. About 40 of the control group members did hear about the voucher and the training program (presumably from co-workers or neighbors) and asked if they could join it. Their request was refused.

The Proempleo voucher entitled a hiring employer to a wage subsidy of $150 per month for workers aged above 45 years and $100 for younger workers. This subsidy was paid directly to the beneficiary as a part of his or her salary, and the employer had to deduct the amount of the subsidy from the gross wages paid to the worker. The minimum wage rate in Argentina at the time was $200 per month.

The subsidy was received for 18 months, conditional on the employer registering the worker formally, and so incurring the government's social security charges for that worker. The latter represent 30% of the gross wage.[11] The subsidy was only paid if the social security charges had been paid. The level of the subsidy was set to avoid the displacement effect that would have occurred if hiring firms had simply fired a current registered worker in order to hire and register a Proempleo worker. Legally binding severance pay obligations are sufficiently high in Argentina to have discouraged employers from taking such action.[12] Those assigned the Proempleo voucher also received instruction lasting 2–3 hours to explain the program and how to use the voucher. The voucher had the participant's name on it, and was non-transferable.

The training had two components. The first was a three-day "labor market orientation" workshop that included presentations on labor demand in the area, how to look for work, and how to become self-employed. This component was mandatory. Once this workshop was completed, training coupons were issued for the second component, which provided training in a specific skill and required 200–300 hours of attendance. In this second part, the participants were given working materials and received a stipend set at a rate 10% lower than that for the Trabajar program.[13] According to their personal interest and the available quota, participants proceeded to select from a list of 12 subjects, chosen in the light of local labor demand and the profiles of participants. Two of the courses were on the management of small-scale enterprises, two on industrial welding, two on home-build-

[11]For example, an employer hiring someone over age 45 with a voucher at the minimum wage of $200 and registering this fact with the Ministry of Labor would incur a social security charge of $60 but receive a subsidy of $150, implying a net wage of $110; for a worker under 45, the net wage would be $160.

[12]The normal severance pay (at the time of our baseline survey) amounted to roughly the monthly wage (for the last month of employment) times the number of years worked by the employee.

[13]The courses were part of a larger training program of the Ministry of Labor, called "Project Joven." This was financed by the Inter-American Development Bank.

Table 1. Sample Breakdown between Treatments and Control and Assigned versus Actual Participation.

		Actual			
		Voucher + Training	*Voucher*	*Control Group*	*Total*
Assignment	Voucher + Training	210	90	0	300
	Voucher	3	264	0	267
	Control Group	0	0	281	281
	Total	213	354	281	848

ing, one on professional cooking, one on raising pigs, one on greenhouse cultivation, and one on skills needed to become an electrician.

The data collection began in December 1998 with a baseline survey. The invitation to join the treatment groups for those selected was made in January–February 1999, either by a house visit or at the place of work. The sample was re-interviewed for three waves at intervals of 5–6 months (June 1999, December 1999, and May 2000).

The last six months or so of the study period saw a sizable retrenchment of national ALMP's (including the Trabajar program) to help keep within macroeconomic targets. Aggregate spending on the Trabajar program in the first five months of 2000 was only 29% of its level in the last five months of 1999. Existing projects were completed, but the number of new projects approved (the best indicator available of new work opportunities in the program) shrank sharply in the early part of 2000, to bring down the center's outlays.[14] This contraction also occurred in the study towns. The first five months of 2000 saw only 56% as many Trabajar participants in Cutral Co and Plaza Huincul as had the last five months of 1999. There was also a reduction in the benefit level under Trabajar, from $200 to $160 per month. These observations will be important when we interpret the results of the experiment.

The rest of this paper focuses on the comparison of the results for the last wave (May 2000) with the baseline survey (December 1998). We repeated the analysis using the intermediate rounds, and compare the results.

Baseline Characteristics and Attrition

Table 1 compares the randomized assignment with observed treatment status.[15] It can be seen that about 30% of those offered "voucher and training" did not take up the training component. These were reassigned to the voucher-only component, though our results for that group were robust with respect to excluding these reassigned workers. The large percentage of people who did not take up the training suggests that there may be a serious problem of selective compliance with the randomized assignment of training, as discussed above in general terms (see "Evaluation Methods for Active Labor Market Programs"). We return to this problem in the next section.

The top panel of Table 2 provides summary statistics on the employment status of the various treatment groups and the con-

[14] The average number of new projects approved was 439 per month in the period February–June 2000, as compared to 872 per month in February–November 1999 (the peak summer months of December and January have unusually low project approval rates, so we dropped them from this calculation).

[15] We excluded five workers who were assigned to the comparison group but mistakenly ended up receiving the treatment. We also excluded 32 members of Fogoneros, a political activist group whose behavioral responses might be expected to be quite unusual, and those who dropped out between the baseline and wave four.

Table 2. Descriptive Statistics from the Baseline Survey.

	Assignment To:				Difference: Assignment Versus Control (p-value)[a]		
Independent Variable	*Control*	*Either Treatment*	*Voucher Only*	*Voucher + Training*	*Either Treatment*	*Voucher Only*	*Voucher + Training*
Employment Status							
Unemployed/Inactive	0.075	0.102	0.107	0.094	(1) 0.19	0.16	0.45
	(0.26)	(0.30)	(0.31)	(0.29)	(2) 0.10	0.08	0.22
Self-Employed	0.011	0.007	0.008	0.005	(1) 0.58	0.46	0.78
	(0.10)	(0.04)	(0.09)	(0.07)	(2) 0.29	0.23	0.39
Employed in the Private Sector	0.011	0.019	0.023	0.014	(1) 0.35	0.25	0.73
	(0.10)	(0.14)	(0.15)	(0.12)	(2) 0.17	0.13	0.37
Temporary Employment Program	0.900	0.871	0.862	0.887	(1) 0.22	0.14	0.64
	(0.30)	(0.33)	(0.34)	(0.31)	(2) 0.11	0.07	0.32
Other Characteristics							
Age	32.33	32.20	32.24	32.14	(1) 0.88	0.92	0.85
	(12.12)	(11.63)	(11.92)	(11.15)			
Sex (Proportion Female)	0.470	0.437	0.437	0.437	(1) 0.37	0.42	0.46
	(0.50)	(0.49)	(0.49)	(0.49)			
Primary Education	0.492	0.526	0.527	0.524	(1) 0.36	0.39	0.51
	(0.30)	(0.21)	(0.26)	(0.35)			
Household Size	4.29	4.36	4.33	4.41	(1) 0.68	0.84	0.55
	(2.05)	(2.25)	(2.24)	(2.26)			
Individual Income ($/Month)	188.4	182.6	181.4	184.5	(1) 0.27	0.24	0.52
	(67.1)	(76.9)	(81.6)	(68.7)			
Household Income ($/Month)	406.6	424.8	411.7	446.5	(1) 0.42	0.55	0.15
	(267.2)	(332.8)	(315.5)	(359.5)			
No. Observations	281	567	354	213			

[a]$H_0 = 0$ vs. (1) H_1 NotEqualToSign 0 (2) H_1 >/<0 (one sided-test); standard deviations in parentheses.

trol group in the baseline survey. In each case, the vast bulk (85–90%) of sampled workers were in temporary employment programs. While all were in such programs at the time the sample frame was constructed, there was an unavoidable two-month delay in completing the baseline survey. Some Trabajar projects terminated during that period. There is weak evidence of higher participation in the temporary employment programs among the control group. This might reflect some contamination of treatment effects in the baseline survey. We will check the robustness of our results with respect to using a double difference estimator.

Table 2 also compares the various subsamples in terms of other worker characteristics, such as age, sex, and household size. There are no statistically significant differences, suggesting that the randomization has adequately balanced observed characteristics. Similarly, observable worker and household characteristics were (individually and jointly) statistically insignificant in regressions for whether an observation was in a treatment group or the control group.

As one would expect, sample attrition occurred over the study period. Of those interviewed in the baseline, 77.5% stayed through the fourth round. We followed Fitzgerald, Gottschalk, and Moffitt (1998) in testing whether the selection out of the

Table 3. Aggregate Impact Estimates.

Outcome Variable	*Means* Treated $Y(D=1)$	*Means* Control $Y(D=0)$	*Difference* $Y(D=1)-Y(D=0)$
Either Treatment[a]			
Any Employment	0.478	0.452	0.026
Wage Employment	0.143	0.085	0.057**
Self-Employment	0.035	0.021	0.014
Private Employment (Wage/Self-Emp.):			
Permanent Employment	0.075	0.057	0.018
Temporary Employment	0.106	0.050	0.056**
Temporary Employment Program	0.296	0.345	–0.049
Labor Income	120.59	119.27	1.32
Voucher Only			
Any Employment	0.469	0.452	0.017
Wage Employment	0.147	0.085	0.061**
Self-Employment	0.037	0.021	0.015
Private Employment (Wage/Self-Emp.):			
Permanent Employment	0.076	0.057	0.020
Temporary Employment	0.110	0.050	0.060**
Temporary Employment Program	0.282	0.345	–0.063*
Labor Income	123.18	119.27	3.91

[a]$D = 1$ if received either voucher or voucher plus training; $D = 0$ if Control. All estimates refer to 18 months effects.

*Statistically significant at the .10 level; **at the .05 level.

sample is based on observable characteristics in the baseline (b), including the outcome indicator (Y_b). The attrition indicator (whether an observation stays in the sample) is regressed on Y_b and other baseline characteristics of the worker and household (gender, age, and schooling). The test for attrition bias is equivalent to testing whether Y_b is statistically significant. In the present case, the fact that the beneficiary group consists mainly of workfare participants attenuates the observed variance in outcome indicators in the baseline survey. However, we have seen that there is still some variation in incomes and private sector employment (Table 2); indeed, the coefficient of variation of incomes is 35–45%. If there is a serious problem of endogenous attrition, this test should be able to pick it up.

We found no sign of attrition bias. Neither individual income nor employment in the private sector was a statistically significant predictor of attrition (t-ratios of 0.5 or lower). The only statistically significant variables were age (attrition was more likely among the young) and education (those with better than primary schooling were more likely to stay in the survey). We also tested whether attrition depended on the treatment, which would change the experimental nature of the sample, but found no evidence that it did.

Participant Effects and Employer Take-up

The impact estimates for those receiving any treatment (voucher, or voucher plus training) can be found in the top panel of Table 3, while the lower panel gives the results for those who only received the voucher. We find a statistically significant effect of either treatment (voucher or voucher plus training) on the probability of becoming employed in the private sector

Table 4. Impact Estimates across Different Demographic and Education Groups.

	Difference in Means: Y(D=1)–Y(D=0)					
Outcome Variable:	*Men*	*Women*	*Age(30*	*Age>30*	*Primary Education*	*Secondary Education*
Either Treatment[a]						
Any Employment	0.044	0.013	–0.007	0.062	0.079	–0.025
Wage Employment	0.034	0.076**	0.092**	0.020	0.015	0.102**
Self-Employment	0.034	–0.001	0.0003	0.029	0.018	0.009
Temporary Employment Program	–0.028	–0.065	–0.103**	–0.010	0.042	–0.136**
Labor Income	2.009	2.345	14.639	–14.00	2.95	–1.17
Voucher Only						
Any Employment	0.028	0.009	–0.004	0.043	0.050	–0.049
Wage Employment	0.042	0.078**	0.088**	0.029	0.024	0.093**
Self-Employment	0.040	–0.003	0.004	0.029	0.021	0.008
Temporary Employment Program	–0.060	–0.065	–0.102**	–0.014	0.000	–0.107**
Labor Income	3.656	11.18	17.829	–12.464	1.02	6.63

[a]$D = 1$ if received either voucher or voucher plus training; $D = 0$ if Control. All estimates refer to 18 months effects.

*Statistically significant at the .10 level; **at the .05 level.

on the order of six percentage points. Wage employment clearly displaced temporary employment programs (Table 3). For wage employment or self-employment, the survey data allow us to separate what is identified by the respondents as "permanent employment" from "temporary employment," most of which is likely to be in the informal sector. We find that most of the employment gains took the form of temporary employment (Table 3). There is no statistically significant effect on other outcomes: being self-employed, being employed in a temporary employment program, or wage earnings. We will discuss these results in the next section.

When we repeated the above calculations for the intermediate survey rounds, we found a statistically significant impact on employment after six months for voucher recipients only, but not after 12 months, and all other results were similar.[16] The wage employment effect after six months was 0.033 and was significant at the 5% level. Thus, there is an indication of a U shape in the employment impact, in which an initial positive impact soon subsided, but returned after 18 months. However, there are no signs of a U-shaped employment effect for the voucher recipients with the training component, for whom the impact materialized only after 18 months. (We discuss this finding in the next section.)

In Table 4 we provide a breakdown of the impact estimates by demographic group, splitting the sample by gender, by whether the worker was under 30 (the median age), and by education level (primary versus higher education). The statistically significant effects are confined to the wage employment of women, those under 30, and those with secondary education.

As we saw in Table 1, there is a potential problem of endogenous compliance with the training component. There may be some latent correlate of the outcome measure that influenced the choice to take up the program among those who were given access. The bias could go in either direction. Workers with intrinsically lower employment prospects (at given program placement) might have been more likely to take up the training to try to compensate. If so, then correcting for endogenous com-

[16]The results by round are available upon request.

Table 5. Voucher and Training Effects Allowing for Endogenous Take-Up of Training.

Employment Category	*Treated Y(D=1)*	*Control Y(D=0)*	*Difference in Means Y(D=1)–Y(D=0)*	*Treatment on the Treated Difference 2SLS Y on D (Z as IV)*	*Assigned Y(Z=1)*	*Not Assigned Y(Z=0)*	*Intention to Treat Difference Y(Z=1)–Y(Z=0)*
Any Employment	0.493	0.458	0.035	0.057	0.490	0.451	0.039
Wage Employment	0.136	0.102	0.033	0.075**	0.140	0.088	0.052**
Self-Employment	0.033	0.032	0.0005	0.032	0.043	0.021	0.022*
Private Employment (Wage/Self-Emp.):							
Permanent Employment	0.087	0.061	0.026	0.029	0.077	0.056	0.020
Temporary Employment	0.116	0.075	0.041	0.083**	0.110	0.528	0.057**
Temporary Employment Program	0.319	0.323	–0.004	–0.055	0.303	0.345	–0.038
Labor Income	121.36	116.28	–5.08	1.247	119.927	119.067	0.860

Notes: $D = 1$ if received voucher + training, $D = 0$ if Control; $Z = 1$ if assigned to voucher + training, $Z = 0$ if Control. All estimates refer to 18 months effects.
*Statistically significant at the .10 level; **at the .05 level.

pliance would give higher estimates of the impact of training. Alternatively, those who chose to take the training may have been those with high latent ability, which would also lead to a higher probability of employment. If so, then correcting for selection bias will give lower estimates of the employment gain from training.

Table 5 gives the estimates for those who received both the voucher and training. We present both the ordinary estimates for the actual treatment (analogous to Table 3) and the 2SLS estimates allowing for endogenous take-up of the training component (as discussed above, under "Evaluation Methods for Active Labor Market Programs"). The 2SLS estimates suggest a greater effect than the single difference in means for the actual treatment—a gain to the proportion in wage employment of 7.5 percentage points. However, the extra impact of the training (an increment to the employment rate of 7.5 percentage points versus 6.1 for the voucher only) is not statistically significant at the 5% level (z-score = 0.70).[17] When we repeated this analysis for the survey rounds at 6 and 12 months, we found no signs of a statistically significant impact on employment.

Table 6 gives a breakdown by the same groups as in Table 4. Again, statistically significant effects on wage employment are confined to women and those under 30, though significant effects of the voucher and training on the self-employment of men and those over 30 do emerge. Furthermore, only those with secondary education are successful in moving out of welfare to find employment in the private sector. The simple difference in outcomes (Table 6, upper panel) underestimates the impact for workers with secondary schooling. We find a statistically significant impact of the training for those with secondary education. The increment in the employment rate for the more educated workers goes

[17]The estimated standard error (s) of the difference between the two incremental employment rates (treatment minus comparison) is 0.020. This was calculated using the formula $s^2 = \bar{p}(1 - \bar{p})(n_1^{-1} + n_2^{-1})$, where $\bar{p} = (n_1p_1 + n_2p_2)/(n_1 + n_2)$ and n_1, p_1 are the sample sizes and estimated incremental employment probabilities for the voucher-only sample, while n_2, p_2 are the corresponding numbers for the voucher + training sample (see, for example, Hamburg 1977). This calculation treats the comparison group employment probability as non-stochastic; factoring in the sampling variance in the latter estimate will make the difference in the incremental employment probabilities between the two treatment samples even less statistically significant.

Table 6. Voucher and Training Effects across Demographic and Education Groups Allowing for Endogenous Take-Up of Training.

Employment Category	*Men*	*Women*	*Age(30*	*Age>30*	*Primary Education*	*Secondary Education*
Outcome of Interest: Difference in Means for Actual Treatment, Y(D=1)–Y(D=0)						
Any Employment	0.061	0.018	–0.009	0.071	0.130*	–0.072
Wage Employment	0.024	0.039	0.076*	–0.004	0.002	0.079*
Self-Employment	0.002	0.006	–0.008	0.007	–0.005	–0.002
Temporary Employment Program	0.039	0.036	–0.076	0.058	0.124**	–0.149**
Labor Income	13.69	–17.32	3.81	–16.67	7.67	–19.8
Outcome of Interest: Intention to Treat Difference, Y(Z=1)–Y(Z=0)						
Any Employment	0.064	0.021	–0.009	0.097	0.082	0.011
Wage Employment	0.0007	0.096**	0.082**	0.017	0.015	0.130**
Self-Employment	0.052**	0.001	–0.001	0.050**	0.017	0.022
Temporary Employment Program	0.012	–0.080	–0.090*	0.023	0.074	–0.163**
Labor Income	4.90	–1.24	12.09	–12.43	0.09	0.76
Treatment on the Treated Difference, 2SLS Y on D (Z as IV)						
Any Employment	0.098	0.029	0.016	0.121	0.119	–0.016
Wage Employment	–0.001	0.133**	0.137**	0.021	–0.022	0.194**
Self-Employment	0.080*	–0.001	–0.002	0.062**	0.025	0.032
Temporary Employment Program	0.019	–0.111	–0.151*	0.028	0.106	–0.243**
Labor Income	7.55	–1.72	20.22	–15.59	0.13	1.14

Note: All estimates refer to 18 months effects.
*Statistically significant at the .10 level; **at the .05 level.

from 7.9% in the actual treatment to 19.4% when we take account of the endogeneity in compliance. This suggests that the relatively better-educated Trabajar workers tended to take up training when their employment prospects were intrinsically low (creating a negative correlation between compliance with the training assignment and latent determinants of employment prospects). Less well-educated workers did not behave this way. The results suggest a complementarity between schooling and the perceived gains from extra skill training. This interpretation is also confirmed by the fact that although there are no income gains for the treated workers, those who took up the training gained significantly higher hourly wages relative to the control group (conditional on employment).

Recall that due to a slight delay in collecting the baseline survey, workers in the control sample were more likely than those in the two other samples to be employed in temporary employment programs in the baseline survey (Table 2). Thus our results so far might overstate the employment gains from the program. To address this concern, we use a double difference ("difference-in-difference") approach in which we net out the baseline differences when calculating mean effects. Table 7 gives these estimates. The effect on private employment holds in the double difference estimates. Again, too, there are no statistically significant effects on other outcomes. The close correspondence between the double-difference and single-difference results is consistent with success in randomizing the assignment.

Turning to administrative records of the Ministry of Labor, we find that take-up of the wage subsidy by firms hiring a worker with a voucher was very low. Indeed, only three of the workers in the treatment group who were hired by private firms were in fact registered by their new employer—and a single firm was responsible for registering all three of these workers. We offer an explanation for this finding in the next section.

Table 7. Double Difference Estimates of Impact.

Outcome of Interest:	*Double Difference: ΔY(D=1)–ΔY(D=0)*			
	Treated	*Control*	*Difference*	
Either Treatment				
Any Employment	–0.420	–0.473	0.054*	
Wage Employment	0.123	0.075	0.049*	
Self-Employment	0.028	0.011	0.018	
Temporary Employment Program	–0.575	–0.555	–0.020	
Labor Income	–58.38	–66.99	8.61	
Voucher Only				
Any Employment	–0.424	–0.473	–0.050	
Wage Employment	0.124	0.075	0.050	
Self-Employment	0.028	0.011	0.018	
Temporary Employment Program	–0.579	–0.555	–0.024	
Labor Income	–53.54	–66.99	13.45	
Voucher and Training				
Any Employment	–0.413	–0.464	0.050	
Wage Employment	0.122	0.084	0.039*	
Self-Employment	0.028	0.022	0.007	
Temporary Employment Program	–0.568	–0.566	0.002	
Labor Income	–66.42	–67.08	0.66	
Outcome of Interest:	*Double Difference: ΔY(Z=1)–ΔY(Z=0)*			*2SLS ΔY on D (Z as IV)*
Any Employment	–0.417	–0.475	0.059	0.085
Wage Employment	0.120	0.074	0.046*	0.067*
Self-Employment	0.037	0.011	0.026*	0.038
Temporary Employment Program	–0.577	–0.556	–0.020	–0.029
Labor Income	–67.76	–65.96	1.8	2.61

Note: Wave 4 (18 months) relative to the baseline.
*Statistically significant at the .10 level; **at the .05 level.

Interpretations

Some of our results require care in interpretation, particularly in drawing conclusions about the likely effects of scaling up the program. Although employment improved among voucher recipients, there are no signs of an impact on labor incomes, at least over the 18 months of the experiment. The program helped smooth the transition of workfare workers into temporary employment in the informal labor market at wages similar to those in the Trabajar program. The fact that current earnings were no higher than for the control group, and at a level below the minimum wage (and market wages), is also suggestive that voucher holders were willing and able to undercut the going wage so as to get a job. Although the program did not help the participants secure permanent jobs in the short run, there might be gains from the program that take time to materialize. Qualitative questions added to the survey suggest that the participants felt that the program improved their technical skills, though not necessarily that it improved their employment opportunities in a situation of depressed labor demand. It appears that voucher recipients took up private sector jobs in the expectation of higher or more stable incomes in the future.

Take-up of the wage subsidy by hiring firms was low. This finding echoes results for wage subsidy schemes in the United States. For example, Woodbury and Spiegelman (1987) reported that only 12% of the employers who were eligible for the wage subsidy took it up (see also O'Neill 1982; Burtless 1985). In the present context, there is a possible explanation for low employer take-up of the subsidy. Registering a worker so as to receive the subsidy was not costless, since it entailed administrative costs and also meant incurring the government's social charges. While the subsidy was greater than the social charges for as long as the subsidy lasted (18 months), after that period the employer would have faced severance payments to fire the worker. Many potential employers were also outside the formal sector, and did not register any workers. (This applies to about half of Argentina's work force.) For such firms, registering one worker to receive the subsidy may well have seemed risky, since other workers might then have demanded to be registered, with possible legal action against the firm by workers and the government.

So the impact of the voucher was clearly not through access to the wage subsidy by firms. That is consistent with the fact that signs of a statistically significant impact from this experiment emerged initially during a period of economic expansion (after six months) for voucher recipients but then faded with the subsequent economic downturn, only to re-emerge 18 months after the baseline survey. This re-emergence is suggestive of a supply-side channel of impact, given that there was a sharp contraction in new demand for work under the national workfare program in the few months prior to the last survey. The apparent willingness of voucher recipients to undercut the minimum wage and the incentives to look actively for a job in the private sector are expected to be higher with lower expected prospects of public income support. The voucher may well have encouraged workers to make more effort to find work. By this interpretation, it had an "empowerment" effect, in making these workers more confident in approaching employers. Here it should be recalled that the eligible participants came mainly from the poor (through the self-targeting of workfare participation). Our informal interviews with Trabajar participants indicated that many of them had little or no experience in approaching employers for regular private sector jobs, and relied heavily on more casual labor markets and informal networks. For such workers, the voucher may well have served as a useful "letter of introduction" to prospective employers.

But why did employers hire workers with the voucher? One clue can be found in the fact that the Trabajar workers in these company towns had developed a reputation locally as "trouble-makers" due to their involvement in various protests about economic conditions in the towns and the perceived inaction by the government. This reputation may well have made them less employable as a group. In principle, holding the voucher could have changed that perception either way. A worker's possession of the voucher might have been perceived by employers as a further negative signal, on the assumption that the government was targeting the trouble-makers to help get them into regular jobs.[18] However, we find the opposite. Possibly receipt of the voucher made employers feel that these workers were more trustworthy than typical Trabajar workers in these towns, since the government was willing to help get them jobs. (Recall that efforts were made to avoid any local knowledge about how the vouchers had in fact been assigned.)

Conclusions

We find that 18 months after the baseline survey, the proportion of the sample of

[18]This is also what one would expect if the voucher had a stigmatizing effect. In a similar wage-subsidy experiment for the United States, Burtless (1985) found a *negative* impact of a wage subsidy experiment on employment, and he interpreted this as a stigma effect of the wage subsidy.

workfare participants getting a private sector job was 14% for randomly selected voucher recipients versus 9% for the control group. This difference is statistically significant. On disaggregating the impact, we find that the gain in wage employment was largely confined to women, younger workers, and more educated workers. We find no statistically significant impact of the wage subsidy on current incomes; presumably the extra workers taking up private sector employment expected higher future incomes, suggesting the existence of longer-term effects beyond the study period.

There was a slightly higher impact on employment for those who took up both the offer of training and the voucher than for those who took only the voucher. An impact of training only emerges once we correct for the endogeneity of take-up among those randomly assigned to the treatment. For the sample as a whole, we cannot reject the null hypothesis that the training had no impact even after we correct for endogenous compliance. However, with the correction for compliance there is a statistically significant extra impact on employment of the training among those with better initial education.

Clearly, Proempleo did not succeed in achieving a major transition to private sector employment among workfare recipients in the study area. Arguably, without greater labor demand in these company towns, such an outcome would have been difficult. Nonetheless, the program did help some of those on workfare make the transition into (primarily) informal-sector wage employment. Given employers' low take-up of the subsidy, the gain in private sector employment attributed to the voucher was achieved at very little cost to the government. Since the workfare wage was roughly the same as the subsidy paid to firms, the government saved about 5% of its expenditure on workfare wages for those receiving the voucher in return for an outlay on wage subsidies that represented only 10% of that saving.

In considering the policy implications, one must acknowledge the possibility that scaling up could increase the subsidy take-up rate among firms, or reduce the empowerment effect of the voucher, or both. The signal value to employers of the voucher could well be different in a national version of the program. And scaling up might occur at the expense of the non-participants (by displacing their jobs); the general equilibrium costs of a larger program would need to be considered.

From a methodological perspective, the Proempleo Experiment illustrates a number of concerns about both the internal and external validity of randomized trials. With regard to internal validity, we find that selective compliance with the randomized assignment can bias the results, though this problem can be readily corrected econometrically. More worrying, probably, are the external validity concerns we have highlighted. It can be argued that even the limited success we find for this pilot program may well vanish on scaling up to a national program.

REFERENCES

Angrist, Joshua, Guido Imbens, and Donald Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association,* Vol. 91 (June), pp. 444–55.

Bell, Brian, Richard Blundell, and John Van Reenen. 1999. "Getting the Unemployed Back to Work: The Role of Targeted Wage Subsidies." *International Tax and Public Finance,* Vol. 6, No. 3, pp. 339–60.

Besley, Timothy, and Stephen Coate. 1992. "Workfare vs. Welfare: Incentive Arguments for Work Requirements in Poverty Alleviation Programs." *American Economic Review,* Vol. 82, No. 1 (March), pp. 249–61.

Besley, Timothy, and Ravi Kanbur. 1993. "Principles of Targeting." In Michael Lipton and Jacques van der Gaag, eds., *Including the Poor.* Washington, D.C.: World Bank.

Bloom, Howard S. 1984. "Accounting for No-Shows in Experimental Evaluation Designs." *Evaluation Review,* Vol. 8, No. 2 (April), pp. 225–46.

Blundell, Richard. 2001. "Welfare to Work: Which Policies Work and Why?" Keynes Lectures in Economics, University College of London and Institute of Fiscal Studies.

Burtless, Gary. 1985. "Are Targeted Wage Subsidies Harmful? Evidence from a Wage Voucher Experiment." **Industrial and Labor Relations Review**, Vol. 39, No. 1 (October), pp. 105–15.

Dehejia, Rajeev H., and Sadek Wahba. 1999. "Causal Effects in Non-Experimental Studies: Re-Evaluating the Evaluation of Training Programs." *Journal of the American Statistical Association,* Vol. 94, No. 448 (December), pp. 1053–62.

Dubin, Jeffrey A., and Douglas Rivers. 1993. "Experimental Estimates of the Impact of Wage Subsidies." *Journal of Econometrics,* Vol. 56, No. 1/2, pp. 219–42.

Favaro, Orieta, ed. 1999. *Neuquén: la construcción de un orden estatal.* Centro de Estudios Históricos Estado, Política y Cultura de la Universidad Nacional del Comahue, Neuquén.

Fitzgerald, John, Peter Gottschalk, and Robert Moffitt. 1998. "An Analysis of Sample Attrition in Panel Data: The Michigan Study of Income Dynamics." *Journal of Human Resources,* Vol. 33, No. 2 (Spring), pp. 300–344.

Galasso, Emanuela, Martin Ravallion, and Agustin Salvia. 2002. "Assisting the Transition from Workfare to Work: A Randomized Experiment." Washington, D.C.: Policy Research Working Paper 2738, World Bank.

Hamburg, Morris. 1977. *Statistical Analysis for Decision Making.* New York: Harcourt Brace Jovanovich.

Heckman, James. 1996. "Randomization as an Instrumental Variable." *Review of Economics and Statistics,* Vol. 78, No. 2 (May), pp. 336–41.

Heckman, J., H. Ichimura, and P. Todd. 1997. "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program." *Review of Economic Studies,* Vol. 64, No. 4, pp. 605–54.

Heckman, James, Robert LaLonde, and Jeffrey Smith. 1999. "The Economics and Econometrics of Active Labor Market Policies." In Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics,* Vol. 3A. Amsterdam: North-Holland.

Jalan, Jyotsna, and Martin Ravallion. 2003. "Estimating the Benefit Incidence of an Anti-Poverty Program by Propensity-Score Matching." *Journal of Business and Economic Statistics,* Vol. 21, No. 1 (January), pp. 19–30.

Katz, Lawrence F. 1996. "Wage Subsidies for the Disadvantaged." Cambridge Mass.: NBER Working Paper 5679.

LaLonde, Robert. 1986. "Evaluating the Econometric Evaluations of Training Programs." *American Economic Review,* Vol. 76, No. 4 (September), pp. 604–20.

Lipton, Michael, and Martin Ravallion. 1995. "Poverty and Policy." In Jere Behrman and T. N. Srinivasan, eds., *Handbook of Development Economics,* Vol. 3. Amsterdam: North Holland.

O'Neill, Dave M. 1982. "Employment Tax Credit Programs: The Effects of Socioeconomic Targeting Provisions." *Journal of Human Resources,* Vol. 17, No. 3 (Summer), pp. 449–59.

Rosenbaum, Paul R., and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika,* Vol. 70, No. 1 (March), pp. 41–55.

Salvia, Agustín. 2001. "Sectores que ganan, sociedades que pierden: reestructuración y globalización en la Patagonia Austral." *Estudios Sociológicos,* Vol. 19 (mayo–agosto), CES-El Colegio de México.

Smith, Jeffrey, and Petra Todd. 2001. "Reconciling Conflicting Evidence on the Performance of Propensity-Score Matching Methods." *American Economic Review,* Vol. 91, No. 2 (May), pp. 112–18.

Woodbury, Stephen, and Robert Spiegelman. 1987. "Bonuses to Workers and Employers to Reduce Unemployment." *American Economic Review,* Vol. 77, No. 4 (September), pp. 513–30.