

Assisting the Transition from Workfare to Work: A Randomized Experiment

Emanuela Galasso, Martin Ravallion and Agustin Salvia¹

December 2001

Abstract: Randomly sampled workfare participants in a welfare-dependent region of Argentina were given a voucher that entitled an employer to a sizable wage subsidy. A second sample also received the option of skill training while a third sample formed the control group. Double-difference and instrumental-variables methods were used to deal with potential experimental biases, including selective compliance with the randomized assignment. Compared to the control group, voucher recipients had a significantly higher probability of employment, though their current incomes were no higher. The impact was largely confined to women and younger workers. Labor supply effects appear to have been important. Training had no significant impact. The experiment was cost effective, given that take-up of the subsidy by employers was low.

Keywords: workfare, wage subsidies, training programs, randomization

JEL codes: I38, J20

¹ Galasso and Ravallion are with the Development Research Group of the World Bank and Salvia is with 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' thanks go to staff of the Trabajar project office in the Ministry of Labor, Government of Argentina, who have helped in countless ways, and to the Bank's Manager for the project, Polly Jones, for her continuing support of the evaluation effort. An important acknowledgement goes to Julio Zelarayan who managed the operations of the experiment and the data processing. For their comments we are grateful to Polly Jones, Michael Lokshin, Biju Rao, Dominique van de Walle and participants at presentations given at the Ministry of Labor Argentina, the World Bank and the Network on Inequality and Poverty of the Latin American and Caribbean Economics Association. These are the views of the authors, and need not reflect those of the Government of Argentina or the World Bank. Correspondence: Martin Ravallion, World Bank (MSN MC3-306), 1818 H Street NW, Washington DC, 20433 USA; mravallion@worldbank.org.

I. Introduction

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 impact.² It has proved difficult to get robust estimates of impact using non-experimental methods.³

This paper reports on the “Proempleo Experiment” in Argentina in 1998-2000. This 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/or training programs seem obvious responses. But will they work, and at what cost?

The Proempleo Experiment was designed to assess the efficacy of providing a wage subsidy and specialized training in assisting the transition from workfare to regular work. The

² An overview of the arguments for and against wage subsidies can be found in Katz (1996), and 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).

³ A classic study by Lalonde (1986) found large biases in non-experimental methods when compared to a randomized evaluation of a US training program. On 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 (again using the same data set), Smith and Todd (2000) question this finding, arguing that Dehejia and Wahba’s PSM estimates are sensitive to choices made in sample selection and model specification.

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 on workfare programs, mainly Argentina's "Trabajar" program. The design features of this program assured that it was well targeted. The wage rate on the program is deliberately set at a low level.⁴ Taking account of foregone income from participation, 80% of participants coming from the poorest quintile nationally on the basis of estimated pre-intervention income per person (Jalan and Ravallion, 2001). Thus the Proempleo Experiment was implicitly targeted to low-wage workers who tended to come from poor families.

Under the experiment, one random sample received a voucher that entitled a private-sector employer to a wage subsidy that covered 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, there was a series of follow-up surveys spanning 18 months. We compare employment and incomes over time across these three samples. To our knowledge this is the first time that a randomized experiment has been done to assess options for promoting the private-sector employment of workfare participants, and it is one of only a handful of randomized evaluations of active labor market programs (ALMP's) more generally.

The following section discusses evaluation methods for ALMPs. We then describe the setting for our study (section III) and provide an overview of the data (section IV). The results are then presented in detail in section V, while section VI offers some interpretations. Section VII concludes.

⁴ Earnings data from the October 2000 Permanent Household Survey for Argentina indicate that 95% of workers in full-time jobs (35 hours or more per week) earner more than the prevailing wage rate on the Trabajar program.

II. Evaluation methods for active labor market programs

There can be no presumption that either a targeted wage subsidy or a training program will have any impact on the employment prospects of the unemployed or those on workfare programs. The literature points to a number of reasons why public expenditures on such interventions could be largely or wholly wasted. Take the case of a wage subsidy. There are possible substitution effects, whereby an employer simply replaces a current worker with a subsidized worker, and deadweight losses arising from the possibility that the worker would have been employed either way (Katz, 1996; Bell et al., 1999). Similar arguments can be made about training programs.

In the light of these uncertainties about impact, 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 amongst those eligible does not reveal well the likely circumstances of 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 that outcomes are independent of participation given the observable covariates. Instrumental-variables methods require an alternative conditional independence assumption, namely the exclusion restriction that the IV is uncorrelated with outcomes given participation and the control variables.

In some (rather rare) cases, evaluations of ALMP’s have used randomized assignment. In the case of training programs, two examples are the US Job Training and Partnership Act (see, for example, Heckman et al., 1997), and the US National Supported Work Demonstration (studied by Lalonde, 1986, and Dehejia and Wahba, 1999, amongst others). For wage subsidy programs, randomized evaluations have been done by Burtless (1985), Woodbury and Spiegelman (1987) and Dubin and Rivers (1993) — all for targeted wage subsidy schemes in the US; indeed, Woodbury-Spiegelman and Dubin-Rivers studied the same experiment by the Illinois Department of Employment Security in the mid-1980s.

Randomization is the theoretical ideal in that it balances the distributions of all (observed or unobserved) covariates between the treatment and control groups. If everyone who was given access to the training automatically took it up and access was assigned randomly then we would have no difficulty estimating the impact on those treated by comparing measured outcomes with those observed for the randomized-out control group. Under weak conditions, an unbiased estimate of mean impact for those treated can then 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 because this dummy variable is exogenous given randomization and full compliance.

However, it is often the case in randomized policy interventions that some of those randomly selected for the program do not want to participate. Then actual treatment ceases to be exogenous, even when this is true of the assignment to treatment. A common method of dealing with this problem in the evaluation literature is to first calculate what is referred to as the “intention to treat” (ITT) given by the difference in mean outcomes between those who are assigned the program (whether they take it up or not) and those not assigned. The ITT effect is then deflated by the “compliance rate,” given by the proportion of those assigned the program who take it up. This method appears to be due to Bloom (1984).⁵

An alternative econometric method views selective take-up as a classic endogeneity problem, for which the now standard solution is to find an instrumental variable that is correlated with actual treatment but uncorrelated with outcomes given treatment. While finding a valid IV is often difficult in practice, the randomized assignment is a natural choice. The impact estimator is the Two-Stage Least Squares (2SLS) regression coefficient of the outcome measure on a treatment dummy variable, using a dummy variable for assignment as the IV. As with all IV estimators, this requires an exclusion restriction, namely that the fact of being randomly assigned to the program only affects outcomes via actual participation. Angrist, Imbens and Rubin (1996) provide a precise statement of exclusion restriction and other conditions under which this 2SLS estimator gives a consistent estimate of the impact of treatment on the treated.⁶

We will use the 2SLS method to correct for endogenous compliance. However, it is not difficult to show that these two methods of dealing with selective compliance give identical results under seemingly weak conditions; the Appendix provides details.

⁵ An example in the context of assessing the Illinois experimental wage subsidy program can be found in Dubin and Rivers (1993).

⁶ For an example, see Katz et al. (2001), in the context of a randomized mobility experiment with selective take up.

III. The Proempleo Experiment

Workfare programs impose work requirements on welfare recipients, typically at benefit levels that are lower than 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, in that they turn to regular work when available.⁷

Workfare wages are typically fixed across participants and across geographic areas. This is a defensible design feature, given both 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 sub-set of workfare participants who become dependent on the scheme even though this is not generally so.

The Proempleo Experiment was done in two adjacent towns, Cutral Co and Plaza Huincul, making up the bulk of the department of Confluencia in the province of Neuquén. 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, an oil company.⁹ Even five years later, the Trabajar participation rate appeared to be unusually high in

⁷ Besley and Coate (1992) provide 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, section 6).

⁹ The state-owned oil company (Yacimientos Petrolíferos Fiscales) had facilities for scouting, extracting and refining oil in Plaza Huincul and Cutral Co, which made them the logistic and demographic center of an ample area of 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).

Confluencia, relative to expected demand for the program. In late 1998, the average number of Trabajar participants per month represented 28% of the estimated number of people living in poor households that 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 who had a total 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 is the set of beneficiaries of temporary employment programs, managed by the Ministry of Labor. The main program is the aforementioned Trabajar program. (There was a similar but much smaller provincial program; the beneficiaries of the latter 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 households with workfare participants answered the baseline questionnaire in full.

Three roughly equal random samples were then drawn by a lottery, one getting just the voucher, one getting the voucher and training, and one reserved as the control group. All were given the same questionnaire as for the Permanent Household Survey, done twice a year by the statistical office of the Government of Argentina. Some questions were added specific to temporary employment programs. All interviews were done at the beneficiaries' homes. The experiment was not announced publicly by the Ministry of Labor. Nor were any of the

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 May and October 1998.

beneficiaries (in either treatment group or the control group) told that they were part of an experiment, and they only believed that they were selected for a special survey. 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 informed of how the selection of beneficiaries had been done. So there does not appear to be any way in which program participants or potential employers knew about the fact that the assignment was random. Efforts were taken to avoid members of the one treatment group 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 somehow hear about the voucher and the training program (presumably from co-workers or neighbors) and asked if they could join it. This 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/her salary, and the employer had to discount the amount of the subsidy from the gross wages paid to the worker. The minimum wage rate in Argentina at this 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 firm only received the subsidy if the social security charges had been paid. The level of the subsidy was set such that hiring firms would be unlikely to perceive an incentive to simply fire a current registered worker to hire and register a Proempleo worker, taking account of the severance pay requirements for firing regular workers.

¹¹ For example, an employer hiring someone over 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.

Those assigned the Proempleo voucher also received an instruction lasting 2-3 hours which explained 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 which 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 had to fulfill labour practices. They also received economic support benefits, paid at a 10% lower rate than for the Trabajar program.¹² 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-building, one on professional cooking, one on raising pigs, one on greenhouse cultivation and one on skills needed for 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 and February of 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, 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

¹² 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.

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 on the program) shrank sharply in the early part of 2000, to bring down the center's outlays.¹³ This contraction also occurred in the study towns. The total number of Trabajar participants in Cutral Co and Plaza Huincul in the first five months of 2000 was only 56% of that for 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 come to 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 those of the baseline survey (December 1998). We repeated the analysis using the intermediate rounds, and we note any differences with the results for the period as a whole.

IV. Baseline characteristics and attrition

Table 1 compares the randomized assignment with observed treatment status.¹⁴ There is perfect take-up for the “voucher only” case. However, 30% of those offered “voucher and training” did not take up the training component and so were re-assigned to the voucher-only component. Our results for that group were robust to excluding these re-assigned workers. The fact that so many people did not take-up the training suggests that there may be a serious compliance problem, as discussed in section II in general terms. We return to address this problem in the next section.

¹³ The average number of new projects approved was 439 per month in the period February to 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).

¹⁴ We excluded five workers who were assigned to the control group but mistakenly ended up getting the treatment. We also excluded 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.

The top panel of Table 2 provides summary statistics on the employment status of the various treatment groups and the control 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 doing the baseline survey. Some Trabajar projects terminated in that two-month period. There is a slight sign of higher participation in the temporary employment programs amongst the control group. This might reflect some contamination of treatment effects in the baseline survey. We will check robustness of our results to using a double difference estimator.

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

As one would expect, there was sample attrition over the study period. 77.5% of those interviewed in the baseline stayed to the fourth round. We followed Fitzgerald, Gottschalk and Moffitt (1998) in testing whether the selection out of the sample is based on observable characteristics in the baseline, including the outcome indicator (Y_0). The attrition indicator (whether an observation stays in the sample) is then regressed on Y_0 and other baseline characteristics of the worker and household (gender, age and schooling). The test for attrition bias is equivalent to testing whether Y_0 is significant. In our case, the fact that the beneficiary group is mainly 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 CV of incomes is 35-45%. If there is a serious problem of endogenous attrition then 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 were significant predictors of attrition (t-ratios of 0.5 or lower). The only significant variables were age (attrition is more likely among the young) and education (those with better than primary schooling are more likely to stay in the survey).

V. Participant impacts 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 that there is a significant effect of either treatment (voucher or voucher plus training) on those becoming employed in the private sector of the order of six percentage points. The gain in wage employment clearly came from temporary employment programs (Table 3, lower panel). There is no significant effect on other outcomes: being self-employed, being employed in a temporary employment program or wage earnings. We will discuss this in the next section.

When we repeated the above calculations for the first and second survey rounds we found a significant impact on employment after six months, but not 12 months, and all other results were similar. (The wage employment effect after six months was 0.033 and was significant at the 5% level.) So 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.

In Table 4 we provide a breakdown of the impact estimates by demographic groups. We split the sample by gender and whether or not the worker is under 30, the median age. The significant impacts are confined to the wage employment of women and those under 30.

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 amongst those who are assigned access. A possible argument in the present setting is that those with intrinsically lower prospects of employment amongst workfare participants are the ones with relatively lower skills. In turn, this leads them to expect higher returns from taking up the training — to help bridge their skill deficiency, relative to those they compete with in the labor market. In this situation, ignoring the endogeneity of take-up would lead one to underestimate the true impact of the training on employment.

Table 5 gives the estimates for those who received both the voucher and training. Here we give both the ordinary estimates (analogous to Table 3) as well as the 2SLS estimates allowing for endogenous take-up of the training component (as discussed in section II). The 2SLS estimate gives higher impact — a gain to the proportion in wage employment of 7.5 percentage points. This is consistent with the expectation that people with lower skills and lower prospects of employment perceived higher gains from taking up the training. However, the extra impact of the training (an increment to the employment rate of 7.5% points versus 6.1% for the voucher only) is not statistically significant at the 5% level (z-score=0.70).¹⁵ When we repeated this analysis for the survey rounds at six and 12 months, we found no signs of a significant impact on employment.

¹⁵ The estimated standard error (s) of the difference between the two incremental employment rates (treatment minus control) is 0.020. This was calculated using the formula: $s^2 = \bar{p}(1 - \bar{p})(n_1^{-1} + n_2^{-1})$ where $\bar{p} = (n_1 p_1 + n_2 p_2) / (n_1 + n_2)$ 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 control 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 significant.

Analogously to Table 5, we give a breakdown by demographic group in Table 6. Again significant impacts on wage employment are confined to women and those under 30, though significant impacts of the voucher and training on the self-employment of men and workers over 30 do emerge.

Recall that some workers in the control sample were more likely to be employed in temporary employment programs in the baseline survey (Table 2). Thus our results so far might over-state 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 impacts. Table 7 gives these estimates. The effect on private employment holds in the double difference estimates. And again there are no 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 found 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 one firm registered all three workers.) We offer an explanation for this finding in the next section.

VI. Interpretations

Some features of our results require care in interpretation, and in drawing conclusions about the effects of scaling up the program. Although employment improved amongst voucher recipients, there are no signs of an impact on their incomes, at least by 18 months. Here it appears that voucher recipients took up private sector jobs in the expectation of a higher and/or more stable stream of incomes in the future. Given that opportunities for workfare jobs were

contracting, this would not be an unreasonable expectation. 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.

Take up of the wage subsidy by hiring firms was low. This finding echoes results for wage subsidy schemes in the US (O'Neill, 1982; Burtless, 1985; Woodbury and Spiegelman, 1987). (For example, Woodbury and Spiegelman report that only 12% of the employers who were eligible for the wage subsidy took it up.) In the present context, there is a plausible explanation for low employer take-up of the subsidy. Registering a worker so as to receive the subsidy was not costless, since it also meant incurring the government's social charges and assuming administrative costs. While the subsidy was greater than the social charges for as long as the subsidy lasted (18 months), the employer would then 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 workforce.) For such firms, registering one worker to receive the subsidy may well have been seen as risky, in that other workers would demand 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 significant impact from this experiment emerged initially (after six months) but then faded, only to re-emerge after 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. Of course, the apparent willingness of voucher recipients

to under-cut the minimum wage would have been attractive to employers, though this willingness to under-cut probably took time to develop.

The voucher may well have encouraged workers to make more effort to find work. By this interpretation, the voucher had an “empowerment” effect, in making these workers — notably young and female 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.

It should also be noted that the Trabajar workers in these company towns had 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. It might be expected that being one of those holding the voucher would 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.¹⁶ 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.)

¹⁶ This is also what one would expect if there was a stigmatizing effect of the voucher. In a similar wage-subsidy experiment for the US, Burtless (1985) found a negative impact of a wage subsidy experiment on employment and interpreted this as a stigma effect of the wage subsidy.

VII. Conclusions

We find that 18 months after the baseline survey, the proportion of the sample of 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 aggregate impact we find that the gain in wage employment was largely confined to women and younger workers. There was a slightly higher impact on employment for those who also took up the offer of training, in addition to the voucher. An impact of training only emerges once one corrects for the endogeneity of take-up amongst those randomly assigned to the treatment. However, we cannot reject the null hypothesis that the training had no impact even after correcting for endogenous compliance.

Clearly, Proempleo did not succeed in achieving a major transition to private sector jobs amongst workfare recipients in the study area. Arguably, without greater labor demand in these company towns, this would have been difficult. However, even in this setting, there appears to be scope for assisting the transition to regular jobs. Given the 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 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 of these findings, one must acknowledge the possibility that scaling up could increase the take-up rate amongst firms, and/or reduce the empowerment effect of the voucher. 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.

Nonetheless, this experiment is at least suggestive that wage-subsidy vouchers can provide cost effective assistance to low-wage workers for finding regular paid employment. The cost-effectiveness stemmed in part from low take up by firms. This also points to the inadequacy of basing assessments of the impact of wage subsidies on the extent of take up by firms. Given the evident supply-side impact of this experiment, it is clear that low take up is consistent with significant benefits to participating workers.

Appendix: Correcting for Selective Compliance

Let D be an indicator of observed treatment (=1 if treated, 0 if control), Y the observed outcome, and Z an indicator of exposure to the program (=1 if assigned to treatment, 0 if control). Randomization of the assignment allows unbiased estimation of the average effect of Z on D , and of Z on Y . In a regression setting we can write these equations as:

$$D_i = Z_i \mathbf{p}_1 + \mathbf{h}_{1i} \quad (\mathbf{p}_1 \neq 0) \quad (1.1)$$

$$Y_i = Z_i \mathbf{p}_2 + \mathbf{h}_{2i} \quad (1.2)$$

where the \mathbf{h} 's represent other determinants of D and Y . Assuming that $E(D|Z=0) = 0$ (treatment is only possible if one is assigned), \mathbf{p}_1 in equation (1.1) is simply the treatment take-up rate. (Note that randomization implies that $E(\mathbf{h}_1|Z=0) = E(\mathbf{h}_1|Z=1)$ while $E(D|Z=0) = 0$ implies that $E(\mathbf{h}_1|Z=0) = 0$.) The intention-to-treat (ITT) effect (section 2) is given by $\mathbf{p}_2 = E(Y|Z=1) - E(Y|Z=0)$. For a pure randomization, Z is exogenous in (1.1)-(1.2), i.e., $\text{cov}(Z_i, \mathbf{h}_{1i}) = \text{cov}(Z_i, \mathbf{h}_{2i}) = 0$. So both equations are consistently estimated by OLS, giving $\hat{\mathbf{p}}_1$ and $\hat{\mathbf{p}}_2$. The ratio:

$$\frac{\hat{\mathbf{p}}_2}{\hat{\mathbf{p}}_1} = \frac{\sum (Y_i - \bar{Y})(Z_i - \bar{Z})}{\sum (D_i - \bar{D})(Z_i - \bar{Z})} \quad (2)$$

is both the ITT effect deflated by the compliance rate and the 2SLS regression coefficient of Y on D with Z as the IV, i.e., the 2SLS estimator using the randomized assignment as the instrument for treatment gives the mean treatment effect on the treated. So these two methods of dealing with endogenous compliance are equivalent.

References

- Angrist, Joshua, Guido Imbens and Donald Rubin, 1996, "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association*, Vol. XCI, pp. 444-455.
- 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(3), pp. 339-360.
- Bloom, Howard S., 1984, "Accounting for No-shows in Experimental Evaluation Designs," *Evaluation Review* Vol. 8, pp.225-246.
- Besley, Timothy and Stephen Coate., 1992, "Workfare vs.Welfare: Incentive Arguments for Work Requirements in Poverty Alleviation Programs", *American Economic Review*, Vol. 82, pp. 249-261.
- Besley, Timothy, and Ravi Kanbur, 1993, "Principles of Targeting", in Michael Lipton and Jacques van der Gaag (eds) *Including the Poor*, Washington DC: World Bank.
- Burtless, Gary, 1985, "Are Targeted Wage Subsidies Harmful? Evidence from a Wage Voucher Experiment" *Industrial & Labor Relations Review*, Vol. 39, pp. 105-115.
- 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, pp. 1053-1062.
- Dubin, Jeffrey A., and Douglas Rivers, 1993, "Experimental Estimates of the Impact of Wage Subsidies," *Journal of Econometrics*, Vol.56(1/2), pp. 219-242.
- Favaro, Orieta, 1999 (Comp.), *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(2), pp. 300-344.
- Hamburg, Morris, 1977, *Statistical Analysis for Decision Making*, New York: Harcourt Brace Jovanovich.
- 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(4), pp. 605-654.
- 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, Volume 3A* (Amsterdam: North-Holland).
- Jalan, Jyotsna and Martin Ravallion, 2001, "Income Gains to the Poor from Workfare: Estimates for Argentina's Trabajar Program," *Journal of Business and Economic Statistics*, forthcoming.
- Katz, Lawrence F., 1996, "Wage Subsidies for the Disadvantaged," NBER Working Paper 5679, Cambridge Mass: NBER.
- Katz, Lawrence F., Jeffrey R. Kling and Jeffrey B. Liebman, 2001, "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment," *Quarterly Journal of Economics*, Vol. 116(2), pp. 607-654.
- Lalonde, R., 1986, "Evaluating the Econometric Evaluations of Training Programs," *American Economic Review*, Vol. 76, pp. 604-620.
- Lipton, Michael and Martin Ravallion, 1995, "Poverty and Policy," in *Handbook of Development Economics* Volume 3 (edited by Jere Behrman and T.N. Srinivasan) Amsterdam: North Holland.
- O'Neill, Dave M., 1982, "Employment Tax Credit Programs: The Effects of Socioeconomic Targeting Provisions," *Journal of Human Resources*, Vol. 17(3), 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, pp. 41-55.
- Salvia, Agustín, 2001, "Sectores que ganan, sociedades que pierden: reestructuración y globalización en la Patagonia Austral", en *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(2), pp. 112-118.
- Woodbury, Stephen and Robert Spiegelman, 1987, "Bonuses to Workers and Employers to Reduce Unemployment," *American Economic Review*, Vol. 77, pp. 513-530.

Table 1: Sample breakdown between treatments and control and assigned versus actual participation

		Voucher + Training	Voucher	Actual 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	<u>848</u>

Table 2: Descriptive statistics from the baseline survey

	Control	Either treatment	Voucher only	Voucher + Training	Difference: Treatment minus control (p-value) ^{††}		
					Either treatment	Voucher only	Voucher + training
Employment status							
Unemployed/inactive	0.075 (0.26)	0.102 (0.30)	0.107 (0.31)	0.094 (0.29)	(1) 0.19 (2) 0.10	0.16 0.08	0.45 0.22
Self-employed	0.011 (0.10)	0.007 (0.04)	0.008 (0.09)	0.005 (0.07)	(1) 0.58 (2) 0.29	0.46 0.23	0.78 0.39
Employed in the private sector	0.011 (0.10)	0.019 (0.14)	0.023 (0.15)	0.014 (0.12)	(1) 0.35 (2) 0.17	0.25 0.13	0.73 0.37
Temporary employment program	0.900 (0.30)	0.871 (0.33)	0.862 (0.34)	0.887 (0.31)	(1) 0.22 (2) 0.11	0.14 0.07	0.64 0.32
Other characteristics							
Age	32.33 (12.12)	32.20 (11.63)	32.24 (11.92)	32.14 (11.15)	(1) 0.88	0.92	0.85
Sex (proportion female)	0.470 (0.50)	0.437 (0.49)	0.437 (0.49)	0.437 (0.49)	(1) 0.37	0.42	0.46
Household size	4.29 (2.05)	4.36 (2.25)	4.33 (2.24)	4.41 (2.26)	(1) 0.68	0.84	0.55
Individual income (\$/month)	188.4 (67.1)	182.6 (76.9)	181.4 (81.6)	184.5 (68.7)	(1) 0.27	0.24	0.52
Household income (\$/month)	406.6 (267.2)	424.8 (332.8)	411.7 (315.5)	446.5 (359.5)	(1) 0.42	0.55	0.15
No. observations	281	567	354	213			

Note: ^{††} $H_0=0$ vs. (1) $H_1 \neq 0$ (2) $H_1 > < 0$ (one sided-test); standard deviations in parentheses.

Table 3: Aggregate impact estimates

Outcome variable	Means		Difference $\bar{Y}(D=1) - \bar{Y}(D=0)$
	Treated $\bar{Y}(D=1)$	Control $\bar{Y}(D=0)$	
Either treatment			
any employment	0.478	0.452	0.026
wage employment	0.143	0.085	<u>0.057</u> **
self-employment	0.035	0.021	0.014
employment program	0.296	0.345	<u>-0.049</u> *
labor income	120.591	119.271	1.320
Voucher only			
any employment	0.469	0.452	0.017
wage employment	0.147	0.085	<u>0.061</u> **
self-employment	0.037	0.021	0.015
employment program	0.282	0.345	<u>-0.063</u> *
labor income	123.184	119.271	3.913

Note: $D=1$ if received either voucher or voucher plus training; $D=0$ if control;
* indicates significance at 10% level, ** at 5% level.

Table 4: Impact estimates across different demographic groups

Outcome variable	Difference in means: $\bar{Y}(D=1) - \bar{Y}(D=0)$			
	Males	Females	Age \leq 30	Age $>$ 30
Either treatment				
any employment	0.044	0.013	-0.007	0.062
wage employment	0.034	0.076**	<u>0.092</u> **	0.020
self-employment	0.034	-0.001	0.0003	0.029
employment program	-0.028	-0.065	<u>-0.103</u> **	-0.010
labor income	2.009	2.345	14.639	-14.00
Voucher only				
any employment	0.028	0.009	-0.004	0.043
wage employment	0.042	0.078**	<u>0.088</u> **	0.029
self-employment	0.040	-0.003	0.004	0.029
employment program	-0.060	-0.065	<u>-0.102</u> **	-0.014
labor income	3.656	11.18	17.829	-12.464

Table 5: Voucher and training allowing for endogenous take-up of training

	Treated $\bar{Y}(D=1)$	Control $\bar{Y}(D=0)$	Difference $\bar{Y}(D=1) - \bar{Y}(D=0)$
any employment	0.493	0.452	0.035
wage employment	0.136	0.085	<u>0.051</u> *
self-employment	0.033	0.021	0.012
employment program	0.319	0.345	-0.026
labor income	119.271	116.282	2.989

	Assigned $\bar{Y}(Z=1)$	Not assigned $\bar{Y}(Z=0)$	Difference $\bar{Y}(Z=1) - \bar{Y}(Z=0)$	2SLS Y on D (Z as IV)
any employment	0.490	0.451	0.039	0.057
wage employment	0.140	0.088	<u>0.052</u> **	<u>0.075</u> **
self-employment	0.043	0.021	<u>0.022</u> *	0.032
employment program	0.303	0.345	-0.038	-0.055
labor income	119.927	119.067	0.860	1.247

Note: $D=1$ if received voucher + training, $D=0$ if control; $Z=1$ if assigned to voucher + training, $Z=0$ if control; *indicates significance at 10% level, ** at 5% level.

Table 6: Voucher and training allowing for endogenous take-up of training: impact across demographic groups

Outcome of interest:	Difference in means: $\bar{Y}(Z=1) - \bar{Y}(Z=0)$			
	Males	Females	Age \leq 30	Age $>$ 30
any employment	0.064	0.021	-0.009	0.097
wage employment	0.0007	0.096**	<u>0.082</u> **	0.017
self-employment	0.052**	0.001	-0.001	0.050**
employment program	0.012	-0.080	-0.090*	0.023
labor income	4.904	-1.244	12.088	-12.434

	2SLS Y on D (Z as IV)			
	Males	Females	Age \leq 30	Age $>$ 30
any employment	0.098	0.029	0.016	0.121
wage employment	-0.001	0.133**	<u>0.137</u> **	0.021
self-employment	0.080*	-0.001	<u>-0.002</u>	0.062**
employment program	0.019	-0.111	-0.151*	0.028
labor income	7.554	-1.722	20.22	-15.59

Table 7: Double difference estimates of impact

Outcome of interest:	Double difference: $\Delta\bar{Y}(D=1) - \Delta\bar{Y}(D=0)$			
	Treated	Control	Difference	
Either treatment				
any employment	-0.420	-0.473	0.054*	
wage employment	0.123	0.075	<u>0.049*</u>	
self-employment	0.028	0.011	0.018	
employment program	-0.575	-0.555	-0.020	
labor income	-58.377	-66.993	8.615	
Voucher only				
any employment	-0.424	-0.473	-0.050	
wage employment	0.124	0.075	<u>0.050*</u>	
self-employment	0.028	0.011	0.018	
employment program	-0.579	-0.555	-0.024	
labor income	-53.540	-66.993	13.453	
Voucher and training				
any employment	-0.413	-0.464	0.050	
wage employment	0.122	0.084	<u>0.039*</u>	
self-employment	0.028	0.022	0.007	
employment program	-0.568	-0.566	0.002	
labor income	-66.418	-67.081	0.663	
				2SLS
	Double difference: $\Delta\bar{Y}(Z=1) - \Delta\bar{Y}(Z=0)$			ΔY on D (Z as IV)
any employment	-0.417	-0.475	0.059	0.085
wage employment	0.120	0.074	<u>0.046*</u>	<u>0.067*</u>
self-employment	0.037	0.011	<u>0.026*</u>	0.038
employment program	-0.577	-0.556	-0.020	-0.029
labor income	-67.764	-65.963	1.801	2.612

Note: wave 4 with respect to the baseline; *indicates significance at 10% level, ** at 5% level.