

How to Improve Women's Employability and Quality of Work in Developing and Transition Economies

Petra E. Todd¹

January, 2013

¹Todd is the Alfred L. Cass Term Professor of Economics at the University of Pennsylvania, an Associate of the Population Studies Center at the University of Pennsylvania, NBER, and IZA. She wrote this paper as a consultant to the United Nations Foundation. It builds on an earlier study entitled "Effectiveness of Interventions Aimed at Improving Women's Employability and Quality of Work: A Critical Review," that she did for the Women's Economic Empowerment project of the Women and Population Department in the World Bank. The paper benefitted from Mayra Buvinic's, Elena Bardasi's, Andrew Morrison's and Anne Golla's helpful comments and Brian Collopy's excellent research assistance.

Contents

1	Introduction	3
1.1	Progress, but some persistent gaps	3
1.2	Survey coverage	4
1.3	Questions of interest	6
2	A review of alternative evaluation approaches	8
2.1	The Evaluation Problem and Key Parameters of Interest	9
2.2	Solutions to the Evaluation Problem	11
2.2.1	Randomization	11
2.2.2	Nonexperimental Estimators	12
3	Employment Creation and Job Training Strategies for Adults	14
3.1	Active labor market (ALMP) policies	14
3.1.1	Latin America	16
3.1.2	Transition Economies	21
3.1.3	China	25
3.2	Summary	27
4	Programs to Support Youth Employment	29
4.1	Combined classroom training and internship programs	30
4.1.1	Latin America	30
4.1.2	Middle East	36
4.2	Summary	36
5	Child care programs	38
5.0.1	Latin America	40
5.0.2	Middle East	45
5.0.3	Asia	47
5.1	Child-care subsidies	48
5.1.1	Transition Economies	48
5.2	Summary	51
6	How elder-care demands affect women’s work	52
6.1	Latin America	54
6.2	China	55
7	Other types of programs	56
7.1	Workplace health programs in Bangladesh	56
7.2	Workplace gender equity programs	59
7.3	Land titling programs in Peru and Argentina	64

1 Introduction

1.1 Progress, but some persistent gaps

One of the eight goals set forth at the 2010 Millennium Development Goal Summit called for gender parity in education and health, economic opportunities and decision making. Although there has been significant progress over recent decades in shrinking gender gaps, most notably, in educational enrollment, in life expectancy, and in labor force participation, the progress has not been universal and some persistent gaps remain. According to the recent World Bank Development Report on Gender Equality and Development (2012), women now make up 40 percent of the global labor force, with large increases in participation between 1980 and 2008 in regions with historically low rates, such as Latin America, the Caribbean, the Middle East and North Africa. However, women still have significantly lower rates of labor force participation than men, higher rates of unemployment, and higher rates of participation in informal sector jobs. Even when working, women tend to be concentrated in lower productivity activities and in lower paid industries and occupations.¹

Gender differences may to some extent reflect differences in preferences, but it is also clear that in many countries women have different economic opportunities and face different constraints than men. Expectations of future labor market prospects are likely an important determinant of women's preferences and aspirations. In some regions of the world, such as in many Sub-Saharan African and South Asian countries, women, particularly those from low income families, are severely disadvantaged in comparison to men in both educational and work opportunities. In other regions, such as in Latin America, women's educational opportunities have increased greatly and gender gaps in school enrollment have even reversed, but gaps in labor market outcomes persist.

Policy interventions have the potential to reduce male-female differences in educational and economic opportunities. To understand the scope for intervention, it is useful to consider the many different demand-side and supply-side factors that contribute to lower rates of labor

¹See The World Bank Development Report on Gender Equality and Development (2012).

force participation, lower wages, and higher unemployment rates among women. In regions with high fertility rates, such as in Latin America, India, the Middle East, and Africa, an important factor limiting the supply of female labor is lack of affordable child care. There is often a cultural expectation that women be the main care providers not only for their children but also for elderly parents. In some societies, women who want to work may also encounter direct opposition from family and male partners to their working. In addition, women often have relatively lower levels of schooling than men as well as lower levels of job market experience, which both contribute to a lower earnings capacity that discourages work.

On the demand side, labor market discrimination against hiring women is also not uncommon, and surveys of employers have found that many employers openly express an explicit preference for hiring men over women.² Women also tend to have less well developed job networks, making it more difficult for them to find good jobs. Some studies have shown that women and men go about searching for jobs in different ways, with men making greater use of contacts outside their family. In some countries, cultural prohibitions on the degree to which women can interact with others outside their family severely restrict work opportunities.

1.2 Survey coverage

Policy interventions in developing countries often aim to improve the employment prospects of low income individuals and their families, particularly those facing severe labor market shocks or credit constraints. Some interventions specifically target women, under the view that women have an inequitable share of power in household decision-making or that women face barriers to labor market entry or to advancement. Other policies, such as active labor market programs, do not explicitly target women but may affect men and women in different ways.

This paper examines the effectiveness of a variety of policy interventions that have been tried in developing and transition economies in improving women's employability and quality

²See Katz, 2008.

of work. The programs include active labor market programs for adults, training programs for youth, child care programs, workplace health programs, workplace gender equity programs, and land titling programs. The interventions studied are heterogeneous, differing in scope, content, targeting and eligibility criteria, as well as in objectives. Common objectives are to reduce unemployment, facilitate formal sector employment, increase wages, promote gender equity, provide social protection, and/or increase women's empowerment in the household. Child care programs often have important secondary goals related to improving the health and educational development of children. Some programs are government sponsored, while others were introduced by international organizations, NGOs or private firms. The populations served by the programs are heterogeneous in terms of demographics and labor markets, residing in rural and urban settings in Latin America, Africa, Europe and Asia.

Because the coverage of this survey is broad, we consider only relatively recent programs that have the subject of rigorous impact evaluations. Also, we focus on programs that aim to directly affect employment and earnings and do not consider programs oriented more generally towards education or health, which may indirectly influence labor market prospects. For example, conditional cash transfer (CCT) programs that encourage children's enrollment in school are now very common in the developing world. CCT programs have been found to increase educational achievement of children/adolescents and to influence the time use of their mothers, but such programs do not aim to directly influence women's labor market outcomes and are therefore not discussed here.³ Also, programs that provide reproductive health services to women likely affect their capacity to work, but they are also not reviewed here.⁴

The focus of this survey is on programs that aim to improve women's prospects for finding paid work. We consider the following:

- Active Labor Market Policy (ALMP) programs that provide job search and training services and sometimes also wage subsidies to the unemployed

³For a comprehensive survey of CCT programs around the world and summaries of evaluation results, see Fiszbein et. al. (2009).

⁴With one exception of a workplace health program that also provides family planning services.

- Youth oriented training/internship programs
- Child care and child care subsidy programs
- Effects of elder care on women's working behavior
- Workplace health programs
- Workplace gender equity programs
- Land titling interventions

In reviewing the literature, we discuss the key findings from alternative evaluation studies as well as the methodologies used to generate these findings and attempt to synthesize results across multiple studies to infer which programs are most effective and in what settings.

1.3 Questions of interest

There are many questions that are potentially of interest in evaluating the effects of a social policy intervention. Foremost is the question of whether people affected by the policy or program intervention benefit from it, at least on average. Most of the evaluation literature, including most of the studies examined in this paper, focus on estimating the so-called *average effect of treatment on the treated*, which is the average program impact for people who were exposed to some treatment. The treatment may represent active participation in some program, such as a job training program, or passive exposure to some policy, such as being eligible for a subsidy or being given access to a child care program. Another question is how program benefits are distributed across people, for example, what fraction of people experience a positive benefit from the intervention and how do the benefits vary according to the characteristics of the participant? A third question that usually merits consideration is whether the program benefits outweigh the costs. A program that generates benefits that are less than its costs might be deemed unsuccessful, unless there are other compensating

factors.⁵ A fourth question is how program impacts and costs would differ if some features of the program were changed. For example, if a policy intervention provides child care subsidies, we may want to know how an increase in the level of the subsidy would affect mother's labor supply. Answering these types of questions requires assessing the effects of programs that have never been tried, by extrapolating from previous experience with an existing program. Relatedly, it might be of interest to explore how program benefits would vary if the program were extended to new segments of the population. For example, the program eligibility criteria might be relaxed to be more inclusive or the program may be introduced to new regions.

Most program evaluation studies compare the performance of a group that participates in a program to that of a group that does not participate. It is common to target social programs selectively at families or individuals that are deemed most in need of them or that are likely to get the largest benefit from participating. For example, a family planning program might be targeted at high fertility regions or an unemployment program at areas with high unemployment rates. Individuals cannot usually be coerced to participate in programs and typically self-select into them, so that only more motivated individuals or those that expect to benefit most from the programs may choose to participate. Selective targeting and program self-selection can promote efficient use of program resources; but these mechanisms also pose challenges for evaluating the impact of the program, because they generate differences between the groups that participate and do not participate in the program. Such differences need to be taken into account in evaluating the program's impact through a comparison of program participants and nonparticipants. In weighing the evidence on program effects that is reported in different evaluation studies, an important consideration is the reliability of the statistical methods used to take into account preexisting differences between the participant and nonparticipant groups.

This paper is organized as follows. Section two provides an overview of experimental

⁵For example, some training programs in the U.S. that are targeted at older displaced workers generate income benefits that are less than the cost of the program, but it may still be desirable to provide training and employment services for such workers if a high value is placed on their employment per se.

and nonexperimental evaluation methods, focusing on the methods most commonly used in the evaluation studies that are reviewed in this paper. Section three summarizes results of evaluations of so-called Active Labor Market Policy (ALMP) programs that have been tried in Latin America, Eastern Europe and China. The aim of ALMP programs is usually to mitigate the effects of severe macroeconomic shocks. Most programs are not explicitly targeted at women, but women participate in them. The types of programs considered in this paper include job search assistance programs, wage subsidy programs, public works programs, occupational retraining programs, and internship programs. Participation in ALMP programs is sometimes mandatory to receive unemployment benefits.

Section four reviews evidence the effectiveness of education/training programs that are targeted at youth. These programs are similar to ALMP programs, but they are explicitly targeted at younger people and often include an educational component and an internship component. Section five considers the effectiveness of programs that are designed to facilitate women's work by increasing the availability and affordability of child care. Section six reviews the evidence on how elder care demands affects women's work. Lastly, section seven discusses the effectiveness of a few other kinds of programs. Section eight summarizes and provides policy recommendations.

2 A review of alternative evaluation approaches

There are two main approaches for evaluating social programs in a way that addresses the problem of noncomparability between participants and nonparticipants. One approach, considered by many to be the gold standard, is to use a randomized experimental design.⁶ Under a randomized design, some fraction of individuals that satisfy program eligibility criteria are randomly excluded from the program and serve as the control group. Randomization ensures that the group that is offered the program is comparable to the group that is not offered the program along both observable and unobservable dimensions, which is the major virtue

⁶For a discussion of the use of randomized social experiments in developing country contexts, see Duflo and Kremer (2004). For a more critical discussion of the use of randomization in economic development, see Deaton (2009)

of experiments. When randomized experiments are successful, they provide a relatively straightforward way to estimate the average effect of treatment on the treated.

An alternative evaluation approach is the nonexperimental approach, which uses nonexperimental data (sometimes called observational data) on program participants and nonparticipants and employs statistical methods to adjust for noncomparability between the groups. There is an extensive literature that develops nonexperimental methods for evaluating the impact of social interventions and assesses their efficacy. The statistical estimator used and the quality of the data available for modeling program participation decisions are important factors affecting the reliability of inferences from nonexperimental evaluation studies.

2.1 The Evaluation Problem and Key Parameters of Interest

We begin by defining some notation for describing common parameters of interest in an evaluation. Suppose there are two states of the world, corresponding to the state of being with and without some treatment intervention. For example, the outcome of interest could be an indicator for whether employed or unemployed and the treatment could be participating in a job training program.

Let $D = 1$ for persons who receive the intervention and $D = 0$ for persons who do not receive it. Associated with each state is a potential outcome, which may or may not be realized. Y_0 denotes the potential outcome in the untreated state (e.g. without training) and Y_1 the potential outcome in the treated state (with training). Each person has associated a (Y_0, Y_1) pair that represents the outcomes that would be realized in the two states of the world. Because a person can only be in one state at a time, at most one of the two potential outcomes is observed at any point in time. The observed outcome can be written as

$$Y = DY_1 + (1 - D)Y_0.$$

The gain from moving an individual from the state “without treatment” to the state “with treatment” is the treatment effect for that individual:

$$\Delta = Y_1 - Y_0.$$

Because only one of the states is observed, the gain from treatment is not directly observed for anyone. Inferring gains from treatment therefore requires solving a missing data problem. The evaluation literature has developed a variety of approaches to solve this problem. The literature has focused mainly on estimating direct effects of the program on program participants under the assumption that the indirect effects of the program on nonparticipants are negligible. This allows nonparticipants to be used as a source of comparison group data and to represent the “no treatment” state.

Because treatment impacts are not directly observed, researchers usually aim instead to uncover some features of the treatment impact distribution, such as the mean or median program impact. Much of the evaluation literature focuses on methods for estimating two key parameters of interest:⁷

- (i) the average gain from the program for persons with characteristics X , commonly referred to as the *average impact of treatment (ATE)*

$$E(Y_1 - Y_0|X) = E(\Delta|X).$$

- (ii) the average gain from the program for program participants with characteristics X , known as the *average impact of treatment on the treated (TT)*:

$$E(Y_1 - Y_0|D = 1, X) = E(\Delta|D = 1, X).$$

The ATE parameter is the gain from the program that would be experienced on average if a randomly chosen person with characteristics X were assigned to participate in the program. The TT parameter is the average gain for individuals who actually participated in the program (for whom $D = 1$). If individuals who take the program tend to be ones that receive the greatest benefit from it, then we would expect $TT(X) > ATE(X)$.

A comparison of the average gain accruing to participants, expressed in monetary terms, and the average costs of a program is informative on whether the program covers its costs.

⁷See, e.g., Rosenbaum and Rubin (1985), Heckman and Robb (1985), or Heckman, Lalonde and Smith (1999) for discussions of different parameters of interest.

In determining the average gain to participants, any opportunity costs of participating in a program also need to be taken into account. For example, while a worker is participating in a three month job training program, she may not be able to work. The gain from the program might be calculated as the sum of earnings in the 18 months following program participation, inclusive of the zero earnings during the 3 month training period, minus the predicted sum of 18 month earnings that the individual would have experienced in the absence of the program. This net gain can then be compared with the cost of the program to come up with a benefit-cost ratio that is informative on whether the program at least covers its costs.

2.2 Solutions to the Evaluation Problem

2.2.1 Randomization

Randomized social experiments are considered by many researchers and policy makers to be the ideal design for evaluating the effects of a treatment. Under a randomized experimental design, a group of individuals is randomly selected to receive a treatment and another group is randomly denied the treatment and serves as the control group. The main advantage of random assignment is that it ensures that program participants and nonparticipants are comparable both in terms of observables and unobservables. Randomization can be performed either before or after application to a program. The discussion below assumes that randomization is performed after people apply to a program and are determined to be eligible for it.

In terms of the previously described parameters of interest, randomization provides a way of estimating the average effect of treatment on the treated (TT). To see why, let $D = 1$ denote having applied and been deemed eligible for a program, else $D = 0$. Also, let $R = 1$ if randomly assigned to the treatment group and $R = 0$ if randomly assigned to the control group. From the treatment group, we obtain

$$E(Y_1|R = 1, D = 1, X)$$

and from the control group

$$E(Y_0|R = 0, D = 1, X).$$

The difference in means gives

$$\begin{aligned} & E(Y_1|R = 1, D = 1, X) - E(Y_0|R = 0, D = 1, X) \\ &= E(Y_1|D = 1, X) - E(Y_0|D = 1, X) \\ &= TT(X) \end{aligned}$$

were the conditioning on R can be dropped by virtue of random assignment (R is uninformative about Y_1 or Y_0). Thus, a well designed randomized experiment delivers one of the key parameters of interest in evaluations.⁸

Although there are many advantages to using randomized experimental designs to evaluate effects of program interventions in terms of assuring comparability between the treatment and control groups, there are also some potential drawbacks to randomized experiments that are described in Appendix A.

2.2.2 Nonexperimental Estimators

In the absence of a randomized experiment, evaluations must be based on nonexperimental (or observational) data. Also, when a randomized experimental design is compromised in some way, nonexperimental methods can also be used to increase comparability between the treatment and control groups. Nonexperimental estimators of program impacts typically use two types of data to impute the missing counterfactual (Y_0) outcomes for program participants: data on participants at a point in time prior to entering the program and data on nonparticipants. The following types of methods are commonly used in evaluation work and in the studies described in this review:

(i) *Cross-section or difference-in-difference regression*: These methods evaluate the effects of the program by comparing the outcomes for a treated group to that of a nonexperimental comparison group, using regression adjustment to control for preexisting dif-

⁸We have so far discussed randomization that is performed after assessing eligibility for a program. An alternative randomized design randomizes eligibility for the program. For example, some randomly determined fraction of the population is told that they are eligible for the program. The eligibles may then choose whether to apply or not. This alternative randomization strategy, which is less commonly used, is discussed in Heckman, Lalonde and Smith (1999).

ferences in observed characteristics. The cross-section estimator compares participant and nonparticipant outcomes at some point in time after the program start date, whereas the difference-in-difference estimator compares the change in outcomes between a post-program and pre-program time period (for example, the change in earnings).

(ii) *Cross-section or difference-in-difference matching*: Matching methods evaluate the impacts of a program by matching individuals in the treatment group to observably similar individuals in the control group on the basis of a set of observable characteristics. The overall effect of the program is obtained by averaging over the differences in participant and matched nonparticipant outcomes. A commonly used metric for matching the individuals is the propensity score, which is the predicted probability of participating in the program conditional on a set of observed characteristics.

(iii) *Control function methods*: These methods evaluate the impacts of the program by a comparison of treated individuals and comparison group individuals, using statistical adjustment to control for both observed and potentially unobserved differences between the groups. These estimators usually require assumptions on the distribution of unobservables and on their relationship to observables.⁹

(iv) *Instrumental Variables or LATE estimators* : These estimators require that there be some factor that influences the program participation decision but not the outcome directly, for example, an administrative rule that affects whether individuals are admitted in the program but that is not correlated with individual outcomes. When treatment effects are heterogeneous across people, they provide an estimate of the average impact of the program for the subgroup whose participation status is affected by the factor.

(v) *Behavioral modeling*: In some contexts, particularly in evaluating ex ante the effects of hypothetical treatments that have never been tried, researchers proceed by fully specifying a behavioral model that can then be used to extrapolate from historical observations on behavior to a new environment. For example, the model representing women's choices about

⁹For example, it might be assumed that the unobservables affecting the program participation decisions and outcomes are jointly normally distributed and statistically independent of the observables.

labor supply given current availability and pricing of day care could be estimated and then used to analyze the effect of changing the availability or pricing.

Todd (2008) describes these methods and the assumptions needed to justify them in greater detail. The performance of alternative nonexperimental methods depends on whether these assumptions necessary to apply them are justified and on the quality of the data used in implementing them. For example, matching-on-observables approaches usually perform best in situations where the data are rich enough to capture the key determinants of the program participation process. Another important consideration in weighing the evidence from any particular evaluation study is whether the sample used is representative of the types of people expected to participate in the program.¹⁰

3 Employment Creation and Job Training Strategies for Adults

3.1 Active labor market (ALMP) policies

We next review the evidence on the effectiveness of a variety of Active Labor Market Policy (ALMP) programs in affecting the employment, wages and poverty status of participants. Two predominant types of ALMP programs are wage subsidy programs that subsidize wages in either public or private sector jobs, and training programs that provide formal training designed to help participants develop occupational skills. The training can be provided by either public or private training institutions and can either take the form of classroom vocational training or on-the-job training, which is often an unpaid internship at a private sector firm. Some programs combine classroom training for a few weeks with an internship at a firm afterwards. Other types of ALMP programs are basic education programs, or short-term interventions that enhance job search skills. ALMP programs sometimes also provide other sorts of benefits, such as work clothing, child care and transportation expenses. ALMP programs are popular in many European countries (e.g. Germany, Netherlands, Norway,

¹⁰If so, then the study is considered to have high external validity.

Switzerland), where program participation is often mandatory to receive unemployment benefits. However, the discussion here focuses on the use of ALMP programs in developing and transition economies.

There are multiple channels through which ALMP programs might be expected to influence workers' employment and earnings outcomes. One is that they may increase the productivity of the worker and therefore her offered wage by augmenting her skill set or by providing opportunities to gain work experience (e.g. through internships). ALMP programs may also affect the process by which workers match with firms, for example, by reducing the costs of searching for a job or by increasing the arrival rate of job offers.

Wage subsidy programs could induce some worker-firm matches to take place that might otherwise not take place. For example, in the absence of any subsidy, a worker might only be willing to accept a wage offer above a certain threshold or else keep searching. If the subsidy is given to the worker, the worker might then be willing to accept wage offers that would otherwise have been deemed too low. When the subsidy is removed, the worker-firm relationship might dissolve, unless the worker has gained enough experience on the job to increase the wage offer and depending on the costs of searching. Alternatively, the subsidy could be given to the firm and the firm might then be willing to hire workers that it otherwise would not have hired. One concern about any sort of wage subsidy/voucher program is that the subsidy/voucher facilitates job matches for workers who receive the subsidy possibly at the expense of other workers who do not receive it. The costs imposed on other workers are often not taken into account in evaluations of subsidy programs.

Most evidence on evaluations of training programs comes from the U.S. and from Europe, where training programs have been rigorously evaluated. The impacts of job training programs in the U.S. and in Europe have generally been found to be modest and there is a tendency to find larger impacts for women than for men. Kluve (2012) does a meta-analysis of evaluations of European programs and finds that certain types of programs are more effective than others. He concludes that, in the European context, wage subsidies and job search services are among the most effective types of programs in increasing participants'

employment probability. Training programs also have modestly positive effects. However, direct employment programs that provide jobs in the public sector are detrimental. Card et. al. (2010) reports the results a meta-analysis of European programs and concludes that programs serving youth are less likely to show positive impacts than programs serving adults. Betcherman et. al. (2004) finds that estimated training program impacts in Latin America are generally more positive than the impacts of programs in the U.S. and in Europe.

In this section, we first consider a number of ALMP programs in Latin America (the majority in Argentina) and then discuss other programs in Eastern European transition economics and in China. Many of these programs do not focus exclusively on women, but women are included among the participants and we highlight how the programs affect women.

3.1.1 Latin America

Proempleo Program in Argentina: A number of large-scale ALMP programs have been introduced in Latin America, as a way of alleviating the effects of severe labor market shocks affecting the region. One of these is the *Proempleo* program in Argentina, which was studied by Galasso, Ravallion and Salvia (2001) using a randomized experimental design. The program provided vouchers for workfare participants to give to prospective employers. The voucher entitled employers to a sizable wage subsidy, \$150 per month for workers age 45 and older and \$100 per month for younger workers, which lasted for up to 18 months. The experiment randomly allocated individuals into two treatment groups and one control group. One treatment group received only the voucher program, whereas the other received the vouchers plus had an additional option of skill training. The controls received neither. The program was made available to beneficiaries of temporary employment programs managed by the Ministry of Labor, the main program being *Trabajar*.

Galasso et. al. (2001) find that the voucher program reduced the probability of unemployment, despite the fact that few firms actually took up the voucher subsidy. Voucher recipients had a significantly higher probability of employment but had no higher current income. Women and younger workers experienced the largest treatment impacts. Only 30%

of those assigned to the voucher plus training treatment arm took advantage of the training option. Impact estimates based on a comparison of the two treatment groups indicate that the additional option to take training had no additional impact.

Galasso et. al. (2001) hypothesize that the treatment effect of the vouchers may have been an "empowerment effect" in that workers who received vouchers seem to have been more comfortable in approaching private employers. They might also have been perceived by employers to be different from regular *Trabajador* workers. One possible reason for the low employer take-up rate is that taking up the voucher requires formalizing the employment arrangement, which could imply additional costs for the firm (such as severance payments to fire the worker).

The Galasso et. al. (2001) study reports both intent-to-treat estimates and local average treatment effect (LATE) estimates of program impacts. The intent-to-treat estimates give the effect of the program offer to participants (irrespective of whether participants took advantage of the program). The LATE estimates use the randomized group assignment as an instrument for program participation status. The LATE estimate represents the average program effect for the group induced to participate in the voucher program as a result of being assigned to the treatment group (known in the literature as group of "compliers").¹¹ The outcome variables are the changes in earnings and employment between the last follow-up wave (May, 2000) and a baseline survey in December of 1998. A limitation of the analysis is that there was some attrition in the experimental samples, with only 77.5% of those interviewed at baseline staying until the fourth round. Galasso et. al. (2001) find that private sector employment improved among voucher recipients, with an employment rate of 14% for voucher recipients and 9% for the control group. The differences were statistically significantly different after 18 months. Disaggregations revealed that the impacts were largest for female workers, younger workers and those with more education and most of the jobs gained through the program were in the informal sector. The program was deemed cost-

¹¹For discussion and development of LATE estimation methods, see Imbens and Angrist (1994) and Heckman and Vytlačil (2005).

effective, because it yielded employment impacts at very low cost, because of the low take-up of the subsidy by employers.

Trabajar II program in Argentina: In response to the macroeconomic crisis in the mid 1990s, the government of Argentina introduced in May 1997 the *Trabajar II* program, which provided short-term work opportunities at relatively low wages and targeted unemployed workers from poor families. Under the program, local governmental and nongovernmental organizations submit proposals for socially useful projects, such as projects to repair local infrastructure. The proposals have to be viable with respect to a set of criteria and are given priority according to how well they target poor areas, what benefits they are likely to bring to the community, and how much the area has already benefitted from the program. To be eligible for program benefits, workers had to be hired on to a successful proposal project and cannot be receiving unemployment benefits or be participating in another employment or training program. The projects lasted a maximum of six months but a worker could continue in the program if he/she switched to working on a new project. The wage rate was set at a maximum of \$200 per month, which was deemed low enough to assure good targeting and to help assure that workers prefer regular work when it becomes available.

Jalan and Ravallion (2003) analyze the impacts of the *Trabajar II* program household income using a nearest neighbor propensity score matching methodology. The average gain accruing to program participants is \$103 dollars, about half of the average *Trabajar* wage. The gains for female participants are not much different from the gains of male participants, but female participants tend to be from less poor backgrounds. Income gains are greatest for younger people (in the 15-24 age range).

Jalan and Ravallion (2003) do not report a benefit-cost analysis of the program. If the productivity of the workers on the socially useful projects exceeded the wages, then the program could be considered to provide a benefit that exceeds the cost. However, a clear aim of the program was also redistribution towards the poor. In that case, the program might be deemed a success even if the worker productivity did not exceed the program's expenses, depending on how the government valued the income redistribution achieved through the

program.

Jefe program in Argentina: A subsequent study by Galasso and Ravallion (2004) analyzes effects of another more recent Argentinian program, called *Jefes*, that replaced the earlier *Trabajar* program and was designed to provide direct income support for heads of households with dependents who became unemployed as a result of Argentina's economic crisis of 2002. At that time, the poverty rate soared from 37% to 58%. The *Jefes* program had work requirements, instituted to ensure that it reached those in greatest need, and it covered about 2 million households. Program participants were required to do 20 hours per week of community work, training, school attendance or employment in a private company with a wage subsidy. A major concern with regard to program implementation was program leakage, because administrators did not closely monitor whether the people signing up for the program were truly heads of households. It was also difficult to verify unemployment status, because many Argentinians work in the undocumented informal sector.

The impact analysis carried out by Galasso and Ravallion (2004) is based on the Oct, 2001 and Oct, 2002 rounds of the *Encuesta Permanente de Hogares* survey, which covers urban areas. With respect to leakage, the study finds that one third of those receiving the program were ineligible and that 80% of individuals who were eligible did not receive the program. In particular, more than half of the program participants were women who were probably not heads of households. Despite the problems in imposing eligibility criteria, however, the program was fairly well targeted at poor households. About half of program participants came from poorest fifth and 80% came from the poorest 20% of the population.

Galasso and Ravallion (2004) evaluate program impacts using cross-sectional and difference-in-difference propensity score matching approaches. The treatment group includes those who applied and were admitted into the program, and the comparison group includes persons who applied for the program but had not yet joined. Galasso and Ravallion (2004) find it difficult to predict participation status among program applicants, and the predictive power of the propensity score model is not very high, raising some concerns as to whether the observables included in the propensity score model adequately control for differences between the

treatment and comparison groups.

The matching analysis reveals that program participants experienced a smaller drop in real income on average than the comparison group, suggesting net gains on average between half and two-thirds of the gross wage, depending on the estimator used. Galasso and Ravallion (2004) argue that given the level of income support the observed income gains should have been in the range (0,150) and that negative estimates or estimates that exceed 150 should therefore be excluded. On these grounds, they prefer the impact estimates derived from the cross-sectional estimator, which indicate that 26% of Jefes participants would have been unemployed were it not for the program and 23% would have been inactive (primarily women).¹² On the whole, Galasso and Ravallion (2004) find that the program reduced Argentina's aggregate unemployment rate by about 2.5% and contributed to social protection during the economic crisis by supplementing the income of poor families.

Assessing the cost effectiveness of the *Jefes* program requires comparing the program costs to the value associated with the income redistribution under the program.

PROBECAT program in Mexico: Revenga, Riboud, Tan (1994) evaluate the effects of short-term vocational training in Mexico provided by the PROBECAT program, which was offered to more than 250,000 unemployed people. Program participants were selected according to an eligibility index that gives weight to factors such as the number of economic dependents, whether the individual attained a basic level of education, whether the individual was unemployed for less than 3 months, and prior work experience. Also, to be eligible, individuals had to be within the age range of 20 and 55 and be registered at the unemployment office. The impact analysis is based on longitudinal data on PROBECAT trainees combined with a separate dataset on a control group of unemployed people who did not join PROBECAT, drawn from national labor force surveys, which surveys individuals for 5 quarters. Women were 49% of the trainee group but only 33.8% of the comparison group.

¹²Ruling out negative estimates or estimates that exceed 150 is potentially problematic. Income support could have been used for productive purposes, such as a small business, making gains in excess of 150 possible. See, for example, the discussion of the Kaboski and Townsend (2007) study in section 4. Also, negative income gains are also feasible, if, given the transfer, some women might have withdrawn from the labor force.

The average female trainee was 29 years old and 46% were married. Women were less likely than men to have completed secondary education (grades 10-12).

The key outcomes analyzed in the study are employment, monthly salary, and number of hours worked. Impacts of the program on the length of unemployment spells are assessed using a Cox proportional hazards model, estimated on the treatment group and on a subgroup of controls predicted to have a high probability of participating in the program based on their characteristics. Potential drawbacks of the analysis are that participants and controls were given different surveys, so that measurements across surveys may not be comparable, and the control sample is relatively small. On average, program trainees found jobs more quickly (program participation reduces unemployment spell by 1.9 months for females and 2.5 months for males). Subgroup analyses reveal that the program positive program impacts in terms of shortening unemployment spells are concentrated on trainees older than 25 and those with work experience. Female trainees with work experience are more likely to be employed at 3,6 and 12 months after the training than are similar controls. Training increases the number of hours worked for both women and men, but only increases monthly earnings for men. Revanga, Riboud and Tan (1994) conclude that the program is cost effective, in the sense of covering its costs, for women over age 25 but not for younger women.

3.1.2 Transition Economies

Transition economies typically undergo large shifts in the demand for different kinds of labor as they move from a centralized to a more market-based economy. Because of labor market frictions and because it takes time to acquire new skills, transition economies typically undergo a period of high unemployment rates and large stocks of long-term unemployed persons. There have been a number of large scale ALMP programs implemented in transition economies aimed at equipping workers with skills that are in greater demand in the market economy and at facilitating their job search process. Here we review the results of programs implemented in Russia, Romania, Slovakia, and Poland. None of these programs were specifically targeted at women, although women in each case made up a substantial fraction of

the participants. The training programs are often a prerequisite for receiving unemployment benefits, so the impacts obtained are for the training program plus the stream of associated unemployment benefits and not for the program alone.

ALMP programs in Russia and Romania: Benus, Brinza, Cuica, Denisova, and Kartseva (2005) analyze the effects of ALMP programs in Russia and Romania. The program eligibility criteria and the populations served differ somewhat by country. In Russia, training services are available only to individuals who are officially registered as unemployed with an employment center and receipt of unemployment benefits is conditional on making efforts to gain employment and on being available for work and taking suitable jobs as they become available. The unemployment benefit for persons who worked at least 26 weeks over last year is equal to 75% of former wage at first and declines to 45% or the minimum wage (whichever is greater) over time. The benefit for other categories of workers is the minimum wage. A person who has been unemployed for a year and whose family income does not exceed two minimum wages also qualifies for social assistance benefits. The analysis sample used in the evaluation study consists of a group of program participants and a control group that were selected on the basis of 2002 administrative data on training program clients.

In Romania, eligibility for ALMP programs depends on being registered unemployed, having income less than 50% of the minimum wage, being unemployed due to layoffs, having been employed at least 6 months of the last 12 months or being a recent graduate from school or university. There are various types of training available to program participants, including a public service component, whereby local government and other eligible organizations can propose public projects with a maximum cost of up to \$50,000 and hire ALMP participants to work on them. The analysis sample is comprised of individuals who entered the register not earlier than January 2001 and got off no later than December 2002.

Benus, et al. (2005) evaluate the impacts of both the Russian and Romanian ALMP programs overall and separately by gender, using a propensity score methodology. The propensity score model is based on a fairly limited set of predictors of program participation, that include gender, age and education. Nonparticipants are people who applied for training

but were not selected for it, so they would be expected to differ in some respects from participants. The outcomes of interest in the evaluation are the likelihood of being employed at the time of the follow-up survey, the likelihood of being employed at least once after the program, the likelihood of having a high salary, and the length of the current unemployment spell. The impact evaluation finds no significant effects of the ALMP programs in Russia, on the whole. In Romania, however, the program is found to have a statistically significant impact for three of the four outcomes (the likelihood of employment, the likelihood of being employed at least once, and the level of wages are all higher among participants in Romania). Subgroup analyses reveal some gender, age, and education heterogeneity in the impacts for Romania. Re-training increases the probability of employment and decreases the wage for females. That is, re-training appears to help find employment but at a lower wage than the individuals would find on their own. Middle age and lower education level individuals experience the biggest program impacts. For men, the retraining is found to have no effect.

The authors hypothesize that the difference in impact findings between Russia and Romania may be attributed to differences in the characteristics of program participants affecting their labor market prospects; the Russian sample is better educated (45% have university degree), older and has more labor force experience than the Romanian sample. In both countries, the program was not found to be beneficial for highly educated workers. Also, program participants may have been more negatively selected in Russia, because not everyone unemployed get registered in Russia.

ALMP programs in Slovakia: Lubyova and Ours (1999) uses administrative data from 20 Slovak districts to analyze whether it is beneficial for unemployed workers who want a regular job to accept a temporary ALMP job or enter a retraining program. Specifically, they study the effects of two ALMP programs in Slovakia targeted at registered unemployed workers on their exit rate from unemployment. The programs were targeted especially towards older workers, disabled workers and long-term unemployed. The program provided retraining and counseling services as well as wage subsidies in two types of jobs: Socially purposeful jobs (SPJ) and publicly useful jobs (PUJ). The subsidy in SPJ, which could be at private sector

firms, had a minimum duration of 2 years and the subsidy at PUJ, which were typically public works jobs, had a maximum duration of 6 months.

Lubyova and Ours (1999) analysis is based on administrative records of 100,000 individuals who entered unemployment in 1993. The records allow construction of detailed labor market histories. Using multivariate duration analysis, the authors jointly model the duration of unemployment and duration of stay in an ALMP program, controlling for observable and unobservable heterogeneity among people. The focus of the study is on whether participation in ALMP programs affects the exit rate from unemployment to regular jobs. For women, 40% exit unemployment by finding jobs, 9% exit by entering ALMP and 51% had right censored spells. For males, 47% exit unemployment by finding a job, 8% exit it by entering an ALMP program and 45% had a right censored spell. On average, workers that enter the ALMP programs are found to have a 150% increase in the exit rate into a regular job, with similar estimated program impacts for men than women. From additional analyses allowing the ALMP program effect to depend on the type of program, the authors conclude that there are positive benefits of retraining and publicly useful jobs on exiting unemployment into a job. For socially purposeful jobs, however, they find a negative effect.

ALMP programs in Poland: Kluve, Lehmann and Schmidt (1998) study the effectiveness of ALMP programs in Poland. The program took three forms: publicly financed training and retraining, intervention works (wage subsidies for workers in private or public firms), and public works. The aim of training and retraining was to increase the skill set of individuals in demanded fields such as data processing, accounting, secretarial work and welding, through courses lasting on average 2.-3 months. Individuals receive unemployment benefits during the course of their studies. Workers in the training component tended to have higher education levels and to be female. The wage subsidy component of the program was structured so that the subsidy was increasing in the time the worker stays with the firm. The public works component was targeted at longer term unemployed and many of the jobs available were low skill jobs working on infrastructure improvements. Workers in either the wage subsidy program or the public workers program had an incentive to participate in the wage subsidy

program at least 6 months to qualify for another 12 months of unemployment benefits.

The ALMP programs are evaluated using a difference-in-difference matching approach, where the main outcome of interest is labor force status. The samples are drawn from the 18th wave of the Polish Labor Force Survey that includes a supplement with four years of historical information on individual labor market histories (monthly from 1992-1996). The treated group are individuals who were offered participation in the programs by their local labor office and who accepted the offer. Sample sizes in the three types of programs (training or retraining, intervention works or public works) were 241, 532 and 93. The control group consists of 7784 individuals who had been registered at least once as unemployed since January 1992. The matching procedure pairs treated individuals with control individuals who have the exact same labor force history and are matched on certain demographic characteristics (gender, marital status, education, region and age). The matching impact estimates turn out to be sensitive to which variables are used in the matching analysis, for example, to whether the selection of matches also takes into account local labor market conditions.

The impact estimates indicate that the training/retraining program increases the average employment probability for both men and women. Participation in the non-training ALMP programs does not affect women's employment probabilities but has a negative effect on men's employment probabilities, which the authors attribute to benefit churning rather than stigmatization of intervention and public works participants. That is, males appear to take intervention works and public works jobs in between two spells of unemployment benefit receipt. Overall, the study concludes that ALMP training/retraining programs in Poland raise women's employment rates over the short and medium term.

3.1.3 China

Bidani, Goh and O'Leary (2002) analyze the effects of a retraining program in China called the "Reemployment Project," which was designed to promote labor market entry of so-called *xiagang*, who are people that were laid off from state-owned-enterprises but remain attached to their former employer for unemployment stipends, health insurance, pensions and some-

times also housing. The Reemployment project was administered by labor bureaus in local areas and included a range of active labor market policies such as job search assistance, counseling, training, wage subsidies, tax incentives for firms and assistance for self-employment. Individuals are allowed to be registered with the Reemployment service for up to three years.

The impact evaluation study was carried out in a city with very high unemployment, Shenyang in northeastern China, and in another city with moderate unemployment, Wuhan, in central China. The laid off workers are 47% female, tend to be less educated and to have ages concentrated in the <35, and 35-46 year-old range. The training intervention was relatively short-term (one-month, 132 hours of classroom training) and included courses in computer training, beauty and massage, hair cutting, sewing, toy making, cooking, repair training and driver education. Class sizes during the training sessions were often large with 200 to 300 workers in a small classroom.

Analysis samples were drawn from a census which required each state-owned enterprise to provide a list of workers laid off at different times. The sample of trainees was selected from the training registers from the training institutes (in the case of Shenyang) and from a master list supplied by the Wuhan Labor Bureau.¹³ Three different treatment/comparison group samples are analyzed using multiple methodologies that include prop score matching, matching on odds-ratios, and OLS. The outcomes of interest are employment and earnings, and, with a few exceptions, most of the impact estimates are robust to the use of different methods. Training is found to have a negative impact in Shenyang on employment probability and no effect on earnings. In Wuhan, however, training is found to have a positive impact on employment probability. The estimated impact on earnings is difficult to ascertain in the case of Wuhan, because the estimates are not robust to changes in the sample used for the analysis and/or estimator used.

Subgroup analyses show, somewhat surprisingly, that training impacts do not differ much by age, marital status, gender, educational attainment and home ownership. In Shenyang,

¹³No-shows were excluded from the treatment group and in some cases included in the control group and cross-overs were excluded from the control group and in some cases added to the treatment group.

training appears to have a stronger negative impact on men and those with lower education. In both Shenyang and Wuhan, individuals who contributed personally to the cost of training had higher reemployment rates. It is expected that individuals who expect to benefit the most from an intervention would be willing to contribute more to its cost.

To understand the reason for the observed discrepancy in impact estimates across the two program sites, it is useful to consider the details of the program's operation. It appears that the Reemployment program may not have been properly implemented in Shenyang, as many workers seemed not to receive the layoff stipends that they were supposed to have received under the program. In Wuhan, the program appears to have been better implemented. Quality of training remains an issue in both sites, though, because it may have been difficult for workers to learn new skills in large, overcrowded classrooms.

3.2 Summary

The studies described above generally find that ALMP programs tried in the context of Argentina, Mexico and Peru have been effective in increasing employment rates and that women are often major beneficiaries of such programs. Some of the programs have been observed to also increase wages and income, but the observed effects on employment tend to be more robust than those on wages. Wage subsidy programs such as the Argentinian *Proemplo* voucher program increased the employment rate of participants on average but did not affect their income levels.¹⁴ Public works programs, of the kind made available by *Trabajar II*, led to job creation for workers and to increased income, with similar estimated gains for men and women. The *Jefes* income support program also increased employment and income and was highly demanded by women. It provided support to poor families during a particularly difficult economic crisis. The Mexican PROBECAT program was successful in augmenting employment of women over age 25 with previous labor market experience, but

¹⁴A standard theoretical job search model (such as Burdett and Mortensen (1998)) would predict that firms would be more willing to hire workers with a voucher subsidy but would not necessarily pay a higher wage, which appears to be borne out in the data. Some worker-firm matches would take place with the voucher that would not otherwise be profitable. For the matches that would take place regardless, whether the worker's wage increases would depend on the bargaining between the firm and the worker.

not in increasing their earnings.

On the whole, the empirical evidence suggests that many of the ALMP programs were effective in increasing employment rates. The evidence on whether the programs also increase wages is more mixed. The pattern of higher employment without higher wages might be expected for two reasons. First, it is difficult to bring about large changes in an individual's earnings capacity with any short-term program intervention. Rather, it seems that many ALMP programs operate by facilitating the worker-firm matching process, for example, by introducing workers to firms through internships. The worker-firm matches sometimes result in lower wages than the worker might have obtained independently from a longer job search. Where jobs are very scarce, for example, during a particularly downturn in the economy as was the case at the time of the introduction of *Jefes* in Argentina, ALMP programs do appear to increase wages and income of program beneficiaries and to alleviate poverty.

The evaluation studies of ALMP programs in transition economies tend to find positive program benefits for women, although the evidence differs somewhat depending on the country context. The main way that ALMP programs alter women's employment outcomes is by increasing their probability of employment and their exit rate from unemployment into jobs. Again, there is less support for an effect of these programs on the level of wages received by the employed. The few studies that have examined longer term effects, such as the Kluge, Lehmann and Schmidt (1999) study, find that the positive program effects for women tend to be sustained over an 18 month time frame. More evidence is needed on the costs of the various programs to allow a study of their cost effectiveness.

Some of the ALMP programs we have reviewed operate on a very large scale, and there remains the question of to what extent people who are not participating in these programs suffer adverse consequences, such as job displacement or lower wages. It is also possible that such large governmental training programs crowd out training that firms or individuals might carry out privately. These questions warrant further examination before any full accounting of the effectiveness of these programs can be made.

Betcherman, Olivas and Dar (2004) provide an overview of the recent international ex-

perience with active labor market programs (ALMPs), focussing on the impacts of ALMPs on the employment and earnings of participants and considering the impacts of ALMPs in developed, developing country and transition economy settings. The evidence is reviewed for seven different ALMP categories: employment services, training for the unemployed, training for workers in mass layoffs, training for youth, wage and employment subsidies, public works, and micro-enterprise development/self-employment assistance. Betcherman et. al. (2004) also review the evidence on the impact of training and job placement programs for the adult unemployed population and conclude that the programs tend to increase employment but not wages. They also emphasize the importance of incorporating the substitution effects towards beneficiary workers and away from other workers into any analysis of program effects, which is not commonly done.

4 Programs to Support Youth Employment

High rates of youth unemployment are a problem in many countries around the world. Even when employed, youth are more likely to hold informal sector jobs, to be employed in minimum wage industries, and to work in jobs with few fringe benefits. The fact that young workers with little labor market experience have a wage that is close to their marginal product also means that they are more likely to be fired during an economic downturn, when their marginal product falls, than other more experienced workers.

An important factor that contributes to high youth unemployment rates is that many youth are in the process of making the transition from school to work, which naturally leads to higher unemployment rates during the job search period when they may be trying out different jobs/occupations in an effort to find a good match. Some extra unemployment and turnover may be positive if its leads to better job matches, but extensive periods of unemployment that are not associated with extra schooling are surely counter-productive.

A variety of programs have been implemented with the goal of reducing youth unemployment. Some programs aim to increase skills and thereby increase their marginal product and the potential wages that the youth can earn. Other programs aim to facilitate job search by

increasing the rate at which they meet potential employers or providing young people with better information about job opportunities. For example, wage subsidy programs lower the wage that employers have to pay a given worker, making it more likely that the worker gets hired. Usually, the hope is that the worker accumulates human capital from working on the job and that the firm will ultimately want to keep the worker after the subsidy expires. Many programs combine both training and job search features.

Betcherman et. al. (2007) review the evidence from 289 interventions in 84 countries for which information was compiled in the World Bank Youth Employment Inventory (YEI). As they note, many of these interventions were not subject to rigorous interventions, so it is hard to draw conclusions about their impact relative to cost. They find that only one in ten interventions has had an evaluation that measured both net impact and cost. The Betcherman et. al. (2007) survey concludes that the majority of the 172 interventions for which employment and earnings impacts on participants were available have positive labor market impacts for participants. However, when the analysis is restricted to the subset of studies that use a control/comparison group methodology and measure "net impacts" relative to the control/comparison group, then there were positive impacts for only 60%. Of those studies where cost effectiveness was measured, 56% were found to be cost effective and 44% were not. In general, the study concluded that youth interventions tend to be more effective in developing countries than in developed countries. The Betcherman et. al. (2007) study also used meta-analysis to examine which program features contributed to a program's effectiveness. The meta-analysis found no statistically significant differences in the impact of the different program types.

4.1 Combined classroom training and internship programs

4.1.1 Latin America

Programa Joven in Argentina: Aedo and Nunez (2004) study the effectiveness of another training program introduced in Argentina called *Programa Joven* that was targeted at low income individuals aged less than 35. The program targeted young people from poor house-

holds with low education levels, little or no working experience, who were either unemployed or inactive. It provided an average of 200 hours of training, a monetary subsidy for females with young children, transportation expenses, medical checkups, books, material and work clothing. The duration of training varied from 14-20 weeks and was divided into a technical knowledge phase, in which participants were taught occupational skills, and an internship phase in which participants completed an eight week internship at a firm. The impact evaluation study is based on two analysis samples: (i) a sample of 139732 so-called "Acreditados" who qualified and registered to take training and who were at different points in their training at the time of answering the survey (some had not started, some were in the technical knowledge phase, some in the internship phase and some had finished or dropped out), and (ii) a subsample of 3340 program beneficiaries and matched comparisons, drawn from the Acreditados sample, who were surveyed at the time of registration and then again one year later.

The evaluation study examines whether the program increased the labor income of trainees and their probability of employment, using a cross-sectional propensity score matching methodology to control for preprogram differences between program participants and nonparticipants.¹⁵ The propensity score model depends on current labor force status of the individual, a poverty measure, sociodemographics, education, marital status, and geographic region, and is estimating separately for four groups: young males, adult males, young females and adult females. A potential drawback of the propensity score model is that there is no historical data available on earnings or employment history at the time of program registration. Recent labor force history is a good predictor of participation in training programs, and lacking such data the predictive power of the propensity score models is not high.

The impact estimates show statistically significant effects of the program on earnings of adult women (age 21-35) and young males (age less than 21) but not for young females or adult males. Statistically significant effects on employment were found only on adult women,

¹⁵An individual was defined as a program beneficiary if he/she has completed the technical knowledge phase, else was designated a nonbeneficiary.

in the range of 9-12 percentage points. Estimates obtained using alternative propensity score models and nearest matching are fairly robust relative to changes in the variables included in the propensity score model, to differences in the source of data used and to variation in the number of neighbors used.¹⁶ The estimated impact for adult women and young males on earnings is around US \$20-\$25 / month. Rate of return estimates, obtained under alternative assumptions on the discount rate, show that the program would have a positive return only if the benefits are fairly long-lasting (9 years or more).

ProJoven program in Peru: Ñopo, Robles and Saavedra (2007) analyze the impact of a Peruvian youth labor training program, called *ProJoven*, on female and male youth living in urban areas. The program provided classroom training and internships lasting about three months for youths from poor families. Trainees receive stipends during their training period, with mothers of young children receiving a double stipend. More than 20,000 youth participated in the program. An interesting feature of the program was that one of its explicit goals was to train female youth for traditionally male occupations so as to reduce gender segregation in the labor force.

Ñopo et. al.'s evaluation of the program is based on a sample of beneficiaries and a sample of matched controls that were selected on the basis of gender, age, geographic proximity, poverty status, income, schooling, numbers of children and employment status. These individuals were administered a baseline survey and three follow-up surveys at 3, 6 and 18 months. The impact estimates are derived from a somewhat unusual two-stage matching procedure that first selects for each treated individual three matched controls on the basis of similarity in hourly wages. The motivation for this two-stage procedure was to closely align the treatment and comparison groups in terms of preprogram earnings so as to account for the so-called "Ashenfelter Dip" problem, namely, that program participants often exhibit a preprogram dip in their earnings that is not seen in comparison group data. The two-stage

¹⁶Recent research has shown that bootstrapping does not lead to valid inference about standard errors for nearest neighbor matching estimators, so the standard errors reported in this study would not be valid. However, correcting them using the alternative standard error estimators suggested in Abadie and Imbens (2005) or using kernel smoothed estimators for which bootstrapping is valid probably would make little difference.

matching procedure generates a similar preprogram earnings dip pattern in both the treatment and comparison groups, although aligning the groups in terms of preprogram earnings would not necessarily guarantee that post-program comparisons are valid.

The outcome measures of interest in the study are labor supply, hourly earnings, monthly earnings, and occupational segregation. Employment impacts for women are found to be greater than for men, with women experiencing positive impacts of 6% at 12 months and 15% at 18 months and men experiencing negative employment impacts. The impacts on hours worked, hourly earnings and monthly earnings, are positive for both women and men. The program has especially pronounced effects on monthly income from the main job. After 18 months, beneficiary females generate 92.88 percent more labor income than their control counterparts, in comparison with an increase of 10.92 percent for males. As a result of participation in ProJoven, the levels of occupational segregation, measured by the Duncan Index, are noticeably lower among program beneficiaries.

Jovenes en Accion program in Colombia Attanasio, Kugler and Meghir (2008) analyze the impact of a training program for disadvantaged youth age 18-25 living in urban areas of Colombia in 2001-2005. The program provided 3 months of classroom training followed by 3 months of on-the-job training in the form of an unpaid internship at a company, with 1009 companies participating. The program was targeted at young people who were unemployed and who came from the two lowest deciles of the income distribution. Training was provided by private training institutions that were paid at market rates for their services, with payment being made when an individual completed the program. The training program provides played in key role in determining what courses would be offered and the design of the courses. The average cost of training was \$ 875 per person. Although the internship was unpaid, the program also provided a stipend of about %2.20 per day to trainees for meals and transportation and % 3.00 per day to women with children to help cover the costs of child care.

The impact evaluation used a randomized experimental design. Training institutions in the seven largest cities were asked to select a larger number of applicants than slots they

could fill and about 2/3 were then accepted into the program and 1/3 randomized out. The average age of the evaluation sample was 21 with an average education level of 10 years. Compliance with the treatment protocol was very high, with almost everyone assigned to training taking and completing the training. Moreover, very few people who were assigned to the control group were able to get training (i.e., few cross-overs). Thus, the program could be evaluated using intention-to-treat estimates that simply compare means for the treatment group to means for the control group. Data were collected at baseline, before the program started, and at follow-up, after the six months training was completed. Although attrition was somewhat high at %18, this is a highly mobile group and attrition was found to be unrelated to the randomized program assignment.

The evaluation found that the program increases earnings and employment for both men and women, with the largest effects observed for women. Women offered training earned on average 18 % more and men earned on average 8 % more. Both men and women offered training were more likely to be employed and to have jobs that offered benefits and had a formal wage contract. The study finds the benefits of on-the-job training to be higher than that of classroom training.

Juventud y Empleo program in Dominican Republic: Card et. al. (2007) analyze the effects of the *Juventud y Empleo* program, which was a training program targeted at less educated youth. The program consisted of up to 350 hours of classroom training followed by an internship at a private sector firm.¹⁷ Participants were not paid during the program, but they received a daily stipend of roughly two dollars per day. As described in Card et. al. (2007) , the program was implemented by first selecting private training institutions through a competitive process that required that they provide written statements from firms that would be willing to hire their trainees for two month internships. This process was intended to ensure that the types of training provided would be relevant to the needs of local employers.

The evaluation of the program is based on a sample of program applicants in 2004 who

¹⁷The program was funded by IADB.

were randomized into treatment and control groups. Of 8391 eligible applicants, 5802 were originally assigned to the treatment group and 2589 to the control group. However, the program providers were told to reassign people from the control group to receive treatment in the event that treatment group members failed to show up for training or dropped out early in the course of training. No-show and drop-out rates were relatively high. 1011 individuals who were originally assigned to treatment did not show up or dropped out and 966 members of the control group were then reassigned to treatment. The "realized" treatment and control groups are potentially no longer comparable, if dropout or reassignment decisions are non-random. Card et. al. (2007) argue that the reassignment process for controls was essentially random, but they do find some differences in the baseline characteristics of the realized treatment and control groups, particularly in the age distribution. They adjust for these differences using a reweighting approach that is equivalent to using a matching estimator and they report unadjusted and adjusted estimates.

Individuals in the evaluation sample were on average age 22.3 years at baseline and 22.8 at follow-up. A comparison of the treatment and control outcomes 10-14 months after training showed no statistically significant effects on employment. 57 % of treated individuals were employed in comparison to 56 % of controls. Disaggregated results by gender, age, and education also show no effects on employment. An examination of monthly earnings and of hours worked per week also yielded statistically insignificant findings. The program did have a marginally significant 10% impact on hourly wages of participants and the jobs of male program participants were more likely to provide health insurance coverage. Overall, though, the authors of the study conclude that the program "had (at best) relatively modest effects," and that "it is unlikely that programs of this nature, operating under similar financial and operational constraints, can fully address the many barriers and problems faced by disadvantaged youths in this region."

4.1.2 Middle East

The Jordan New Opportunities for Women (NOW) program is a pilot program, sponsored by the World Bank, that aims to evaluate the effectiveness of short-term wage subsidies and skills training. The program is targeted at young female community college graduates and aims to help them find work. In 2009, Jordan was ranked by the World Economic Forum 122 out of 134 countries in gender equality in economic participation and opportunity. The female labor force participation rate 23 percent, ranking is 177 out of 185 countries. The female youth unemployment rate is 38 percent in comparison to 17 percent for male youth. Among recent community college graduates, though, 93 percent of women express an interest in working after they graduate and 91 percent express an interest in working after they are married.

Through the NOW program, 600 young women were selected either to receive 45 hours of employability training (to improve their communication and basic job skills) or to receive job vouchers, which provide a short-term incentive for firms to take a chance on hiring them. A subset were selected to receive both and a control group received neither. Each voucher has a face value of 150 Jordanian Dinar (US\$210) per month for a maximum of six months, and is transferred to the employer upon verification of employment contract and salary.

Although the NOW project is still in its early stages, initial results are encouraging.¹⁸ Employment rates among graduates who received vouchers alone or vouchers plus training are 55-57 percent in comparison to 17-19 percent in the group that only received training or received neither training nor vouchers.

4.2 Summary

We reviewed impact findings for two programs that are targeted at youth or young adults. The *Programa Joven* program, which is both a job training and an internship program, led to statistically significant impacts on earnings for adult woman and younger males but not for

¹⁸Preliminary evaluation results are discussed in the World Bank Gender Equality and Development Report, 2012.

adult men or younger females. The impacts were relatively small in magnitude, though. The Peruvian *ProJoven* program had fairly large impacts on employment and hours worked of women as well as modest positive impacts on earnings. It also affected women's occupational choices and decreased gender segregation, through the focus on training women to work in traditionally male dominated occupations.

Betcherman et. al. (2004) review five studies of youth training programs in developing countries (Jovenes programs in Argentina, Chile, Peru, and Uruguay), which include many of the studies described above. All of the evaluations found positive employment impacts and two of the three computing earning effects also found positive impacts. The positive impacts found for developing countries contrast with the mainly negative estimated impacts for youth oriented ALMP programs in developed and transition economies. Betcherman et. al. (2004) argue that for youth oriented programs to be effective, they need to offer a comprehensive set of services that include basic education, employment services and social services.

The recent evidence reported in the Attanasio, Kugler, and Meghir (2008) study of the Jovenes en Accion program in Colombia generally supports the conclusion that youth training is effective in increasing earnings and employment. The evaluation is based on a high quality randomized experiment. The job training component is found to be more important than the classroom training component. However, the evidence from the Card et. al. (2007) study, is less encouraging about the effects of youth training programs, at least in the setting of the Dominican Republic. That study was also based on a randomized experiment and found very little effect of the program on youth employment and hours worked (female or male), although some positive impacts were noted on wage levels and on whether the job held by the youth offered benefits. There was high attrition that may have to some extent compromised the experimental design.

The NOW program in Jordan is unusual in being targeted at women with community college degrees in an Arab setting. That program documented that these well educated women express a very strong desire to work but face very limited work opportunities, in

large part due to cultural practices of firms not filling positions with women. Although the evidence is preliminary, it appears that the job vouchers, which gave firms monetary incentives to hire these women at least for six months, were effective in helping them gain employment. The vouchers may have also led to different search behavior for the women if it gave them more confidence in searching for a job.

Only a few of the evaluation studies we reviewed here carried out a rigorous cost-effectiveness analysis. One study was Aedo and Nunez (2004), which found that program impacts have to be sustained over 9 years or more for the program to be cost-effective. Evaluation studies do not typically follow individuals for such long periods of time, so whether benefits can be sustained over such a long time horizon is unclear.

In evaluating the effectiveness of any youth oriented training program, an important question to consider is whether the program funds might be better spent on alternative programs that keep youth from dropping out of school. Formal schooling, especially secondary school, is often found to have a relatively high wage return of 10% or more in developing countries.

5 Child care programs

The next section reviews the results of some evaluation studies of programs designed to facilitate women's work, either by providing better access to affordable and reliable child care, by providing child care subsidies or by providing publicly provided child care. The accessibility of affordable and reliable daycare is a potentially important determinant of women's LFP decisions. Formal sector jobs, such as factory work, often require long hours of work and do not easily accommodate the presence of children. Informal sector jobs, such as making things at home or caring for other mothers' children in your home, typically pay less and are less likely to include health care benefits, but may be more flexible in terms of hours and allowing mothers to supervise children while working. Public provision of child care services is relatively rare in developing countries, where mothers are often expected to either stay home with their children or make private arrangements with extended family

members or domestic workers. Poor families frequently cope with childcare needs by leaving their younger children with older siblings, especially girls, which in some countries greatly affects the educational opportunities for girls. For example, in Indonesia, 40 % of working women care for their children while working, 10% leave them with older girls, and 37% leave them with female relatives.¹⁹ In Nairobi, 54% of poor mothers bring their children to work with them, whereas 85% of mothers of higher socioeconomic status used the services of “house-girls.”²⁰ The availability of child care affects the decision about whether to work, whether the mother engages in formal market work, the early childhood environment for her young children (e.g. whether young children are exposed to hazardous substances or are engaged in learning activities), and the schooling opportunities and employment prospects of older siblings. The proximity of child-care can also affect decisions such as how often to feed younger children.

In recognition of the importance of daycare to a mother’s working decision, some countries have introduced mandates or programs aimed at increasing the supply and lowering the costs of daycare. For example, In Brazil, Chile and India, there are mandates that employers with a sufficient number of female workers offer workplace childcare (sometimes the mandate only applies, though, for breastfeeding mothers). Although employer mandates can increase the supply or childcare programs, they can also potentially reduce the incentives to hire female workers and result in discrimination against female workers. Other countries have introduced social programs, such as community daycare programs, which have been implemented in Peru, Colombia, Bolivia, and Venezuela as well as in most Central American countries. These programs aim to (i) foster human capital accumulation of children through better nutrition, better hygiene and through activities that promote child development and socialization and (ii) facilitate mothers’ working outside the home and at higher wage jobs. The largest program, *Hogares Comunitarios* in Colombia has nearly one million beneficiaries.

Hein and Cassirer (2010) provide a nice overview of childcare provisions and services in

¹⁹Kamerman (2000)

²⁰Lakati et. al. (2002)

Brazil, Chile, France, Hungary, India, Kenya, South Africa, Thailand, the UK and the US. Many childcare programs are simply evaluated by asking people who use the services how they benefit from them, without making systematic use of comparison groups of "untreated" people. Although this kind of evidence helpful in determining which kinds of features are most valued, it does not allow quantification of the benefits of people having access to the service verses not having access to it and is not informative about workers who do not take up the services.

Economic theory would suggest that the effect of a child care program is theoretically ambiguous. A program that subsidizes child care, has both an income and a substitution effect on women's labor supply. If the mother's wage stays the same, then the income effect would be expected to reduce the number of hours she works, assuming that leisure is a normal good. However, if the mother now gets higher wage offers, then she might increase her hours worked. The substitution effect would increase consumption of the subsidized good, i.e. increased use of child care and increased hours worked. Because subsidizing child care lowers the cost of having children, there is also a possible concern that such programs might increase fertility.

5.0.1 Latin America

Community daycare programs in Guatemala: Ruel et. al. (2006) present the results of an impact evaluation of Guatemala's Community Day Care Program. The evaluation focussed on Guatemala City, although the program is available throughout the country in both rural and urban areas. Under the program, a group of parents selects a women from the community to serve as an in-home daycare provider for up to 10 children (under the age of seven), Monday through Friday from 6am to 6pm. As part of the program, children are involved in developmental activities and receive food and snacks. The caretaker typically receives furniture, educational materials and money for compensation and for the children's food. The program is designed to provide 80% of the children's nutrition and 40% of the program cost goes to food. In addition to the compensation the caretaker receives from the program

providers, families are expected to make some supplemental contributions. In Guatemala, only 3% of eligible families participated in the program, in part because spaces were limited.

The impact evaluation found that the program improved children's diets and increased mother's income. The impact evaluation was carried out using the method of matching, comparing a cross-section of beneficiaries to a matched set of control households with children in the age 2-5 range. The controls were selected by matching beneficiary children to other children from the same neighborhood, of a similar age and gender and whose mothers were working. Impact estimates were obtained by propensity score matching.

Because all mothers had to be working to be a part of the evaluation, it is impossible to estimate the effect of the program on mother's rate of working. However, to get an idea of the potential effect, Ruel and Quisumbing (2006) compare the labor force participation rate for a random sample of mothers who have children aged newborn to 6 years with that of women who do not. The comparison suggests that the program could potentially increase labor force participation by 25 percent.

The program's impact on mother's earnings and job choice was assessed by comparing beneficiary mothers to matched control mothers. This comparison indicates that beneficiary mothers had 30% higher earnings than mothers using alternative child care arrangements and were more likely to be employed in the formal sector. The program had the largest benefits for younger and older women with low levels of education. A limitation of the analysis is that all mothers in the control group were restricted to be working. It is likely that some of the mothers in the treated group would not have been working had the program not been available to them. For this reason, the evaluation probably understates the effect of the program on mother's income. The results from interviews with focus groups show that the program was very well received and much appreciated by beneficiaries.

A comparison of Guatemala and Ghana: Quisumbing, Hallman and Ruel (2003) analyze the determinants of mothers' joint labor supply and daycare utilization choices in Guatemala City and in Accra, Ghana. The analysis for Guatemala is based on a random sample of mothers with preschool age children from one zone of Guatemala City. The analysis for

Accra is based on a household survey.

The urban settings of Guatemala and Ghana are quite different with regard to the type of work that women typically do. In Guatemala City, the landscape is becoming increasingly urbanized and women often engage in formal work, such as factory work. About one fifth of households are headed by single women, about half of which are poor or indigent. In Accra, 71.9% of female employment is in the informal sector. For cultural reasons, there is a high percentage of female headed households - 35.1%.²¹

Quisumbing et. al. (2003) estimate a model of mother's working and childcare decisions for these two different samples of women. From their analysis, they conclude that the supply of daycare is an important influence on mother's working decisions only in Guatemala City, where women are more likely to work in the formal sector. In Accra, the supply measures, such as the distance to the nearest formal daycare provider, have no effect on labor supply choices. The findings suggest that enhancing formal daycare options for women is likely to be most effective in increasing female labor supply where a large percentage of women work in formal sector jobs.

Hogares Comunitarios in Colombia: Attanasio and Vera-Hernandez (2004) analyze the impacts of a large community daycare program in rural Colombia on children's nutrition, female labor supply and on school achievement. The operational aspects of the program are similar to the Guatemalan program described above, except that caretaker mothers can have up to 15 children and food was delivered weekly to their house. The *Hogares Comunitarios (HC)* program is the largest welfare program in Colombia. It is targeted at poor households, as measured by an eligibility index. As noted above, the program has extensive coverage, but there are still many children not participating that can serve as a comparison group for the purpose of evaluating the impacts of the program.

In evaluating the impact of the HC program both on mother's labor supply and on child outcomes, Attanasio and Vera-Hernandez (2004) compare beneficiary families with nonbeneficiary families. They argue that the use of cross-sectional matching on observables

²¹Men and women from the indigenous Ga population traditionally live in separate houses after marriage.

methods would be inappropriate, because they believe the participation decision to be based in part on unobserved attributes. Indeed, when they apply propensity score matching, they get negative estimated impacts of the program. To allow program selectivity to be based in part on unobservables, they implement an instrumental variables estimator. To do so, they require a variable that influences the program participation decision but has no direct influence on the outcomes. They maintain that distance of the household to the nearest HC and distance averaged at the community level can serve as instruments. This requires an assumption that households are not choosing their location with regard to the location of the HC, but Attanasio and Vera-Hernandez (2004) present evidence that the location of the nearest HC frequently changes, making this assumption more likely to be satisfied.

The IV estimates indicate that the program has extremely large, positive impacts on female employment and hours worked as well as beneficial impacts on child outcomes. The probability of female employment increases from 0.12 to 0.37 and the number of hours worked increases by 75 hours per month. The study also finds statistically significant effects of the program on children's height and also, over the longer-term, on school going and grade achievement.

Preschool building program in Argentina Berlinski and Galiani (2007) analyze the impact of a large pre-primary school building program in Argentina on pre-primary school attendance and maternal labor supply. The program was mainly targeted at middle income households living in urban areas. It attempted to compensate for geographic differences in the availability of preschool facilities by rolling out the program first in areas with the lowest level of facilities. Between 1994 and 2000, the program created about 175,000 places, which expanded the number of places available at baseline by 18%, with different regions and different cohorts were differentially exposed to the program. Berlinsky and Galiani (2007) use a difference-in-difference approach to evaluate the effects of the program on material employment and weekly hours worked, exploiting differences across regions in the numbers of facilities built and in exposure of cohorts due to program timing. The study finds evidence of full take-up of new facilities and also that the likelihood of maternal employment

increases between 7 and 14 percentage points. The effect of the program on weekly hours is not precisely estimated.

Publicly provided child care in Brazil: Brazilian women have rapidly increased their labor force participation rate over the last decade, from 41.3% in 1990 to 64% in 2007. Their rate of participation is higher than other Latin American countries at comparable income levels and is just below the OECD average of 65 percent. Despite these gains, however, there is still the concern that women are more likely than men to be employed in the informal sector and that unemployment rates among women are higher than for men. 22 percent of women age 15-24 are unemployed in comparison to 13 percent of young men.

The rapid increase in female labor force participation has been accompanied by an increase in the demand for child care services. Barros et. al. (2011) analyze the impact of a child care program that gives mothers living in low income neighborhoods access to publicly provided child care in the municipality of Rio de Janeiro. The program is a early childhood development program targeted at children age 0-3 that provides full time daycare services, health services, food, and instructional toys and material for the children. The program encourages parental involvement.

Barros et. al. (2011) evaluate the causal impact of the program on (a) the female labor force participation ratio, (b) on net costs of the policy, taking into account extra tax revenue from newly employed women, and (c) on whether publicly provided care crowds out private care. They use a special dataset collected in 2008 for the purpose of evaluating the child care program, which provides free services to families living in poor neighborhoods. Their analysis makes use of the fact that the initial assignment of 1000 children to receive the program was selected by lottery from an applicant pool of 24,000. The eligible children who were not selected were placed on a waiting list. The study analysis data on 4348 children, half who ere lottery winners and half who were on the waiting list.

Barros et. al. (2011) use an intent-to-treat estimator and also a LATE estimator, to take into account that compliance with the treatment assignment was not perfect. 94 percent of the lottery winners enrolled in day care centers and 51 percent of control children enrolled

as well. From the intent-to-treat analysis, Barros et. al. (2011) find that providing access to publicly provided child care services led to a large increase in the use of care, from 51 to 94 percent. It also increased mothers' employment from 36 to 46 percent, with an even larger increase in the subgroup of mothers who were not working in the six months before the lottery for whom employment rates increased by 97 percent. However, the program did not significantly increase hours worked among mothers who were working before the program. Also, although the rise in mother's employment is associated with an increase in household incomes, from R\$569 to R\$661 (an increase of 16 percent), the cost of providing the child care was approximately R\$250. Thus, the increase in income was less than the cost of child care, raising questions as to whether the program was cost effective. In addition, mothers who received the publicly provided care were much less likely to make use of private care services than the comparison group, so the publicly provided care seems to crowd out private childcare services. Barros et. al. (2001) conclude that the program was effective in improving labor market outcomes for women, but the use of alternative programs such as subsidizing the use of private child care through vouchers should be investigated as a potentially more cost effective means of achieving similar gains.

5.0.2 Middle East

Free public preschool for children aged 3 and 4 in Israel

Women in North Africa have relatively low LFP rates. Factors that likely contribute to low rates are low levels of education, high fertility rates, limited employment opportunities, religious prohibitions on certain kinds of interactions with men and lack of child care services. There has been relatively little research on whether LFP rates of women in North Africa might be increased through child care policies. An interesting recent study by Schlosser (2011) analyzes the effect of the Compulsory and Free Preschool Law for Ages 3 and 4, which provided free preschool to all Israeli children age 3 and 4 on children's enrollment and mothers' labor supply among Israeli Arabs. The law was introduced in 1999 but was only gradually implemented.

For Israeli Arab women aged 25-54, the LFP rate in 2001 was 22 percent, in comparison to a LFP rate of 78 for same age Jewish women. Schlosser (2011) notes that the LFP rate of Israeli Arab women is similar to that of the neighboring Arab populations. The LFP rate for the same time period was 17 percent in Jordan, 26 percent in Lebanon, 17 percent in the West Bank and Gaza Strip, 20 percent in Syria and 24 percent in Egypt (ILO, 2002). The low LFP rates for Israeli Arab women are accompanied by low rates of preschool enrollment.

The Preschool Law, which provides free preschool education for all children aged 3 and 4, was introduced in different towns at different times and, even in 2011, many towns had not yet been incorporated. The preschool was made available first in towns of lower average socioeconomic status and greatly enhanced access of the Arab population to publicly funded preschool education. The enrollment rates reached 83 percent in 2003 in towns that had access to the free preschools, in comparison to around 30 percent for towns in which the program had not yet been introduced.

Schlosser (2011) evaluates the program using a difference-in-difference strategy that compares the change in preschool enrollment, mother's labor supply and fertility in towns where the policy had been introduced to that in towns that had not yet begun to supply free preschool. The study focuses on 11 treated towns and 13 untreated towns that are identified in the Israeli Labor Force Surveys.²² According to population estimates provided by the Israeli Central Bureau of Statistics, in 2003 there were 12,956 children aged 3-4 living in the treated towns and 12,656 children in the comparison towns. Preschool enrollment rates were similar in the treated and untreated towns prior to the program, despite the treated towns being more disadvantaged on average. However, mothers in treated towns had on average significantly lower levels of education and come from lower socioeconomic backgrounds, indicating the importance of controlling for preexisting differences in the analysis.

Empirical estimates show that preschool enrollment and mothers labor supply both increase sharply after the introduction of the new policy in the treated towns, with the increase in labor supply occurring mainly among more educated mothers. Schlosser's preferred es-

²²Town codes are not available for towns with less than 10000 inhabitants.

timate (based on the most comprehensive model) suggests that free preschool provision increased the labor force participation of mothers of children aged 2-4 by 0.071 (s.e.=0.034) percentage points, which is reasonably large considering that the labor force participation of this group in the post-law period is only 17.1 percent. The impact of the free preschool provision among mothers whose youngest child is 2-4 years old is even larger, 0.117 (s.e.=0.035). The fertility of these mothers does not change, at least over the time period of the evaluation. There is no evidence of any change in the untreated towns. Schlosser (2011) also provides evidence that the treatment effects grow over time and become significant only after 3 years of the program, suggesting that it takes time for the changes in female labor supply to occur.

5.0.3 Asia

The SEWA Program in India: The Self-Employed Women's Association (SEWA) is a trade-union with members in seven Indian states, who tend to be poor working women working in the informal economy. SEWA, which started in 1972, makes a variety of services available to these women, including health care, child care, banking and insurance.

The Balwadi Study assesses the impact of the SEWA Child Care Centers on the lives of the children and their family. The child care centers are open to all SEWA members and provide day care and nutrition for children aged up to 5 years old. The evaluation study examines impacts both on the child and on the mother by comparing outcomes of children that participated in the program to those of children who had never participated and by soliciting feedback from the mothers about their satisfaction with the child care centers.

The study of outcomes looks at the development of children who attended the centers in two districts about 10 years ago on the children's lives. Children between the ages of 11-17 who attended the center were sampled along with their mothers (of any age). The control group was selected from the same geographic areas from children who never attended the SEWA and had similar socioeconomic backgrounds. It is not clear what methods were used to control for differences in observables, raising concerns that people who participate and do not participate may not be comparable. Also, migration poses a potential threat as it

was not clear that the study followed families who no longer lived in the area. The cost of participating in SEWA and of the child care components of the program were also not mentioned.

Mothers reported that using the child care center improved their lives improved. About 42-53% reported that their productivity increased, that their number of days of employed increased, that they had to spend less money on children's food/snacks and medicine for illnesses and that their family savings increased. It is impossible to attribute these benefits all to the child-care program, though, because of the extensive other services that SEWA also provides. The mothers generally expressed high rates of satisfaction with only 5-7% saying they would not recommend the child care services to others.

5.1 Child-care subsidies

5.1.1 Transition Economies

In Communist economies, such as the former economies of Russia and Romania, almost all women participated in the labor force. The high rates of participation were feasible in part because of the wide availability of government sponsored child care centers (nurseries, preschool, kindergartens and after school programs). The transition to a market economy has diminished the availability and increased the cost of child care in many of these countries. Fong and Lokshin (2000) analyze how mother's demand for paid child care, mother's labor force participation and working hours respond to changes in the cost of care and to changes in wage offers in Romania. Between 1989 and 1995, Romania saw a sharp decline in public funding for child care services. Over the same time period, legislation was passed that provided mothers with 65% of their previous salary if they cared for their own child during the first year.

The approach taken by Fong and Lockshin (2000) to understand the effects of these and other policy interventions is to jointly model households' decisions about child care and mothers' labor supply. To this end they develop an economic model of household decision making about consumption of child care quality, of market goods and of leisure. Child

care arrangements are classified into six categories defined by combinations of the mother's employment status, mode of care (formal and informal), and employment status of other household members. The model assumes that households pay a flat fee for child care services, where the fees charged by kindergartens are a function of the quality of care provided and the total level of child care prices within the locality.²³

The empirical model consists of a discrete choice equation for the child care mode and mother's labor supply, an equation for a mother's hours at work, and an equation for children's hours in paid care. The effects of unobserved variables is incorporated by imposing a factor structure on the error components of the model. The estimation is based on household survey data from the Romanian Child Care and Employment Survey (RCCES) linked with data on child care providers from the Romania Child Care Facilities Survey (RCCFS). Both datasets were collected by the World Bank in the same geographical areas during the same time period. The surveys were designed to allow matching data on child care fees and child care quality with the households surveyed in the communities. After the model parameters are estimated, the model is used to simulate responses to changes in the policy environment, for example, to changes in the price of child care and the level of mother's wages. The model is used to extrapolate to policy variation that is outside the range of that in the data.

Model simulations indicate that a 10% increase in mother's wage offers increases the rate of mother's labor force participation by 10.9% and increases her use of formal care arrangements by 4.3%. Changes in the price of child care have a smaller effect on the level of maternal employment and on the use of formal care; an increase in the price of care by 10 percent leads to a 1.2 percent decline in the number of working mothers and a 2 percent decrease in the number of households that use formal care. A policy that fully subsidizes formally provided child care, however, would increase the rate of women's labor force participation by as much as 12-15 percent. The elasticities of mother's labor supply with respect to child care cost are found to be -0.17, which is in line with estimated elasticities

²³In actuality, households pay fees that depend on their total household income and the number of children enrolled. However, to avoid complications in estimation, Fong and Lokshin (2000) treat child care fees as exogenous to the households, using regional variation in fees as the main source of identifying information.

reported in the related literature based on data from the U.S. and Canada.²⁴ Fong and Lokshin (2000) conclude that government subsidies for child care are an effective means of increasing the number of mothers who work, increasing the incomes of poor households and lifting some families out of poverty, but that the effects of such policies are less significant for the poorest households.

A similar study by Lockshin (1999) studies mother's participation in the labor force, working hours and demand for child care in Russia. In the 1980s, most women in Russia worked and the government heavily subsidized child care programs that were widely available. A decline in GDP in the 1990s led to a sharp decrease in the availability of state-run child care facilities and an increase in the cost of sending children to these facilities. As described in Lockshin (1999), Russia moved from a country in which child care was provided by the government and almost all households with children had access to affordable or free child care to one in which few households have access and the cost of day care significantly affects labor force participation decisions. Lockshin (1999) builds a static utility maximizing model of household's decisions about labor force participation, working hours, and choice of child care mode to motivate an econometric model that he uses to assess the effects of three different kinds of policy interventions: family allowances, child care cost subsidies and wage subsidies. The model is estimated using panel data from the Russian Longitudinal Monitoring Survey (RLMS).

The simulations show that child care subsidies increase maternal employment by almost twice as much as comparable wage subsidies. Also, child care subsidies are more effective than wage subsidies or family allowance transfers (transfers to families with children) in increasing family income. Child care subsidies increase the amount of time working women spend at work and increase the proportion of mothers who choose to work.

A limitation of the analysis is that it does not examine changes in utility of households under the alternative types of policies. Increasing mother's work would likely reduce her

²⁴Cleveland, Gunderson, and Hyatt (1996) for Canada, Connelly (1992) and Blau and Robins (1988) for the U.S.

time spent in leisure, which leads to a loss in utility that offsets the benefit of higher income. It would be useful to compare the alternative policies on a utility basis. Also, in both the Lockshin (1999) and Fong and Lockhin (2000) studies, it is not clear how the quality of formal child care provided compares to quality of home or informal care, leaving open the question of how children's development is affected by policies that encourage the use of more formal care.

5.2 Summary

The studies of Latin American community day care programs described above generally find the programs to have substantial positive impacts on women's propensity to work, on their numbers of hours worked, and on the health, nutrition, and educational outcomes of children. Assessing the benefit-cost ratio of these programs requires an assessment of the monetary value of the child outcomes. One assessment for a community daycare program in Bolivia that is reported in Behrman, Cheng and Todd (2004) finds benefit-cost ratios of roughly 2:1 (for a discount rate of 5%), only taking into account the benefits accruing to the child participants and not taking into account any benefits for the mothers. Some additional support for the effectiveness of these programs comes from the evidence that the take-up for them is high.

Berlinski and Galiani's (2007) study of a preschool building program in Argentina also found full take-up of the new places for children created by the program. They show that simply expanding the availability of preschool led to an increase in women's labor force participation.

Formerly communist countries have undergone large changes in the pricing and availability of child care, which provides valuable information on how women's labor supply and hours worked responds to price changes. The modeling frameworks of Lokshin (1999) and Fong and Lokshin (2000) permit comparisons of the relative effectiveness of wage subsidy programs, child care subsidy programs and income transfer programs in affecting female labor supply. That analysis finds that child care subsidies are more cost effective than comparable

wage subsidies or family allowance transfers in increasing the proportion of mothers working and the amount of time they spent at work. Unfortunately, not that much is known about whether and to what extent children in the context of these transition economies (Russia, Romania) benefit from being in formal daycare rather than other forms of care (e.g. home care or relative provided care). More evidence needs to be accumulated to better understand how these subsidy programs affect both the mothers and their children.

Both Barros et. al. (2011) and Schlosser (2011) analyze the effect of free publicly provided daycare for preschool age children. Barros et. al. (2011) is the strongest of the evaluation designs, because participation in the program was in part determined by lottery assignment. Both studies provide evidence that providing free preschool significantly increases women's LFP, with larger increases found among Brazilian women than among Arab Israeli women. Barros et. al. (2011) notes, however, that the average increase in household income for the women is less than the cost of the preschool and that the publicly provided program tends to crowd out demand for private preschool. For this reason, he suggests looking at vouchers as a potentially more cost effective way of achieving similar gains at lesser cost.

6 How elder-care demands affect women's work

Many countries around the world are experiencing population aging with increasing time demands for elder care that typically fall on women. Concerns about the effects of providing such care on employment has prompted legislation such as the *The Family and Medical Leave Act*, which was passed in the US in 1993 and allows workers to take time off from work to care for relatives without fear of losing their jobs..²⁵ In the U.S., there has been a trend towards older individuals living longer on their own and then living in old age facilities, that are typically covered by Medicaid after the individual's assets run down..²⁶ In developing countries, however, the private market for elder care is often less well established and it is common for older age individuals to live at home with a daughter or daughter-in-law.

²⁵Workers can take up to 12 workweeks of unpaid leave in any 12 month period.

²⁶Although long-term care insurance is available, only a small fraction of individuals buy it.

A number of research studies have documented a negative correlation between caring for the elderly and female labor force participation and hours worked, but the evidence is mixed on whether providing care actually causes women to have lower labor force participation and lower hours work. An alternative to providing own care is market care, and it is possible that the need to pay for market care could increase female labor force participation.

It is difficult to measure a “causal effect” of caregiving on labor market outcomes, in part because the need to care for parents is never randomly allocated and any empirical analysis of the effect of caregiving on female labor supply has to rely on observational data. Early research in this area assumed that the decision to provide informal care for the elderly did not depend on the same factors affecting labor supply, so that caregiving could be taken as exogenously given. However, this assumption is unlikely to be tenable as the decision to provide care likely depends on unobservable factors that also influence the decision to work. Subsequent work developed ways of addressing the endogeneity of caregiving. Wolf and Soldo (1994) instead estimate a simultaneous equations model of employment, hours work, and elderly parent care provision using US data. They do not find evidence of a reduced propensity to be employed or of reduced hours due to provision of parent care.

A recent paper by Skira (2012) is the first paper to examine effects of caregiving on current and future labor force participation, at both the intensive and extensive margins, and on wages within a dynamic discrete choice modeling framework. The model is estimated on US data from the Health and Retirement (HRS) study. Skira (2012) finds that women face low probabilities of returning to work or of increasing work hours after a caregiving spell. A major factor inhibiting work is that the chances of receiving a wage offer fall greatly after a spell of nonemployment. Skira (2012) uses the estimated model to simulate the caregiving, employment, and welfare effects of a longer unpaid work leave than currently available under the Family and Medical Leave Act, of a paid leave, and of a caregiver allowance. She finds that longer term leaves, either paid or unpaid, allow women to return to their previous jobs and lead to higher rates of women returning to work. On the other hand, a payment given to caretakers leads to greater provision of elder care but to lower rates of working.

Lilly, Laporte and Coyte (2007) note that a number of studies indicate that women with young children at home are likely to provide elder care, possibly because their opportunity costs of providing such care is lower. This suggests that child care and elder care decisions may be linked, although there are no studies as of yet that incorporate both decisions within one framework.

The literature on the effect of elder-care on labor supply of women in developing and transition economies is in its infancy. We review the results of a few studies below. In Latin America, where there are strong cultural expectations that women be main or sole caregivers for elderly parents, the empirical studies of the association between caregiving and female labor supply have yielded somewhat mixed results depending on the empirical methodology used.

6.1 Latin America

Chile: Bravo and Puentes (2012) analyze female labor force participation and care-giving in Chile, where 8.9% of women provide informal care for elderly, chronically ill and disabled people living at home. Using an instrumental variables approach, Bravo and Puentes (2012) estimate the causal effect of care-giving on female labor supply, where the instrument used is the number of siblings present in the household when the individual was growing up (at age 15), assumed to affect care-giving but to not be directly affect labor supply. Like Soldo and Wolf (1994), Bravo and Puentes (2010) do not find a significant effect of elder-care provision on labor force participation after controlling for the endogeneity of care provision.

Mexico: Mroz et. al. (2012) analyze the effect of elderly parents on adult women's labor supply and labor income, within the framework of an economic model of labor supply, caregiving and goods consumption. They develop conditional demand functions for leisure and caregiving time as a function of wages, prices, income and health shocks that can affect the marginal utility of leisure and caregiving time. For non-workers, for whom wages are unobserved, they substitute determinants of wages, so their equation is a reduced form conditional demand function. Mroz et. al. (2012) argue that estimating standard labor

supply models that do not condition on the simultaneous choice of caring for an elderly parent provide better estimates than those based on iv approaches with potentially problematic instruments.

Their analysis is based on data from the 2002 and 2005 waves of the Mexican Family Life Survey. In their sample of women age 21-64, women who provide care for family members (elderly or children or sick adults), tend to be younger, married with more children, have lower job skills, have higher levels of education, and to express a higher perceived pressure to allocate their time between caregiver and other activities than women who do not provide care. The empirical analysis examines the association between parents' health and parents' presence on women's time spent working and time spent caregiving. The reduced form conditional demand functions lead to probit and linear probability fixed effects models. The estimates indicate that the presence of elderly parents significantly affects a women's decisions to be a caregiver, to work and how much income she earns. The estimates imply a 3 percentage point increase in the probability of caregiving for each parent having a chronic condition when no siblings are in the household. Each additional sibling reduces the propensity for a woman to be a caregiver by 0.4 percentage points when the mother has a chronic illness and by a full percentage point when the father has a chronic illness. The estimates also show that the presence of elderly parents reduces the probability that a woman works. Having a mother with a chronic condition reduces it by 4 to 7 percentage points, and having a father with a chronic condition also reduces it, but by about half the magnitude. Overall, the Mroz et. al. (2012) concludes that the presence of elderly parents more so than their health status leads to lower rates of female labor force participation.

6.2 China

China historically had very high rates of female LFP at the time when the state guaranteed workers a job and all individuals were expected to work. Mauer-Fazio et. al. (2011) observe that female LFP was 86.7% in 1982. By 2000, after the reform of the labor system when the state no longer assigned workers to jobs, the LFP of married, urban, prime-aged,

nonimmigrant women had fallen to 77.3%. Most of the decline occurred from 1990-2000. Maurer-Fazio et. al. (2011) examine the determinants of female LFP decisions in urban China and how these decisions have been affected by China's changing economic and demographic environment.

Using census data on married, urban women, they estimate a model of the probability that a woman participates in the labor market that depends in part on whether the household includes a parent age 75 or older (a parent, parent-in-law or other person). In China, coresidence with elders is common and does not necessary imply a need for elder care. Maurer-Fazio et. al. (2011) note that the percentage of elderly men and women living with their children was 60 percent and 69 percent in 2000. The analysis finds that the presence of preschool age children has a strong negative effect on women's labor force participation, but coresidence with older household members increases her likelihood of working.²⁷

7 Other types of programs

This section describes other types of interventions that have been found to influence women's labor market outcomes. These include workplace health programs, gender equity programs, and land titling programs.

7.1 Workplace health programs in Bangladesh

Many women in developing countries have poor access to health care, which can affect their capacity for employment and their productivity. Providing workers with better health care services can be effective not only in helping them cope with illness, but also in providing better access to family planning services and medications and in educating workers about job safety, such as how to minimize exposure to dyes or solvents in factory settings.

Extending Service Delivery (ESD) Project in Bangladesh: The ESD project, funded by US Aid, was a demonstration project examining of the cost-effectiveness of providing work

²⁷Marenzi and Pagani (2005), using Italian data, also report that adult children often benefit from childcare services provided by parents and parents-in-law.

site health services. The program was designed, in part, to address unmet need for family planning and to increase the use of reproductive health/family planning services. The program operated in a garment factory in Chittagong, Bangladesh in 2005. The garment industry has problems of high staff turnover, high absenteeism, poor working conditions, and depends mostly on young female workers (age 16-34) for its workforce. Profit margins are typically low, as the factories in Bangladesh face intense competition from China. The female workers are often in their prime reproductive years and factory work, that may involve the use of harsh chemicals, has the potential to place additional health risks on these women.

The on-site health services provided by *ESD* consisted of a health clinic with a doctor, a nurse-counselor, and an attendant. The clinic operated three hours per day, one day per week. It served about 15-20 patients per day at a cost of about \$47 per week plus the cost of the space. Visits to the clinic required a co-pay of 5 taka (about 7 cents) (about the cost of a cup of tea or two bananas) for which the patients also received medicine. The clinic treated general health conditions (e.g. asthma, diarrhea, stomach ailments), and also offered family planning counseling, reproductive health services, antenatal and prenatal counseling, and HIV counseling.

An evaluation of ESD reported in Chowdhury et. al. (2007) studied the costs and benefits of the program. The evaluation was based on a cross-sectional survey that was administered to a random half of the factory workers on payroll in the survey month and random half of the managers and on administrative data on absenteeism. Qualitative data were also collected through focus groups with workers who had and had not used the on-site clinic. The study compared work days missing in 2004 with work days missed in 2005 to determine whether the ESD program lead to a change in absenteeism. It also compared the number of new recruits between Jan-Dec 2004 to new recruits in Jan-June 2006 to determine if the program influenced turnover. The main threats to the validity of this before-after study design are that there may be other factors unrelated to the program that can lead to major year to year fluctuations in absenteeism or turnover (such as general strikes, outbreaks of

flu etc.)

The evaluation found the following benefits: (i) 11 percent fewer days lost to absenteeism in the first year of the program and a 18 percent decline in the first 18 months, (ii) A 43 percent decrease in staff turnover (from 40 per month to 23) in the first year of the program and a 46 percent decrease in the first 18 months. A comparison of the cost savings from reduced turnover and absenteeism to cost of the program found an impressive benefit cost ratio of 2.4:1 in the first year and 3:1 over 18 months. These estimates probably understate the benefits, as they do not account for changes in the worker's well-being. Also, if workers value the on-the-job amenities, they may be willing to work at lower wages at that job, leading to further cost savings. Whether the program had any effect on wages was not examined.

Focus group interviews indicated that female workers highly valued the clinic services. It is notable that workers self-reported a relatively low use of family planning/reproductive health methods but the clinic records showed a much higher use, suggesting that young women were reluctant to admit using these services. Survey results indicated that 51% of workers expressed a willingness to pay double the co-pay and that workers especially valued the free medications provided by the clinic. Female workers also valued there being female doctors/nurses and mentioned that they were not available at the government clinics. Workers expressed a desire for the clinic to be open more days per week and to have services expanded to include their children. Managers also said the clinic availability made a difference in their own decisions to come to work and in whether to consider leaving for other jobs. Management recommended that workers use the services and they used it themselves, which was deemed an important determinant of the program's success. A peer educator program helped to spread information about the clinics.

Interventions such as ESD might be expected to work in other settings where individuals do not otherwise have access to similar health services and where large numbers of workers work at a location (such as at a medium or large size factory) to be able to support the operation of a clinic. This study did not examine, however, the issue of how to ensure that

doctors and nurses hired by the clinic provide high quality services and what to do if any conflicts of interest might arise between the clinic staff and the employer. Also, it is clear that the workers in the ESD study would have liked services to be extended to family members, especially children, but there would need to be more consideration of whether bringing sick children to work to be seen at the clinic would be in the interest of the employer. It could possibly be, if treating childhood illnesses resulted in lower absenteeism by adult workers or if the adult workers worked harder to keep their job with these extra amenities.²⁸

7.2 Workplace gender equity programs

In many countries, workplace discrimination is a problem and potentially a major barrier to women's employment and advancement. Becker's theory of discriminatory firm behavior showed how competition can be a force driving out discrimination, but only if there are enough low cost nondiscriminatory firms willing to enter the market.²⁹ In situations where discrimination is widespread and/or where firms have some market power, due, for example, to workers having to engage in costly search to find a job, discrimination would be expected to lower levels of female employment and lower wages for women. Instituting laws against discrimination and seeking better enforcement of them is usually not a viable option over the short-term.

There are a few studies that attempt to measure the degree of discrimination against women in developing countries. For example, Moreno et. al. (2012) perform a pseudo-audit pair study in which they follow the job search process of applicants to a job intermediation service in Lima, Peru, a labor market characterized by a high degree of gender occupational segregation. Moreno et. al. (2012) know the requirements attached to every job posting listed with the job intermediation service and they gather information on the characteristics of the applicants, including all the jobs to which they apply and the job search outcomes.

²⁸Efficiency wage theory suggests that workers who are paid higher wages, for example, by getting better amenities on the job, might choose to work harder.

²⁹The nondiscriminatory firms could pay lower wages and therefore have lower costs and expand their production and, if there are enough of them, ultimately drive out the discriminators.

The goal of their study is to quantify the extent to which hiring decisions are influenced by gender and racial characteristics, focussing on job postings for accounting and administrative assistants, secretaries, and salespersons. Overall, their study finds no overall significant differences in the hiring rates for different gender-race groups, with the exception of some evidence of discrimination in hiring for female indigenous applicants for secretarial positions. They also find that females expect to receive about 7% lower wages on their new job than males.

A potential difficulty with any nonexperimental study of hiring behavior is that applicants may choose not to apply to jobs where they anticipate a low chance of a successful hire. For example, the set of women who apply for a job where the chances of being hired are low may only be those who are unusually well qualified. Nonrandomness in the pool of applicants makes it difficult to assess the degree of discrimination at those jobs, because the males and females may not be otherwise comparable. The reliability of results from any nonexperimental audit pair study depend crucially on whether the observables available in the data to control for differences between men and women accurately capture all the characteristics that are relevant for the job.

Galaza and Yamada (2012) perform an experimental audit pair study to test for labor market discrimination in hiring/recruitment in Peru. They sent out 4,820 fictitious resumes in response to 1,205 real job vacancies advertised in a major newspaper, with different surnames randomized on the resumes that indicate the gender and indigenous status of the applicant. They find that males receive 20 percent more callbacks than females, and whites receive 80 percent more calls than indigenous applicants.³⁰

The *Gender Equity Model (GEM)* was developed, tested and implemented by the World Bank as a way of promoting gender equity in private forms, public institutions and civic organizations. The model provides assistance to firms in achieving gender equity goals and institutes a certification process to publicly recognize good practices. The program was first

³⁰For a critique of audit pair studies and for a discussion of the inferences that can be drawn from them, see Heckman (1998)

introduced in Mexico in 2003 in a pilot project with participation of 20 firms. GEM has the following main objectives: i) eliminate gender discrimination in hiring; ii) eliminate practices that can prevent female participation in training and facilitate access to management positions; iii) eliminate factors that inhibit promotion of females; iv) promote affirmative actions that allow for family-work balance (e.g. maternity rights), and v) introduce mechanisms for handling sexual harassment. The certification process, which takes about a year, proceeds as follows. First, there is a call for firms to voluntarily participate in the program. Then, there is a firm self-assessment of whether and how opportunities differ for men and women, in consultation with workers and staff. Firms then develop an action plan in consultation with the GEM advisers for how to decrease the inequalities. Lastly, there is a review of the firm's compliance with the plan and a certification that is performed by an independent firm.

In Mexico, GEM was evaluated through a qualitative survey and firms reported high rates of satisfaction with the program. By 2009, 255 firms with about 300,000 employees had been certified. 42% of these were government agencies. GEM has since been replicated for Chile, Argentina, Dominican Republic and Colombia. In Mexico and in Chile, GEM policies were adopted as official policies for government agencies.³¹

The *Gender Equity Model* has also been recently extended and adapted to Egypt. The *The Promoting Gender Equity and Productivity in Private Firms: the Gender Equity Model, Egypt (GEME)* was started as a partnership between the World Bank, UNIFEM and the International Center for Research on Women (ICRW). As described in Golla et. al. (2011), the program was implemented within ten private sector firms in the Cairo area in from March, 2008-June, 2010. The goals were to promote gender equity in working conditions, access to jobs, professional development and training, and participation of women in decision-making. Egypt faces problems of high unemployment and large gaps in female and male unemployment rates. Labor force participation of women is fairly low, with only 24.3% of women aged 15-64 participating in the formal labor market in 2008. Female employment

³¹In Argentina (where the program is called MEGA) the model was extended to include a requirement that firms present a non-sexist public image.

tends to be concentrated in government and public sector work, and women often drop out of the labor force when they marry and have children.

The program selected ten medium to large-scale, private sector firms from the Cairo area for the intervention. 300 large exporting firms were notified about their eligibility for the program. Initially the response was low and the GEME program implementers had to contact firms and encourage them to participate. By March 2008, seven firms had been recruited and later three more firms joined. Another nine were identified as a group of firms called “Friends of GEME” who agreed to participate as controls. The GEME program tried to get support for their activities from managers and stake-holders, but initially support was low, with some firms not seeing a need for the changes, expressing concern that the program would favor women over men, and finding it problematic that the program might violate social norms. Some firms already had gender policies in place. By the end of the program, however, Golla et. al. (2011) report widespread support for the intervention in participating firms, with firms recognizing that the program could help them improve existing gender equity policies.

The GEME program was evaluated using quantitative methods. The evaluation used a matched pair design, matching treated firms to control firms, preselected to match the characteristics of the treated firms. Panel data were collected on firms and employees both before and after the program. The program impact evaluation was based on eight treatment and eight control firms.³² The firm-level indicators examined are the proportion of women among all hired in the last 6 months, the proportion of women among all job applicants in the past 6 months, the proportion of women among all trained in the last 6 months, the proportion of women promoted in the last 6 months and the proportion of women among all in any management position. In addition, the evaluation examined indicators related to individuals’ perception of equality of opportunity, measures of job satisfaction, and days absent at the job. The sample of employees included 3490 workers total in the treatment

³²A good match could not be found for one of the treated firms and another declined to participate in the final data collection.

and control groups. The individuals were on average very well educated, with more than 90 % having secondary degrees and over half having university degrees.

Treatment effect estimates are obtained from a panel data regression, controlling for some characteristics of the firm (such as industry, firm size). The firm-level analysis did not find any statistically significant effect of the program for any of the firm-level indicators. Although all the indicators with the exception of proportion of women promoted were positive, none of the treatment impact estimates are statistically different from zero. The lack of significance of the treatment effects on the firm level indicators can easily be due to the low sample size and therefore low power, as the authors of the study note.

At the individual level, many of the measures of job satisfaction and of women's perception with regard to gender equity were positive and statistically significantly different from zero at conventional levels. However, there is also a positive effect of the program on women reporting that they experienced discrimination based on gender in the past year. This finding could of course be due to the program generating greater awareness about discrimination or to women being more willing to report discrimination. There was an observed negative effect on harassment indicators (such as inappropriate joking). Curiously, employees in treatment firms are less likely to report that their firm engages in "gender sensitive" behavior and are less likely to think that top personnel treat female workers with respect. They also did not report being better informed about gender equity policies or about whether there were workshops in their firm on the topic of gender equity. The study also did not find any program effects on absenteeism, considered as a measure of productivity.

Two potential limitations of this evaluation study are the small sample size, for the firm-level analysis, and the non-random selection of firms. Another limitation that is particular to this type of intervention is that gender equity programs could make employees more aware of gender inequities, leading to report differently and making it very challenging to measure the effects of such programs.

The initial evidence from implementing the *Gender Equity model* in Mexico and other Latin American countries is encouraging, particularly the fact that the program led to a

formalization of gender equity public policies in Mexico and Chile. Firms in those settings generally report high rates of satisfaction with the program. However, the evidence from the quantitative evaluation study performed in Egypt is less encouraging about the effects of the gender equity workplace interventions in improving working conditions for women, at least in the setting of Egypt. The fact that it was difficult to recruit firms to take part in the study already suggests that these types of programs may not be easy to implement in that setting and that firms are not easily perceiving the benefits from participating and may instead anticipate possible higher costs in terms of female workers becoming more aware of discrimination in the workplace.

7.3 Land titling programs in Peru and Argentina

In this section, we review evidence on the effects of Peruvian and an Argentinian land titling programs, which provide land titles to household squatting on public or private land. By providing legal protection for property, land titling programs significantly affect the wealth of these households and reduce the incentives for women and children to stay home. Although these types of programs do not directly target women's work, we review them here because previous research has found that they have substantial impacts on women working.

Land titling programs typically transfer property rights to poor households occupying the land. These types of programs do not aim to directly influence female employment. However, in transferring wealth to the household and in securing property rights, these programs often do have a significant influence on women's working decisions.

Field (2003a, 2003b) studies the effect of an urban land titling program in Peru that was targeted at urban squatter households. The program greatly decreased the administrative burden of obtaining a land title, which had required application at a large number of offices, and distributed over 1.6 million property titles over a five year period. One of the explicit aims of the program was to improve gender inequality of property ownership and to this end the program rules stipulated that, among common law and legally married households, both spouses' names had to appear on government issued property documents. The allocation of

the program was not random, although the neighborhoods that received the program early are highly similar in terms of observable characteristics to those of those that received the program later.³³ The evaluation approach taken in the study is a difference-in-difference comparison of eligible women in treated neighborhoods to eligible women in as yet untreated neighborhoods, with an adjustment for differences between noneligible households living in the same neighborhoods. Noneligible households are those that possess a title prior to the program and therefore had nothing to gain from participating in it.

Field (2003a) finds that program beneficiary women are significantly more likely to appear on property documents and are more likely to report participating in household decision-making. Field also examines whether the program affected their fertility behavior. An effect of land titling on birth rates may come through multiple channels and the direction of the effect is theoretically ambiguous. On the one hand, land titling represents a wealth effect that may increase fertility, assuming children are a normal good. On the other hand, including women's names on property documents can change the balance of power in a relationship and, if women have preferences for lower fertility and fertility decisions are made through family bargaining, can decrease fertility. Land titling can also influence the value of children, for example, by affecting parents' options about where to live in old age. Lastly, when people have a legal right to their property and no longer have to squat to retain control of it, their time can be used for other purposes. Changes in the value of mother's time may also affect fertility. Field (2003a) finds a strong effect of the land titling program on fertility behavior, with eligible households who had been exposed to the program exhibiting roughly a 20% reduction in annual birth rates in the few years after the program.

In a separate study of the same program, Field (2003b) analyzes the effects of the program on hours of work, location of entrepreneurial activity and child labor force participation, using a similar evaluation strategy. The impact estimates indicate that land titling increases labor hours, shifts labor supply away from work at home and towards work in the outside market

³³The characteristics include rates of malnutrition, illiteracy, fraction of school-aged children not in school, residential crowding, proportion of the population without access to water, sewer or electricity services.

and leads to substitution of adult for child labor. On average, labor hours increase by 17%, the probability of working inside the home decreases by 47% and the probability of child labor declines by 27%.

Another study of land titling is that of Galiani and Schargrodsky (2009), which examines the effects of land titling in an urban area of Buenos Aires. When the squatters had originally settled on the land, they thought it was public land, but it was actually private land. In 1984, a law was passed expropriating the former owners' land (with compensation) and entitling current occupants. Some original owners accepted the governmental decision and compensation, while others challenged the decision in drawn-out lawsuits (some of which are still pending). Galiani and Schargrodsky exploit the variation in the owners' decision to accept or challenge the law to identify the effects of land titling on occupants' behaviors. Although it is conceivable that owners with more favorable land quality would have been more likely to contest the appropriation, Galiani and Schargrodsky show that this is not the case and that the parcels of land for which owners contested or did not contest the appropriation are actually highly comparable and basically next to each other. For this reason, their analysis considers the land titling "treatment" exogenous from the point of view of the squatters.

Galiani and Schargrodsky find substantial effects of the land titling program on household behaviors using data from two surveys, performed in 2003 and 2007, and focusing on 245 families that were identified as having arrived on the land prior to the intervention. Their main findings are that entitled families increased their housing investment, reduced household size, and increased children's education relative to the control group. The entitled households also show a reduction in the number of offspring to the household head.

Both the Field (2003a,b) and the Galiani and Schargrodsky (2009) evaluations of land titling programs find substantial benefits of the programs that include increases in mother's working, decreases in fertility, increases in children's education, decreasing child labor, and increases in housing investment. Although not that much evidence has been accumulated yet on the effects of land titling programs, they appear to be a promising way of significantly

increasing a poor household's wealth over the short-term and inducing an array of changes in behavior. Requiring that spouse's names appear on the title, as in the Peruvian program, significantly increased the probability that a woman's name was included on the title.

8 Synthesis and Policy Recommendations

This paper has studied the effectiveness of a variety of policy interventions and social programs at improving the quantity and quality of women's work. All of these programs were subjected to quantitative impact evaluations of different kinds and some also to rigorous cost-benefit analyses. Many programs were found to be effective in increasing women's quantity of work as measured by increased rates of labor market participation and number of hours worked. In some cases, the programs also increased quality of work, for example, by increasing the capacity for women to work in the formal rather than the informal sector where wages are higher on average and where women are more likely to have access to health, retirement, and other benefits. Here, we summarize the key lessons learned about the effectiveness of each of the types of programs.

ALMP programs These programs are often adopted by countries as a way of reducing unemployment or ameliorating the effects of macroeconomic shocks. These programs are not usually targeted at women, although women often seek out the program services and benefit from them. As noted in the introduction, ALMP programs take a variety of forms that may include job search assistance, training, internships with firms, public works jobs, or wage subsidies. Often, programs have more than one of these features. Table 1 summarizes the features and the impacts of the ALMP programs that were reviewed in this study.

The following broad conclusions emerge from the survey for designing effective training policies:

- ALMP program impacts for women often exceed those of men. Studies differ, though, in their findings about whether more educated or less educated women or older or younger women benefit most from participating.

- The majority of ALMP programs increase women’s employment rates and increase their exit rate out of unemployment, but do not lead to substantial wage increases. The evidence suggests that ALMP programs help workers find jobs more quickly but sometimes at a lower wage than they would otherwise be able to obtain searching on their own.
- Short term classroom training interventions (e.g. 3-6 months) do not bring about large changes in earnings capacity.
- The effectiveness of training programs is enhanced when there are explicit ties to private sector firms, with firms agreeing to provide jobs or internships for people after they receive the training. Such ties help to ensure that there will be future placements for the trainees and that the type of training being provided is the type demanded in the market place.
- Programs should be designed in a way that creates incentives for training program providers to help trainees find jobs. For example, training program providers can be paid for their services only upon successful placement of the individual in a firm for a certain number of months. However, such provisions will also typically lead to training program providers being more selective in who they admit into training and special provisions may need to be made to get them to serve so-called “hard-to-serve” individuals, if they are a priority of the program.
- Asking trainees to contribute a small amount towards their training can be an effective way of selecting among applicants those who expect to benefit the most from training.
- Highly educated workers with specialized skills (such as those who worked in former communist countries like Russia) do not appear to benefit from basic training/retraining programs. More evidence is needed on how the programs can be tailored to the meet the needs of these kinds of workers.

- Wage subsidies/vouchers can be effective in settings where women have relevant skills, but there are obstacles to their meeting prospective employers, for example, due to discrimination. They provide an “empowerment” effect, giving women more confidence in approaching employers, and can also help overcome firms’ barriers to hiring workers.
- Public works programs (sometimes called Socially Useful Projects) can be effective in alleviating poverty and mitigating the effects of macro-economic shocks in settings where there is widespread unemployment and job creation is a priority. But, otherwise, programs that attempt to place trainees at existing jobs at private sector firms are more effective than public works programs.

What evidence is needed? In many countries, ALMP programs operate on a very large scale. An issue that was not addressed in the evaluation studies that bears further consideration is to what extent large government-sponsored training programs crowd out private training. Additionally, it is possible that the benefits observed for program beneficiaries are at the expense of nonbeneficiaries, because of displacement, substitution effects or lower wages. For example, a wage subsidy might encourage a firm to hire a particular worker rather than some other worker. It would therefore be useful to more systematically analyze the effects of these programs within a job search modeling framework to gain a better understanding of the mechanisms through which the programs operate, that is, how they affect the costs of searching for a job, the arrival rate of offers and the distribution of wage offers for all workers, not just participating ones.

Also, none of the studies reviewed analyzed whether providing adult women with job training/job search assistance affects their marriage or fertility choices. If women have better labor market prospects, they may delay marriage, find better marriage partners and postpone fertility, although higher income might also lead to an increase in fertility.³⁴ These questions warrant more investigation to more fully assess of the effectiveness of these programs on the women who participate in them.

³⁴It is usually assumed that children are a normal good.

Youth-oriented training programs Table 2 summarizes the evidence on the effectiveness of programs targeted towards youth, which are generally more uniformly positive on employment and wages. The programs tend to combine classroom training with an internship component. Some key findings are the following:

- Evaluations of programs in Argentina, Peru, and Colombia found evidence that youth training programs (classroom training followed by OJT) led to higher employment and higher earnings, although one study of a similarly designed program in Dominican Republic found it to be ineffective in increasing employment or earnings.
- Benefits of the on-the-job training component have been found to be higher than that of the classroom training component. The link between the training provider and the private firm willing to take on the worker after training is important to a program's effectiveness.
- Evidence is preliminary and study is on-going, but the job voucher program in Jordan (Jordan NOW) was in the short-term effective in increasing employment rates for young female college graduates. The program operates in a setting where many young women express a desire to work but they have few work opportunities, in part because the firms were not accustomed to hiring women.

What evidence is needed? With any youth employment program, there is the risk that the program may draw youth out of school. Targeting programs at low skill youth can make it more attractive to be a low skill youth, which can discourage skill investment. Most of the programs examined were targeted at youth who were already out of school and were not intended to affect youth in school, but if in-school youth know about the availability of these programs, there is the possibility that these programs encourage youth to drop out of school.³⁵ There may also be effects on other early life decisions, such as decisions regarding

³⁵Keane and Wolpin (2000) study a wage subsidy program in the US context and find that it induces youth to drop out of school.

marriage and fertility. The studies reviewed here did not consider how these programs affect schooling levels, marriage, or fertility decisions.

Very few of the studies analyze cost-effectiveness. The one study that does by Aedo and Nunez (2004) (for Argentina) found that program impacts have to be sustained over the longer term (9 years or more) for the program to be cost-effective, and none of the evaluation studies follows participants over that long a period.

Child care programs Another class of programs we reviewed are programs that affect the availability and pricing of child-care. Table 3 summarizes the impact evaluation results from the variety of child care programs described in section 5. In some cases, the program built and staffed new child care centers or preschools. In other cases, it payed for one women to be an at-home community day care provider, which facilitates work for other women. A key concern with child care programs is that they might stimulate fertility by lowering the costs of rearing children. Of the studies reviewed here, only the Schlosser (2011) study examined potential effects on fertility and did not find any effects on fertility.

In addition to the potential benefits for mothers, child care programs have potential health and education benefits for children participating in them, who receive nutrition and education, and for older children in the family (typically girls), who are able to attend school rather than having to care for younger siblings. Below, we summarize broad conclusions and policy recommendations with regard to child care programs

- Evaluations of child care programs in Guatemala, Colombia, Argentina and Brazil found strong effects of child-care availability of mother's rates of working and on the the number of hours worked. Several of the community daycare programs also showed substantial positive benefits on the nutrition and development of the young children participating in them, that imply high benefit-cost ratios.
- Providing free preschool for Arab mothers in Israel, a group for whom labor force participation rates were low, led to a substantial increase in their labor supply, with no observed changes in fertility. Thus, the programs can have potentially large effects

even in settings where women working is not common.

- Strong evidence for a substantial effect of child-care costs on working behavior comes from formerly Communist countries like Romania and Russia that have undergone very large changes in the costs of child-care. Some countries went from child-care being free and widely available to a situation where child-care is expensive and a major determinant of whether women work. The two studies reviewed here by Lockshin (1999) and Fong and Lockshin (2000) found women's labor supply to be fairly elastic with respect to the price of childcare.
- Child-care subsidies are more effective than wage subsidies or family income subsidies in increasing family income levels, in part because of better targeting at women whose labor force behavior is potentially affected by the policy.
- The effectiveness of child care programs in increasing mother's labor supply and hours worked is likely to depend on the country context. In the context of Accra, Ghana, the local supply of child-care was found not to be a significant determinant of mother's working decisions, because mothers worked in jobs where they found it relatively easy to combine work with taking care of children.

What additional evidence is needed? With any child care program, there is a concern that publicly provided child care may crowd out the supply of privately provided care and this is something that so far has not been rigorously examined. In settings where there are quality private sector child care options available, it may be better to provide women with child care vouchers rather than building new child care facilities.

Additionally, it is desirable to know how putting the child in the child care program compares to the child staying home with the mother. Some of the studies reviewed examined the effects of the child care program on child health and development and the effects were generally found to be positive, but there needs to be more systematic investigation of child outcomes along with adult labor market outcomes as well as of effects on fertility to fully assess the effectiveness of child care programs.

Elder care The empirical evidence on whether elder care inhibits women's work is mixed, with Mroz et. al. (2012) for Mexico finding that elder care reduces the probability of women working, Bravo and Puentes (2012) for Chile finding no effect, and Mauer-Fazio et. al. (2012) for China finding that coresidence with elderly persons facilitates women's work. Whether the presence of elderly persons in the household affects women's capacity for work is likely to be context specific, depending on the reasons elderly coincide with younger or middle aged persons (usually, their children). In country settings where coresidence with elderly is common and is not closely tied to the elderly needing care, then coresidence may enhance the capacity for women to work, as older parents or parents-in-laws assist with child-caring duties. But in other settings, such as the U.S., where co-residence is relatively uncommon unless the older person needs care, then the presence of elderly likely imposes demands on women that compete with their capacity to work, without leading to many benefits. More research is needed on the decisions that women jointly make about child care and elder care and coresidence, to better understand how the factors jointly influence women's labor market outcomes. At this stage, there is not enough evidence to support elder care programs as a way of increasing female employment and earnings.

Other programs: Section seven reviewed a variety of other types of programs that have been found to influence women's labor market outcomes, including a workplace health program, a gender equity program, and land titling programs. The evaluations from these programs yielded the following main results:

- The ESD workplace health program, which was a USAID demonstration project, was successful in reducing worker absenteeism and turnover and was deemed cost effective. Similar interventions might work in other settings with there are clearly inadequate public health services for women and/or where the work environment includes chemical hazards.
- Workplace gender equity programs try to raise awareness about gender equity issues and to reduce workplace discrimination. There is evidence from process evaluations

and from qualitative evaluations that these programs have been successful in Mexico, Chile, and other Latin American countries where they introduced an in some case they have led to the adoption of formal government policies concerning gender equality. However, when the program was adapted and implemented in Egypt, a quantitative evaluation of the program indicates that it was largely unsuccessful, in part because few firms were interested in participating. For those firms that did participate, its effect is hard to measure. If a program raises awareness of what constitutes discriminatory behavior, then it is likely that women in the treatment group will be more likely to report experiencing discrimination in the workplace, which seems to have happened in the evaluation.

- Land titling programs are an innovative approach to increasing the wealth of poor households and, when the title is in the name of the woman, to potentially have a large effect on family bargaining arrangements. They have been found to change an array of household behaviors, including labor supply of household members. The land titling programs we reviewed (in Peru and Argentina) increased female labor supply, increased the tendency for women to work outside the home, and decreased fertility.

Thus, the workplace health program and the land titling programs had promising positive effects both on women's LFP and on women's hours of market work.

References

- [1] Abadie, Alberto and Guido W. Imbens (2005): “Large Sample Properties of Matching Estimators for Average Treatment Effects,” *Econometrica*, 74:1, 235-267.
- [2] Adato, Michelle, Bénédicte de la Briere, Dubravka Mindek and Agnes Quisumbing (2000), “The Impact of Progresa on Women’s Status and Intrahousehold Relations,” IFPRI Report, Washington, DC.
- [3] Aedo, Cristian and Sergio Nunez (2004): “The Impact of Training Policies in Latin America and the Carribean: The Case of Programa Joven,” IDB working paper #R-483.
- [4] Angelucci, Manuela (2008): “Love on the rocks: aggressive behavior and alcohol abuse in rural Mexico,” working paper, University of Arizona.
- [5] Aiyagari, S. R., “Uninsured Idiosyncratic Risk and Aggregate Saving.” *Quarterly Journal of Economics*, 109 (August 1994): 659-684.
- [6] Attanasio, Orazio and Marcos Vera-Hernandez (2004): “Medium and Long Run Effects of Nutrition and Child Care: Evaluation of a Community Nursery Programme in Rural Colombia,” Institute for Fiscal Studies, Working Paper 04/06, University College, London.
- [7] Attanasio, Orazio, Adriana Kugler and Costas Meghir (2008): “Training Disadvantaged Youth in Latin America: Evidence From a Randomized Trial,” NBER working paper #13931.
- [8] Barros, Ricardo, Pedro Olinto, Trine Lunde, Mirela Carvalho (2011): ”The Impact of Access to Free Childcare on Women’s Labor Market Outcomes: Evidence from a Randomized Trial in Low-Income Neighborhoods of Rio de Janeiro,”

- [9] Banerjee, Abhijit, Esther Duflo, Rachel Glennerster, and Cynthia Kinnan (2009): “The miracle of microfinance? Evidence from a randomized evaluation,” working paper, MIT.
- [10] Behrman, Jere R., Yingmei Cheng and Petra E Todd (2004): “Evaluating Preschool Programs When Length of Exposure to the Program Varies: A Nonparametric Approach,” Vol. 86, No. 1, Pages 108-132.
- [11] Behrman, Jere R., Piyali Sengupta and Petra E. Todd (2005): “Progressing through PROGRESA: An Impact Assessment of a School Subsidy Experiment,” *Economic Development and Cultural Change*, Volume 54, No. 1, 237-275.
- [12] Behrman, Jere R., Susan W. Parker and Petra E. Todd (2010): “Do Conditional Cash Transfers for Schooling Generate Lasting Benefits? A Five-Year Follow-Up of PROGRESA/Oportunidades, forthcoming *Journal of Human Resources*.
- [13] Behrman, Jere R. and Elizabeth M. King (2008): “Program Impact and Variation in Duration of Exposure,” in Samia Amin, Jishnu Das and Markus Goldstein, eds. *Are You Being Served: New Tools for Measuring Service Delivery*, Washington, DC: World Bank, 147-172.
- [14] Behrman, Jere R. and Elizabeth M. King (2009): “Timing and Duration of Exposure in Evaluations of Social Programs,” *World Bank Research Observer*, forthcoming.
- [15] Benus, Jacob, Raluca Catrinel Brinza, Casilica Cuica, Irina Denisova, and Marina Kartseva (2005): “Retraining Programs in Russia and Romania: Impact Evaluation Study,” CEFIR research papers, <http://www.cefir.ru/index.php?l=eng&id=34&yf=2004>.
- [16] Berlinski, Samuel and Sebastian Galiani (2007): “The effect of a large expansion of pre-primary school facilities on preschool attendance and maternal employment,” in *Labour Economics*, 14, 665-680.

- [17] Betcherman, Gordon, Martin Godfrey, Susana Puerto, Friederike Rother, and Antoneta Stavreska (2007): “A Review of Interventions to Support Young Workers: Findings of the Youth Employment Inventory,” SP Discussion Papers, No. 0715.
- [18] Betcherman, Gordon, Karina Olivas and Amit Dar (2004): “Impacts of Active Labor Market Programs: New Evidence from Evaluations with Particular Attention to Developing and Transition Countries,” Social Protection Discussion Paper Series No. 0402, The World Bank.
- [19] Bidani, Benu, Chorching Goh and Christopher J. O’Learly (2002): “Has Training Helped Employ Xiagang in China? A Tale from Two Cities,” World Bank manuscript.
- [20] Bravo, David and Esteban Puentes (2012): “Female Labor Force Participation and Informal Care of Adults: Evidence for a middle-income country,” working paper, University of Chile.
- [21] Burdett, Kenneth and Dale Mortensen (1998): “Wage Differentials, Employer Size, and Unemployment,” *International Economic Review*, Vol. 39, No. 2, pp. 257-273.
- [22] Card, David, Pablo Ibarrraran, Ferdinando Regalia, David Rosas, Yuri Soares (2007): “The Labor Market Impacts of Youth Training in the Dominican Republic: Evidence from a Randomized Evaluation,” NBER working paper #12883.
- [23] Card, David, Jochen Kluve, Andrea Weber (2010): “Active Labour Market Policy Evaluations: A Meta-Analysis,” *The Economic Journal*, Vol. 120, 548, pp. F452F477.
- [24] Chowdhury, Sorowar, David Wofford, Veronique Dupont (ESD) (2007): “Effects of a workplace health program on absenteeism, turnover, and worker attitudes in a Bangladesh garment factory,” research report, Extending Service Delivery Project, Washington, DC.

- [25] Cleveland, G., Gunderson, M., Hyatt, D., (1996) “Child care costs and the employment decision of women: Canadian evidence.” *Canadian Journal of Economics*, XXIX, No. 1.
- [26] Deaton, Angus (2009): “Instruments of Development: Randomization in the Tropics, and the Search for the Elusive Keys to Economic Development,” NBER working paper No. W14690.
- [27] Dulfo, Esther and Michael Kremer (2004): “Use of Randomization in the Evaluation of Development Effectiveness” in *Evaluating Development Effectiveness*, ed. George Pitman, Osvaldo Feinstein, Gregroy Ingram, World Bank.
- [28] Field, Erica (2003a): “Fertility Responses to Urban Land Titling Programs: The Roles of Ownership Security and the Distribution of Household Assets,” working paper, Harvard University.
- [29] Field, Erica (2003b): “Entitled to Work: Urban Property Rights and Labor Supply in Peru,” manuscript, Harvard University.
- [30] Fiszbein, Ariel, Norbert Schady, Francisco H. G. Ferreira, Margaret Grosh, Nial Kelleher (2009): “Conditional Cash Transfers: Reducing Present and Future Poverty,” *World Bank Publications*.
- [31] Fong, Monica and Michael Lokshin (2000): “Child Care and Women’s Labor Force Participation in Romania,” Policy Research Working Paper #2400, the World Bank.
- [32] Galasso, Emanuela and Martin Ravallion (2004): “Social Protection in a Crisis: Argentina’s *Plan Jefes y Jefas*,” *The World Bank Economic Review*, Vol. 18, No. 3.
- [33] Galasso, Emanuela, Martin Ravallion and Agustin Salvia (2001): “Assisting the Transition from Workfare to Work: A Randomized Experiment in Argentina,” World Bank manuscript.

- [34] Galarza, Francisco B. and Gustavo Yamada (2012): “Labor Market Discrimination in Lima, Peru: Evidence from a Field Experiment,” Working paper, Universidad del Pacifico, Department of Economics and Centro de Investigacin (CIUP), Lima, Peru
- [35] Galiani, Sebastian and Ernesto Schargrodsky (2009): “Property Rights for the Poor: Effects of Land Titling,” Ronald Coase Institute Working Paper Series, No. 7.
- [36] Golla, Anne Maried, Mona Selim, Ramadan Mohamed, and Faten Abd El-Fattah (2011): “Egypt Results Based Initiative: Promoting Gender Equity and Productivity in Private Firms in Egypt: The Gender Equity Model, Egypt”, International Center for Research on Women working paper, Washington, DC.
- [37] Hall, John A. (2010): “The ILO’s Better Factories Cambodia Program: A Viable Blueprint for Promoting International Labor Rights?” *Stanford Law and Policy review*, Vol. 21:427-460.
- [38] Heckman, James and Richard Robb (1985): “Alternative Methods for Evaluating the Impact of Interventions,” in James Heckman and Burton Singer, eds., *Longitudinal Analysis of Labor Market Data* (Cambridge, England: Cambridge University), 156-246.
- [39] Heckman, James (1998). “Detecting Discrimination. *The Journal of Economic Perspectives*, Vol 12, Issue 2, 101-116.
- [40] Heckman, James, Robert Lalonde and Jeffrey Smith (1999): “The Economics and Econometrics of Active Labor Market Programs” in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics Volume 3A* (Amsterdam: North-Holland), 1865-2097.
- [41] Heckman, James and Edward Vytlacil (2005): ”Structural Equations, Treatment Effects, and Econometric Policy Evaluation,” *Econometrica*.
- [42] Hein, Cathering and Naomi Cassirer (2010): *Workplace solutions for childcare*, ILO Publications, International Labor Office, Geneva, Switzerland.

- [43] Hussein, Maliha and Shazreh Hussein (2003): “The Impact of Micro Finance on Poverty and Gender Equity: Approaches and Evidence from Pakistan,” Pakistan Micro Finance Network working paper.
- [44] Imbens, G., and J. Angrist (1994): “Identification of Local Average Treatment Effects,” in *Econometrica* 62, 467-475.
- [45] Jalan, Joytsna and Martin Ravallion (2003): “Estimating the Benefit Incidence of an Anti-Poverty Program using Propensity Score Matching,” *Journal of Business and Economic Statistics*, Vol. 21, No. 1.
- [46] Kaboski, Joseph P. and Robert M. Townsend (2007): “Testing a Structural Model of Credit Constraints Using a Large-Scale Quasi-Experimental Microfinance Initiative,” manuscript, MIT.
- [47] Kamerman, S.B. (2000): “Early childhood education and care (ECEC): An overview of developments in OECD countries,” New York, Columbia University, Institute for Child and Family policy, <http://www.columbia.edu/cu/childpolicy/kamerman.pdf>.
- [48] Karlan, Dean and Martin Valdivia (2006): “Teaching Entrepreneurship: Impact of Business Training on Microfinance Clients and Institutions,” manuscript, Yale University.
- [49] Karlan, Dean and Jonathan Zinman (2007): “Expanding Credit Access: Using Randomized Supply Decisions to Estimate the Impacts,” Economic Growth Center Discussion Paper No. 956, Yale University.
- [50] Katz, Elizabeth (2008): “Programs Promoting Young Women’s Employment: What Works?,” The Adolescent Girls Initiative, World Bank Working Paper.
- [51] Keane, Michael P. and Kenneth I. Wolpin (2000): “Eliminating Race Differences in School Attainment and Labor Market Success,” *Journal of Labor Economics* , Vol. 18, No. 4, pp. 614-652

- [52] Khandker, S. R., H. A. Samad and Z. H. Khan (1998): "Income and Employment Effects of Micro-credit Programmes: Village-level Evidence from Bangladesh," *Journal of Development Studies*, 35, 2, 96-124.
- [53] Kluge, Jochen, Lehmann, Hartmut and Christoph Schmidt (1998): "Active Labor Market Policies in Poland: Human Capital Enhancement, Stigmatization or Benefit Churning," *Journal of Comparative Economics*, 27, 61-89.
- [54] Lakati, A., C. Binns and M. Stevenson (2002): "Breast-feeding and the working mother," in *Public Health and Nutrition*, Dec, 5(6):715-8.
- [55] Landsberger, Henry (1948): *Hawthorne Revisited*, Cornell University Press.
- [56] Lise, Jeremy, Seitz, Shannon and Jeffrey Smith (2003): "Equilibrium Policy Experiments and the Evaluation of Social Programs," IZA discussion paper No. 758.
- [57] Lockshin, Michael M. (1999): "Household Child Care Choices and Women's Work Behavior in Russia," Policy Research Working Paper #2206, The World Bank.
- [58] Lubyova, Martina and Jan C. van Ours (1999): "Effects of Active Labor Market Programs on the Transition Rate from Unemployment into Regular Jobs in the Slovak Republic," *Journal of Comparative Economics*, Volume 27, Issue 1, March 1999, Pages 90-112.
- [59] Maluccio, John A. (2005): "Coping with the 'Coffee Crisis' in Central America: The Role of the Nicaraguan Red de Proteccion Social (RPS)," IFPRI Discussion Paper BRIEFS No. 188.
- [60] Meredith B. Lilly, Audrey Laporte and Peter C. Coyte (2007): "Labor Market Work and Home Care's Unpaid Caregivers: A Systematic Review of Labor Force Participation Rates, Predictors of Labor Market Withdrawal, and Hours of Work," *The Milbank Quarterly*, Vol. 85, No. 4 (Dec., 2007), pp. 641-690.

- [61] Marenzi, Anna, and Laura Pagani (2005): “The Impact of Elderly Parents on the Labour Force Participation of Italian Women. *Revista di Politica Economica*, 95(34):15589.
- [62] Margaret Maurer-Fazio, Rachel Connelly, Lan Chen, and Lixin Tang (2011): “Childcare, Eldercare, and Labor Force Participation of Married Women in Urban China, 1982-2000,” *J. Human Resources*, March 20, 2011 46:261-294.
- [63] Margaret Maurer-Fazio, Rachel Connelly, Lan Chen, Lixin Tang (2011): “Childcare, Eldercare, and Labor Force Participation of Married Women in Urban China, 1982-2000,” *Journal of Human Resources*, Volume 46, Number 2, pp. 261-294.
- [64] Maurer-Fazio, Margaret, Rachel Connelly, Lan Chen, Lixin Tang (2011): “Childcare, eldercare, and labor force participation of married women in urban China, 1982-2000,” *The Journal of Human Resources*, Vol. 46, Issue 2, p 261-294.
- [65] Morduch, Jonathan (1994): “Poverty and Vulnerability,” *American Economic Review*, 84(2):221-25.
- [66] Morduch, Jonathan (1995): “Income Smoothing and Consumption Smoothing,” *Journal of Economic Perspectives*, 9(3), 103-14.
- [67] Moreno, Martin, Hugo Nopo, Jaime Saavedra, and Maximo Torero, “Detecting Gender and Racial Discrimination in Hiring Through Monitoring Intermediation Services: The Case of Selected Occupations in Metropolitan Lima, Peru,” *World Development*, Volume 40, Issue 2, 2012, Pages 315-328
- [68] Mroz, Thomas A., Antonio Trujillo, Claudia Piras, Gustavo A. Angeles (2012): “The Economics of and Econometrics for Caregiving and Labor Supply Functions with an Application to the Effects of Elderly Parents and their Health on Adult Daughters’ Labor Supply and Caregiving in Mexico,” manuscript.

- [69] Ñopo, Hugo, Miguel Robles and Jaime Saavedra (2007): “Occupational Training to Reduce Gender Segregation: The Impacts of ProJoven,” Inter-American Development Bank Working Paper No. 623.
- [70] Paes de Barros, Ricardo, Pedro Olinto, Trine Lunde, Miretla Cavalho (2011): “The Impact of Access to Free Childcare on Women’s Labor Market Outcomes: Evidence from a Randomized Trial in Low-Income Neighborhoods of Rio de Janeiro,” working paper.
- [71] Parker, Susan and Emmanuel Skoufias (2000): “The Impact of PROGRESA on work, leisure and time allocation,” International Food Policy Research Institute, Final Report.
- [72] Pitt, Mark and Shahidur R. Khandker (1998): “The Impact of Group-Based Credit Programs on Poor Households in Bangladesh: Does the Gender of Participants Matter?” in *The Journal of Political Economy*, Vol. 106, No. 5, p. 958-996.
- [73] Pungiluppi, Juliana, Mara Elena Castro and Ana Mara Muoz-Boudet () “A Model for Promoting Gender Equity in Private Companies and in Government Agencies: The GEM Certification Process in Latin America and the Caribbean (The cases of Mexico, Chile, Argentina, and the Dominican Republic),” World Bank Note, www.worldbank.org/enbreve.
- [74] Quisumbing, Agnes R., Kelly Hallman and Marie T. Tuel (2003): “*Maquiladoras* and Market Mamas: Women’s Work and Childcare in Guatemala City and Accra,” IFPRI Discussion Paper No. 153, Food Consumption and Nutrition Division.
- [75] Revenga, Ana, Michelle Riboud, and Hong Tan (1994): “The Impact of Mexico’s Retraining Program on Employment and Wages” *The World Bank Economic Review*, Vol. 8, No. 2, p.247-277.

- [76] Roodman, David and Jonathan Morduch (2009): “The Impact of Microcredit on the Poor in Bangladesh: Revisiting the Evidence,” Working paper, Center for Global Development and New York University.
- [77] Rosenbaum, Paul and Donald Rubin (1983): “The Central Role of the Propensity Score in Observational Studies for Causal Effects,” *Biometrika*, 70,41-55.
- [78] Rosenbaum, Paul and Donald Rubin (1985): “Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score,” *American Statistician*, 39, 33-38.
- [79] Ruel, Marie T. and Agnes R. Quisumbing with Kelly Hallman (2006): “The Guatemala Community Day Care Program,” IFPRI Research Report No. 144, Washington, DC.
- [80] Schlosser, Analia (2011): “Public Preschool and the Labor Supply of Arab Mothers: Evidence from a Natural Experiment,” working paper, The Eitan Berglas School of Economics Tel Aviv University.
- [81] Schultz, T. Paul (2000): “Impact of PROGRESA on school attendance rates in the sampled population,” February. Report submitted to PROGRESA. International Food Policy Research Institute, Washington, D.C.
- [82] Schultz, T. Paul (2004): “School Subsidies for the Poor: Evaluating the Mexican ProgresA Poverty Program,” *Journal of Development Economics* 74:2 (June), 199-250.
- [83] Skira, Meghan (2012): “Dynamic Wage and Employment Effects of Elder Parent Care,” working paper, University of Georgia.
- [84] Skoufias, E. and B. McClafferty (2001): “Is PROGRESA working? Summary of the results of an Evaluation by IFPRI.” Report submitted to PROGRESA. Washington, D.C.: International Food Policy Research Institute, <http://www.ifpri.org/themes/progresA.htm>.

- [85] Smith, Jeffrey and Petra Todd (2005): "Does matching overcome LaLonde's critique of nonexperimental estimators?" *Journal of Econometrics*, Vol. 125, p. 305-353.
- [86] Todd, Petra and Kenneth Wolpin (2006): "Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility," *American Economic Review*, Volume: 96, Issue: 5, 1384-1417.
- [87] Todd, Petra (2008): Evaluating Social Programs with Endogenous Program Placement and Selection of the Treated, in *Handbook of Development Economics*, Vol. 4, ed. T. Paul Schulz and John A. Strauss, chapter 60, Elsevier, p. 3847-3894.
- [88] Todd, Petra (2010): "Effectiveness of Interventions Aimed at Improving Women's Employability and Quality of Work: A Critical Review," World Bank working paper.
- [89] Wolf, Douglas A. and Beth J. Soldo (1994): "Married Women's Allocation of Time to Employment and Care of Elderly Parents, *The Journal of Human Resources* , Vol. 29, No. 4, pp. 1259-1276
- [90] Wolf, Douglas A. and Beth J. Soldo (1994): "Married Women's Allocation of Time to Employment and Care of Elderly Parents," in *The Journal of Human Resources*, Vol. 29, No. 4, Special Issue: The Family and Intergenerational Relations (Autumn, 1994), pp. 1259-1276.
- [91] Woolcock, J. V. (1999): "Learning from Failures in Micro-Finance," *The American Journal of Economics and Sociology*, p. 17.
- [92] World Bank Development Report on Gender Equality and Development, 2012, <http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/EXTWDRS/EXTVSitePK:7778063 contentMDK:22851055,00.html>

Appendix: Potential problems that can arise in randomized experiments

The following types of problems may arise:

(i) *Randomization bias* or so-called *Hawthorne effects*: This problem occurs when introducing randomization changes the way the program operates. For example, individuals might choose not to apply to a program if they know they will be subject to randomization, which could change the mix of individuals receiving treatment and therefore change the observed treatment effects. Individuals may also behave differently when they know they are being observed as part of an experiment.³⁶ In these cases, the outcomes that are observed do not necessarily represent the outcomes that would be observed in the absence of the treatment intervention, calling into question the external validity of the experiment.

(ii) *Contamination* or *cross-over effects*: The problem of contamination occurs if some of the controls that were randomly excluded from treatment are nonetheless able to receive the treatment and/or some members of the treatment group do not receive treatment.

(iii) *Dropout*: Some of the treatment group may drop out before completing the program, sometimes at a very early stage. In that case, the offer of treatment was randomized but not whether individuals completed the treatment.

(iv) *Sample attrition*: Random assignment ensures that the treatment and control groups are comparable at the start of the experiment. However, people cannot usually be compelled to participate in the program over a longer term or to respond to surveys. For this reason, the treatment and control groups may become less comparable over time due to nonrandom attrition. Oftentimes, people receiving the treatment have higher response rates on surveys than people who were excluded from the treatment, because they are happy to have been included in the program. When there is nonrandom program attrition that differentially affects the treatment and control groups, a nonexperimental evaluation method usually needs to be used address bias concerns.

Another important issue concerning experimental evaluations is the issue of internal verses external validity. If the experimental protocol was followed and the potential problems

³⁶Landsberger (1968).

described by (i)-(iv) are not that significant, then the experiment could be considered to be internally valid. However, extrapolating the results of the experiment to a larger population of interest requires external validity. That is, the sample participating in the experiment should be representative of the population of interest, especially if it is expected that people will respond to treatment in different ways. If the sample participating in the experiment is not similar, for example, if the sample in the experiment is younger, poorer or more likely to be female, then statistical adjustment can sometimes be used to extrapolate from the experimental results to the larger population of interest. A difficulty arises, though, if the sample in the experiment is dissimilar in unobservable ways, for example, if the sample that signed up to participate in the experiment subsidy experiment is more motivated.

Lastly, most field experiments in economics are run for fairly short periods of time (at most 2-3 years). They usually do not permit an evaluation of programs for longer periods, and may, in addition, be affected by “pioneer” effects, stemming from the program not having been in operation for some time. (On this point, see Behrman and King, 2008, 2009). For a recent critical view on the value of randomized control trials in economic development, see Deaton (2009). In this paper, we present evidence from both experimental and nonexperimental studies, in recognition of the fact that both approaches have relative strengths and limitations.

Table 1
Summary of Impacts for Employment Creation and Job Training Programs for Adults

Program	Program Description	Targeting	Evaluation approach	Impacts
<i>Proempleo Program</i> in Argentina (Galasso, Ravaillion and Salvia, 2001)	Provided skill training and/or vouchers for workfare participants that they could give prospective employers (wage subsidy that lasted 18 months)	Beneficiaries of temporary employment programs	Randomized experiment, with two treatments (training offer plus voucher or voucher only) and a control group.	Voucher reduced probability of unemployment, despite fact that few firms made use of the voucher. Women and younger workers had largest impacts. Private sector employment was 15% for voucher recipients compared to 9% for controls.
<i>Trabajar II</i> program in Argentina (Jalan and Ravaillion, 2003)	Local governments and nongovernmental organizations submit proposals for “socially useful projects,” which are then ranked according to some criteria. Projects last at most six months.	Workers had to be affiliated to a successful project and could not be receiving unemployment benefits of participating in some other training program.	Nearest neighbor propensity score matching	Average gain in household income for participants was \$103 dollars, with greater gains for younger people.
<i>Jefe</i> program in Argentina (Galasso and Ravaillion (2004))	Program participants work 20 hours per week in community work, training, school attendance or employment at a private company and in exchange received direct income support.	Heads of households with dependents who became unemployed as a result of Argentina’s economic crisis in 2003.	Cross-sectional and difference-in-difference propensity score matching	Study found substantial leakage to ineligible, but the program was still well targeted at the poor. 26% of Jefes participants would have been unemployed were it not for the program and 23% would have been inactive.
<i>Probecat</i> program in Mexico (Revenga, Riboud, and Tan (1994))	Short-term vocational education training	Program offered to 250,000 registered unemployed people age 20-55, people selected on basis of eligibility index	Cox proportional hazards model estimated on participants and matched nonparticipants (based on propensity score)	Program trainees found jobs more quickly. Impacts mainly for trainees older than 25 with work experience. Cost effective for women over age 25 but not for younger women.
ALMP Programs in Russia (Benus, Brinza, Cuica, Denisova and Kartseva (2005))	Training and unemployment benefits	Registered unemployed. Unemployment benefit requires working at least 26 weeks last year. 45% of sample had university degree.	Propensity score matching, where nonparticipants are those who applied for training but were not selected	No significant effects of ALMP programs.

ALMP Programs in Romania (Benus, Brinza, Cuica, Denisova and Kartseva (2005))	Training, including a public service component, where local government and other eligible organizations propose public projects and hire ALMP participants to work on them.	Registered unemployed, having income less than 50% of min wage, employed 6 months of last 12 or recent graduate	Propensity score matching, where nonparticipants are those who applied for training but were not selected.	Statistically significant effects on the likelihood of employment, the likelihood of being employed at least once, and on wage levels. Retraining increases the probability of employment and decreases the wage for females. Middle aged had biggest impacts. Program not beneficial for highly educated.
ALMP Programs in Slovakia (Lubyova and Ours (1999))	Two ALMP Programs providing retraining and counseling services and wage subsidies at either socially purposeful jobs (up to 2 years duration) or publicly useful jobs (up to 6 months duration at public works type job)	Unemployed workers	Model the duration of unemployment and duration of stay in an ALMP program.	Workers that enter ALMP have a 150% increase in the exit rate into a regular job. Benefits of retraining were only observed for socially purposeful jobs.
ALMP Programs in Poland (Kluve, Lehmann, Schmidt (1998))	Three forms of training: publicly financed training and retraining, wage subsidies for workers in private or public firms, and public works. Courses lasted 2-3 months. People receive unemployment benefits during training.	Unemployed workers offered program at their local labor office.	Difference-in-difference matching, where outcome is labor force status. Control group consists of people who had been registered unemployed over same time period.	Find training/program increase average employment probability for both men and women over both the short and medium term. Non-training ALMP did not have a positive benefit.
<i>Reemployment Project program in China</i> (Bidani, Coh and O'Leary (2002))	Retraining and job search assistance to promote labor market entry of people laid off by state-owned enterprises. One month training courses in computer training, beauty and massage, hair cutting, sewing, toy making, cooking, repair training and driver education. Class sizes during the training sessions were often large with 200 to 300 workers in a small classroom.	Administered by local labor bureaus. In two cities: Shenyang and Wuhan.	Three different treatment/comparison group samples are analyzed using multiple methodologies that include prop score matching, matching on odds-ratios, and OLS.	Program found to have a negative effect on employment probability and no effect on earnings in Shenyang, but a positive impact on employment probability in Wuhan.

Table 2
Summary of Impacts of Programs to Support Youth Employment

Program	Program Description	Targeting	Evaluation approach	Impacts
<i>Programa Joven</i> in Argentina (Aedo and Nunez (2004))	Provides an average of 200 hours of training over 14-20 weeks (classroom followed by internship), a monetary subsidy for females with children, transportation expenses, medical checkups, books, materials and work clothing.	Young people from poor households with low education levels, little work experience, who were unemployed or inactive	Cross-sectional propensity score matching, using program beneficiaries and eligible nonparticipants	Effects on employment for women age 21-35 in the range of 9-12 percentage points, but not for younger females (age less than 21) Impacts on earnings of around US \$20-25 per month. Positive rates of return only if benefits are long-lasting (9 years or more).
<i>ProJoven</i> program in Peru (Nopo, Robels and Saavedra (2007))	Provides classroom training and internships lasting three months. Trainees receive stipend during training period, with mothers of young children receiving double stipend. Focus on training females for traditionally male occupations.	Youths from poor families.	Two-stage matching procedure that first selects for each treated a matched control on the basis of similarity in preprogram hourly wages.	Positive employment impacts for women of 6% at 12 months and 15% at 18 months (in contrast to negative impacts for men). After 18 months, beneficiary females generate 93 percent ore labor income than their control counterparts. Decrease in measures of occupational segregation.
<i>Jovenes en Accion</i> program in Colombia (Attanasio, Kugler and Meghir (2008))	3 months of classroom training followed by 3 months of OJT in the form of an unpaid internship, with 1009 companies participating.	Youth age 18-25 living in urban areas of Colombia in 2001-2005 and who were unemployed and came from families in the two lowest deciles of the income distribution.	Randomized experiment – Training institutions selected larger number of applicants than they had slots for and one-third were randomized out into a control group.	Increased earnings and employment for both men and women, with the largest effects for women. Earnings for women increased on average 18%. Both men and women who were offered training (intent-to-treat estimates) were more likely to be employed and to have jobs that offered benefits and to have a formal wage contract. Benefits of OJT found to be higher than that of classroom training.
<i>Juventud y Empleo</i> program in the Dominican Republic (Card et. al. (2007))	Up to 350 hours of classroom training followed by an unpaid internship at a private sector firm. Participants received a stipend of \$2 per day during the program.	Less educated youth (average age 22.3 at baseline).	Randomized experiment. Program applicants in 2004 were randomized into treatment and control groups.	Comparison a10-14 months after training showed no statistically significant effects on employment, earnings or hours worked, even when disaggregated by gender, age and education. Find a 10% impact on hourly wages and that the jobs of male participants were more likely to provide health care coverage.

Table 2 Continued
Summary of Impacts of Programs to Support Youth Employment

Program	Program Description	Targeting	Evaluation approach	Impacts
<i>Jordan New Opportunities for Women (NOW) program</i>	Provides employability training (to improve communication and basic job skills) and/or job vouchers, which are short-term monetary incentives for firms to hire workers (US \$210 per month for maximum of six months)	Young female community college graduates	600 women randomly selected to receive employability training, job vouchers (incentives for firms to hire them) or both.	Employment rates among graduates who received vouchers alone or vouchers plus training are 55-57 percent, compared to 17-19 percent in the group that received only training or neither training nor vouchers.

Table 3
Child Care Programs

Program	Program Description	Targeting	Evaluation approach	Impacts
<i>Guatemala Community Day Care & Nutrition Program</i> (Ruel and Quisumbing (2006))	Parents select one mother to be day care provider for up to 10 children, enabling other mothers to work	Guatemala city, with poverty-based eligibility criteria. Mothers had to work to qualify for the program.	Matching	Improved children's diets and increased mothers' incomes by 30%. Largest impacts observed for younger and older women with low levels of education.
<i>Hogares Comunitarios</i> in Colombia (Attanasio and Vera-Hernandez (2004))	Parents select one mother to be day care provider for around 15 children	Poor households, based on an eligibility score.	Instrumental Variables using distance to nearest center as an instrument	Large, positive impacts on female employment and numbers of hours worked. Prob of employment increases from 0.12 to 0.37 and number of hours worked increased by 75 per month. Positive impacts on children's heights, school-going and grades achieved.
Preschool building program in Argentina (Berlinski and Galiani (2007))	Between 1994-2000, created 175,000 new preschool places.	Middle income households living in urban areas	Difference-in-differences, exploiting differences across regions and cohorts in program exposure	Full take-up of new facilities and an increase in the probability of employment from 7 to 14 percentage points.
Publicly provided child care in Brazil Barros et. al. (2011)		Study of Rio de Janeiro municipality	Program assignment originally based on a lottery, with lottery winners getting program and losers being put on waiting list	Large increase in the use of care, from 51 to 94 percent. Increase in mothers' employment from 36 to 46 percent, with an even larger increase in subgroup of mothers who were not working in the six months before the lottery for whom employment rates increased by 97 percent.
Free preschool for children age 3-4 in Israel	Analyzes effect of the Compulsory and Free Preschool Law	Program for all children, but preschools were first made available in poor areas. The analysis sample consists mainly of Arab mothers.	Difference-in-differences comparing 11 treated towns and 13 untreated towns.	Preschool enrollment and mother's labor supply both increase sharply after introduction of the policy, especially for more educated mothers. LFP rate increased by 0.07 percentage points for mothers of children age 2-4 (who participate at a rate of 17.1 percent) with no change in fertility.
Baldawi study of Self-employed Women's Association (SEWA) program in India	Program provides health care, child care, banking and insurance services	Study examines outcomes of mothers and children who attended centers in two districts 10 years ago between ages 11-17.	Cross-sectional comparisons of participants and nonparticipants and participants' reports on how program changed their lives.	About 42-53% of mothers reported that program improved their lives in terms of increasing their productivity, increasing employment days, increasing family savings.