UNIVERSITY OF LONDON THESIS

Degree PhD Year 2007 Name of Author DE GIORGI, Giacomo

DIRECT AND INDIRECT EFFECTS OF PUBLIC POLICIES

by
Giacomo De Giorgi

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in The University College London
2006

Doctoral Committee:

Professor Richard Blundell, Chair
Professor Martin Browning
Professor John van Reenen
Assistant Professor Pedro Carneiro

I, Giacomo De Giorgi, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

I would like to dedicate this Dissertation to my Mother, my Father, Paola my wife and my sisters Giusy and Marta.

ACKNOWLEDGEMENTS

A special thank goes to my advisor Richard Blundell whose continuous suggestions and encouragement have been invaluable. Richard has been incredibly supportive in the various phases of my doctoral studies and hopefully I have learnt a great deal from him. I have deeply benefitted from several conversations with Pedro Carneiro. I would also like to thank Orazio Attanasio for his substantial insights across continents. Costas Meghir motivated me in a number of occasions. A particular acknowledgement goes to Manuela Angelucci, indeed Chapther IV is the result of our collaboration. Imran Rasul has generously and thoughtfully reviewed part of this work, his comments have greatly improved this dissertation. I am grateful to Wendy Carlin for her encouragement. I also wish to thank the IFS for hosting me during part of this project. I gratefully acknowledge financial support from the Fondazione Einaudi and the Royal Economic Society. I would also like to thank my colleagues at UCL, especially Andreas, Giovanni, Ifty, Katrien, Marco, Mario and the UCL staff. A special thank goes to the Clapham Uncommons (my football mates) and to my friends Alessandra, Chris, Elena, Mihalis, Paolino, Piergiorgio and Vladi.

I have to express here my deepest gratitude to my wife Paola for the incredible support and fundamental remarks on my work and life.

Finally, I am most indebted to my father and my family who made this possible.

ABSTRACT

Direct and Indirect Effects of Public Policies

by

Giacomo De Giorgi

Chair: Richard Blundell

This dissertation analyzes two important recent public policies and sheds light on the direct and indirect impacts of both; and more in general on the need of sound evaluations to carefully consider possible spillovers as a fundamental component of programs.

In the first two essays I focus on the "New Deal for Young People (NDYP) in the UK", a multiple treatments policy launched in 1998. Such analysis is conducted on two main margins: i. employment; and ii. job quality as wage returns. I discuss possible indirect effects as substitution of workers and general equilibrium effects on wages. The main results are that the policy enhances the (re)employment probability of participants by 5% and the effect lasts over several cohorts. Further, no evidence of displacement as well as general equilibrium effects is found. On the relative effectiveness of the different treatments: IV estimates of the treatment effect of the subsidized employment option clearly point towards a negative and significant

penalty of roughly 20%, due to stigma or worsening of the matching function. The other treatments do not seem to have a significantly differential impact.

Chapter IV analyzes an aid policy implemented in rural Mexico: *PROGRESA*. The focus of the chapter is the identification of the spillovers of the policy to those households who just happen to live in treated communities, but are not eligible for the program. A formal definition of the Indirect Treatment Effect (ITE) parameter is given. Focusing on consumption the mechanisms generating such indirect effects are discussed. Those mechanisms need not be limited to the specific policy. In particular, in developing Countries policies are likely to affect all residents of the areas where they are implemented, especially when village economies and social networks create strong links between a limited number of households. The large liquidity injections into small communities increase the consumption of the non-treated through changes in the credit and insurance markets. Thus, the total effect of the policy is larger than its effect on the treated. Further, the results confirm that a key identifying assumption - that the program has no effect on non-treated individuals (SUTVA)- is likely to be violated in similar policy designs.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

CHAPTER I

Introduction

This dissertation analyzes two important recent public policies and sheds light on the direct and indirect impacts of both; and more in general on the need of sound evaluations to carefully consider possible spillovers as a fundamental component of programs. Furthermore, whenever possible the mechanisms through which such external effects occur are analyzed. It is crucial for policy evaluation and design to understand what are the complex effects of the policy at hand. There are a number of reasons why one should not focus only on the direct or targeted outcomes of a given program. Firstly, in case of substantial spillover or general equilibrium effects the validity of the evaluation might be undermined by the violation of a fundamental identifying assumption, i.e. Stable Unit Treatment Value Assumption or SUTVA formulated by Rubin (1980 and 1986).[1] If such assumption is violated and wrongly assumed to hold the parameter estimated might be meaningless. Furthermore, a thorough understanding of the process generating particular spillovers is crucial in the design of public policies. In addition, policies judged not to be worthwhile in a cost and benefit analysis based on direct impacts might be worth implementing if positive spillovers are found; obviously the reverse argument might be true in other

[1] Briefly, SUTVA rules out any interaction between treated and control units, i.e. the potential outcome of a unit only depends on her treatment status and not on the other units.

circumstances. Indeed, in this work I present two possibly opposite cases of indirect effects: one positive, one negative. In Chapter II the type of externality which might arise is negative to those individuals who did not participate in the policy, i.e. an employer who employs a program participant might receive a subsidy, which renders such individual cheaper than an otherwise identical non-participant; this might in turn generate a displacement effect to the non-treated individual since she would have possibly got the job had the program (subsidy) not existed. On the other hand, in Chapter IV, I analyze a case of a positive externality given by a substantial cash injection in poor rural communities. In that particular case credit constraints are eased and therefore Consumption increases for those households who do not participate in the program.

Calmfors (1994) highlights the importance of general equilibrium effects in a cost-benefit analysis of a public policy. I test for general equilibrium effects without imposing structure on the data, while on the other hand I am only able to pin down a limited number of mechanisms in a not so long time-horizon. I am not able to analyze two different equilibria, but rather a possible transition between two different states of the world. Therefore the exercises I propose here are somewhere between a partial approach and a general equilibrium setting, although not as ambitious from the structure side as Heckman *et al.*, 1998; or Lise *et al.*, 2005a and 2005b.

In general, public policies have been evaluated either using a randomized experiment or a quasi-experimental approach often claiming that the parameters estimated through small scale programs could be considered valid if the policy were to be extended to a much larger scale. However, it is likely that policies of global implementation have general equilibrium effects and in such sense it is plausible that the SUTVA might not hold when the program is extended. On the other hand the

SUTVA might not hold even in small scale programs if the policy at hand hits particular mechanisms or relaxes pre-existing binding constraints, as discussed in Chapter IV.

There is a large and growing literature on the direct effects of public policies; Heckman *et al.* (1999) provide a number of references on such strand of the literature.

There are however only a limited number of papers in the indirect effects literature. Such literature can be roughly grouped into papers that use sources of exogenous variation to identify peer effects, and papers that set up and calibrate structural models to estimate general equilibrium effects. The first group includes Philipson (2000), Katz *et al.* (2001), Sacerdote (2001), Duflo and Saez (2002), and Miguel and Kremer (2004). Bobonis and Finan (2005) and Lalive and Cattaneo (2005) use PROGRESA data to estimate peer effects in schooling decisions. The second group includes, among others, Blundell *et al.* (2003), the already cited, Heckman *et al.* (1998) and Lise *et al.* (2005a and 2005b).

In the first two essays I focus on the "New Deal for Young People (NDYP) in the UK", the battle-horse employment policy launched by the newly elected labor government in 1998 and still ongoing to date; Blundell *et al.*, 2004. Such analysis is conducted on two main margins: i. employment; and ii. job quality as wage returns. Possible indirect effects are discussed and analyzed in Chapter II. The hypotheses tested in such respect, are substitution of workers and general equilibrium effects on wages arising from a substantial increase in the effective labor force participation.

Chapter IV analyzes a public policy intervention to alleviate poverty, improve health and foster human capital accumulation in Mexico (Skoufias *et al.*, 1999a and 1999b): *PROGRESA*. The focus of the chapter is the identification of the spillovers of the policy to those households who just happen to live in treated communities, but are not eligible for the program. A formal definition of the Indirect Treatment Effect (ITE) parameter is given; and the analysis centers on the consumption spillover while a detailed analysis of the mechanisms generating such indirect effects is presented.

The analysis of the first two essays fits in the large and still growing literature on Active Labor Market Policies (ALMPs) evaluation. In the past fifteen years there has been a renewed interest both in Europe and the US on such policies. Many European countries have experienced a vast array of programs addressed to fostering human capital, search intensity and job attachment. In Sweden (Sianesi, 2004) a series of policies, contemplating all the above features, were implemented in response to the deep crisis of the beginning of the 1990s. In Switzerland (Gerfin and Lechner, 2000) there are examples of subsidized employment and job creation schemes. The US experience on ALMP's is also vast. Hotz *et al.* (2000) in revising the influential GAIN program in California shed some light on the effectiveness of human capital accumulation and job first approach in a long term perspective. Other US evidence focuses on subsidies to those unemployed who move out of welfare as the EITC (see Eissa and Liebman, 1996; or the survey of different subsidies to employment by Katz, 1998). Other policies, such as the JTPA, focus on the improvement of human capital; the work of Bloom *et al.* (1997) provide evidence on the effectiveness of classroom and on-the-job training as well as of job search assistance. Nonetheless, despite the large amount of work on ALMPs there is still no consensus on whether such policies help youths to get out of welfare and back to work (Heckman *et al.*, 1999). Furthermore, it is not clear which one of the above mentioned approaches is the most successful.

An earlier paper on the NDYP by Blundell *et al.* (2004) covers the initial launch of the program focusing mainly on the job search assistance component of the policy.

They found a positive effect of about 10% on the (re)employment probability within the Gateway for the first cohort of participants, while half the effect was found for some later cohorts. It is then crucial to extend such analysis to a long-term horizon considering as well the various aspects of the policy. Therefore, in the first essay I analyze the multiple treatments offered as a whole (while the different components are the topic of Chapter III).

The main questions answered in Chapter II are: is the policy really improving employment prospects for young males? Are the effects of the policy lasting over different cohorts? Are there substantial spillover to untreated individuals in terms of substitution on the job or equilibrium wages?

The NDYP is the major welfare-to-work program in the UK, and targets the 18-24 year old unemployed (on Job Seeker Allowance, JSA for at least 6 months). The policy is mandatory, and the sanction for non compliers is the withdrawal, at least temporarily, from the benefit. It is composed of two main parts: i. intensive job search (Gateway), common to every participant; and ii. an option period spanning from education/training to an employment subsidy or some public placement to be compulsorily taken if a regular job is not found within four months in the Gateway stage.[2]

According to politicians and program administrators we are looking at a success story in roughly all its components. Here is part of a piece written by Andrew Smith (2004) the former Secretary of State at the Department for Work and Pensions:

"The Government investment in the New Deal and Jobcentre Plus has helped to deliver one of the most effective labour market programmes in the World....".

While a program participant states (www.newdeal.gov.uk):

[2] The details of the policy are given in Section 2.2.

"If it weren't for New Deal, I wouldn't be here now. They helped me and they pushed me when I needed it. I've got a lot more confidence and I've got skills."

The analysis is limited to males as the vast majority (75%) of participants are; furthermore, the NDYP is basically the only program available to young males, while there are other programs for females difficult to tell apart with the available data. There are several potential outcomes, one could in principle look at, i.e. probability of being employed at some point in time or probability of gaining employment in a given interval. I concentrate on the (re)employment probability within one year of entering the program. Such an outcome allows to evaluate the effectiveness of the entire program not distinguishing among different types of treatment. Treatment is here understood as a combination of job search assistance, training, subsidies and some work experience (voluntary sector or environmental services).

As standard in the evaluation literature the problem reduces to that of missing outcome (Heckman *et al.*, 1999; Blundell and Costa-Dias, 2000). A given individual cannot be in two different states at the same time. He/she is either in the program or out of it. Therefore I have to identify a suitable missing counterfactual. In a non experimental study exercise, such a problem is exacerbated due to the nonexistence of an administered control group and in the specific case due to the global implementation of the program: everyone in the UK who is younger than 25 after 6 months in open unemployment is mechanically pushed through the program. The approach (Section, 2.3) followed in this Chapter is that of a Regression Discontinuity (RD) design since it seems to be rather appropriate given the sharp eligibility rule (six months of JSA, plus younger than 25). The intuition behind such an approach is that participation changes according to a known deterministic function at a discontinuity point. Unemployed slightly younger than 25 are in, while those slightly older are out. There are not other differences, apart from the treatment status, between treated and untreated in the neighborhood of the discontinuity. This gives rise to a natural comparison, nonparametrically recovering a Local Average Treatment Effect parameter (LATE), i.e. the effect of the policy for those near the cut-off point, under a very weak continuity assumption.[3]

Consistently with the non parametric identification strategy, the estimation (Section, 2.5) is implemented by Local Linear Regression (LLR) known to have desirable boundary properties (Fan, 1992; Porter, 2003). I am ultimately estimating at a boundary point, where the size of the discontinuity is the parameter of interest. The difference between the two conditional mean functions from both sides of the discontinuity will recover the LATE.

There is convincing evidence (Tables 2.2 and 2.3) that individuals who were born only few days apart one another are fundamentally identical, and would have had the same performance in the labor market had the program not existed (Figure 2.5).

The appealing feature of the RD design, in this work, is that it allows comparing similar individuals who only differ slightly in their date of birth. However, the high comparability on the one hand permits to relax the identification strategy and on the other hand might raise concerns on the possible substitutability of those individuals, i.e. participants could displace controls. If it is the case that participants do substitute controls the SUTVA would be violated and the parameter estimated would not recover the causal effect of the program, but rather an upward biased estimate of it as in equation 2.3. A considerable part of the essay is devoted to the analysis of possible indirect effects as the substitution bias (Section 2.3 and 2.8) and no evidence of possible substitution effects is found (Section 2.7).

[3]Non-treatment potential outcome to be continuous at the cut-off point.

In the large class of indirect effects one could consider standard General Equilibrium (GE) effects of a policy, i.e. the effective labor supply could increase (participants really look for a job) easing the wage pressure and therefore delivering a lower equilibrium wage and as a consequence an increase in overall employment. However, such phenomenon requires a rise in the overall labor supply. Although the NDYP is a policy of global implementation, the number of individuals involved relative to the active male population does not seem capable of initiating important GE mechanisms (Figures 2.12 and 2.13). For the particular parameter estimated the GE effect mentioned should not be relevant given that for those individuals at the margin between participation and non-participation the possible rise in employment should be symmetric. Furthermore, given the cohort specific approach, I am able to add evidence on negligible GE effects; the GE mechanisms should kick in the more individuals are involved overtime, however the point estimates (Table 2.5) do not support the claim of relevant GE effects since they are quite stable overtime though the number of individuals involved changes.

The main result is that the combination of treatments offered in the NDYP enhances the (re)employment probability by 5%, at least in the neighborhood of the cut-off point. The effect lasts over several cohorts of new dealers even after 5 years from its launch. Further, no evidence of displacement as well as general equilibrium effects is found.

Having analyzed the program as a whole in terms of its effectiveness in enhancing employment chances for young unemployed, and discussed possible spillovers of the policy in such context, it seems natural to extend the analysis in order to shed some light on the impacts of the different treatments offered as well as on the quality of the jobs gained by the participants. Such task is accomplished in chapter III. The

NDYP is a prototypical example of a multiple treatment program. It combines, at different stages, job search assistance, training/education, subsidies and reinstatement in the labor force through governmental or voluntary sector jobs (Section 3.2). Such particular framework has been employed in several labor market programs (see Katz, 1998; Sianesi, 2001 and 2004; Frölich, 2004). It is widely acknowledged that Active Labor Market Policies (ALMPs) have to be flexible and capable of improving employability. However, it remains an open issue whether different treatments are generally beneficial or only some of them are, while others are a pure burden to the system. Chapter III addresses the following questions: are the different treatments offered by the NDYP equally valuable in terms of returns? Or it is rather the case that a particular option is delivering a higher return in terms of wage once the treated is out of the program? It is crucial to understand how to shape a successful ALMP given the government budget constraint and the limited availability of resources. On the other hand given the heterogeneity in the nature of the unemployed population a certain degree of flexibility is necessary in order to give a valuable treatment to the particular individual.

There is some evidence that the subsidy to employment is the 'star' option of the NDYP with respect to enhancing the (re)employment probability of participants (Dorsett, 2006) however there isn't any available study that looks into job quality and namely wages. The mere employment is not per se an indicator of the success of a program, especially when there might be concerns over the quality of jobs. In the particular case the star option could be successful in getting unemployed out of dole, but the point is: are these jobs worthwhile? A subsidy to employment is relevant when the productivity of the worker is possibly below his/her cost. However, such subsidy cannot last forever. If the subsidy takers who found a job are of lower quality than the average in the population then this should be reflected in their salary once a regular job is found. On the other hand it is also possible that the subsidy is wrongly signalling to the market the type of the agent and therefore such lower wage should disappear once the employer learns about the quality of the participant; unfortunately it is not possible with the available data to test such hypothesis. Since option assignment cannot be considered a random process a convincing identification strategy has to be devised if one wants to infer the causal impact of the different options on regular wages. In principle the assignment process should have been a joint decision between the caseworker and the participant. However, this was not the case for two main reasons: firstly, certain options were simply not available in certain areas, i.e. some local units of delivery did not have the possibility of placing a participant in the environmental task force, while others did not have voluntary sectors job available and so on (*rationing*). Secondly, there is a clear pattern of preference for a particular treatment in certain units, this might depend on the fact that placing someone in a subsidized job is simply more expensive than sending someone to school, both in terms of effort to be exerted by the caseworker and monetary cost (*costs*). There is a large variation in option take-up across different UoD's substantiated by anecdotal and formal evidence later in the paper (Section IDENTIFICATION). Such variation remains even when a number of confounding factors are partialled out. Furthermore, there is evidence of non-random option allocations in Table 3.2. It is clear how better quality participants were assigned to the subsidized employment option. In fact, they are significantly better in terms of schooling, ability (reading/math problems), work history (although they have surprisingly longer unemployment) than their counterpart who engaged in an extended job search treatment. If one where to take option assignment as a random process and consequently estimate the option effects through a simple OLS, the results would be that the average effect of a subsidized placement is not distinguishable form an extended search period. However, when non-random selection both at the options stage and in the employment node is taken into account a dramatically different picture emerges. IV estimates clearly point towards a negative and significant penalty of roughly 20%. Such results is in line with earlier findings on subsidized employment by Katz (1996). The intuition for a negative return is that of a stigma effect attached to those participants who got a subsidy in order to be employable. It seems therefore that the specific option is rather signalling to the market (might be wrongly) the low productivity type. It might also be that such option forces unemployed into low wage jobs or jobs that are not particularly suited for the given unemployed altering then the matching process. Unfortunately, I could not test the hypothesis of a temporary effects given the available data.

The third essay analyzes the indirect effects of PROGRESA, a public aid policy implemented in poor rural villages in Mexico. It sheds light on the relevant mechanisms generating such externalities; those mechanisms might not only be limited to the specific policy, they might be relevant in a number of other contexts where certain conditions apply as detailed later.

Policy interventions in developing countries are likely to affect all residents of the areas where they are implemented, especially when village economies and social networks create strong links between a limited number of households. The unique randomized design of PROGRESA is exploited to estimate its indirect effect on consumption for non-eligible households who live in treatment areas, and to understand the mechanisms through which this indirect effect occurs. Liquidity injections into small rural communities increase the consumption of the non-treated through changes in the credit and insurance markets. Thus, the total effect of the policy is larger than its effect on the treated. Further, the results confirm that a key identifying assumption - that the program has no effect on non-treated individuals (SUTVA)- is likely to be violated in similar policy designs.

Conditional cash transfers are a popular type of aid program, which provides monetary transfers to eligible recipients, provided they send their children to school, attend nutrition classes, and have periodic health checks. Programs with this format are currently implemented in numerous countries, including Bangladesh, Bolivia, Brazil, Colombia, Honduras, Jamaica, Mexico, and Nicaragua. There are talks of implementing similar programs in Cina and New York City as well.[4]

The design of the experimental trial and the data collected for the evaluation have some unique features. First, the randomization was implemented at the village level. Second, program administrators collected data on all households, both poor and non-poor, although only poor households were eligible for the treatment. Thus, we have information on four groups: poor and non-poor households in treatment and control villages. Non-poor households in control villages provide a valid counterfactual for the non-poor in treatment ones.

We focus on consumption because it provides an indicator of household well-being. We find that there is a positive, significant, ad sizeable indirect program effect on consumption for non-eligible families. Further, we study the mechanisms that lead to this increase in consumption. For example, the implementation of PROGRESA may modify labor supply, altering equilibrium wages, or it may increase goods prices through higher demand. We find that there are no significant indirect effects on labor earnings, prices (with the exception of increases in few food items in 1998),

[4]See recent article (October 9, 2006) by Bob Herbert in the New York Times: Cash with a Catch, http://topics.nytimes.com/top/opinion/editorialsandoped/oped/columnists/bobherbert/index.html?8qa.

and welfare receipt, and that sales of agricultural products decrease. Therefore, we rule out the hypothesis that the indirect program effect on consumption is generated by an increase in current income. Instead, we show that non-poor households in treatment villages consume more by receiving more transfers, by borrowing more money - almost exclusively from family, friends, or informal moneylenders - and by reducing their stocks of grains and animals. In addition, we show that the indirect program effects on consumption and loans are larger for households hit by a negative idiosyncratic shock. Thus, we conclude that cash transfers in treatment villages indirectly benefit non-treated households by improving consumption smoothing. These results correspond to our knowledge of developing countries, where credit and insurance occur through informal networks of family, friends, and neighbors. Positive income shocks to some households benefit the whole network, whose other members receive larger loans and transfers, especially the ones hit by negative shocks. The availability of additional liquidity in the network enables households to reduce their savings.

While it is often difficult to predict the effects of a nationwide program using data from limited geographic areas, the effects on the credit and insurance market should not be a function of the number of treated villages, as long as social networks are village-specific.

It is possible to learn from this exercise that when the distance (economic, social or geographic) between treatment and control group is small, and when the treatment group is a large fraction of the local economy, the SUTVA may be less likely to hold.

In sum, the essay contributes to different literatures: i. consumption smoothing and credit and insurance markets in low-income economies. Where the main references in the risk-sharing literature are the work of Deaton (1991), Townsend (1994, 1995a, 1995b), Udry (1994, 1995), Banerjee *et al.* (2003), and Banerjee (2004), among others. ii. The essay contributes to the program evaluation literature in multiple ways: first, it shows that a class of widely implemented aid policies has important positive externalities; second, a substantial attempt is made to extrapolate the indirect effects of a nationwide conditional cash transfer program in the credit and insurance market. Third, the essay provides an example of the failure of the SUTVA, which is usually non-testable.

CHAPTER II

Long-Term Effects of a Mandatory Multistage Policy: the New Deal for Young People in the UK

2.1 Introduction

In the past fifteen years there has been a growing interest both in Europe and the US on active labor market policies (ALMPs). Many European countries have experienced a vast array of policies addressed to fostering human capital, search intensity and job attachment. In Sweden (Sianesi, 2004) a series of policies, contemplating all the above features, were implemented in response to the deep crisis of the beginning of the 1990s. In Switzerland (Gerfin and Lechner, 2000) there are examples of subsidized employment and job creation schemes. The US experience on ALMP's is also vast. Hotz et al. (2000) in revising the influential GAIN program in California shed some light on the effectiveness of human capital accumulation and job first approach in a long term perspective. Other US evidence focuses on subsidies to those unemployed who move out of welfare as the EITC (see Eissa and Liebman, 1996; or the survey of different subsidies to employment by Katz, 1998). Other policies, such as the JTPA, focus on the improvement of human capital; the work of Bloom et al. (1997) provides evidence on the effectiveness of classroom and on-the-job training as well as of job search assistance. Nonetheless, despite the large amount of work on ALMPs there is still no consensus on whether such policies help youths to get out of welfare and back to work (Heckman et al., 1999). Furthermore, it is not clear which one of the above mentioned approaches is the most successful.

In the UK, the newly elected Labor government launched in January 1998 the New Deal initiative, of which the New Deal for Young People (NDYP) is the largest component: so far it has involved over 1 million youths at a total cost of more than 2 billion British pounds (about 3.6 billion USD). The NDYP was initiated in selected areas (Pilot Period) and extended to the entire UK by April 1998 (National Roll-Out). The policy is targeted at 18 to 24 year old unemployed who have been receiving Job Seeker Allowance (JSA) for at least 6 months[1]. It is a mandatory program in all its components imposing a significant sanction to non compliers: the withdrawal, at least temporarily, of the unemployment benefit. The design of the program (Figure 3.1) illustrates how the policy was conceived to be flexible enough for tackling specific difficulties in getting a job. In the NDYP design (Section 2.2) individuals should have been in principle screened and helped according to the particular needs. The option period should have accomplished this duty, e.g. training and education supposedly devoted to those youths who were lacking basic skills.

An earlier paper by Blundell et al. (2004) covers the initial launch of the program and focuses on the job search assistance component of the policy, while I analyze the multiple treatments offered as a whole. Here treatment is understood as the combination of job search assistance, education and training, subsidized employment as well as job experience through voluntary sector or governmental (environmental task force) placements. The main question addressed in this work is whether the combination of the treatments above is effective in getting young unemployed back to work. Therefore I focus on the (re)employment impact of the NDYP within 12

[1]JSA is basically the only unemployment insurance in the UK, available to anyone who is able to work, ie. no previous employment history is required. Prior to the introduction of the New Deal there were no stringent conditions applied in order to receive it in principle indefinitely.

months since entering the program. The analysis is developed for males, the large majority of participants (75%), and does not cover the Pilot period[2]. Blundell et al. (2004) found a positive and significant impact of the job search component of the policy: the scheme enhanced the (re)employment chance of participant males by 10% for the Pilot group while such estimate halved to 5% for the first months of the National Roll-Out. The decaying effect could suggest a significant introductory effect due to vanish over time. Therefore I consider 5 different cohorts of 9 months each, defined according to the date of entry in the NDYP, spanning from April 1998 to December 2001.

The eligibility rules informing the program show a clear discontinuity in the participation function: only those unemployed younger than 25 by the time they reach the sixth month of JSA claim are eligible and treated[3]. The "sharp" discontinuity (Hahn et al., 2001) is exploited for identification of a meaningful policy parameter under very weak assumptions (see Section 2.3). It is possible, by comparing unemployed arbitrarily close to the discontinuity point, to (non-parametrically) identify the causal impact of the program for at least those individuals in the neighborhood of the cut-off point[4].

The intuition behind the identification strategy in the RD design is pretty simple: since treatment changes discontinuously at a threshold, as a step function of a continuous underlying variable (age), the only difference between those unemployed marginally below or above the cut-off is the treatment status and therefore in the neighborhood of the discontinuity point assignment to treatment is almost random. There is convincing evidence (Tables 2.2 and 2.3) that individuals who were born

[2]The long-term nature of the outcome considered would not allow a meaningful use of the initial Pilot.

[3]The program is mandatory in all its components.

[4]The local parameter could be "the" parameter if the idea under scrutiny is that of extending the program marginally or if the interest lies exactly on that subgroup of unemployed. The local parameter extends to an average treatment effect under the assumption of constant treatment effect.

only few days apart one another are fundamentally identical[5], and would have had the same performance in the labor market had the program not existed (Figure 2.5).

The appealing feature of the RD design, in this work, is that it allows comparing similar individuals who only differ slightly in their date of birth. However, the high comparability on the one hand permits to relax the identification strategy and on the other hand might raise concerns on the possible substitutability of those individuals, ie. treated could displace controls. If it is the case that participants do substitute controls the Stable Unit Treatment Value Assumption (SUTVA), as defined in Rubin (1980 and 1986), would be violated and the parameter estimated would not recover the causal effect of the program, but rather an upward biased estimate of it as in equation 2.3. A consistent part of the paper is therefore devoted to the analysis of the substitution bias (Section 2.3 and 2.8) and no evidence of possible substitution effects is found (Section 2.7).

Often neglected in the program evaluation literature is the discussion on possible General Equilibrium (GE) effects of a policy[6], eg. the effective labor supply could increase (participants really look for a job) easing the wage pressure and therefore delivering a lower equilibrium wage and as a consequence an increase in employment. However, such phenomenon requires a rise in the overall labor supply. Although the NDYP is a policy of global implementation, the number of individuals involved relative to the active male population does not seem capable of initiating important GE mechanisms (Figures 2.12 and 2.13). For the particular parameter estimated the GE effect mentioned should not be relevant given that for those individuals at the margin between participation and non-participation the possible rise in employment

[5]As we will see later there are a large number of observations close to the cut-off point. The above statement has to be considered in an average sense.

[6]There are few exceptions as the dynamic GE model proposed in Heckman et al. (1998) or to somewhat a lesser extent the recent work by Lise et al. (2005) or on a different approach Angelucci and De Giorgi (2005).

should be symmetric. Furthermore, given the cohort specific approach, I am able to add evidence on negligible GE effects; the GE mechanisms should kick in the more individuals are involved overtime, however the point estimates (Table 2.5) do not support the claim of relevant GE effects since they are quite stable overtime[7] though the number of individuals involved changes.

The main result is that the combination of treatments offered in the NDYP enhances the (re)employment probability by 5%, at least in the neighborhood of the cut-off point. The effect lasts over several cohorts of new dealers even after 5 years from its launch.

The rest of the paper is organized as follows: Section 2.2 describes various features of the program; Section 2.3 covers the identification strategy adopted; 2.4 carefully describes the data used; 2.5 describes the estimation strategy; 2.6 provides some montecarlo evidence on the performance of the estimator; 2.7 presents the results; 2.8 addresses the substitution puzzle and 2.9 concludes.

2.2 The Program

As from Figure 3.1 the NDYP is a sequential program, where different treatments are offered to the participants. Following a period of 6 months[8] in open unemployment 18 to 24 year olds (JSA recipients) are automatically transferred into the program in order to be still eligible for the benefit. It is therefore a mandatory policy administered to everyone in the UK who, after 6 months of unemployment, are aged between 18 and 24. The age composition of the participants (Figure 2.3) is fairly uniform between the established age bracket[9]; job seekers with at least 6

[7]But for the last one.

[8]Only a very small number of unemployed, not included in the analysis, can access the program earlier than the sixth month. This particular group is composed by ex-offenders, disable and unemployed lacking very basic skills (writing and reading difficulties).

[9]The age distribution of participants is constructed from administrative data (NDED) containing virtually all participants (Section 4.2), 18 year old unemployed included in the figure are however excluded from the empirical

months of unemployment history are placed in the program even shortly before their twenty-fifth birthday. During the open unemployment period there are not strict requirements imposed on the job seeker. However, at the sixth month deadline, it is not possible to avoid program participation and still receive the unemployment benefit. For the group of unemployed considered in this work, i.e. young males, the opportunity cost of loosing the JSA is rather high, given that this is the main available subsidy[10].

The first four months of treatment (Gateway period) are nominally[11] devoted to intensive job search assistance and some basic skill training, eg. CV writing. Participants are obliged to meet a personal mentor once every two weeks and they have to report and prove the actions taken in order to gain employment. Such actions typically consist of job applications, direct contact between possible employers and caseworker. Failure to comply with any of the program requirements may result in a benefit sanction and eventually the withdrawal of it.

While in the gateway participants receive a benefit equal to the JSA (about 40 British pounds[12] per week). If a regular job is not found during the gateway, a second phase follows: the options. On the basis of personal considerations, given individual characteristics, the caseworker agrees with the participant on the option to be taken[13]. The option period can last from 6 to 12 (full time training or education) months and is compulsory, ie. participants can not refuse to enter an option. Common practice among units of delivery was to try placing the unemployed in a subsidized job during the second month of treatment. In case of a subsidized place-

analysis given that they might still be in high-school.

[10]There are some forms of social assistance, whose relevance is however not comparable with the JSA.

[11]Nevertheless, in the data, some individuals enter an option during the gateway period. The first guidelines given by the government stated clearly that one could exit the gateway period only toward a regular job. Later, they were adjusted according to the *de facto* behavior.

[12]Roughly 74 USD.

[13]This is not always the case since certain units of delivery tend to "favor" a particular option.

ment (at least a 30 hours per week job), the treated receives the salary paid by the employer who gets, for a maximum of 6 months, a subsidy of 60 British pounds (about 110 USD) per week plus 750 British pounds (about 1400 USD) as a one-off payment for the compulsory (minimum) one day a week training to be provided[14]. The second option, education or training, is targeted at youths lacking basic skills and it can last up to 12 months. While attending such courses the unemployed still receives his JSA payment. Typically a program participant attends the local college and, in most cases, there is no distinction between a vocational course attended by a treated or anyone else in the community.

A third option is that of a voluntary sector job where the participant receives an amount at least identical to the JSA plus 400 British pounds (about 740 USD) spread over the 6 months. A typical placement would be shop assistant in a charity shop.

The same monetary treatment is granted in the fourth option: Environmental Task Force, basically a governmental job, meant to be the last possible placement. A participant would typically be involved in the maintenance of public parks.

Participants are allocated to these last two options in the third and fourth month of the gateway.

Eventually a third phase follows: the follow-through, essentially maximum of 13 weeks similar to the initial gateway. It consists of intensive job search as well as training courses to maintain the skills acquired during the option period.

[14]Such subsidy seems quite generous when compared to the sort of hourly rate (close to the minimum wage 4.5 British pounds or 8 USD) a typical participant would get. In a crude computation, the weekly subsidy plus the one-off payment would amount to about 50% of a weekly pay for a minimum wage worker, however the 750 British pounds would have to repay for the loss of production due to the minimum of one day training. Under very simple assumptions (perfectly competitive markets) those 750 British pounds would not be enough to compensate for that loss. In fact, taking the latter into account the subsidy would not be greater than 30%, but still generous though. However, job turnover could be itself quite costly making such an option not as appealing as it looks like at a first glance. This point seems to be confirmed by the low take up rate in the data, only a sixth of those entering an option would go for the subsidized job.

The program was launched in January 1998 in selected areas (pilot period) and then extended to the rest of the nation in April of that year (national roll-out). About 1 million young britons have been involved since the beginning to December 2003, of which roughly 75% are males.

As mentioned earlier the aim of this work is to quantify the long run impact of the program in terms of (re)employment probability. The outcome of interest is defined as a treatment effect in the "Black box" (the shaded area in Figure 3.1). This is because I do not distinguish among different stages of the program (gateway or options), but I concentrate on the program effect as a whole. Therefore, the focus is on the (re)employment probability within 12 months since entering the program (or 18 months since claiming JSA given at least 6 months of unemployment). The choice of such an outcome is determined by the interest in the causal impact of the whole program[15] and its long run effects. While Blundell et al. (2004) focused on the effect of job search assistance during the first year of the policy implementation. The 12 months limit arises from the fact that the control group I exploit is forced in a similar program (New Deal for Long Term Unemployed) after 18 months of open unemployment[16]. The latter would make a comparison on a longer time interval misleading.

Another important aspect to notice is that the program is one of global implementation and therefore there could be concerns about possible general equilibrium effects, dictated by the increase in the overall labor supply, denied by a partial equilibrium approach. However, if such effects are relevant they should be increasingly so as the program broadens and involves more and more individuals. I tackle this

[15]The vast majority (about 90%) of participants would have completed their first option by the twelfth months.

[16]Possible anticipation effects have been investigated for the controls and based on the results of a before/after the program comparison of the first 18 months of unemployment, I could not find evidence of such behavior in the control group. The strategy followed for testing such hypothesis is the same followed for the treated individuals as from Figures 2.8 and 2.9. Results are available on request.

issue relying upon a cohort specific approach, namely I analyze the impact of the program for 5 evenly spaced cohorts (of three quarters each) entering the program from April 1998 to December 2001[17]. An initial test of the importance of general equilibrium effects is given by the simple time path of the program impacts. On the other hand I have also to consider possible substitution effects, if 18-24 year old are good substitute for 25-30 then we should see the former replacing the latter and therefore the program effect would be amplified by the substitution effect. I approach this potential source of bias by looking at treated and controls before and after the program (Sections 2.3 and 2.8).

2.3 Identification Strategy

As explained in the previous sections, participation in the program is compulsory and established by a deterministic rule: 6 months of JSA plus younger than 25. It is not possible to remain in open unemployment further than 6 months and still receive the benefit. While no one who is older than 24 is allowed in as from Figure 2.3. At the completion of the sixth month of unemployment the eligible job seeker is automatically transferred in the program. This gives an immediate comparison group or a so called "sharp" Regression Discontinuity (RD) design (Hahn et al., 2001) where the discontinuity in the treatment is given by the age rule informing the program[18]. The RD design was introduced in the evaluation literature by Thistlethwaite and Campbell (1960), who analyzed the effect of student scholarships on later career and has been recently employed in Hahn et al. (1999) in the study of financial

[17] This limit is imposed by the available data.

[18] Provided that the individual behavior does not change due to the existence of the program, ie. this rules out any anticipatory effects. This issue is analyzed in Figures 2.8 and 2.9 where the non-parametric survival probabilities are plotted for potentially eligibles before and after the start of the program. As evident from those figures they cannot be consistent with anticipatory behaviors of any relevance. The exit rates from JSA are not different for potential participants after the program began. Formal tests on the equality of the survival functions cannot reject the null at conventional significance levels.

aid on college attendance and in DiNardo and Lee (2004) to identify the effect of unionization on several firm's outcome: business survival, productivity, wages, etc. The RD approach is a quasi-experimental design where the known discontinuity is exploited for identification. If we believe that unemployed in the neighborhood of 25 years of age are pretty similar in their characteristics (Tables, 2.2 and 2.3), both observable and unobservable, we can safely assume that without the program they will perform the same in the labor market (Figure 2.5). This allows to identify the effect of the program at least for those near the discontinuity under a local continuity assumption.

In a very simple chart the treatment function in a "sharp" RD design would look something like Figure 2.2, where below the threshold the probability of treatment is equal to 1 and above is exactly 0. In the specific case of the NDYP the sharp design is guaranteed by the mandatory nature of the program.

The advantage of such a method relies on the minimal set of assumptions required for the identification of a local parameter[19].

Formally, let D be the program participation status: $D = 1$ for participants, $D = 0$ for non-participants. Y^1, Y^0 be two potential outcomes, resulting from participation/non-participation respectively and $Y = Y^0 + D(Y^1 - Y^0)$ the observed outcome. The impact from participation is defined as $\beta = Y^1 - Y^0$. The eligibility rule $D = 1(A < a)$ is a known deterministic step function of A (age, continuous) and steps from 1 to 0 at a (25 years).

Taking the mean outcome difference for those marginally below (a^-) and above

[19]The local parameter can be on its own right an interesting parameter or even 'the' parameter of interest if the idea under scrutiny is that of extending the program marginally or to capture the effect of the program on the particular subgroup. Obviously it does not translate into an ATE unless constant treatment effect is assumed or under some particular smoothness conditions. It is also worth noticing that the parameter identified in the current context is not the same as the LATE defined by Angrist and Imbens (1994), their parameter is analogous to the one presented by Hahn et al. (2001) in the case of a fuzzy design.

(a^+) the threshold a:

$$(2.1)\ E[Y|a^-] - E[Y|a^+] = E[Y^0|a^-] - E[Y^0|a^+] + E[D\beta|a^-] - \underbrace{E[D\beta|a^+]}_{=0 \text{ by design}}$$

ASSUMPTION(1): $E[Y^0|A]$ continuous at a. Then the mean program effect on the treated[20]

$$(2.2) \qquad E[\beta|a^-] = E[Y|a^-] - E[Y|a^+]$$

is identified in the neighborhood of the threshold a.

However, I might observe Y^2 (non treated outcome) instead of Y^0 (non program outcome) since there might be substitution effects[21], treated might substitute controls at the threshold because they might be "cheaper".

Replacing Y^2 to Y^0 in the observed outcome and proceeding as before, instead of (2.2), by adding and subtracting the same quantity (ASSUMPTION (1)), I get:

$$(2.3)\ E[Y|a^-] - E[Y|a^+] = \underbrace{E[Y^1|a^-] - E[Y^0|a^-]}_{E[\beta|a^-]} + \underbrace{E(Y^0|a^+) - E(Y^2|a^+)}_{SB}$$

Where $E[\beta|a^-]$ is the parameter of interest and SB the substitution bias. The substitution bias is potentially important if the subsidized employment option has a large take-up and if treated are effectively cheaper than controls. However, I can provide some evidence on the absence of any substitution bias.

By considering a cohort approach. Let me rewrite (2.3) as:

$$E(Y|a^-, c) - E(Y|a^+, c) = E(Y^1|a^-, c) - E(Y^0|a^-, c) + E(Y^0|a^+, c) - E(Y^2|a^+, c)$$

[20] If the interest is on the average treatment effect in the neighborhood of the discontinuity, assumption (1) has to be extended to: $E[Y^i|A]$ for $i = 0, 1$ continuous at a.

[21] I left aside the discussion on possible general equilibrium effects because for the parameter I am identifying those effects should not be relevant. In the neighborhood of the discontinuity, even if there is an increase in the effective labor supply (given the number of participants involved) easing the wage pressure and the equilibrium wage, such an effect should be common to treated and controls and therefore should roughly cancel out.

where c is a cohort after the program. Let me rewrite $E(Y^2|a^+, c) = E(Y^0|a^+, c) - SB$ and assuming: ASSUMPTION(2a): $E(Y^0|a^+, c) = E(Y^0|a^+, c')$ where c' is a cohort before the program. If $E(Y^0|a^+, c') = E(Y^2|a^+, c) \rightarrow SB = 0$. And $E[\beta|a^-]$ is identified[22].

It remains to justify why cohort c' is not affected by substitution. There are a number of reasons why this might be the case. Cohort c' is obviously taken before the program started, the last cohort prior to the program will be the most similar to the one after the program given the economic environment. However, since the outcome I am considering spans over a year after the 6 months of unemployment, c' could in principle compete with cohort c and some of the others. In fact, substitution happens in the first 4 months of treatment, through subsidized placement, among similar individuals, if treated are cheaper than non treated, but for these two cohorts there are not similar individuals, since those in cohort c' have a different unemployment duration then those in cohort c when they are supposed to compete for the same job.

2.4 Data

A ready made dataset does not exists for the purposes of this work. However, it is still possible to recover most of the information needed by combining an administrative dataset (New Deal Evaluation Database, NDED) purposely built and containing virtually all participants, and the publicly available 5% longitudinal sample of UK unemployed (JUVOS). In the latter, it is possible to identify treated and control groups referring to the eligibility rule. Only unemployed aged between 18 and 24 who have received JSA for 6 months constitute the eligible and treated, given the compulsory nature of the policy, population. The JUVOS data contain the exact

[22]It might be argued that substitution could happen as well as a result of the other treatments and not only because of the subsidy, however it might be the case that enhanced job search and the other options improve the matching function by filling in the vacancies more efficiently without negative effects on controls.

date of birth, geographical region of residence, starting and end date of JSA spell, gender, usual and sought occupation and destination on exit from JSA, but has no information after the end of the JSA spell[23]. The sample selected for the analysis only includes job seekers with an unemployment spell of at least 6 months. Since August 1996 the JUVOS data contain a detailed series of exit categories recorded for those who ceased the JSA claim: found a job, other benefit, retired, prison, attending court and education and training. The last two exits are in fact one of the option of the NDYP while no equivalent exists for the control group, at least in the time interval considered. The controls who exit JSA for such destinations are almost certainly involved in small scale programs or simply decided on their own to acquire some training or education. In fact, such exit has less than half of the relevance for controls compared to treated.

Given the presence of such exit categories and the structure of the JUVOS data I would not know whether an unemployed (whose reason for ending the JSA spell is training or education) will find a job within the relevant period. Therefore, for such observations I have to complement the JUVOS data with the administrative data set (NDED). The NDED contains a number of extremely detailed information on participants, i.e. date of entry and termination of New Deal spell, date of birth, region of residence, unit of delivery, type of actions taken to find a job, number of letters sent to potential employers, option attended, status after ending the treatment, reasons for leaving the New Deal and so on. From the NDED, I can recover the exact[24] exit rates to employment for participants (in the particular period of interest). Therefore by using this complementary information, I can input such exit rates for the treated in the JUVOS data. An example might be helpful in clarifying this point, suppose some

[23]All the information on dates are recorded at maximum level of precision, i.e. day/month/year.
[24]As mentioned the NDED records information on all participants.

treated (identified in the JUVOS data) end their JSA spell to improve their education or attend some training (education/training option) I would not know, from JUVOS only, whether they found a job within a year since entering the program. However, I can get such information from the NDED, where I know exactly how many of them actually found a job in such a time interval and I can therefore input such information to the JUVOS data. Unfortunately, such a complementary information is not available for the control group, no controls are included in the NDED, however I can still define three different estimates of the parameter of interest by hypothesizing three alternative scenarios:

1. symmetric exit rates by age and cohorts for treated and controls;
2. all controls, who enrol in a training/education program, get a job in the time horizon considered;
3. none of the controls who attended some education/training course gets a regular job by the time interval of interest.

These strategies will allow to define a best estimate, a lower and an upper bound respectively. The best estimate scenario could be itself a sort of lower bound given that the type of courses attended by treated and controls are basically the same. In fact those individuals might seat in the same class, most of the vocational courses are provided by local colleges and are not differentiated depending on whether an individual is participating in the new deal or not. However, treated individuals are obliged to look for a regular job while attending training/education, the same is not true for controls. Therefore assuming identical exit rates for treated and control might produce a lower bound for the parameter estimated. It is likely that new dealers have a higher chance of getting an employment given the requirements imposed on the education and training option. For similar reasons the scenario classified as lower bound it is an extreme one. It is highly unlikely that otherwise identical individuals (those for which the parameter is defined) would have such different exit rates from very similar vocational courses, especially if one considers the further condition of actively looking for a job imposed by the program on the treated. On the same line the upper bound is an extreme in the other direction.

In order to avoid the inclusion of high-school kids 18 year olds are discarded from the analysis. I define five (post-program) even cohorts, according to the date of entry in the program, spanning from April 1998 to December 2001 (Table, 2.1). Each cohort counts at least three thousands observations and coherently with the RD design there is an almost identical number of treated and controls in each one. As written earlier the key of the identification relies on the discontinuity in the participation rule and on the a-priori belief that in the neighborhood of such point unemployed are almost identical but for the treatment status. Such belief can be confirmed by looking at the occupational (usual and sought) distribution in the proximity of the discontinuity (Tables, 2.2 and 2.3). There is a clear pattern of convergence in those occupational distributions: the closer to the 25 years of age the more similar they are.

As far as the pre-program analysis is concerned only one cohort is available due to the fact that prior to August 1996 the exit categories were not recorded at the same level of precision.

2.5 Estimation

The estimation of the parameter of interest is performed non-parametrically by Local Linear Regression (LLR)[25]. The LLR method consists in running several lo-

[25]Fan, (1992).

cal linear weighted regressions where the weights are assigned according to a kernel function (satisfying some regularity conditions) and a bandwidth. In general, observations close to the estimation point are given larger weights while decreasing weights are assigned to those further away. The estimation in an RD design boils down to estimating at a boundary point, where y^- and y^+ are estimated using observations from the left and right of the discontinuity respectively. The estimate of y^- is given by $\hat{\alpha}$:

$$(\hat{\alpha}, \hat{\beta}) \equiv argmin_{\alpha,\beta} \sum_{i=1}^{n} (y_i - \alpha - \beta(a_i - a))^2 K\left(\frac{a_i - a}{h}\right) 1(a_i < a).$$

Where $K(.)$ is the Kernel function, h an appropriate bandwidth and $a = 25$. Therefore in estimating y^- only observations to the left of the discontinuity are used. There are in principle other estimation methods, in the class of nonparametric estimators[26], available for the exercise proposed in this work, ie. Kernel regressions or Wald estimator. However, it is a known result that constant kernel methods have poor boundary performances due to the lack of observations on one side of the boundary. Such a problem could even be exacerbated in the current context, given that I would compound the bias from both sides of the discontinuity. The LLR method proposed attains the optimal convergence rate due to the local linear approximation (Porter, 2003) under fairly weak assumptions. A standard issue in nonparametric kernel or polynomial methods is that of choosing the "appropriate" bandwidth, or complexity of the model (Fan and Gjibels, (1996)), there is an obvious trade-off between bias and variance of the estimators in such context determined by the choice of the smoothing parameter. A too small bandwidth would cause an increase in the

[26]In an earlier work (De Giorgi, 2005) the estimation is performed parametrically by OLS using flexible functional forms. In the same spirit is the work of DiNardo and Lee, 2004. However, as shown in Section 2.6, the bias arising from a particular parametric assumptions can be substantial.

variance and might capture too much of the noise in the data, reducing the estimate to a simple interpolation of the data. On the other hand a large bandwidth would oversmooth the data, denying important features of the underlying data generating process. Such issue is resolved here by a plugin method for LLR elaborated in Ruppert et al. (1995). The resulting bandwidths are reported in Table 2.6, and a sensitivity analysis is performed to ensure the robustness of the results obtained[27]. The simple montecarlo study (Section 2.6) provides support for the use of the plugin rule-of-thumb over the computationally far more expensive alternative direct plugin. The last estimation step reduces to applying the LLR to the left and right of the discontinuity and taking the difference of the two conditional mean functions estimated. Standard errors have been obtained by bootstrap (300 replications[28]) for each cohort and for the whole sample.

2.6 Montecarlo Study

In this section I implement a simple montecarlo study on the performance of the estimator employed in the paper by comparing it with some alternatives both parametric and non parametric. The size of the discontinuity to be estimated is given in the Table 2.4 as β while the data generating process (dgp) is $y = m(x) + \beta D(x < .5) + \epsilon$. Where $x \sim U[0,1]$, $D = 1$ if the condition between brackets is satisfied, $\epsilon \sim N(0, \sigma_\epsilon^2)$ (σ_ϵ given in Table 2.4).

The proposed estimator (β_{RD}^{rot}) is matched with a very close substitute (β_{RD}^{dp}), the only difference between the two derives from the bandwidth selection criterion: while the former uses a direct rule-of-thumb, the latter relies on the direct plugin method

[27] All estimations are also performed according to a direct plugin method as in Rupert et al. (1995) and to $h_s = 1.06\hat{\sigma}n^{-.2}$, Silverman's rule, and half and twice the bandwidth used. Naturally, the Silverman's rule is not suited for the LLR but it has been used only for a robustness check. The parameter estimates vary very little whatever selection criterion is adopted. Complete set of results is available from the author on request.

[28] The number of replications has been limited to 300 after few checks on the stability of the results. The estimation process for the figures produced in Table 2.5 takes about two weeks on a powerful server.

both defined in Ruppert et al. (1995). Furthermore the performances of two simple OLS estimators (β^l_{OLS}) and (β^q_{OLS}) are analyzed, the superscripts l and q stand for a linear and quartic functional form in x respectively. Finally, a Wald estimator (β_W) on 10% of data around the discontinuity is also presented. Four different data generating processes are employed, whose complexity in estimation is proxied by the noise to signal ratio $\frac{\sigma^2_\epsilon}{\sigma^2_x}$.

The study is based on 500 replications and performed for two different sample sizes ($n = 1000, 3000$). The montecarlo evidence suggests a clear superiority of the proposed estimator in terms of precision with respect to the proposed Wald and OLS estimators[29]. While no ranking can be made between the two LLR estimators, they both perform quite well and are always close to the true parameter. Comparing the order of the bias involved in the use of the Wald estimator, as defined in the experiments, gives striking results: it goes from as little as 11% (first dgp, $n = 1000$) to an astonishing 400 times (second dgp, $n = 3000$).

The simple intuition on the quality of the point estimates obtained by LLR relies upon the locality of the latter. When the underlying function giving rise to the discontinuity is still quite regular but characterized by a highly non linear behavior fitting a local constant in the proximity of the discontinuity, a straight line or a quartic polynomial on the whole sample is not a great idea. On the other hand, the decision to use a Wald estimator on 10% of the observation in the neighborhood of the discontinuity is arbitrary, I could have proposed different candidates all of which would still be based on an arbitrary selection method. In this respect, the advantage of the LLR estimator applied at the discontinuity point is due to the fact that the bandwidth is selected according to a consistent and objective criterion. It

[29] The comparison with the OLS on the entire sample is per se not that meaningful given the idea behind the RD design.

arises from the data generating process itself and it is therefore more reliable and accountable than in the former case. The advantage of selecting the bandwidth through a direct rule-of-thumb with respect to a direct plugin method relies entirely on the computational burden involved in the latter while point estimates are fairly close as confirmed[30] here.

2.7 Results

It is possible to summarize the results by referring to Table 2.5, where I present three sets of estimates named Best, Lower and Upper. As explained in Section 2.4 the three different sets of estimates derive from the fact that I had to "construct" three alternative scenarios given the available data. It is possible to recover the exit rates to employment for those participants who went through the education and training option (about 35% of those who took an option or about 15% of total participants) but for the lack of information on unemployed older than 25 who had left the JUVOS dataset for some training or education I have to rely on some assumptions. The "best" estimates assume exactly the same employment probability for treated and controls when the recorded exit from JUVOS is education and training. This could itself be a lower bound since treated should be expected to have a higher (re)employment chance from that option, given that they are supposed to actively look for a job as part of the policy and are subjected to some form of monitoring from the new deal adviser. Treated individuals who enter the training/education option most typically enrol in a vocational or academic course at the local college and attend the same classes as the rest of the students or apprentices. On this respect there is no difference in what participants and control learn. Therefore assuming an

[30]This result is consistent with the evidence presented in Ruppert et al. (1995). They also found the rule-of-thumb to perform quite well in all their experiments.

identical employment rate for treated and controls seems a conservative option.

The second scenario "lower" relies on the assumption that all controls, who attended some training/education course, found a job in the reference period. It therefore qualifies as an extreme lower bound. In the third scenario, "upper", none of the controls found a job in the reference period, which seems to be an extreme in the other sense. The fact that the "best" estimate scenario is a conservative estimate of the true program effect allows concentrating on those estimates as the main results of the paper, while the lower bound guarantees that in no case there is significant negative impact of the policy. The estimated (re)employment probabilities are plotted, for the "best" scenario case, in Figures 2.6 and 2.7; it is evident a clear jump of those functions exactly at the cut-off point. Furthermore the shape of such conditional means is fairly flat. It is important, as mentioned in DiNardo and Lee (2004), to check whether the jumps obtained by the proposed estimation strategy are true program effects or only arise from the particular nature of the estimator proposed. All figures on pre and post program (re)employment probabilities contain a series of dots representing the mean outcome values by age (yearly averages), the pattern of those dots provide support for the estimates produced. The jumps only appear were they ought to[31].

It does not seem that the program effect is dying out as the results in Blundell et al. (2004) might have suggested. On average over the whole period considered it is possible to estimate a very precise parameter of 5%. The time profile of the estimates does not seem to suggest relevant general equilibrium effects with possible differential

[31] As mentioned, in an earlier work (De Giorgi, 2005), the analysis is performed by OLS using flexible functional forms and the same sort of jumps appear there. Furthermore, as a robustness check, I have also tested whether a significant difference in the post-program (re)employment probabilities appears at any other age (ie. 20,21,22 and so on): I could never find any significant discontinuity but from the true cut-off point (25 years). In principle such test can be performed at any point between 19 and 31 years, this would be obviously unfeasible given the continuous nature of the underlying variable, therefore I only focused my attention to integers (20,21,22..) or half year intervals (19.5,20,20.5...). The test is practically implemented by a dummy variables approach and not by LLR given the extreme computational burden that would be involved by the latter.

impacts on the two groups (at least in local terms). This point is also confirmed by looking at the (re)employment probability for the two groups separately, they do not vary much and certainly not to be consistent with large general equilibrium effects.

On the other hand substitution does not seem to be relevant either. In case of large substitution effect we should see in the conditional mean functions a behavior similar to Figure 2.4. The closer to the discontinuity the more substitutable individuals should be and therefore at the discontinuity the distance between the (re)employment probabilities should be larger[32]. However, this is not the case given Figures 2.6 and 2.7; on both sides of the discontinuity the functions are almost completely flat. This suggests, combined with Figure 2.5 ((re)employment probability before the program), a "global" interpretation of the parameter estimates. However, such an extended interpretation obviously implies a stronger identification structure (i.e. constant treatment effect or particular smoothness). A test on the difference between non treated outcomes before and after the program is also performed formally and in Figures 2.10 and 2.11, the null of equality cannot be rejected at any conventional level of significance. A note of caution is also necessary in interpreting the results from the pre-program analysis for which it is not possible to design any sort of bounds as described above. Nevertheless, this does not undermine the comparability of outcomes between treated and controls in the pre-program case[33], since for none of them a particular education and training program was available. At the same time when controls are compared overtime, before/after the program, as in Figures 2.10 and 2.11, as long as I treat the exit to education/training consistently such comparability is still safe and meaningful. A similar argument can be used when

[32]Obviously, the shape of the (re)employment probability suggested in Figure 2.4 is not the only one compatible with relevant displacement effects, there are other functions of age able to deliver a larger gap at the discontinuity then in any other point.

[33]However it is not possible to safely compare the outcome for controls between Figures 2.5 and 2.6/2.7 since in this case the assumption made to draw the post-program (best estimate) and the pre-program outcome are different.

discussing Figures 2.8 and 2.9, since before entering the program, ie. in the first 6 months of open unemployment, no education/training programs are available.

The lack of evidence of general equilibrium and substitution effects can be explained by a number of factors. Firstly, the sort of general equilibrium effects I have in mind, arising from an increase in the labor supply lowering the equilibrium wage, require a substantial rise in the overall supply of labor, however, though the implementation of the program is global, it is not so massive to affect in a significant way the overall labor supply in the UK. Support for such claim is given in Figures 2.12 and 2.13 where the number of participants, at any one month starting in January 1998, are graphed as a percentage of active males. It is clearly visible an increase in the number of young britons involved in the program since it started, with picks in the spring of 1999 when about 110 thousands young males were administered the policy. However, relative to the number of active males it never overcame 1%. The limited (if at all) influence on the overall labor supply and equilibrium wage is given by the evolution of hourly wages for males' employees (Figure 2.13). I focused only on the lower end of the distribution since those are presumably the sort of rates a typical participant would get. A vertical line denotes the start (national roll-out) of the program, there is no evidence of a differential trend before and after the program. Obviously, these indirect tests cannot rule out some GE effects; however, if anything, they are not supportive of an important general impact of the policy.

As far as the substitution is concerned, it requires that treated individuals are cheaper than untreated, but this might not be the case if the cost of turnover is relatively high. Furthermore, treated are cheaper only in the case of the subsidized employment option, but the take-up rate of such a feature of the program is surprisingly low[34]. In fact only one over 6 treated who went through the option stage were allocated in a subsidized job adding to less than 7% of the new dealers. I have not covered possible general equilibrium effects arising from distortionary taxes devolved to the funding of the program for the simple reason that the program has been funded through the revenues from the privatization processes initiated in those years.

2.8 Is There a Substitution Puzzle?

As mentioned throughout the paper the relevance of possible substitution effects among treated and control individuals is central to the identification structure. The program I consider here has a particular feature (subsidized employment option) that could raise concerns regarding the violation of the SUTVA and therefore the validity of the identification strategy. I have spent a considerable part of the work trying to assess such an issue, and I do not find support for any major concern on the evaluation exercise I propose. Why is it then that there is not any substitution effect? In principle, the presence of a significant subsidy to employment should generate an incentive to substitute workers. Is the subsidy given to participants enough to create such an effect? As explained earlier, by comparing the sort of hourly rate participants should get to the amount of the subsidy granted (weekly plus one off payments) this adds up to about 50% of the salary in the 6 months period for which such subsidy could last for. However, when considering the relevance of the subsidy there are few more things to be accounted for. Firstly, the one off payment has to cover the minimum one day per week of training participants must receive. On its own this would notably lower the previous percentage to 30%. Secondly, the subsidy only last for 6 months and might not be enough to compensate for the turnover costs. Thirdly,

[34]Even the program administrators were surprised by such a low take up.

in a targeted program, as the one considered here, there might be an important stigma effect (Katz, 1998) attached to receiving a subsidy. The only way such a participant is able to get a job is through a discount on the wage received. He is probably not as productive as someone else in the population and while the subsidy could help him getting a job, it would signal to the market his bad type. These are three potential explanations on the absence of relevant substitution effects in the particular program under scrutiny. Are they convincing? I should now go back to the evidence. The very low take up rate for such an option (only 16% of participants who actually went through an option) was surprising even to the program administrators who were expecting a much higher one. The amount of evidence put forward in this respect seems to be clear cut in excluding relevant substitution bias (Figures 2.4, 2.6, 2.7 and 2.10), either comparing cohorts of controls before and after the program, the actual outcomes in terms of employment probability with a prediction of how they should look like in case of any relevant substitution effect.

2.9 Conclusions

Previous US evidence on ALMPs targeted at young unemployed has been rather disappointing (Bloom et al., 1997; Heckman et al., 1999). Those studies did not find any significant impact of job search assistance and training on disadvantaged youths both in terms of employment and wages. On the other hand Heckman et al. (1999) surveyed a series of European studies mainly focusing on young unemployed where in some cases a positive and significant effect was found in terms of employment while an even less clear cut evidence exists on wages.

The evidence presented in this work is somewhat reassuring. A targeted policy such as the NDYP is able to increase the (re)employment chance of young unemployed by a small but significant amount. However, it remains to investigate further whether it is a particular component or the whole structure of the policy that is working effectively.

It has to be pointed out that part of the previous evidence was based on voluntary programs or on policies aimed at particularly disadvantaged youths (ex-offenders). While the target group of the NDYP is constituted by all young unemployed, besides a 6 months unemployment spell for such an age group is not that uncommon in the UK.

It has also been shown (van Den Berg et al., 2004) that policies where non compliers incur significant sanctions are on a theoretical and empirical ground capable of producing beneficial effects in terms of employment, for the simple fact that they push up the level of effort exerted by the unemployed. The mechanism being quite intuitive a worse outside option (withdrawal of the benefit) constitutes a large incentive. Katz (1998) in reviewing different ALMPs found that policies combining wage subsidies with job development, training and job search assistance appear to have been somewhat successful in improving the labor market conditions (employment and earnings) of specific groups.

Table 2.1: Cohorts. Treated 19-24, Controls 25-30 years old

	Treated	Control	Total
Apr. '98-Dec. '98	4,256	3,916	8,172
Jan. '99-Sept. '99	4,261	3,956	8,217
Oct. '99-June '00	3,885	3,563	7,448
July '00-Mar. '01	3,311	3,038	6,349
Apr. '01-Dec. '01	3,282	2,910	6,192
Total	18,995	17,383	36,378

Table 2.2: Usual Occupation by Treatment Status and Age (%)

Usual occupation	*19-30*		*22-27*		*24-25*	
	Treated	Control	Treated	Control	Treated	Control
Managers	0.35	2.17	1.12	2.05	2.10	1.78
Professional	0.42	2.21	0.88	2.21	1.39	2.16
Associate Prof., Technical	2.98	4.14	4.08	6.04	5.34	5.76
Admn. Secretarial	13.16	10.70	14.01	11.27	12.02	11.13
Skilled trades	10.91	16.53	11.84	13.37	13.91	12.65
Personal Service	5.44	4.06	6.00	4.77	5.59	5.76
Sales and Customer service	11.19	3.86	7.04	5.11	6.98	5.60
Process, Plant and Mach. operatives	7.76	11.83	8.40	11.31	9.79	10.75
Elementary occupation	47.80	44.49	46.62	43.86	42.88	44.41

Note: The first two columns compare 19 to 30 year olds; third and fourth 22 to 27 and the last two 24 to 25 year olds. All cohorts are used in this tabulation.

Table 2.3: Sought Occupation by Treatment Status and Age (%)

Sought occupation	*19-30*		*22-27*		*24-25*	
	Treated	Control	Treated	Control	Treated	Control
Managers	0.38	2.15	1.72	2.09	2.61	2.10
Professional	0.71	2.07	1.22	2.26	2.28	2.53
Associate Prof., Technical	3.60	5.06	5.40	6.93	6.06	7.43
Admn. Secretarial	15.18	11.39	14.96	12.33	13.67	12.23
Skilled trades	12.74	16.60	12.39	13.31	13.34	12.61
Personal Service	6.26	4.51	6.35	5.02	5.98	4.90
Sales and Customer service	12.27	4.33	8.00	5.64	7.69	6.22
Process, Plant and Mach. operatives	7.22	12.81	8.91	11.60	9.64	11.08
Elementary occupation	41.64	41.06	41.05	40.80	38.72	40.89

Note: The first two columns compare 19 to 30 year olds; third and fourth 22 to 27 and the last two 24 to 25 year olds. All cohorts are used in this tabulation.

Table 2.4: Montecarlo Experiments

$m(x)$	β	σ_ϵ	noise/signal	$\hat\beta_{RD}^{rot}$	$\hat\beta_{RD}^{dp}$	$\hat\beta_{OLS}^{l}$	$\hat\beta_{OLS}^{q}$	$\hat\beta_{W}^{a}$
$.4x+.2x^2-.7x^3+.1x^4$	.1	.1	.14					
n=1000				.0851	.0865	.0651	.0855	.0835
				(.0192)	*(.0328)*	*(.0144)*	*(.0171)*	*(.0204)*
n=3000				.09154	.0900	.0622	.0809	.0787
				(.0119)	*(.0215)*	*(.0083)*	*(.0099)*	*(.0118)*
$.5\sin(6x)$	.1	.1	.06					
n=1000				.1213	.1110	.6824	.1730	.5826
				(.0266)	*(.0338)*	*(.0231)*	*(.0174)*	*(.0163)*
n=3000				.0988	.0972	.6799	.1672	.5938
				(.0167)	*(.0127)*	*(.0130)*	*(.0101)*	*(.0097)*
$\Phi(.5\sin(6x))$	.1	.1	.28					
n=1000				.1077	.1046	.3217	.1234	.1582
				(.0224)	*(.0133)*	*(.0088)*	*(.0171)*	*(.009)*
n=3000				.0999	.0990	.3197	.1182	.1549
				(.0137)	*(.0086)*	*(.005)*	*(.0099)*	*(.005)*
$\exp(x^{\cdot 2})+.5\sin(6x)$	.1	.1	.13					
n=1000				.1011	.1019	.7457	.1431	.1428
				(.0314)	*(.0333)*	*(.0293)*	*(.0174)*	*(.0347)*
n=3000				.0829	.0949	.7509	.1375	.1291
				(.0231)	*(.0216)*	*(.0167)*	*(.0100)*	*(.0183)*

Note: Φ is the standard normal cdf; β_{OLS}^{l} and β_{OLS}^{q} x controlled for linearly and as a quartic polynomial respectively. a) Wald estimator takes only observations for which $.45 \le x \le .55$.

Table 2.5: Treatment Effects (and Bounds), Bandwidth: Plugin

	Best		*Lower*		*Upper*	
Cohort	Effect	Std.Err.	Effect	Std.Err.	Effect	Std.Err.
Apr. '98-Dec. '98	.0559	*(.0245)*	.0068	*(.0001)*	.0911	*(.0239)*
Jan. '99-Sept. '99	.0499	*(.0223)*	-.0081	*(-.0126)*	.1145	*(.0241)*
Oct. '99-June '00	.0110	*(.0238)*	-.0389	*(-.0431)*	.0588	*(.0234)*
July '00-Mar. '01	.0433	*(.0250)*	.0067	*(.0102)*	.0979	*(.0257)*
Apr. '01-Dec. '01	.1056	*(.0277)*	.0347	*(.0360)*	.1447	*(.0269)*
All	.0499	*(.0107)*	.0019	*(.0112)*	.0995	*(.0118)*

Note: Bootstrap standard errors based on 300 replications. Bandwidth selected with plugin method.

Table 2.6: Bandwidth: Plugin

	Best		*Lower*		*Upper*	
Cohort	Treated	Control	Treated	Control	Treated	Control
Apr. '98-Dec. '98	1.3558	1.0745	1.2502	1.2152	1.2825	1.2548
Jan. '99-Sept. '99	1.1630	.9591	1.1071	1.0027	1.1254	1.2804
Oct. '99-June '00	1.8981	1.6938	1.6272	1.4243	1.2371	1.2958
July '00-Mar. '01	1.7775	1.2709	1.3457	1.1804	1.6974	1.1001
Apr. '01-Dec. '01	1.2906	.9653	1.5243	1.4744	1.1075	.8699
All	1.1805	.9566	1.1135	1.0732	1.1857	1.2357

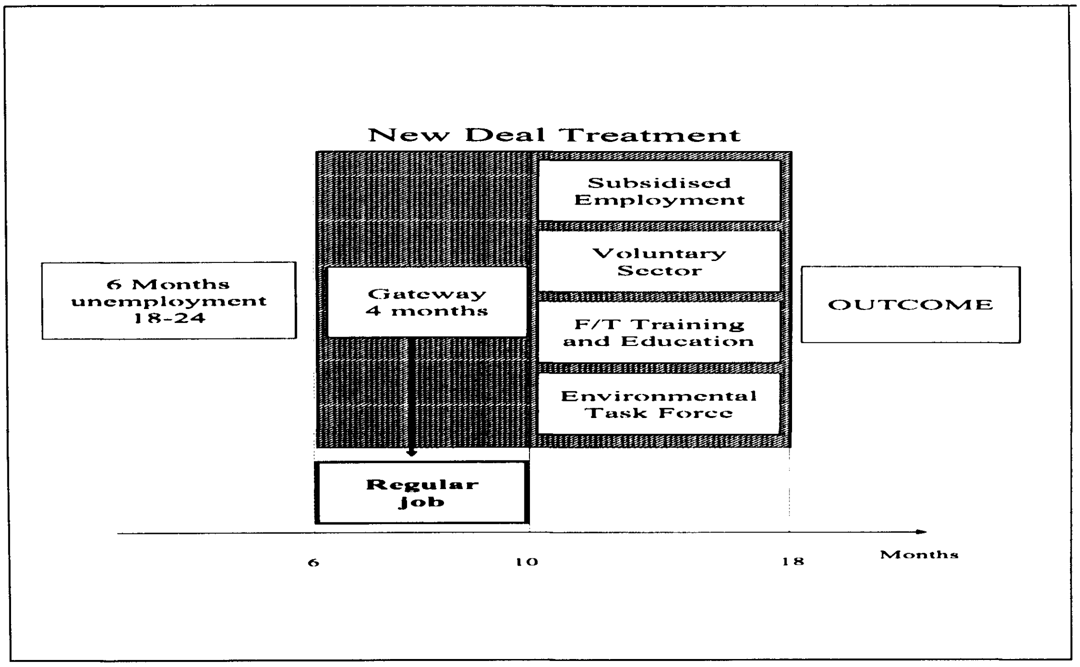

Figure 2.1: The program

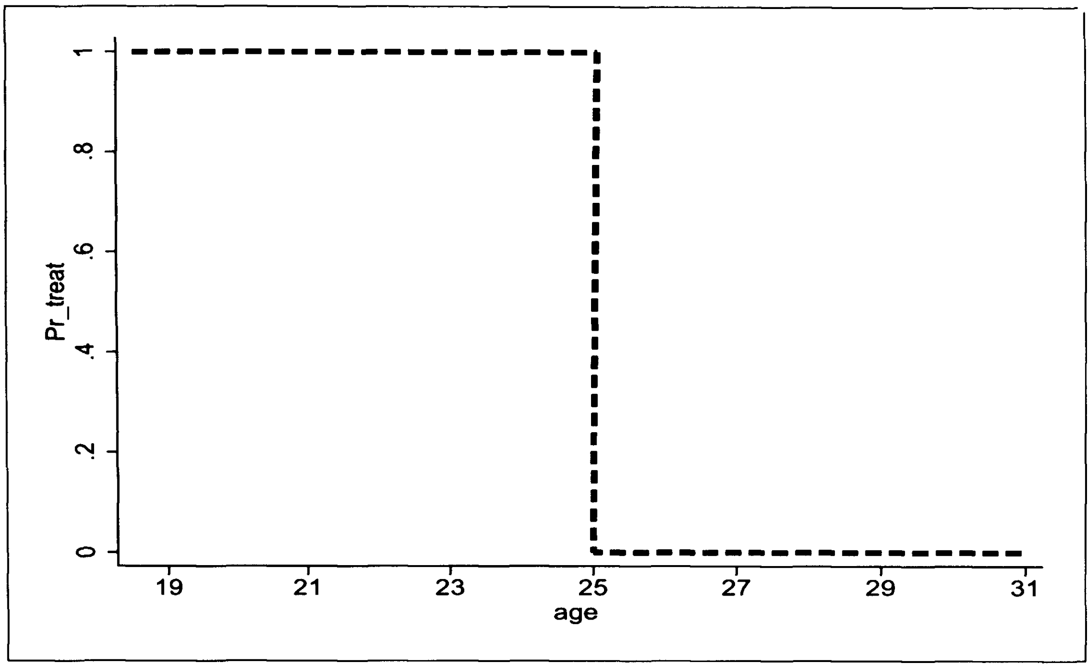

Figure 2.2: The "sharp" RD design

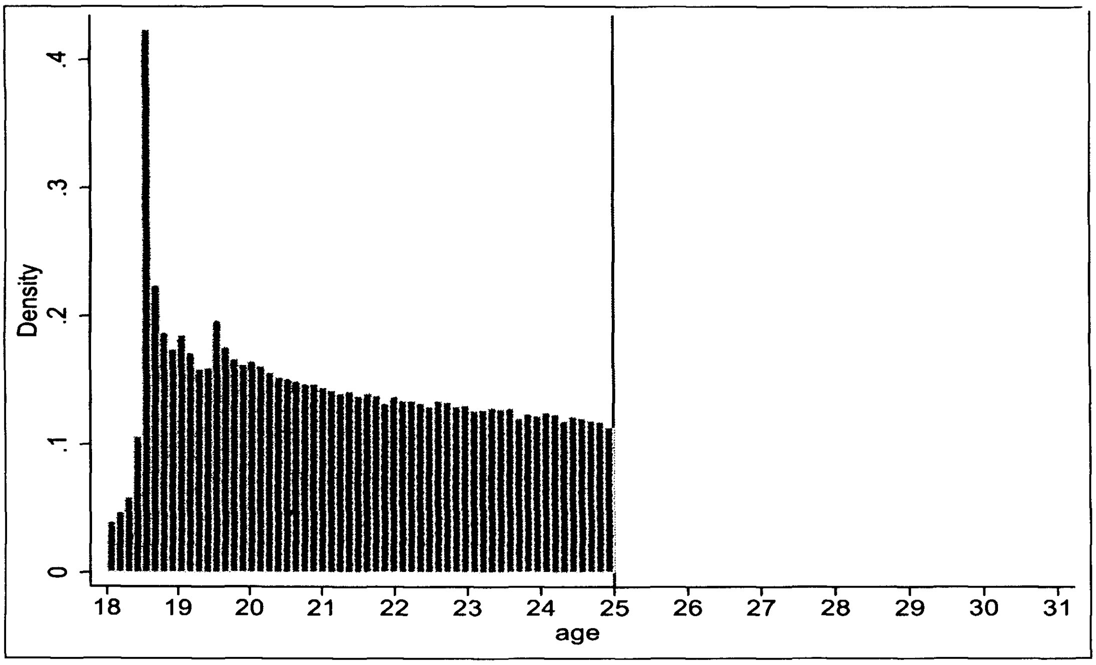

Figure 2.3: Age profile of New Dealers (Males only)

Note: Based on NDED from April '98 to December '01.

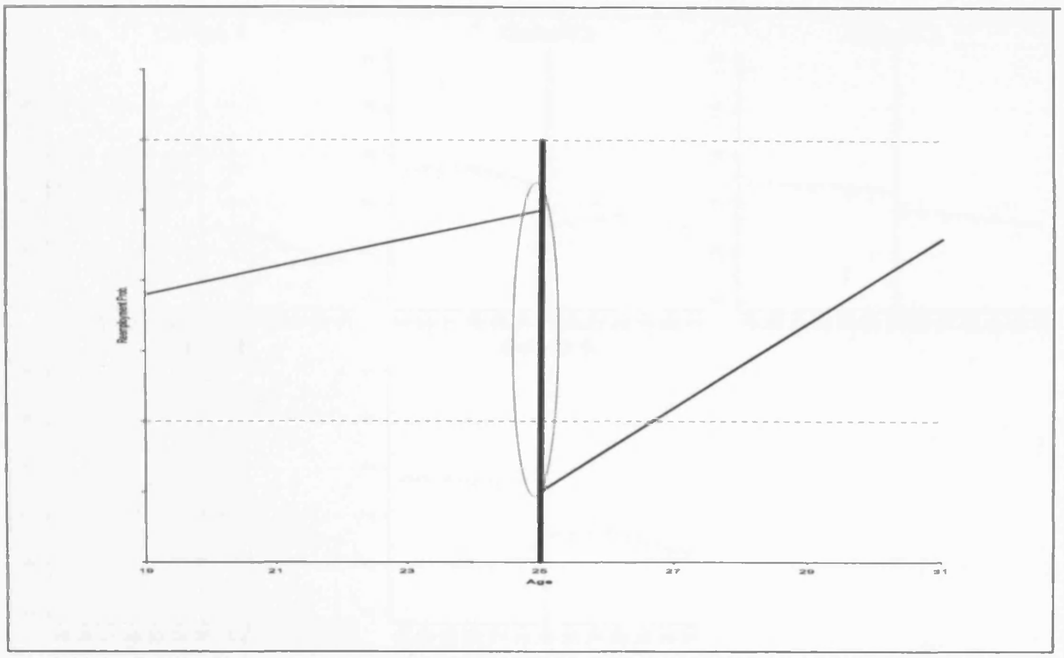

Figure 2.4: Prediction of the (re)employment probability by age in case of relevant substitution effects

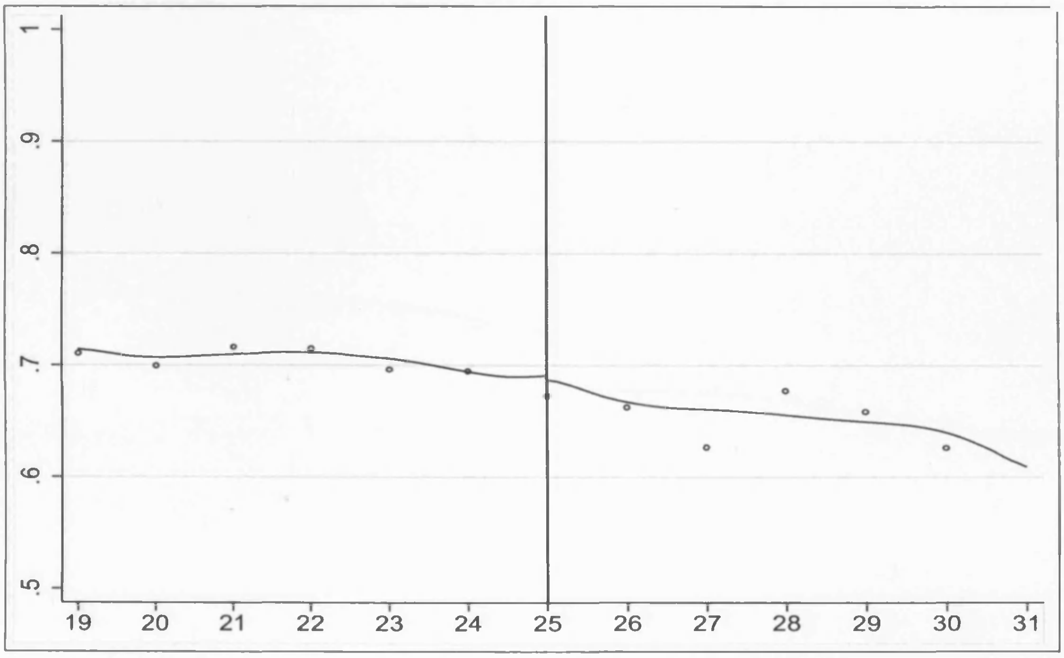

Figure 2.5: Pre-Program (re)employment probabilities by cohort and age
Note: Dots are average (re)employment rates by age in year.

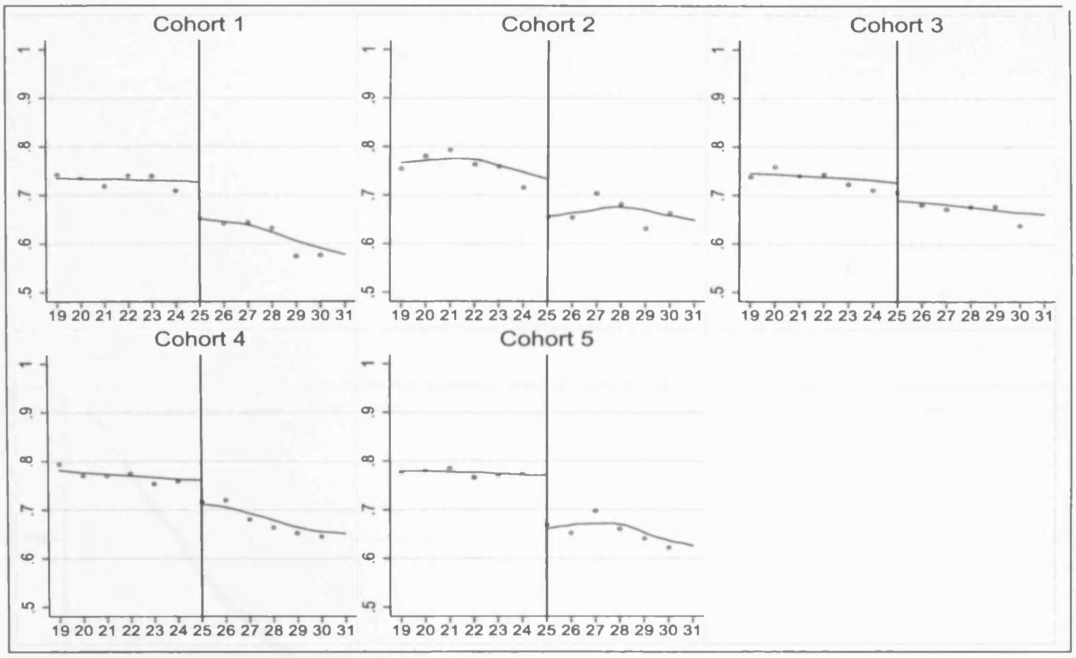

Figure 2.6: Post-Program (re)employment probabilities by cohort and age
Note: Dots are average (re)employment rates by age in year.

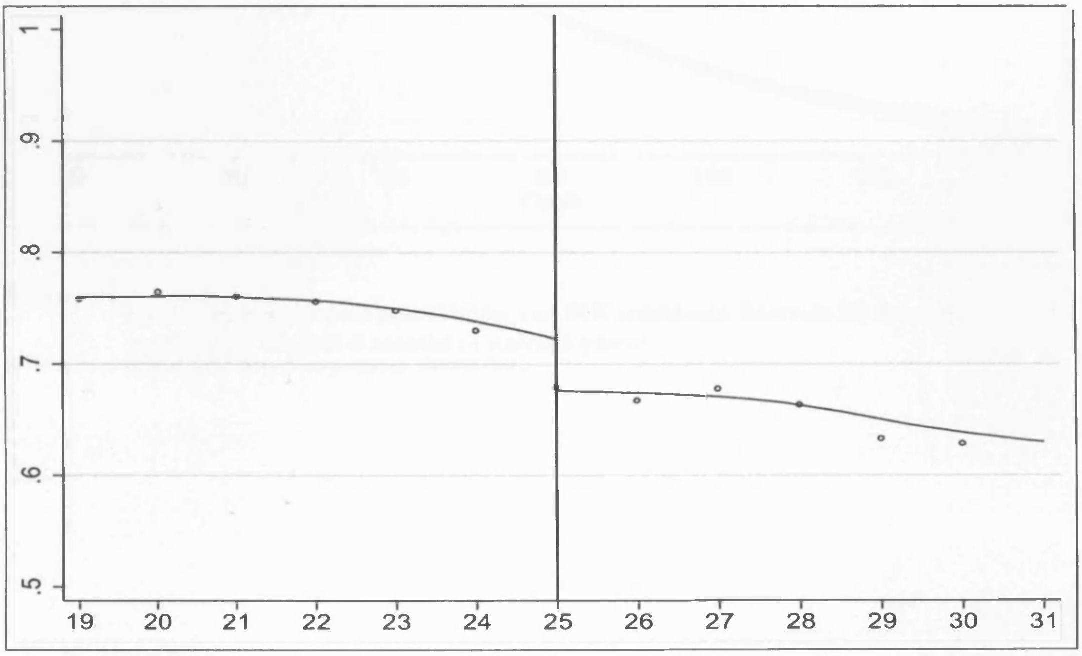

Figure 2.7: Post-Program (re)employment probabilities by age (all cohorts)
Note: Dots are average (re)employment rates by age in year.

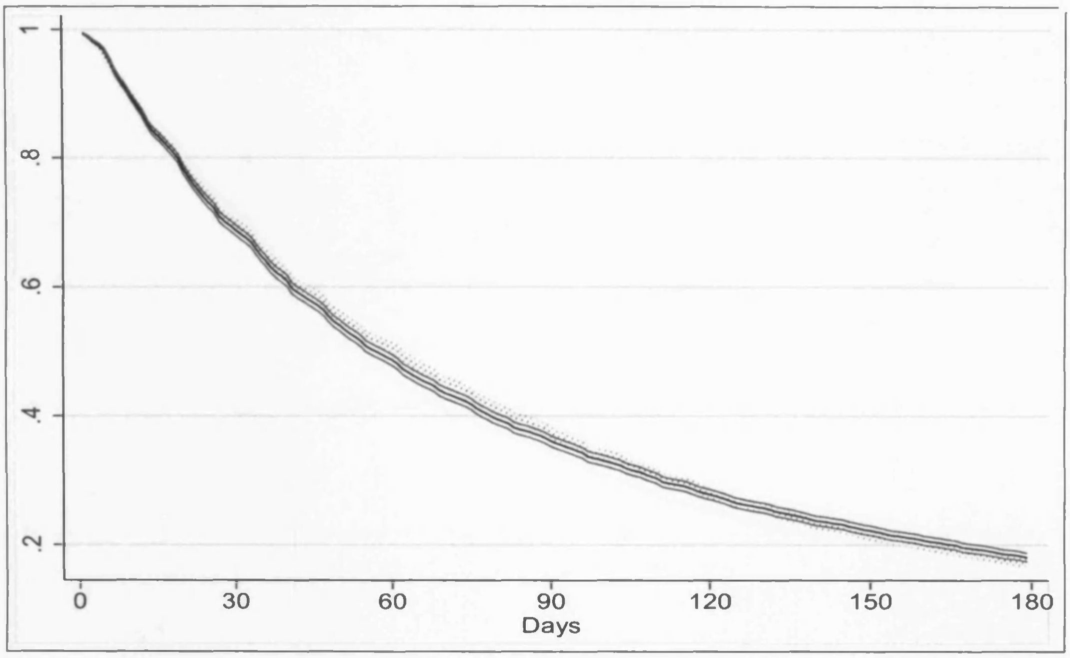

Figure 2.8: Kaplan-Meyer survival probabilities and 95% confidence intervals for 2 cohorts of 19-24 year olds in the first 6 months of unemployment

Note: pre-program, solid line; post-program, dotted line.

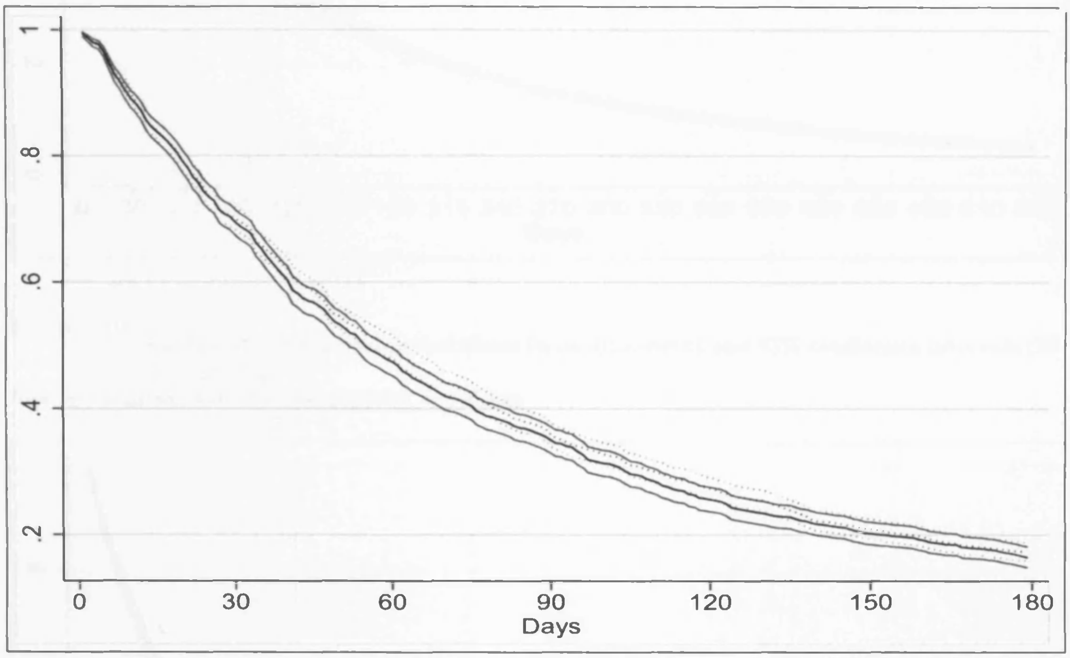

Figure 2.9: Kaplan-Meyer survival probabilities (unemployment) and 95% confidence intervals for 2 cohorts of 24 year olds in the first 6 months of unemployment

Note: pre-program, solid line; post-program, dotted line.

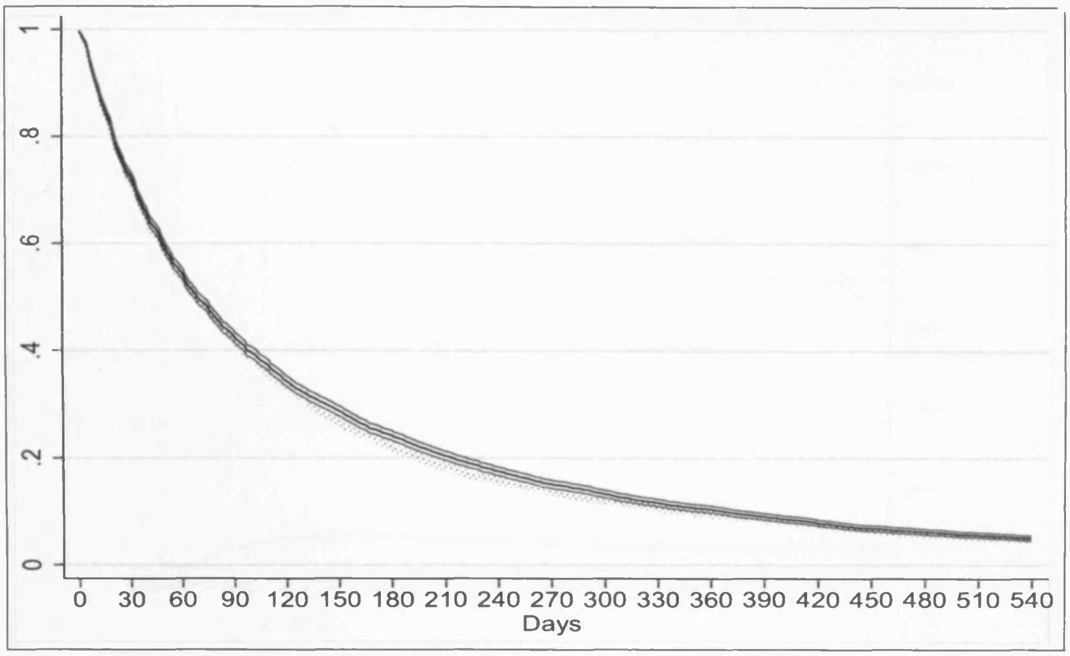

Figure 2.10: Kaplan-Meyer survival probabilities (unemployment) and 95% confidence intervals (25-30 year olds)
Note: pre-program, solid line; post-program, dotted line.

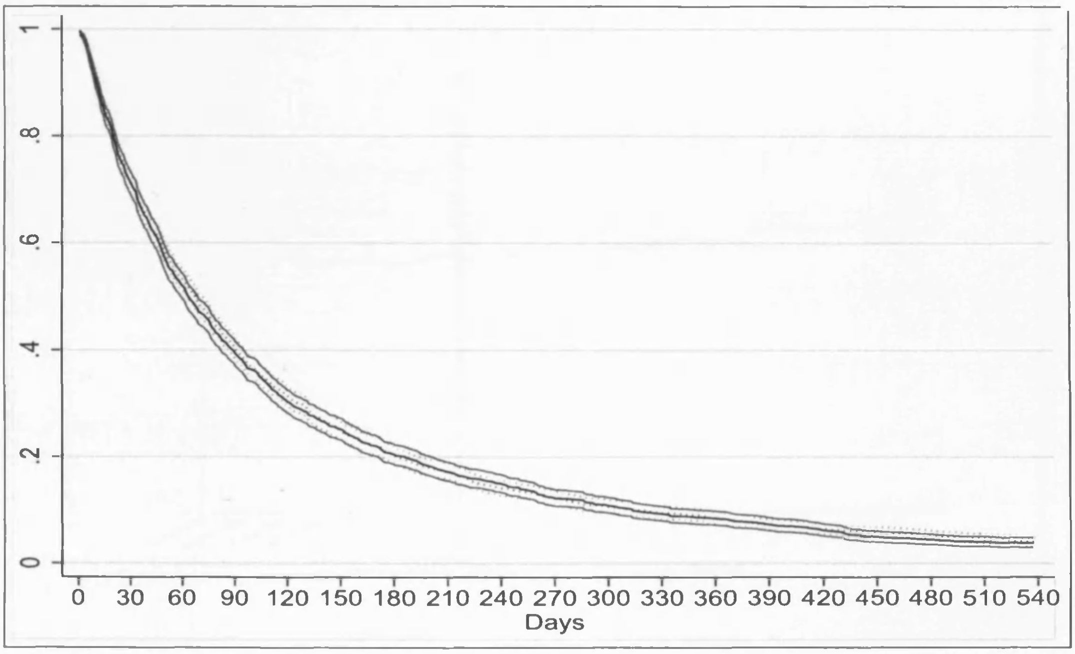

Figure 2.11: Kaplan-Meyer survival probabilities and 95% confidence intervals (25 year olds)
Note: pre-program, solid line; post-program, dotted line.

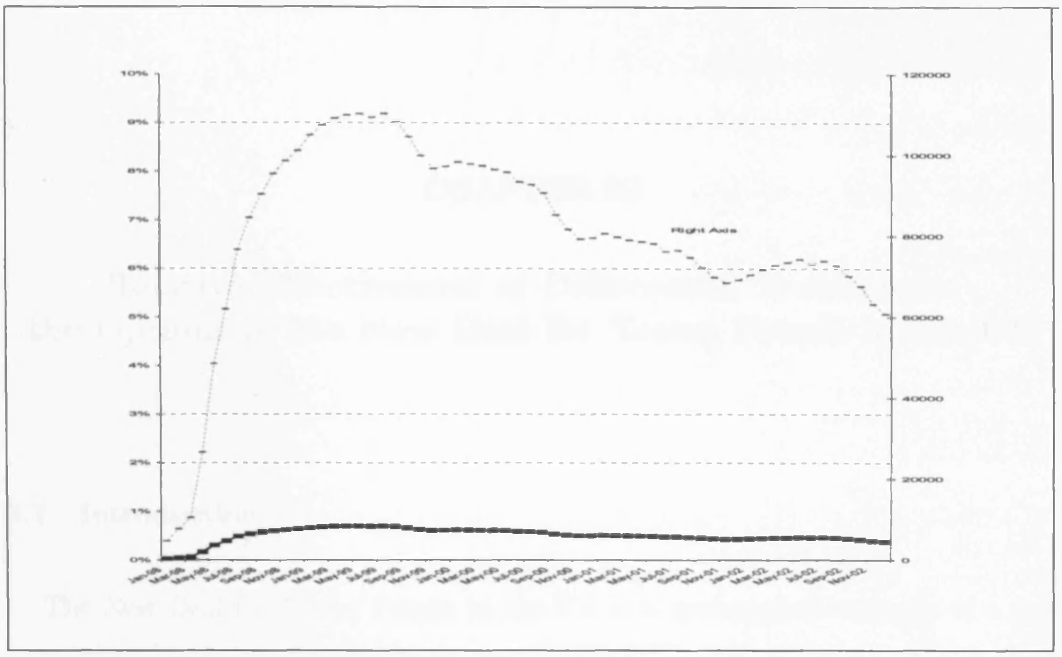

Figure 2.12: New Deal participants (males)
Note: as % of active males, solid line; number, dotted line.

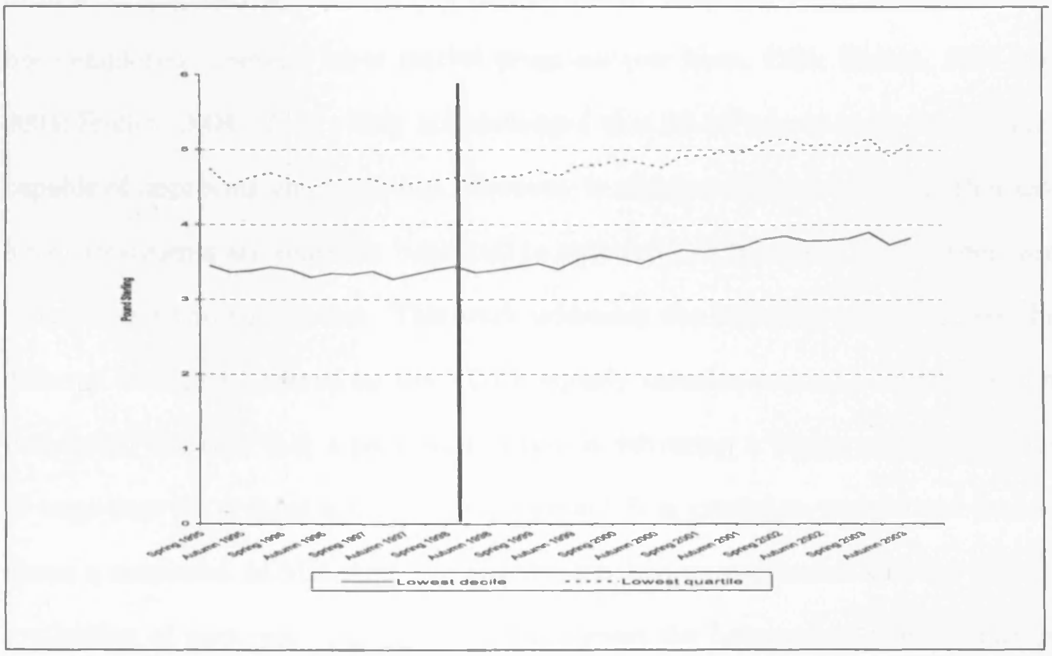

Figure 2.13: Males' hourly earnings (1995Q1 prices), employees only
Note: the vertical line denotes the start of the NDYP. *Source*: Office of National Statistics, Labour Force Survey.

CHAPTER III

Relative Effectiveness of Differential Treatments: the Options in the New Deal for Young People in the UK

3.1 Introduction

The New Deal for Young People in the UK is a prototypical example of a program with multiple treatments. It combines, at different stages, job search assistance, training/education, subsidies and reinstatement in the labor force through governmental or voluntary sector jobs (De Giorgi, 2005). Such particular framework has been employed in several labor market programs (see Katz, 1998; Sianesi, 2001 and 2004; Frölich, 2004). It is widely acknowledged that ALMPs have to be flexible and capable of improving employability. However, it remains an open issue whether different treatments are generally beneficial or only some of them are, while others are a pure burden to the system. This work addresses the following question: are the different treatments offered by the NDYP equally valuable in terms of returns? Or it is rather the case that a particular option is delivering a higher return in terms of wage once the treated is out of the program? It is crucial to understand how to shape a successful ALMP given the government budget constraint and the limited availability of resources. On the other hand given the heterogeneity in the nature of the unemployed population a certain degree of flexibility is necessary in order to give a valuable treatment to the particular individual.

There is evidence that the subsidy to employment is the 'star' option of the NDYP with respect to enhancing the (re)employment probability of participants (Dorsett, 2006) however there isn't any available study that looks into job quality and namely wages. The mere employment is not per se an indicator of the success of a program, especially when there might be concerns over the quality of jobs. In the particular case the star option could be successful in getting unemployed out of dole, but the point is: are these jobs worthwhile? A subsidy to employment is relevant when the productivity of the worker is possibly below his/her cost. However, such subsidy cannot last forever, further it can attach an important stigma to the takers as well as forcing participants into worse match accepting jobs at lower wages than otherwise. I exploit a purposely built survey dataset collecting a large number of information, including wage data, on a particular cohort of participants. As standard in the evaluation literature the main issue is that of having a convincing identification structure of a meaningful policy parameter. This work exploits a particular feature of the treatment process, namely the option assignment, in order to identify the effect of a particular option on those who went through it. In principle the assignment process should have been a joint decision between the caseworker and the participant. However, this was not the case for two main reasons: firstly, certain options were simply not available in certain areas, i.e. some local units of delivery did not have the possibility of placing a participant in the Environmental Task Force, while others did not have voluntary sectors job available and so on (*rationing*). Secondly, there is a clear pattern of preference for a particular treatment in certain units, this depends on the fact that placing someone in a subsidized job is simply more expensive than sending someone to school, both in terms of effort to be exerted by the caseworker and monetary cost (*costs*). There is a large variation in option take-up across different UoD's substantiated by anecdotal and formal evidence later in the paper. Such variation remains even when a number of confounding factors are partialled out. Furthermore, there is evidence of non-random option allocations in Table 3.2. It is clear how better quality participants where assigned to the subsidized employment option. In fact, they are significantly better in terms of schooling, ability (reading/math problems), work history (although they have surprisingly longer unemployment history) than their counterpart who engaged in an extended job search treatment. If we were to believe in the random assignment to different options the simple OLS estimates would suggest that on average no options had a differential impact in terms of hourly wages in regular employment. However, when non-random selection both at the options stage and in the employment node is taken into account a dramatically different picture emerges. IV estimates clearly point towards a negative and significant penalty of roughly 20%, almost 1 GBP (1.9 USD) less. Such result is in line with earlier findings on subsidized employment by Katz, 1996. The intuition for a negative return is consistent with a stigma effect attached to those participants who got a subsidy, while otherwise they would not be productive enough to pay for their wages. Further, it might be that participant would be worse off in the matching process by the acceptance of jobs at lower wages or not particularly suited for them. It seems therefore that the specific option might signal to the market the low productivity type or simply impeding the development of a frictionless matching process. Unfortunately, I could not test the hypothesis of a temporary effects given the available data.

The remainder of the paper is organized as follows: Section 3.2 describes the nature of the program; Section 3.3 details the identification strategy adopted. A description of the data employed is given in Section 3.4; while results and treatment effects are presented in Section 3.5. Finally, Section 3.6 concludes.

3.2 The Program

As from Figure 3.1 the NDYP is a sequential program, where different treatments are offered to the participants. Following a period of six months in "open" unemployment 18 to 24 year olds are forced into the program in order to be still eligible for the unemployment benefit (JSA). It is therefore a mandatory policy administered to everyone in the UK who, after six months of unemployment, are aged between 18 and 24 (further details on the program can be found in Blundell *et al.*, 2004; and De Giorgi, 2005).

The first four months (Gateway period) are nominally devoted to intensive job search assistance and some basic skill training, i.e. CV writing. Participants are obliged to meet a personal mentor once every two weeks and have to report and prove the actions taken in order to gain employment. Failure to comply with any of the program requirements might result in a benefit sanction and eventually the withdrawal from it.

While in the gateway participants receive a benefit equal to the JSA (about 40 pounds per week). If a regular job is not found during the gateway, a second phase follows: the options. This second stage is the focus of this work, in particular the option assignment mechanism will be discussed in great detail in Section 3.3. Reading through the institutional rules for such stage it emerges that: on the basis of personal considerations, given individual characteristics, the caseworker agrees with the participant on the option to be taken. However, as we will see later, this is not always the case since certain units of delivery tend to favor a particular option for two main reasons: i) *rationing*, i.e. not all options are available at a given time in the particular UoD and; ii) *costs*, i.e. some options are more expensive in terms of budget and effort.

The option period can last from 6 to 12 months (full time training or education). Common practice among units of delivery was to try placing the unemployed in a subsidized job during the second month of treatment. In case of a subsidized employment, the treated receives the salary paid by the employer who gets, for a maximum of six months, a subsidy of 60 pounds per week plus 750 pounds as a one-off payment for the compulsory (minimum) one day a week training to be provided[1]. The second option, education or training, is targeted at youths lacking basic skills and it can last up to 12 months (although it is common for such courses to last only 10 months). While attending such courses the unemployed still receives his JSA payment. Typically the courses attended by a program participant would be given by the local college, and in most cases, those are not specially provided to newdealers. There is no obvious distinction between a vocational course attended by a treated and by anyone else in the community.

A third option is that of a voluntary sector job where the participant receives an amount at least identical to the JSA plus 400 pounds spread over the six months. A typical placement would be shop assistant in a charity shop.

The same monetary treatment is granted in the fourth option: Environmental Task Force, basically a governmental job, meant to be the last possible placement. In this option the participant would typically be involved in public parks maintenance.

Participants are allocated to these last two options in the third and fourth month

[1]Such subsidy seems quite generous when compared to the sort of hourly rate (close to the minimum wage) a typical participant would get. In a crude computation, the weekly subsidy plus the one-off payment would amount to about 50% of a weekly pay for a minimum wage worker, however the 750 pounds would have to repay for the loss of production due to the minimum of one day training. Under very simple assumptions (perfectly competitive markets) those 750 pounds would not be enough to compensate for that loss. In fact, taking the latter into account the subsidy would not be greater than 30%, but still generous though. However, job turnover could be itself quite costly making such an option not as appealing as it looks like at a first glance. This point seems to be confirmed by the low take up rate in the data, only a sixth of those entering an option would go for the subsidised job.

of the gateway.

It emerged a fifth, non-contemplated by the program, option: the extended gateway, i.e. a continuing job search assistance. As we will see later this practice was not uncommon.

Eventually a third phase follows: the follow-through, essentially maximum of 13 weeks similar to the initial gateway. It consists of intensive job search as well as training courses to maintain the skills acquired during the option period.

The program was launched in January 1998 in selected areas (pilot period) and then extended to the rest of the nation in April of that year (national roll-out), it is still ongoing and it is by far the largest active labor market policy in the UK. About 1 million young britons have been involved since the beginning to December 2003, of which roughly 75% are males.

3.3 Identification Strategy

The NDYP is a prototypical example of a multiple treatments program. It combines job search assistance with training/education, job creation and subsidies to employment. In this work the focus is on the relative effectiveness of the different treatments offered in terms of wage returns when a regular job is found. As in any evaluation exercise the main difficulty is that of recovering a missing counterfactual outcome.[2] An individual cannot be in two mutually exclusive treatments at the same time. On the other hand simply using a different set of individuals to recover a counterfactual outcome is not advisable given that individuals self-select into different treatments being this a decision and not a random assignment. However, the option assignment mechanism does here provide a meaningful tool to disentangle the causal effect of a particular treatment.

[2]For an extensive discussion see Heckman *et al.*, 1999; and Blundell and Costa-Dias, 2000.

Options in the NDYP are delivered by a local office (Unit of Delivery) that deals as well with benefit payments and a number of activities for different types of unemployed. The assignment to a particular one cannot be assumed as random. In fact there is evidence (Table 3.2) that better quality unemployed were more often assigned to the subsidized employment option than to the extended gateway. Furthermore, the assignment rules themselves contemplated the selection into a particular treatment. In principle, the assignment process is a joint decision of the caseworker and the unemployed. They are supposed to discuss different possibilities and then agree on the one that should have been the most beneficial to the particular subject. This would mean that all possible options should have been at least mentioned during the compulsory meetings, with the limitation (not followed in practice) that the training/education option should not be offered to 'highly' educated unemployed.

However, there is consistent evidence (section 3.5) that some UoD's tend to favor a particular option either because not all the options are available in a given area at a given time (*rationing*) or because of differential effort/cost linked to the specific treatment (*cost*). It is more costly in this respect to contact potential employers (subsidized employment option) than simply sending participants to school (full time education and training option), in the same line the monetary costs of the former are much higher than for the latter.[3]

Therefore after controlling for a number of possibly confounding covariates (both at the individual and location level) the UoD would constitute a suitable instrument: 1) partially correlated with the particular treatment and 2) rightly excluded from the outcome equation.[4]

[3] In a simple back of the envelope computation a full time education/training option would cost roughly 160 GBP a month (40 GBP*4weeks) or 300 USD while the subsidized employment option 365 GBP a month (60 GBP*4 weeks+750/6 months) or 690 USD. Even if we were to consider the possible different durations maximum six months for the subsidy and maximum 10 (formally 12 however courses last maximum 10) months for the education/training the ratio would still be to one half.

[4] Ideally one would like to control for UoD specific characteristics, i.e. tenure of caseworkers, education, budget, etc.

It is here informative to report some quotations from a qualitative survey conducted on NDYP participants (Woodfield *et al.*, 2000):

David has always wanted to train as printer...NDYP was unable to find a subsidized placement for him in the printing field and he did not recall being offered full time education as an Option. He was eventually placed on the Voluntary Sector Option.... He believed that a training course would have provided him with longer term prospects and a possible career.

Amjad was a 24 year old graduate.....He felt that could have been given much better help whilst trying to find work in his specialist area. Eventually he was placed on the FTET Option...

...Julie wanted to pursue a photographic career but found that she would have to wait at least four months before New Deal could start to find her an appropriate course ...

In this line, there a number of other testimonies.[5]

It is also informative in this respect to look at Figure 3.2 and Table 3.5 and 3.6 where the crude variation in the differential take-up rates across UoD's is quite striking. For example *Tower Hamlets* one of the London Boroughs has the lowest take-up (8.3%) for the extended gateway, among positive values in all options (given a duration longer than four months), while *Coventry* has the highest take-up rate (59%). The same argument goes for the subsidized employment option, where the take-up is lowest (2.6%) in *Wearside*, Nort-East of England, and highest (45.4%)

However this has proved impossible after 3 years of negotiations with the UK Department for Work and Pensions (DWP), although since negotiations are still open there is a chance to be blessed and receive such interesting information.

[5] A program administrator told me of a particular case when the caseworker in an attempt to place a participant in a specific training course phoned the local college only the hear that that specific course was not given in the current term and that possibly it would have been available six months later.

in *Dunbarton*, Scotland. Reading through the Tables above consistently confirms a large variation in option take-up rates. However formal testing is required given the mere nature of cross-tabulation. A number of tests on the first stage of the estimation will be presented in Section 3.5 when such variation will be analyzed in a regression framework.

Formalizing the identification strategy in the familiar potential outcome framework (Rubin, 1986; and Heckman *et al.*, 1999) and assuming that the Stable Unit Treatment Value Assumption (SUTVA) holds. In a multiple mutually exclusive treatment model, omitting individual i subscript, the observed outcome Y is written:

$$Y = Y_0 + \sum_{j=1}^{J}(Y_j - Y_0)D_j \tag{3.1}$$

where Y_0 is the potential outcome for the reference treatment (extended gateway) and $D_j = 1$ if treatment j is realized.

There are therefore a number of treatment effects that can be defined, however I will restrict my attention to the comparison between treatment j and 0 in the class of treatment on the treated. Listing them[6]:

$$E(\Delta_j|D_j = 1) = E(Y_j - Y_0|D_j = 1) \tag{3.2}$$

for $j = 1, .., J$. It is also very useful to decompose each potential outcome in two parts: i) a deterministic function of a number of appropriate covariates; ii) and a stochastic one representing the possible heterogeneity in returns.

[6]Whenever not otherwise specified the expectations are taken with respect to X as well throughout this section.

$$Y_j = \mu_j + \epsilon_j \tag{3.3}$$

$$Y_0 = \mu_0 + \epsilon_0 \tag{3.4}$$

where μ is a function of the covariates X here omitted, while ϵ represents the heterogeneity in the returns.

The problem in estimating any treatment effect is fundamentally that of missing data: an individual cannot be in two different states at the same time. Since the parameters I am interested in are defined as pair-wise comparison between option j and the base case 0, only observations on those pair of treatments will be used for identification (Lechner, 1999; and Frölich, 2004). It is therefore possible, without loss of generality, to rewrite what follows in terms of two potential outcomes where the treatment state can be considered binary. So $D = 0$ would simply indicate that treatment is 0 and $D = 1$ that the treatment state under scrutiny is realized. These simplifications would just make the identification section more readable. Therefore, constructing an example based on the comparison between option 1 and 0, rethinking of it as a binary treatment, it is possible to write:

$$\begin{aligned} &E(Y|D=1) - E(Y|D=0) = \\ &E(\Delta|D=1) + E(\epsilon_0|D=1) - E(\epsilon_0|D=0) \end{aligned} \tag{3.5}$$

Therefore by simple comparison, the treatment on the treated would not be identified without further assumptions. The last two terms on the RHS constitute the bias given by the difference in the unobservables in the base state.[7]

[7] For a decomposition of the bias term see Ichimura *et al.* 1998.

This is the standard evaluation problem of possible correlation between the treatment state $D = 0, 1$ and the unobservable ϵ_0. In a "proper" randomized experiment this problem would not arise since the bias would cancel out. However, in a non experimental setting this is not generally the case, i.e. individuals self-select in a particular treatment according to possible gains unobserved to the econometrician.[8]

A possible solution to the problem is that of instrumental variables. However, such variables are quite difficult to find. There are several examples where cleverly devised instrument at a first glance proved not as convincing later on. This has generated a vast literature (Angrist and Imbens, 1994; Heckman, 1997) and some controversial debates on what is ultimately identified and what is not under the weakest possible assumptions. The final goal of this section is to formally state the identifying conditions and discuss the validity of such conditions in the specific case. Rewriting the outcome equation 3.1 in terms of equations 3.3 and 3.4:

$$(3.6)\quad Y = \mu_0 + E(\Delta|D = 1)D + \{\epsilon_0 + D[(\epsilon_1 - \epsilon_0) - E(\epsilon_1 - \epsilon_0|D = 1)]\}.$$

Therefore in order to identify $E(\Delta|D = 1)$ the problem arises from the correlation between D and ϵ_0. If there is a variable Z such that:

1. $E[\{\epsilon_0 + D[(\epsilon_1 - \epsilon_0) - E(\epsilon_1 - \epsilon_0|D = 1)]|Z\}] = 0$
2. $E[D|Z] = Pr[D = 1|Z]$ it is a non trivial function of Z.

The two conditions stated above guarantee the identification of the relevant parameter, however they are non-standard in the sense that the whole conditional expectation of the error term, in curly brackets, in 3.6 has to be equal to 0, while in

[8] There is however some recent evidence on how noisy decisions are and on how program participants are often bad program evaluators (in certain respects). See Smith *et al.* 2006.

a standard dummy endogenous variable model this condition would be in terms of conditional mean independence between ϵ_0 and D given the instrument Z. Notice that the conditions as well as the analysis should be extended to the case where there are J endogenous treatments and therefore for identification at least J instruments would be needed.

It is ultimately a matter of judgement whether a variable is a suitable instrument, i.e. can be excluded from the outcome equation while generating independent variation in the treatment status once the other covariates have been partialled out. In the particular case the instruments I propose are the following: the local units of delivery. As explained earlier each participant has to be registered with a particular unit in her local area. This is generally a governmental employment office, varying in size and resources according to the location. Once the young unemployed reaches the sixth month in "open" unemployment, she is forced to enter the program and attend an interview, in the UoD, with a caseworker. If a regular job is not found in the first four months of treatment the participant should be placed in one of the option, however this is not always the case since there are a number of participants who will be offered a longer job search assistance. Once again, the assignment to different options should be in principle a joint decision between the caseworker and the treated. However, as explained above, there are two main reasons why this is not always the case: (i) *rationing*, i.e. not all options are available to each UoD; (ii) *costs*, i.e. caseworkers or the UoD as a whole might not carefully look for potential employers (low effort, tight budget or a combination of the two) and therefore they would have a limited number of placement available totally unrelated with the unobservables of the treated; a tight budget might not allow to create an Environmental Task Force (ETF) and so forth.

There is substantive and anecdotal evidence that different employment offices tend to favor particular options either because some of the treatments were simply not available in the area or because it was too costly in terms of effort and budget to deliver certain treatment. As standard, assumption (2) can be tested in the first stage of a standard two stage least square estimation. While assumption (1) is generally more controversial, however given the number of instruments at hand it is possible to indirectly test such assumption as well. Once I control for local labor market conditions, i.e. local average wages and a NUTS2 level fixed effect, the unit of delivery should be excludable from the outcome equation, where the outcome is the hourly wage rate in a non "new deal" placement. A second set of issues are that of weak instruments and many instruments (See Staiger and Stock, 1997; Hansen *et al.* 2006.). Furthermore, wages are only observed for those who actually have (or had) a job, after program participation, and therefore this induces the standard selection problem (Heckman, 1979). I will address such concerns in Section 3.5.

3.4 Data

The data used in this work combines an *ad hoc* survey, i.e. The New Deal Survey of Participants, with the New Deal Evaluation Database (NDED), an administrative dataset of the NDYP purposely built and containing virtually the entire population of new dealers.

The former covers a cohort of participants, who entered the program in 1998 between August 31 and November 27 and were interviewed twice. The first interview was held between February 20 and July 30 of 1999; the second and last between February 25 and June 1 of 2000. The first interview was held on average less than one year after program entry, while the second one and a half years after the initial entry.

Such data are combined with the administrative data (NDED) containing virtually all participants and a number of information on activities during the New Deal, i.e. type and durations of various treatments, actions taken, reason for leaving a particular state, etc. While the survey data contain information on some background characteristics, i.e. unemployment history, types of jobs held, education, gender, ethnicity, etc.

Table 3.1 gives the distribution of participants in the various options including the gateway, column (i) includes in the gateway all those who passed by such stage even if they left the program within four months; column (ii) only those who remained in the program long enough to reach the option period according to the stated rules, i.e. longer than 4 months. It is clear that most participants were enrolled in full time education and training options where they had to attend (mostly) vocational courses in the local college. Take-up rates for the other options are quite even. It is worth noticing that the practice of keeping new dealers in longer job search (extended gateway) was not exceptional, roughly 30 percent of participants, with a duration longer then 120 days, never entered a formal option.

The UoD's are located in the local area, here defined as at level NUTS2, e.g. Greater London, Cambridge and East Anglia would be a unique local area. Therefore in a local area, a limited geographic space, there might be multiple offices. In Table 3.3, the distribution of UoD's is presented, the median number of units is 4 with a maximum of 19 (Greater London). There are in total 105 units in 23 geographical areas, therefore the average number of units is equal to roughly 5.[9]

[9]Only areas with at least 2 units have been included in the analysis, given the type of variation exploited. It has to be remembered that throughout the analysis I am controlling for the local area (NUTS2) and average wages in the specific location area of the particular unit, e.g. Hackney, Chelsea, etc. would have specific average wages (for a full time employee) as from the UK Labor Force Survey for the year 2000.

3.5 Results

The source of identification exploited in this work relies on the UoD's where participants are registered. Once again this is not chosen by the participant, but it is tied to the location of residence. The idea is that the UoD's will give a source of exogenous variation in option assignment once confounding factors are taken care of.

As in any sound use of instrumental variables, it is crucial to support the validity of the instrument proposed on the basis of the available evidence and of an economic mechanism exogenous to both the participant and the caseworker. In section 3.3 I do present a number of testimonies supporting the chosen instruments. Furthermore, there is a clear-cut evidence on the large variation in option take-up across UoD's. Although ultimately untestable, being the identifying set of assumptions, further evidence will be here provided in the light of the first step of estimation in a 2SLS estimator.[10] Various robustness checks have been implemented using only a subset of instruments to check for the many instruments and weak instruments well known problems. First stage results support the instrument chosen in the following sense: i) F-statistics are large (above 10 in the vast majority of cases); ii) Hansen-Sargan J-test confirms the validity of the instruments cannot reject the null at 10%; iii) Anderson-Rubin test for endogenous regressors on the same note of the two above (rejects the null at 5%).[11] The above are also consistent whether a subset of the instruments is chosen; notice that there are over 100 UoD's in the data, although at most 60 are used in the estimation. Some UoD's might have very few participants sampled and therefore I conducted also the entire analysis excluding those UoD's with fewer than 10 (5% of the sample) or 20 (15% of the sample) and the results still

[10] Efficient GMM results, available upon request, are qualitatively identical.

[11] The F-statistics above 10 are actually suggested in Stock and Yogo (2003) for the case of a single endogenous variable.

hold true both in the first and second stage. A summary of the first stage results and robustness checks is presented in Table 3.7, where first stage of the main analysis shows large F-statistics (but for the Voluntary Sector option). Further in column 3 and 4, I perform the following exercise: estimation of the treatment effect of the subsidy versus any other treatments, in such case the endogenous variable would be just one. Whether I use all the available IV's or just the one with the largest significance in first stage, the conclusions on the effect of the subsidy still stands robust. Dealing with the number of instruments if added instruments are irrelevant, I select, in column 5 to 7 the eight IV's with highest explanatory power in the first stage (for the 5 endogenous variables), once again the main results are yet confirmed and, if anything, the point estimates are larger suggesting even a larger penalty from taking up the subsidy. However, such point estimates are not statistically different from the main one produced in Table 3.4. Column 8 and 9 confirm the robustness of the results once the smaller units are dropped from the analysis.

Further, throughout the analysis a set of covariates are added to control for confounding factors. In fact the nature of the instruments may induce some concern regarding possible differences in the local labor market conditions that would drive the outcome and clearly bias the analysis. At the local level I am controlling for a location fixed effect at NUTS2 level, local wages in the specific sub-NUTS2 level area and in certain specifications also for the local unemployment rates. At the individual level a number of commonly thought fundamental variables (before program participation) are added: unemployment history, dummy for whether a job was found within three months of leaving school, education level, gender, and finally a number of controls for attitude towards work (this questions where asked in the initial stage of the new deal and certainly pre-option assignment, however excluding them from the analysis does not change the results).

Another issue is that of tackling the non-random employment node: wages are only observed for those who have or had a job since leaving NDYP. Although at a first stance this might not be a major problem since the options did not seem to have a large differential impact on (re)employment probability in my sample[12] However, as a robustness check the analysis is also performed with an initial stage where the employment node is modeled according to a selection equation where the excluded instruments are some of the UoD's that do not appear to have explanatory power in the first stage equations. As confirmed in Table 3.4, results are robust to such a possible non random selection; and the selection term is never significantly different from zero but in one case. Notice then even when the selection parameter is significantly different from zero, the point estimates of the parameters of interest vary very little.

Looking through table 3.4 where results are presented for both samples included in the analysis, it is quite evident that OLS estimates are of very small magnitude and furthermore insignificant.[13] This is the result for all treatment effects as defined earlier. It does not seem that any of the options has a differential impact on the treated compared to the reference 'option' of the extended gateway.

If we were to believe no selection in both option participation and labor supply we would then conclude that none of the options does on average better than the extended gateway in terms of wages once a regular job is found. However once we take into consideration that option assignment is a decisional process instead of a simple random assignment (even based on observables), a whole different picture starts to take shape. Regardless of the sample used the worst performer seems actually to be

[12]This result differs from Dorsett (2006), who finds that the subsidized employment option is the star in this context, this is actually the result you would get in a simple OLS estimation.

[13]Bootstrap standard errors, clustered at the UoD level, are presented in the case of correction for selection.

what has been sometimes termed the 'star' option, i.e. the employment subsidy.[14] It is so much so that a negative return of roughly 22 percent is found.

There is instead no evidence that the other options are worse than the extended job search in a pair-wise comparison as detailed in Section 3.3.[15]

Therefore if we can accept the indirect evidence on the validity of the instruments used, we can conclude that participating in an employment subsidy program is not as beneficial as commonly believed, especially in the anecdotal evidence given the complete absence of any scientific evidence. Why is then that OLS estimates point to a no effect when compared to the gateway? It is more than plausible that the non-random selection in option assignment pushes the more able participants (Table 3.2 in such direction while in fact the particular treatment has a negative return compared to an extended job search period. In the particular exercise performed in this work it is not straightforward to believe that the object estimated is a Local Average Treatment Effect (LATE) as in Angrist and Imbens, 1994. It is quite difficult to relate the type of instruments employed here with the idea of moving only a particular type of treated (*compliers*). The source of variation here would move participants simply because of rationing (or costs) irrespective of their intrinsic characteristics.[16]

3.6 Conclusions

This paper evaluates the relative effectiveness of different treatments offered by the largest UK program devoted to young unemployed who experience a six months unemployment spell while younger than 25 years of age. There are fundamentally

[14] The word has been mentioned to me a number of times from different program administrators.

[15] Direct comparison between the four option would be incorrect given that treatment effects are asymmetric in this context (Lechner, 1999).

[16] Use of LIML method, in certain context better performing in the weak/ many instruments case, has proved unrealistic in terms of estimates in line with the well-known result of Hahn and Hausman (2003) in the analysis of "no-moment" conditions estimators. It is not clear that the use of LIML per se is better than the simple 2SLS unless particular conditions are satisfied.

five types of treatment offered (i.e. extended job search, full time education/training, subsidized employment, voluntary sector placement, and governmental job in the Environmental Task Force), while an initial job search assistance is offered to every participant. The central question is that of identifying the causal impact of each treatment with respect to the extended job search option. What is offered in the program is highly heterogenous while the baseline option is centered on improving the search technology with little commitment to human capital improvement the remaining options focus on different degree of combination between some human capital accumulation (through classroom training or learning on the job) and work experience (to enhance attachment to the labor force and motivation). A high degree of heterogeneity also exist on the costs of each treatment (Section 3.3). It is therefore crucial to understand which one of the options is delivering the best outcome in terms of wage returns once a regular job is found. The challenge is then to disentangle what is the true effect of each options once possible confounding factors are partialled out. The approach followed in this study is that of IV, where a suitable set of instruments is found in the UOD's to which the particular treated is associated. A crucial step is that of providing convincing evidence on the validity of the instruments used. Here I propose and substantiate my claim through a number of formal tests and a series of testimonies consistent with a twofold set of mechanisms through which the set of instrument could be considered valid: i) *rationing*, i.e. not all options are available at a given point in time in a particular UOD; and ii) *costs*, i.e. some options are inherently costlier both in terms of money and effort. The surprising results, however consistent with previous evidence from the US (Katz, 1998), is that in fact the subsidized employment option is delivering a negative return in terms of wages, such effect is actually quite large. Adding up to almost 40 GBP (75 USD) per week. It seems that once the non random selection into different treatments is take care of, participants to the specific option are seen as to be of low productivity and therefore stigmatized by the regular employer. It might also be that the matching process employee jobs gets worsened by the fact that participants in the subsidy option are forced into accepting low wage jobs or jobs not particularly suited for those participants. It however remains an open question whether such effect is due to die out over time, once the employer learns the true productivity of the individual.

3.7 Tables and Figures

Table 3.1: Options

	ALL		*EXTENDED GATEWAY*	
	N	%	N	%
Gateway	1,530	*45.69*	822	*31.30*
Subsidy	296	*8.84*	294	*11.20*
FTET	821	*24.51*	812	*30.92*
VS	345	*10.30*	345	*13.14*
ETF	357	*10.66*	353	*13.44*
N	3349		*2626*	

Note: Figures computed from Survey data.

Table 3.2: **Participant Characteristics by Option**

Option	**Stats**	**Educ.**	**Problem**	**Male**	**Age 2nd**	**Temp-job**	**Work**	**Unemp. Hist.**
GATE	mean	1.93	0.13	0.71	22.50	0.36	0.45	9.76
	sd	0.85	0.34	0.46	1.96	0.48	0.50	13.83
SUBS.	mean	2.08	0.07	0.72	22.56	0.41	0.55	13.60
	sd	0.73	0.25	0.45	1.86	0.49	0.50	16.53
FTET	mean	1.95	0.15	0.71	22.42	0.30	0.36	8.64
	sd	0.73	0.35	0.45	1.86	0.46	0.48	13.09
VS	mean	1.99	0.16	0.58	22.51	0.36	0.34	7.28
	sd	0.88	0.36	0.49	1.97	0.48	0.47	11.34
ETF	mean	1.71	0.19	0.93	22.48	0.35	0.43	10.74
	sd	0.74	0.39	0.25	2.00	0.48	0.50	13.54
Total	mean	1.93	0.14	0.72	22.48	0.34	0.42	9.65
	sd	0.80	0.35	0.45	1.93	0.48	0.49	13.70

Note: Only participants with duration longer then 120 days are included in the table.
Educ. stands for education, a categorical variable taking values 1-5: 1=no qualifications, 5=higher qualifications.
Problem: whether ever had reading/maths problem, 1=yes and 0=no. Age 2nd: age at second interview.
Temp-Job: whether had a series of temporary job after leaving full-time education, 1=yes and 0=no.
Work: whether found a job within 3 months of leaving full-time education, 1=yes and 0=no.
unemp. Hist.: months of unemployment since 1993.

Table 3.3: UoD's in the Local Area

	Median	Min.	Max	N
UoD's	4	2	19	105
Areas				23

Table 3.4: Treatment Effects

	ALL				*EXTENDED GATEWAY*			
	OLS		*IV*		*OLS*		*IV*	
	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)
Subsidy	-.0221	-.0217	-.2204	-.2191	-.0340	-.0336	-.2221	-.2234
	[.0431]	[.0403]	[.1059]	[.1130]	[.0505]	[.0417]	[.0994]	[.1063]
FTET	-.0083	-.0070	.0187	.0374	-.0265	-.0250	-.0737	-.0550
	[.0294]	[.0336]	[.1058]	[.1133]	[.0359]	[.0399]	[.1084]	[.1151]
VS	.0143	.0162	-.0606	-.0327	.0043	.0066	-.1121	-.0822
	[.0420]	[.0369]	[.1378]	[.1418]	[.0465]	[.0398]	[.1473]	[.1829]
ETF	.0341	.0343	.2106	.2446	.0076	.0095	.0944	.1154
	[.0469]	[.0337]	[.1810]	[.2017]	[.0534]	[.039]	[.1613]	[.2133]
λ		-.0446		-.0465		-.0708		-.0752
		[.0476]		[.0366]		[.0623]		[.0368]
Local Area	yes	yes	yes	yes	yes	yes	yes	yes
Wage[a]	yes	yes	yes	yes	yes	yes	yes	yes
Education	yes	yes	yes	yes	yes	yes	yes	yes
Gender	yes	yes	yes	yes	yes	yes	yes	yes
History[b]	yes	yes	yes	yes	yes	yes	yes	yes
Attitude[c]	yes	yes	yes	yes	yes	yes	yes	yes
N	*885*				*635*			

Note: (i) and (ii) without and with sample selection for those who actually work respectively.
Bootstrap [standard errors], columns (ii) clustered at the UoD level obtained with 1000 replications.
[a]: local wage rate for full-time employee (UK LFS, 2000). [b]: months of unemployment since leaving school, whether had a series of temp jobs, whether a job was found within 3 months of leaving school.
[c]: Attitude towards work interviews, luck in the work place, motivation and benefit sanctions.

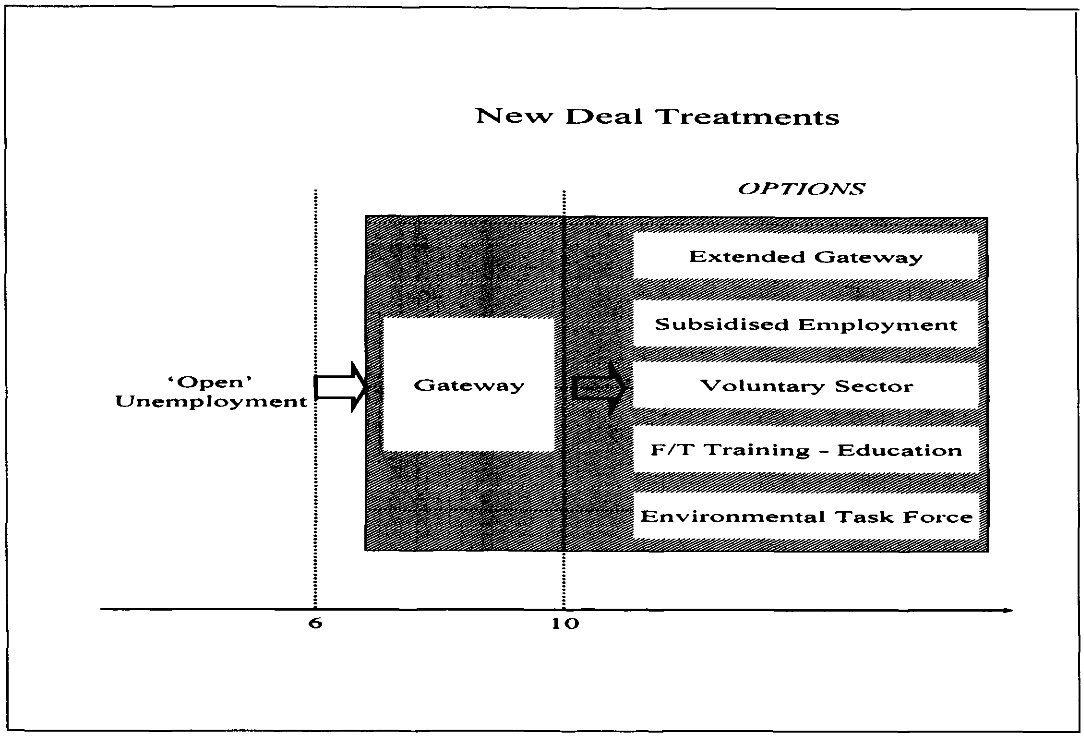

Figure 3.1: The program

3.8 APPENDIX

Table 3.5: Options take-up by UoD's for low educated individuals

UoD	Gate	Subsidy	FTEDT	VS	ETF
ayrshire	50.00	11.11	27.78	11.11	0.00
dunbarton	23.08	38.46	23.08	15.38	0.00
edinburgh, east and mid lothian	50.00	10.00	10.00	20.00	10.00
fife	54.29	11.43	22.86	2.86	8.57
forth valley	50.00	50.00	0.00	0.00	0.00

Continued on next page

Table 3.5 – continued from previous page

UoD	Gate	Subsidy	FTEDT	VS	ETF
glasgow	33.33	7.84	29.41	9.80	19.61
moray, strathspay and badenoch	57.14	14.29	28.57	0.00	0.00
lanarkshire	52.17	10.14	15.94	5.80	15.94
renfrewshire	63.04	8.70	13.04	8.70	6.52
west lothian	38.46	7.69	7.69	30.77	15.38
argyll and the islands	60.00	20.00	0.00	0.00	20.00
ross and cromarty	80.00	0.00	20.00	0.00	0.00
northumberland	20.00	20.00	40.00	20.00	0.00
tyneside north	39.39	3.03	48.48	6.06	3.03
durham north and durham south	27.94	13.24	38.24	5.88	14.71
wearside	32.56	2.33	34.88	18.60	11.63
tees north and south	32.73	12.73	41.82	7.27	5.45
bolton	66.67	20.00	0.00	13.33	0.00
central lancashire	36.36	9.09	18.18	9.09	27.27
city pride (manchester	60.42	10.42	18.75	4.17	6.25
east lancashire	41.67	16.67	33.33	8.33	0.00
knowsley	19.23	7.69	57.69	7.69	7.69
liverpool	35.19	14.81	37.04	4.63	8.33
north lancashire	27.59	13.79	34.48	13.79	10.34
oldham	42.86	14.29	14.29	14.29	14.29
rochdale	46.15	15.38	15.38	0.00	23.08

Continued on next page

Table 3.5 – continued from previous page

UoD	Gate	Subsidy	FTEDT	VS	ETF
sefton	32.14	10.71	32.14	10.71	14.29
st helens	18.18	9.09	36.36	9.09	27.27
stockport	62.50	0.00	12.50	12.50	12.50
w. lancashire	60.00	0.00	30.00	0.00	10.00
wigan	35.71	28.57	14.29	7.14	14.29
south humber	32.14	0.00	21.43	25.00	21.43
calderdale and kirklees	50.00	9.68	27.42	6.45	6.45
barnsley and the dearne	0.00	14.29	57.14	14.29	14.29
bradford	42.22	4.44	35.56	8.89	8.89
hull	30.88	11.76	32.35	7.35	17.65
leeds	50.00	0.00	23.68	18.42	7.89
north yorkshire	38.46	7.69	15.38	23.08	15.38
wakefield and doncaster	27.50	17.50	32.50	10.00	12.50
north wales coast	20.00	0.00	40.00	20.00	20.00
cardiff and vale	53.85	7.69	7.69	23.08	7.69
bridgend and glamorgan valleys	37.50	25.00	18.75	12.50	6.25
heads of the valley and caerphilly	26.92	25.00	30.77	3.85	13.46
newport, torfaen and monmouth	55.00	5.00	20.00	15.00	5.00
birmingham	44.87	8.97	34.62	5.13	6.41
solihull	56.25	6.25	25.00	12.50	0.00
staffordshire	33.33	27.78	16.67	5.56	16.67

Continued on next page

Table 3.5 – continued from previous page

UoD	Gate	Subsidy	FTEDT	VS	ETF
hereford and worcester	55.56	11.11	11.11	22.22	0.00
coventry	65.38	7.69	11.54	7.69	7.69
warwickshire	60.00	0.00	20.00	0.00	20.00
south derbyshire	50.00	0.00	0.00	50.00	0.00
cambridge ttwa	46.15	7.69	30.77	15.38	0.00
north derbyshire	44.00	20.00	32.00	0.00	4.00
leicestershire	53.19	4.26	29.79	10.64	2.13
lincolnshire	25.00	0.00	25.00	50.00	0.00
norfolk	11.76	23.53	35.29	11.76	17.65
northamptonshire	75.00	0.00	0.00	25.00	0.00
greater nottingham	45.61	8.77	21.05	19.30	5.26
north nottinghamshire	31.71	7.32	31.71	12.20	17.07
peterborough	50.00	0.00	25.00	0.00	25.00
suffolk	37.50	12.50	50.00	0.00	0.00
exeter and east devon	27.27	9.09	36.36	9.09	18.18
north devon	0.00	33.33	66.67	0.00	0.00
plymouth	27.78	27.78	27.78	11.11	5.56
lambeth	75.00	0.00	12.50	12.50	0.00
hackney and city	80.00	0.00	20.00	0.00	0.00
brighton	61.54	7.69	23.08	0.00	7.69
canterbury	42.86	7.14	21.43	7.14	21.43

Continued on next page

Table 3.5 – continued from previous page

UoD	Gate	Subsidy	FTEDT	VS	ETF
chatham	36.36	4.55	27.27	13.64	18.18
croydon and bromley	40.00	40.00	20.00	0.00	0.00
guildford	50.00	50.00	0.00	0.00	0.00
lewisham	55.56	11.11	22.22	5.56	5.56
maidstone, dartford and mid kent	80.00	20.00	0.00	0.00	0.00
southwark	58.62	3.45	20.69	17.24	0.00
greenwich	33.33	0.00	66.67	0.00	0.00
sutton, merton esher, kingston, epsom	53.85	7.69	23.08	7.69	7.69
wandsworth	61.54	0.00	30.77	7.69	0.00
west sussex coastal plain	50.00	0.00	0.00	50.00	0.00
south essex	60.00	20.00	20.00	0.00	0.00
north and mid essex	50.00	12.50	25.00	12.50	0.00
edgeware and leaside	44.44	11.11	33.33	11.11	0.00
north east london	52.63	10.53	21.05	10.53	5.26
havering, barking and dagenham	20.00	0.00	20.00	60.00	0.00
newham	50.00	0.00	31.25	12.50	6.25
tower hamlets	50.00	9.09	13.64	27.27	0.00
camden and north islington	42.86	0.00	28.57	14.29	14.29
ealing	57.14	0.00	42.86	0.00	0.00
bedfordshire and luton	55.00	0.00	45.00	0.00	0.00
oxfordshire	25.00	0.00	75.00	0.00	0.00

Continued on next page

Table 3.5 – continued from previous page

UoD	Gate	Subsidy	FTEDT	VS	ETF
portsmouth and se hampshire	66.67	11.11	11.11	11.11	0.00
isle of wight	0.00	0.00	0.00	50.00	50.00
slough	79.17	0.00	12.50	8.33	0.00
wembley	60.00	0.00	30.00	10.00	0.00
hammersmith, fulham, kensington, chelsea	57.14	0.00	42.86	0.00	0.00
Total	43.14	10.13	27.86	9.74	9.14
Total	0.39	1.04	0.60	1.06	1.13

Note: .

Table 3.6: Options take-up by UoD's for low educated individuals if new deal lasts more than 120

UoD	Gate	Subsidy	FTEDT	VS	ETF
ayrshire	30.77	15.38	38.46	15.38	0.00
dunbarton	9.09	45.45	27.27	18.18	0.00
edinburgh, east and mid lothian	37.50	12.50	12.50	25.00	12.50
fife	33.33	16.67	33.33	4.17	12.50
forth valley	50.00	50.00	0.00	0.00	0.00
glasgow	22.73	9.09	34.09	11.36	22.73
moray, strathspay and badenoch	25.00	25.00	50.00	0.00	0.00
lanarkshire	29.79	14.89	23.40	8.51	23.40

Continued on next page

Table 3.6 – continued from previous page

UoD	Gate	Subsidy	FTEDT	VS	ETF
renfrewshire	41.38	13.79	20.69	13.79	10.34
west lothian	27.27	9.09	9.09	36.36	18.18
argyll and the islands	33.33	33.33	0.00	0.00	33.33
northumberland	20.00	20.00	40.00	20.00	0.00
tyneside north	20.00	4.00	64.00	8.00	4.00
durham north and durham south	15.79	15.79	43.86	7.02	17.54
wearside	26.32	2.63	36.84	21.05	13.16
tees north and south	21.28	14.89	48.94	8.51	6.38
bolton	44.44	33.33	0.00	22.22	0.00
central lancashire	30.00	10.00	20.00	10.00	30.00
city pride (manchester	45.45	15.15	24.24	6.06	9.09
east lancashire	36.36	18.18	36.36	9.09	0.00
knowsley	12.50	8.33	62.50	8.33	8.33
liverpool	22.22	17.78	44.44	5.56	10.00
north lancashire	22.22	14.81	37.04	14.81	11.11
oldham	20.00	20.00	20.00	20.00	20.00
rochdale	22.22	22.22	22.22	0.00	33.33
sefton	13.64	13.64	40.91	13.64	18.18
st helens	18.18	9.09	36.36	9.09	27.27
stockport	40.00	0.00	20.00	20.00	20.00
w. lancashire	50.00	0.00	33.33	0.00	16.67

Continued on next page

Table 3.6 – continued from previous page

UoD	Gate	Subsidy	FTEDT	VS	ETF
wigan	18.18	36.36	18.18	9.09	18.18
south humber	24.00	0.00	24.00	28.00	24.00
calderdale and kirklees	35.56	13.33	33.33	8.89	8.89
barnsley and the dearne	0.00	14.29	57.14	14.29	14.29
bradford	36.59	4.88	39.02	9.76	9.76
hull	23.33	11.67	36.67	8.33	20.00
leeds	24.00	0.00	36.00	28.00	12.00
north yorkshire	20.00	10.00	20.00	30.00	20.00
wakefield and doncaster	21.62	18.92	35.14	10.81	13.51
north wales coast	20.00	0.00	40.00	20.00	20.00
cardiff and vale	50.00	8.33	8.33	25.00	8.33
bridgend and glamorgan valleys	28.57	28.57	21.43	14.29	7.14
heads of the valley and caerphilly	11.90	28.57	38.10	4.76	16.67
newport, torfaen and monmouth	35.71	7.14	28.57	21.43	7.14
birmingham	30.65	11.29	43.55	6.45	8.06
solihull	56.25	6.25	25.00	12.50	0.00
staffordshire	20.00	33.33	20.00	6.67	20.00
hereford and worcester	50.00	12.50	12.50	25.00	0.00
coventry	59.09	9.09	13.64	9.09	9.09
warwickshire	33.33	0.00	33.33	0.00	33.33
south derbyshire	50.00	0.00	0.00	50.00	0.00

Continued on next page

Table 3.6 – continued from previous page

UoD	Gate	Subsidy	FTEDT	VS	ETF
cambridge ttwa	30.00	10.00	40.00	20.00	0.00
north derbyshire	30.00	25.00	40.00	0.00	5.00
leicestershire	31.25	6.25	43.75	15.62	3.12
lincolnshire	25.00	0.00	25.00	50.00	0.00
norfolk	12.50	25.00	31.25	12.50	18.75
northamptonshire	50.00	0.00	0.00	50.00	0.00
greater nottingham	26.19	11.90	28.57	26.19	7.14
north nottinghamshire	26.32	7.89	34.21	13.16	18.42
peterborough	50.00	0.00	25.00	0.00	25.00
suffolk	28.57	14.29	57.14	0.00	0.00
exeter and east devon	11.11	11.11	44.44	11.11	22.22
north devon	0.00	33.33	66.67	0.00	0.00
plymouth	18.75	31.25	31.25	12.50	6.25
lambeth	0.00	0.00	50.00	50.00	0.00
hackney and city	75.00	0.00	25.00	0.00	0.00
brighton	44.44	11.11	33.33	0.00	11.11
canterbury	27.27	9.09	27.27	9.09	27.27
chatham	26.32	5.26	31.58	15.79	21.05
croydon and bromley	0.00	66.67	33.33	0.00	0.00
guildford	50.00	50.00	0.00	0.00	0.00
lewisham	46.67	13.33	26.67	6.67	6.67

Continued on next page

Table 3.6 – continued from previous page

UoD	Gate	Subsidy	FTEDT	VS	ETF
southwark	42.86	4.76	28.57	23.81	0.00
sutton, merton esher, kingston, epsom	40.00	10.00	30.00	10.00	10.00
wandsworth	50.00	0.00	40.00	10.00	0.00
west sussex coastal plain	50.00	0.00	0.00	50.00	0.00
south essex	50.00	25.00	25.00	0.00	0.00
north and mid essex	33.33	16.67	33.33	16.67	0.00
edgeware and leaside	37.50	12.50	37.50	12.50	0.00
north east london	43.75	12.50	25.00	12.50	6.25
havering, barking and dagenham	0.00	0.00	25.00	75.00	0.00
newham	36.00	0.00	40.00	16.00	8.00
tower hamlets	8.33	16.67	25.00	50.00	0.00
camden and north islington	20.00	0.00	40.00	20.00	20.00
ealing	25.00	0.00	75.00	0.00	0.00
bedfordshire and luton	50.00	0.00	50.00	0.00	0.00
oxfordshire	25.00	0.00	75.00	0.00	0.00
portsmouth and se hampshire	0.00	33.33	33.33	33.33	0.00
isle of wight	0.00	0.00	0.00	50.00	50.00
slough	66.67	0.00	20.00	13.33	0.00
wembley	50.00	0.00	37.50	12.50	0.00
hammersmith, fulham, kensington, chelsea	40.00	0.00	60.00	0.00	0.00
Total	28.66	12.69	34.79	12.31	11.55

Continued on next page

Table 3.6 – continued from previous page

UoD	Gate	Subsidy	FTEDT	VS	ETF
CV	0.53	1.00	0.52	0.98	1.03

Table 3.7: First Stage and Robustness Checks

	FIRST STAGE				*ROBUSTNESS*				
			Subsidy Vs Remaining			*Selected IV*		*Larger UoD's*	
	R2 Shea	F-stats	All IV	1 IV	All	UoD's 10	UoD's 20	UoD's 10	UoD's 20
Subsidy	0.0950	17.55	-.2328	-.2031	-.3483	-.4254	-.5078	-.2088	-0.2067
			[.1043]	[.0641]	[.1834]	[.1834]	[.2146]	[.1100]	[0.1060]
FTET	0.0807	8.88			-.0714	-.0208	-.0396	-.1072	-0.1267
					[.1917]	[.1989]	[.2351]	[.1163]	[0.1203]
VS	0.0783	1.02			-.2801	-.2708	.-.1668	-.1321	-0.0296
					[.2799]	[.2889]	[.2906]	[.1466	[0.1578]
ETF	0.0723	14.64			.3116	.2287	.3935	-.1587	0.1814
					[.5231]	[.5047]	[.5553]	[.2276]	[0.2734]
Anderson									
P-value			.0751	.0103	.0369	.0338	.0629		
Hansen J									
P-value			.1647		.1391	.1563	.3854		
N			885	885	885	856	752	856	752

Note: Same controls as in Table 3.4, [Standard Errors] are heteroscedasticity robust. More than 10 participants drops 5% (or 5 UoD's) of the sample. More than 20 (or 10 UoD's) drops 15% of the sample.

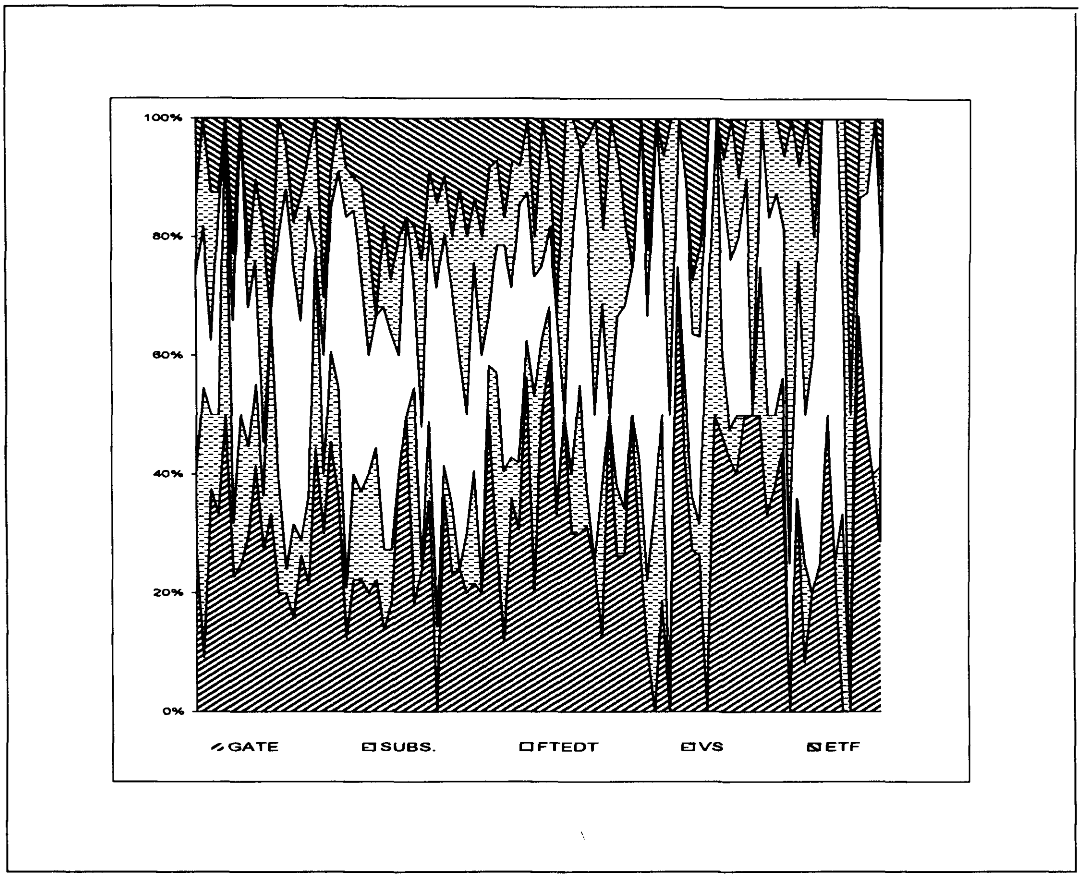

Figure 3.2: Options' Take-Up By UoD's

CHAPTER IV

Indirect Effects of an Aid Program: How do Liquidity Injections Affect Non-Eligibles' Consumption?

with Manuela Angelucci

4.1 Introduction

Policy interventions in developing countries are likely to affect all residents of the areas where they are implemented, especially when village economies and social networks create strong links between a limited number of households. However, the program evaluation literature is mainly focused on estimating the program effects on the treated, rather than the effects on the non-treated or general equilibrium effects. In this paper, we exploit the unique design of a randomized conditional cash transfer program in Mexico, Progresa, to estimate its indirect effect on consumption for non-eligible households who live in treatment areas, and to understand the mechanisms through which this indirect effect occurs. We show that liquidity injections into small rural communities increase the consumption of the non-treated through changes in the credit and insurance markets. Thus, the total effect of Progresa in these communities is larger than its effect on the treated. Our results confirm that a key identifying assumption - that the program has no effect on non-treated individuals - is likely to be violated in similar policy designs.

Conditional cash transfers are a popular type of aid program, which provides monetary transfers to eligible recipients, provided they send their children to school, attend nutrition classes, and have periodic health checks. Programs with this format are currently implemented in numerous countries, including Bangladesh, Bolivia, Brazil, Colombia, Honduras, Jamaica, Mexico, and Nicaragua. Progresa is an on-going conditional cash transfer program targeted at Mexican poor rural households, providing grants to improve education, health, and nutrition.

The design of the experimental trial and the data collected for the evaluation have some unique features. First, the randomization was implemented at the village level. Second, program administrators collected data on all households, both poor and non-poor, although only poor households were eligible for the treatment. Thus, we have information on four groups: poor and non-poor households in treatment and control villages. Non-poor households in control villages provide a valid counterfactual for the non-poor in treatment ones, under the assumptions that assignment is truly random, and that control villages are not indirectly affected by the program. Hence, this particular experimental design enables us to estimate the indirect effects of the program on non-eligible households who live in treatment areas using fairly standard identifying assumptions.

We focus on consumption because it provides an indicator of household well-being. We find that there is a positive, significant, and sizeable indirect program effect on consumption for non-eligible families, amounting to two thirds of the increase in consumption for the poor, in absolute level. Further, we study the mechanisms that lead to this increase in consumption. For example, the implementation of Progresa may modify labor supply, altering equilibrium wages, or it may increase goods prices through higher demand. We find that there are no significant indirect effects on labor earnings, prices (with the exception of increases in few food items in 1998), and welfare receipt, and that sales of agricultural products decrease. Therefore, we rule out the hypothesis that the indirect program effect on consumption is generated by an increase in current income. Instead, we show that non-poor households in treatment villages consume more by receiving more transfers, by borrowing more money - almost exclusively from family, friends, or informal moneylenders - and by reducing their stocks of grains and animals. In addition, we show that the indirect program effects on consumption and loans are larger for households hit by a negative idiosyncratic shock. Thus, we conclude that cash transfers in treatment villages indirectly benefit non-treated households by improving consumption smoothing. These results correspond to our knowledge of developing countries, where credit and insurance occur through informal networks of family, friends, and neighbors. Positive income shocks to some households benefit the whole network, whose other members receive larger loans and transfers, especially the ones hit by negative shocks. The availability of additional liquidity in the network enables households to reduce their savings.

While it is often difficult to predict the effects of a nationwide program using data from limited geographic areas, the effects on the credit and insurance market should not be a function of the number of treated villages, as long as social networks are village-specific. Hence, we can predict positive indirect effects on consumption occurring through changes in local credit and insurance arrangements.[1] Thus, the indirect effects reinforce the direct ones in this class of policies, unlike in many active labor market programs.[2]

[1]As regards the labor market, there should not be any major changes, as long as the decrease in child labor is minimal, compared to the size of the active labor force. The general equilibrium effect on goods prices is instead less clear.

[2]For example, Blundell *et al.* (2004) and De Giorgi (2005) discuss the possibility that subsidized employment for the treated group in the British New Deal program may generate a substitution effects between treated and control

By estimating significant indirect treatment effects, this paper provides a very clear example of the violation of the Stable Unit Treatment Value Assumption (SUTVA). SUTVA states that potential outcomes depend on the treatment received, and not on what treatments other units receive (Rubin, 1980, 1986), ruling out any effect of the program on non-treated households. As such, our exercise highlights the importance of carefully considering the type of data available and the policy at hand before implementing partial equilibrium estimators. For example, when the distance (economic, social or geographic) between treatment and control group is small, and when the treatment group is a large fraction of the local economy, the SUTVA may be less likely to hold.

In sum, the paper contributes to different literatures. We add to the literature that studies consumption smoothing and credit and insurance markets in low-income economies by showing the indirect effects of an exogenous liquidity injection. Some important references in the risk-sharing literature are the work of Deaton (1991), Townsend (1994, 1995a, 1995b), Udry (1994, 1995), Banerjee *et al.* (2003), and Banerjee (2004), among others. We also contribute to the program evaluation literature in multiple ways: first, we show that a class of widely implemented aid policies has important positive externalities, and we establish how they operate. Second, we try and extrapolate the indirect effects of a nationwide conditional cash transfer program in the credit and insurance market. The literature on indirect effects can be roughly grouped into papers that use sources of exogenous variation to identify peer effects, and papers that set up and calibrate structural models to estimate general equilibrium effects. The first group includes Philipson (2000), Katz *et al.* (2001), Sacerdote (2001), Duflo and Saez (2002), and Miguel and Kremer (2004). Bobonis

---
units.

and Finan (2005) and Lalive and Cattaneo (2005) use Progresa data to estimate peer effects in schooling decisions. The second group includes, among others, Heckman *et al.* (1998) and Lise *et al.* (2005). Third, we provide an example of the failure of the SUTVA, which is usually non-testable, and we discuss cases in which this failure is likely to occur.

The paper is organized as follows: section 4.2 describes the structure of Progresa and the characteristics of the data collected for its evaluation. Section 4.3 discusses the identification of the parameters of interest, while Section 4.4 estimates and interprets these parameters. Section 4.5 analyzes the possible channels through which consumption increases, and Section 4.6 investigates the role of idiosyncratic shocks on consumption and credit market. Section 4.7 concludes.

4.2 Progresa: program structure and data characteristics

Progresa is an ongoing Mexican program targeting poor households, providing grants to improve education, health, and nutrition. Started in 1998, this program had about 2.6 million recipient households in more than 2,000 municipalities by the end of 1999, at a cost of approximately 0.2% of Mexico's GDP. Progresa provides grants in the form of nutritional subsidies, as well as scholarships for children attending third to ninth grade. The recipients of the transfers are women. Grants, paid bimonthly, are conditional upon family visits to health centers, women's participation in informal workshops on health and nutrition issues, and verification that children attended classes at least 85% of the time during the previous sixty days. Scholarships are larger for higher school grades and for females attending secondary school, the bimonthly amounts ranging between 160 *pesos* for third grade to 530 and 610 *pesos* for males and female in ninth grade, in November 1999. These payments correspond approximately to one half to two thirds of the wage a child would earn by working full time (Schultz, 2004), and cannot exceed a total of 1500 *pesos* per household (bimonthly), again in November 1999.[3] These grants are quite large, corresponding to 20% of pre-program consumption.

The experimental data for the evaluation of Progresa contain information on households from a sub-sample of 506 poor rural villages from seven different states. The randomization is conducted at the village level, with 186 villages randomized out (about 36% of the sampled localities). Data are collected both before the program starts and during the first 18 months of its implementation. Eligibility depends on poverty status, and households are classified into poor and non-poor according to the information collected in the September 1997 census of localities.[4] There were two rounds of selection of eligible households in Progresa. On average 52 percent of households were initially classified as poor in 1997.[5] The households are informed that, after they are classified as poor and non-poor, their eligibility status will not change until November 1999, irrespective of any income variation. All residents of both control and treatment villages were then interviewed about every six months, first in November 1998 - about a semester after the beginning of the program, and later in May and November 1999. This provides information from three different points in time after the beginning of the program, as well as pre-program data.

Eligible households in control villages were not administered the program until the end of 1999. Thus, our data can be divided in four groups: poor and non-poor households in treatment and control villages. Only poor households in treatment

[3]The size of the scholarships were smaller in 1998 and were later adjusted to keep their real value constant.

[4]For a detailed discussion of the selection criteria both for villages and households see Skoufias et al., 1999b; however poverty status is mainly based on measures of permanent income.

[5]The following year, almost half of the households initially classified as non-poor were added into the beneficiary group. However, most of this latter set of families did not receive the transfers for other exogenous reasons (administrative problems), irrespective of their compliance with the eligibility rules. For this reason, we restrict our sample to the households initially classified as poor when we estimate program effects on the treated.

areas receive the Progresa transfers. Poor households in control villages know that they will be included in the program at the end of 1999, provided that they are still poor and that the program is still in place.[6] The structure of the data is shown in Figure 4.1. In the following section we discuss how we exploit this particular sample design to estimate the effect of the program on non-eligible households living in treatment villages, i.e. our indirect program effect parameter.

4.3 Identification

Experimental data often consist of a sample of eligible individuals randomly assigned to the treatment or to the control group.[7] Instead, the sample for the evaluation of Progresa has some important features that make its design unique: first, it is randomized at the village level. Second, it has data on all households, both poor and non-poor, although only poor households were eligible for the treatment. Thus, we have information on four groups: poor and non-poor households in treatment and control villages. Under the assumptions that the assignment is truly random, and that control villages are not affected by the program, poor in control areas provide a valid counterfactual for treated poor, *and* non-poor in control villages provide a counterfactual for non-poor in treatment villages. Therefore, we can estimate the indirect effect of Progresa on non-eligible households who live in the same locality as treated households using standard identification assumptions. We define this parameter below as the Indirect Treatment Effect (ITE). Estimating significant ITEs implies that the treatment affects potential outcomes of the ineligibles who live in treatment areas. This is a violation of the Stable Unit Treatment Value Assumption (SUTVA). Thus, the sample design enables us to test the SUTVA at the locality

[6]The existence of the program could not be guaranteed beyond 1999 because Progresa may have been discontinued by the new administration, after the 2000 general election.

[7]See, for example, the evaluation of the GAIN project by Hotz *et al.* (2000), or the JTPA randomized trial as in Heckman and Smith (2004).

level, which is not normally possible.

Formally, define Y_{1i} as the potential outcome for non-poor ($NP_i = 1$) in treatment villages ($T_i = 1$) *in the presence of the treatment.* Y_{0i} is the potential outcome for non-poor ($NP_i = 1$) in treatment villages ($T_i = 1$) *in the absence of the treatment.* The observed outcome is: $Y_i = Y_{0i} + T_i(Y_{1i} - Y_{0i})$. The treatment is Progresa transfers to poor households ($NP_i = 0$) in treatment villages ($T_i = 1$). The ITE is the average effect of the program on non-poor households living in treatment villages:[8]

$$\begin{aligned} ITE &= E(Y_{1i} - Y_{0i} | T_i = 1, NP_i = 1) \\ &= E(Y_{1i} | T_i = 1, NP_i = 1) - E(Y_{0i} | T_i = 1, NP_i = 1). \end{aligned} \tag{4.1}$$

Since we do not observe the potential outcome in the absence of the treatment for non-poor households in treated communities, $E(Y_{0i} | T_i = 1, NP_i = 1)$, the identification of indirect treatment effects relies on the assumption that it has the same expected value as the potential outcome in the absence of the program for non-poor households in control villages:

ASSUMPTION (1): $E(Y_{0i} | T_i = 1, NP_i = 1) = E(Y_{0i} | T_i = 0, NP_i = 1)$

Under this assumption, the difference

$$E(Y_i | T_i = 1, NP_i = 1) - E(Y_i | T_i = 0, NP_i = 1) \tag{4.2}$$

identifies the ITE. Note that this is a slight modification of the standard assumption made for the identification of ATT effects.

Despite the randomization, assumption (1) would be violated if outcomes of non-poor households in control villages were indirectly affected by the program. However, if there are indirect program effects for non-poor households in both treatment and control villages, the sign of these effects is likely to be the same for the two groups.

[8] We could rewrite the ITE in its conditional version as $ITE = E(Y_{1i} - Y_{0i} | T_i = 1, NP_i = 1, X)$.

For example, suppose that the increase in school enrollment of treated children reduces child labor. This decrease in labor supply may results in higher employment and earnings for non-poor households in *both* treatment and control villages. Note, however, that the size of these effects is an inverse function of the degree of integration of the village economies.[9] Going back to our example, the fall in relative labor supply may be small enough, compared to the total size of the labor market, to leave employment and earnings virtually unchanged. On the other hand, if the local economies are sufficiently isolated, then the indirect program effects are unlikely to extend to neighboring villages. In the presence of a violation of assumption (1) of the type described above, the difference $E(Y_i|T_i = 1, NP_i = 1) - E(Y_i|T_i = 0, NP_i = 1)$ is a lower bound to the ITE.

We obtain estimates of the ITEs (or of their bounds) from the following equation for non-poor households:

$$Y_i = \alpha + \beta T_i + g(X_i) + u_i \tag{4.3}$$

where the subscript i refers to the $i - th$ household, Y_i is some outcome of interest (e.g. consumption). Under assumption (1), β identifies the ITE. We add a set of conditioning variables, X_i, to increase the precision of the estimates. These are: poverty index, shock dummy, pre-program income and land, household size, dummies for head of household gender, age composition of the household, employment status, language (Spanish, indigenous language or both), and literacy, at the household level; village marginalization index, average number of shocks in the previous 6 months, number of households, number of treated households, geographic region dummies at the locality level. If we estimate (4.3) for poor households, β identifies the ATT

[9] Technically, what is needed is only integration of the product market in the presence of relatively large number of products traded with respect to factors (Mundell, 1966).

under the assumption that $E(Y_{0i}|T_i = 1) = E(Y_{0i}|T_i = 0)$.[10]

4.4 Indirect Treatment Effect on consumption

We now proceed to estimate *ITEs* for food and non-food consumption. We computed measures of monthly food and non-food expenditure per adult equivalent, using an equivalence scale estimated from these data and constant prices. The Appendix provides further details on the creation of these variables.
The means of food and non-food consumption in Table 4.1 show, as expected, that the non-poor consume substantially more than the poor, despite the Progresa transfers. However, the average non-poor household is clearly not very well off, in absolute levels: average food and non-food consumption for non-poor in control areas are about 20 and 7.5 U.S. dollars per adult equivalent per month. Lastly, consumption is higher in treated areas, especially for poor households, but also among non-poor ones (in May and November 1999 for food consumption, and in May 1999 only for non-food consumption).

We obtain estimates of the Indirect Treatment Effect (and *ATT*) for consumption expenditures estimating (4.3) by OLS for each time period, using the logs of food and non-food consumption (expenditures) as dependent variables (Table 4.2).

The ITE is never significant in November 1998, a few months after the program began, while food consumption is significantly higher in 1999, by 5.1% in May and by 6.7% in November. This corresponds to about 10 and 13 *pesos* per adult equivalent, respectively. Non-food consumption is significantly different in May 1999 only, and it is higher by 13.8% in treatment areas, i.e. about 10 *pesos* per adult equivalent, while

[10] Actually, in this case β identifies the average intent to treat effect, i.e. the effect of the treatment on eligible households irrespective of participation. However, in this case the difference between intent to treat and treatment on the treated is negligible, because about 97% of eligible households participate to the program. Therefore, we will continue referring to β as ATT, when estimating (4.3) for poor households.

the effect is negative but not significant in November 1999.[11] Thus, consumption increases overall by approximately 100 and 70 *pesos* per household per month in May and November 1999.[12] Poor food and non-food consumption, instead, increases in all three periods, consistent with the existing evidence (Hoddinott *et al.*, 2000); this impact is proportionally larger over time, although constant in absolute value in 1999, with an average monthly consumption increase of 30 *pesos* per adult equivalent.

The results above are robust to a variety of checks, as detailed in the Appendix. These include alternative ways of dealing with households who reported zero food and non-food consumption (2.9% and 1.5%, respectively), alternative treatment of households with extremely large reported consumption levels, and alternative measures of food consumption. We also fail to detect any pre-program significant difference in food and non-food consumption between households in treatment and control villages.

4.5 Why does Progresa increase non-poor consumption?

We showed that Progresa has a positive spillover effect on non-poor consumption. This externality has multiple potential causes. First, the increased income of treated households, together with the surge in treated children's school attendance, and the decrease in child labor, may cause changes in the goods and labor markets, which could result in higher income for non-poor households. Second, non-poor households may receive additional aid, as we will explain below. Third, the increased liquidity in treatment villages may affect the local credit and insurance markets. Households may borrow more for investment or consumption purposes.[13] Lastly, households may reduce their savings. A decrease in savings would be compatible with better insurance

[11] Total consumption for non-poor in treatment areas is 5.3% and 4.9% higher in May and November 1999. Both differences are statistically significant. Results available upon request.

[12] The average adult equivalent household size is 5.0 and 5.3 in the two 1999 waves.

[13] The former effect would result in higher current consumption through higher current or permanent income.

against risk, which lowers the need for self-insurance, or with a drop in interest rates caused by the increased liquidity. Note that the higher consumption may be financed through a drop in investment, although it is not clear why households would change their preferences in this way. We can summarize the above discussion using the following accounting identity:

$$\Delta Y_i + \Delta L_i = \Delta C_i + \Delta S_i + \Delta I_i \tag{4.4}$$

where Y is income, L are net loans and transfers, C is consumption, S are savings and I investment for household i, and Δ is the indirect program effect for each outcome of interest. In the next sections we will consider these different channels individually, discussing and testing our hypotheses in greater detail.

4.5.1 Labor market effects

The program may affect non-poor consumption levels through higher labor earnings, for example by increasing equilibrium wages or non-poor labor supply. These effects may occur if the increase in poor children school attendance and the related scholarships decrease the overall labor supply by reducing both child and adult labor. Their magnitude depends on the drop in labor supply and on the degree of integration of the local economy: the higher the integration, the smaller the program impact. For example, the potential drop in child labor caused by the increase in treated children's school attendance may increase labor earnings through higher local wages, if treatment villages are economically isolated, while such an effect will be small if the village economies are sufficiently integrated.

We investigate this hypothesis by testing whether households in treatment and control villages earn a different labor income. We compute the household monthly labor earnings variable summing income from both primary and secondary occupations, using the reported wages (which may be daily, weekly, monthly or annual) and hours worked. Table 4.3 reports tobit cross-sectional estimates of differences in per capita monthly labor earnings, where the first column shows pre-program differences. We do not detect any statistically significant difference in labor earnings between non-poor households living in treatment and control localities. The values of the point estimates is also low, compared with the consumption change, with the exception of a 64 *peso* earnings increase in November 1999 estimates, though statistically insignificant. However, this difference becomes smaller once we consider also earnings from informal activities: we considered the effect on earnings from informal work activities such as the provision of transportation, cooking, sewing, repairs, carpentry, and various other paid services, and we found a small and negative program effect, although never significant.[14] Lastly, we tested for differences in hours of work, which never change for the non-poor. Thus, we find no compelling evidence that the increase in consumption is caused by labor-related indirect program effects.

The program effect on labor income has no clear trend for poor households. The difference in earnings is significantly higher in November 1998 and November 1999, but lower in May 1999.

4.5.2 Goods market: effects on prices and sales

It is possible that the higher expenditure induced by the Progresa transfers may increase goods prices. Again, this depends on the degree of integration of the goods market. If prices in treatment villages increase, and non-poor earnings from sales rise, the non-poor in treatment villages may use this extra income to consume more.

[14] We also estimated these regressions by OLS, since tobit estimates are inconsistent in the presence of heteroskedasticity. If we interpret the reported zero labor earnings as very low earnings, or as measurement error due e.g. to illness, then OLS estimates are consistent. The OLS estimates are never significant.

In order to test this hypothesis, we first compare prices in treatment and control localities to see whether they are significantly different.[15] To do so, we consider village prices by good over time. We provide details on the creation of the price variables in the Appendix, as well as estimates of the price differences between treatment and control villages. We noted that some prices differ before the program begins, in March 1998, hence we provide difference-in-difference (DD) estimates of the effect of the Program on village prices. This exercise is possible only for food prices, since there is no pre-program information on non-food prices. We find a small positive effect on 5 out of 36 food prices in November 1998, which we do not expect to increase the cost of the food basket substantially, because prices of staples such as rice, beans, corn, and chicken do not change. We find no food price change in the later waves, nor evidence of price changes from cross sectional variation, when we consider both food and non-food prices.

Even if prices do not differ between control and treatment villages, the non-poor in treated localities may earn a higher income than the non-poor in control villages by selling more goods to poor households. We test whether there is a significant difference in sales of agricultural products and of animals for poor and non-poor households in control and treatment villages. We compute net agricultural sales by subtracting production costs from the gross sales variable, and the net animal sales by subtracting the value of purchases from sales. No data are available in November 1999. Table 4.4 shows estimates of the ITEs and ATTs for these variables, at November 1998 prices. The agricultural ITEs are mainly negative and significant: households reduce production costs by 9 and 18 *pesos*, and sales by even more, namely by 20 and 40 *pesos*, resulting in lower net sales. While alternative explanations are

[15]Note that the higher consumption of treated households may increase local goods prices, resulting in higher nominal prices. However, this would not explain the observed increased consumption, since we use real prices in our measures of consumption.

possible, we suspect that part of this reduction in costs and sales is due to a reduction in the grains buffer stock, which are used to sow the land or as consumption of home produced goods or both. We will provide further evidence consistent with this hypothesis when estimating the program effects on savings. Irrespective of the determinants of this behavior, it is clear that the extra consumption is not financed through higher sales. There is also a small decrease in purchases of animals in 1998, together with a small and positive, yet imprecisely estimated, effect on net sales. Lastly, note the decrease in net sales for the poor, caused by a contemporaneous increase in costs and decrease in sales. We interpret this result as evidence that treated households increase both their agricultural production (and probably also their stock of grains) and their own-consumption, producing more but selling less because a larger share of the harvest is consumed (or saved).

4.5.3 Additional aid program recipience

One additional possible cause for the higher non-poor consumption in treatment areas may be the higher relative supply of transfers from alternative welfare programs. This may occur for two reasons. The first reason is that one of the aims of Progresa is to replace some of the numerous pre-existing welfare program into a single one (DIF, INI, Ninos de Solidaridad, Tortilla, and Liconsa). These programs range from the provision of food (Liconsa, Tortilla) to the assistance of specific sub-samples of the populations (children, in the case of Ninos de Solidaridad, and indigenous households for INI, the National Institute for Indigens). This may leave some agencies located in treatment villages with excess levels of aid (in cash or kind), which they may direct to households classified as non-poor. The second reason is a reduced need of treated households' additional welfare assistance because of Progresa. For example, if Progresa transfers help protect recipients against risk, they may be less in need of emergency assistance. This may be directed to some non-poor in treatment areas.

We test whether non-poor and poor households in treatment and control villages have significant different intakes of alternative welfare programs. We first create some aggregate measures of program receipt. These are the likelihood of participating into at least one alternative program; the number of programs the households participates into, conditional on receiving at least one alternative type of welfare; the total monetary transfers the households received in the previous month. We then test for differences in the likelihood of participating to individual programs. We present aggregate estimates of these differences in Table 4.5, while we report the difference in participation rates program by program in the Appendix. As expected, participation into alternative welfare programs is significantly lower for poor households in treatment communities in all post-program waves (nevertheless, they are still better off with Progresa). Thus, because of a reduction in recipients in treatment villages, it is possible that non-poor households may appropriate some of the resources previously targeted to poor households. Indeed, in May 1999 there is a 1.8 percentage point increase in the likelihood of receiving cash through Solidaridad, and a 0.7 percentage point increase in receipt of free milk for the non-poor. However, these differences are very small. Moreover, there is no significant difference in the overall amount of monetary transfers received from cash programs, nor in participation rates into any other in-kind program in the two village groups.

In the Appendix we provide evidence against the hypothesis that welfare receipt from alternative programs may have increased for non-poor in both treatment and control villages. In sum, we do not find evidence that the higher non-poor consumption in Progresa villages is primarily due to an increase in the receipt of welfare programs.

4.5.4 Credit market

The final transmission mechanism operates through the credit and insurance market. Financial market imperfections in developing countries are well documented in the literature: since formal insurance or credit institutions are almost absent, informal lending and risk-sharing mechanisms often arise, in the form of transfers and loans through social networks (see, for example, Townsend, 1995a; Fafchamps and Lund, 2003; and Munshi and Rosenzweig, 2005). In this sense, credit and insurance markets are in practice merged (see Udry, 1994, among others). As such, in the rest of the paper we consider them jointly, and often refer to them as the "credit market".[16] One advantage of these informal arrangements is their low likelihood of default, or of not reciprocating with transfers, because of the personal relationships between the agents involved, the high amount of available information, and the repeated nature of the interactions, which make exclusion from future transactions a costly punishment for defaulting.

When social networks and informal lending channels are important means to smooth consumption and insure against risk, even small liquidity injections into the network may have substantial spillover effects through increased loans and transfers. Note that the lender/donor here may be both a treated household, which can lend more money and transfer resources to non-treated neighbors because its income has increased, and a non-treated household, which may shift resources from poor households (less needy of help because they receive Progresa grants) to non-poor

[16]Our definition of credit market includes loans and transfers. In developing rural economies, transfers from family and friends may be considered as credit if the receiver reciprocates when the donor is in need.

households within the same network.[17,18] We believe the program may have substantial indirect effects in the credit market because most of the loans are informal: 70% of lenders are friends or relatives of the borrower (and a further 9% are local moneylenders), and because the scale of these informal networks is limited by their being based on personal relationships. Thus, the positive effects of a liquidity injection are shared by a limited number of households.[19]

We have information on receipt of loans in the previous six months, of transfers (monetary and in kind) from family and friends during the previous month, but no information on the identity of lenders and donors, e.g. whether they are from poor or non-poor households. In order to understand the overall program indirect effect on the credit market, we need to observe these variables in all data waves. Unfortunately, this is possible only in November 1998, when very little money had been transferred to treated households.[20] In the remaining waves, we observe loans in May 1999, and transfers in November 1999. We report means and standard deviations in Table 4.6. Only about 12 percent of the non-poor and 8 percent of the poor receive any resources in November 1998. The average monthly receipt amounts to some 400 *pesos* for the non-poor, and to 260 *pesos* for the poor households. Interestingly, this pattern is common for all variables and semesters: a higher proportion of the non-poor receives resources, compared to the poor, and their average receipt is larger, both in treatment and in control villages. A possible explanation for loans may be that the non-poor

[17]Program recipients are unlikely to leave the network and not share their income increase, if they are sufficiently forward-looking, because the Progresa grants may stop in November 1999, and will certainly almost entirely end after all the children complete 9th grade.

[18]The direct effect on loans and transfers for poor households is ambiguous. Public transfers may crowd out private ones, as documented by Albarran and Attanasio (2002), and treated households may have less need to borrow. On the other hand, the higher liquidity in the village and the possibility of using the Progresa transfers as collateral may cause an increase in loans to these households.

[19]Another possible explanation is that the liquidity injection lowers the informal interest rate, resulting in an increase of current consumption levels.

[20]We also observe migrant remittances in November 1998 and November 1999. We add them to loans and transfers to compute our measure of total credit resources. We separately estimated indirect program effects on remittances, and found no significant effect. Results available upon request.

have more assets, which can be used as collateral to borrow against, while the larger transfer size may suggest that the non-poor belong to a wealthier network, where people transfer more resources. In any case, the larger transfer size suggests once more that the so-called non-poor households are indeed not very wealthy, otherwise they would not be in need of transfers from family and friends. Lastly, note that, for the non-poor, both the proportion of recipients and the size of the receipt are larger in treatment than in control areas, while there is no differences for poor households.

We want to test whether Progresa has an indirect impact on both the probability of receiving loans and transfers, and the amounts received, measured in *pesos* per month. For this purpose, we estimate versions of equation (4.3) by probit and tobit, using loans and transfers measures as dependent variables. Table 4.7 provides estimates of $ITEs$ and $ATTs$. There are no significant program effects in 1998: the ITE point estimates are positive but very imprecisely estimated. This lack of significant effects is not surprising, as in November 1998 Progresa had only just started, and very little money had been transferred to the program recipients. Indeed, the effects are positive and significant in 1999, when Progresa had been operating for at least one year. In May 1999, non-poor families in treated villages have a 1.5 percentage point higher chance of having a loan: this is 40% more than in control villages. They also borrow on average 3.8 more *pesos* per month, i.e. one third more than non-poor in control communities. This evidence is consistent with our conjecture that the program liquidity injection may relax a constraint on the lender side, enabling them to lend more. In November 1999, transfers to the non-poor are significantly higher by 6.6 *pesos*, or one third, in treatment villages. These households have also a higher likelihood of receiving transfers, although the coefficient is imprecisely estimated. This evidence suggests that the liquidity injection caused by Progresa may benefit the non-poor through some resource redistribution. However, the proportion of non-poor households receiving in-kind transfers decreases significantly in November 1999. This suggests that there may be a shift in the composition of transfers (e.g. from food and clothing to money), again consistent with the higher liquidity brought by Progresa.[21] To conclude, note that there is a significant decrease in both loans and in-kind transfers to treated households, and no effect on monetary transfers. Our results are robust to different treatments of the outliers and to the use of different estimators.[22]

In sum, we found that both loans and family and friend transfers to the non-poor are significantly higher by one third in treatment villages in 1999. Assuming that the magnitudes of the effects are constant throughout the two 1999 semesters (since we observe loans in May and transfers in November), non-poor households receive on average an extra 10 *pesos* per month. This explains only part of the observed higher consumption, which amounts to about 100 and 70 *pesos* per household (in May and November 1999, respectively).[23]

4.5.5 Savings and investments

So far we have seen that the indirect program effect on consumption is not caused by increases in earnings and welfare receipt, and only partially caused by higher loans and transfers. The difference must be financed through a reduction in savings or investments, as shown in the accounting identity in (4.4).

[21]Unfortunately we do not have any information on the value of in-kind transfers, so we cannot compute the total value of transfers.

[22]We ran the above regressions on the entire sample, trimming the top percentile of the positive values, and dropping the four largest amounts, respectively. We ran all regressions by OLS. In all cases, the point estimates varied very little. We also found that the non-response rates, which vary between 0 and 5.4% for non-poor households, do not differ between treatment and control areas. This may have been an important issue, owing to the relatively small number of households reporting loans or transfers.

[23]Part of the consumption increase may be financed through a reduction in donations from non-poor households in treatment villages. This would be possible if, for example, the Progresa transfers to the poor were crowding out private transfers from the non-poor. However, the difference in resources donated from the non-poor in treatment and control villages is never statistically significant. Note that only a total of 56 non-poor households donated resources to family and friends in the November 1998 and November 1999 data.

Rural households' savings are primarily in the form of grains and livestock: fewer than one percent of non-poor households hold interest-bearing savings (although we do not explicitly observe their monetary value). We test for differences in the stock of grains and animals owned by the non-poor in treatment and control villages, comparing the tons of grains and number of animals owned. The change in stock may differ from the *peso* value of their net sales, which we showed in Table 4.4, if households start consuming part of their stock. This is exactly what we find: for grains, the *peso* value of the stock is lower for non-poor households in treatment villages, though imprecisely estimated. However, in May 1999 we find a significant effect on the likelihood of reducing the grain stock, which drops by 4.5 percentage points, and a 9 *peso* significant increase in consumption of own grains. We also find a decrease in livestock for the non-poor in treatment villages. In particular, the stock of chickens decreases significantly by 0.6 in 1998 (among households who own chickens), and the likelihood of owning pigs drops significantly by 3, 6, and 4 percentage points in November 1998, May 1999, and November 1999, respectively. This is not surprising, as chickens and pigs are the most widely held animals (by 61 and 34% of households in November 1998, respectively). At the same time the number of animals, and both the quantity and the value of the stock of grains owned by the poor in treatment villages increase substantially.[24]

These findings are consistent with both the empirical literature (e.g. Udry (1995) and Lim and Townsend (1998) show that households reduce their stock of grain in response to a shock) and with the predictions of models of incomplete risk sharing in which agents rely partially on self-insurance. These models predict that households hit by positive income shocks (the treated) increase their buffer stock, while house-

[24]Results available upon request.

holds whose income has not changed (the non-poor) decrease savings because poor households' higher income improves insurance against risk. The absence of significant program effects on the *peso* value of grain stock and of net animal purchases suggests that part of the stock is now consumed (or used as productive input in the case of grains). At the same time, the positive program effect on loans and transfers suggests that households do share part of the risk.

Lastly, it is often difficult to separate investment from savings and production costs (as in the case of purchase of fertilizers). While there is no conceptual reason why Progresa would indirectly decrease investment for the non-poor, one could interpret the drop in production costs in such a way.

4.6 Consumption smoothing and insurance against risk: idiosyncratic shocks

We have identified a positive indirect program effect on consumption for the non-poor; we have further shown that these households finance their extra consumption through increased loans and transfers, and through a reduction in both savings and sales of agricultural products. We suspect that one important reason for this consumption increase and reduction in savings may be the improved insurance against risk caused by Progresa's liquidity injection into treatment communities.[25] One way to test this hypothesis is to estimate separate indirect program effects for households who have and have not been hit by a negative idiosyncratic shock. Consider non-

[25]There are additional explanations for the observed increase in non-poor consumption. However, in general these additional explanations are consistent with some, but not all the observed ITEs. One determinant of higher non-poor food consumption may be the effect of better information on the importance of an adequate nutrition: all households in treatment villages were strongly encouraged to attend classes that covered health and nutrition topics, and such classes may have raised awareness on the importance of a nutritious diet, causing higher food expenditure. However, we have also seen an increase in non-food consumption. A further explanation may be the existence of "imitation" effects. Recent work by Bobonis and Finan (2005) and Lalive and Cattaneo (2005) analyze peer effects on schooling decisions in Progresa communities. Both papers show that the school enrollment of ineligible children is a positive function of the enrollment of treated children. Analogously, imitation effects may also increase consumption. These additional causes for the increased consumption explain the reduction in savings, but not the observed increase in loans and transfers.

poor households only. Some fraction of the non-poor is hit by idiosyncratic negative shocks, $S = 1$, in both villages. If Progresa improves insurance against risk, then we will observe all non-poor households in treatment villages increase their consumption through a mix of lower savings (because of the new access to loans and transfers if hit by a negative shock), and higher transfers and loans, compared to the non-poor in control villages. In addition, if shocks are random within village, we will see a larger increase in consumption and transfers for $S = 1$ households than for $S = 0$ households in treatment areas: if Progresa improves risk sharing in treatment villages, the indirect program effect on $S = 1$ households consumption and loans and transfers will be proportionally higher than for households not hit by a shock. This occurs because households hit by a shock will need to smooth consumption more than households not hit by a shock, and they will partly do so by borrowing more (or by receiving more transfers). A clear prediction is that $S = 0$ households should run down on their savings since the need for self-insurance is reduced by the availability of credit in the network, while the double-difference effects on savings is ambiguous. For example, $S = 1$ households may reduce savings equally in treatment and control villages, and finance the extra change in consumption through increased loans and transfers.[26] In other words, a test of whether the ITEs on consumption, loans and transfers, and savings differ by shock between and within village type is an indirect test of the effect of the program on risk-sharing mechanisms.

[26]Suppose that in the absence of Progresa all non-poor households not hit by a shock would consume 100, and households hit by a shock only 80. With Progresa, non-poor households not hit by a shock increase consumption up to 110 through a mix of reduction in savings and higher transfers (each by 5). Households hit by an adverse shock, instead, further reduce their buffer stock, and receive even more assistance from their social network (each by 10), pushing their consumption from 80 to 100. In this case, the effect of Progresa on the change in transfers, savings, and consumption is higher for households hit by a negative shock.

Formally, we define two new parameters of interest:[27]

$$ITE^{S0} = E(Y_{1i}^{S0} - Y_{0i}^{S0} | T_i = 1, S_i = 0, NP_i = 1) \tag{4.5}$$

$$ITE^{S1} = E(Y_{1i}^{S1} - Y_{0i}^{S1} | T_i = 1, S_i = 1, NP_i = 1) \tag{4.6}$$

where S_i is an idiosyncratic shock and the superscripts $S1$ and $S0$ stand for hit or not by the shock respectively. These parameters are the $ITEs$ for households who have and have not been hit by a shock. We want to test the following hypotheses:

$$ITE_C^{S1} > ITE_C^{S0} > 0$$

$$ITE_L^{S1} > ITE_L^{S0} > 0$$

$$ITE_S^{S0} < 0$$

where the subscripts C, L and S refer to consumption, loans and transfers, and savings, respectively. We can identify the above parameters exploiting the randomization by assuming the following:

ASSUMPTION (2):

2.1 $\quad E(Y_{0i}^{S0} | T_i = 1, S_i = 0, NP_i = 1) = E(Y_{0i}^{S0} | T_i = 0, S_i = 0, NP_i = 1),$

2.2 $\quad E(Y_{0i}^{S1} | T_i = 1, S_i = 1, NP_i = 1) = E(Y_{0i}^{S1} | T_i = 0, S_i = 1, NP_i = 1).$

Assumption 2 implies that shocks are random between treatment and control villages. This assumption seems realistic because the village randomization is made within homogeneous geographic regions. For example, if one region is arid, its treatment and control villages are equally likely to suffer from droughts. However, in order to

[27]The structure of the identification problem could be also reformulated in the multiple treatment framework by defining 4 different treatments depending on the combination of T and S.

compare ITE^{S1} and ITE^{S0}, we also need to assume that shocks are random within villages, conditional on household observable characteristics.[28]

We estimate the following equation both for non-poor and poor households:

$$Y_i = \alpha + \beta_1 T_i + \beta_2 S_i + \beta_3 T_i S_i + g(X_i) + u_i \tag{4.7}$$

The parameters of interest in this second step are:

$$\begin{aligned}
\beta_1 &= E(Y_i|T_i = 1, S_i = 0) - E(Y_i|T_i = 0, S_i = 0) \\
\beta_1 + \beta_3 &= [E(Y_i|T_i = 1, S_i = 1) - E(Y_i|T_i = 0, S_i = 1)] \\
\beta_3 &= [E(Y_i|T_i = 1, S_i = 1) - E(Y_i|T_i = 0, S_i = 1)] \\
&\quad -[E(Y_i|T_i = 1, S_i = 0) - E(Y_i|T_i = 0, S_i = 0)].
\end{aligned}$$

Under the above assumptions, β_1 identifies ITE^{S0}, i.e. the indirect program effect for households not hit by an idiosyncratic shock; $\beta_1 + \beta_3$ identifies ITE^{S1}, i.e. the effect for households hit by a shock; β_3 the difference between the two $ITEs$. The set of conditioning variables includes all the variables discussed earlier as well as the average shock intensity at the village level.[29] We estimate (4.7) for both consumption and credit market outcomes.

In our data, we observe whether, in the six months preceding the interview, the household has been hit by any of the following types of natural disasters: drought, flood, frost, fire, plague, earthquake, and hurricane. This is our definition of shock. It is very specific, and it excludes other events that may cause income or wealth losses,

[28]To interpret the parameters in the way discussed above, we further need to assume that the extra liquidity injected by Progresa is sufficient to borrow both to offset the shock and for additional reasons, and that households are willing to lend.

[29]Controlling for the intensity of the shock at the village level allows to net out the aggregate component of the shock, which is uninsurable at the village level, while the village is the reference unit for our test of improved insurance. Once we fix the intensity of the shock in the village, the household-specific shock dummy is then a measure of idiosyncratic shocks, where the parameter of interest is estimated exploiting the within village variation. Note that the parameter is identified for villages where only a sub-set of households are hit by a shock.

such as illness or death of household members. We discuss this variable in greater detail in the Appendix; we provide evidence that there is a substantial within-village variation (Figure 4.2), consistent with Townsend (1994), who shows that natural phenomena - rainfall in his case - are not uniform even within very small villages; we also show that shocks hit households randomly. The average incidence of the shock at the village level is 39, 55 and 30% in November 1998, and May and November 1999. This is a rough measure of how aggregate these shocks are. Note that the timing of the events and the way they are recorded in the data is quite important: the shocks must precede (or be contemporaneous to) the observed outcome of interest. This requirement is satisfied in our data: both shock and loans refer to the 6 months before the interview, while consumption, transfers, and savings data are provided for the previous week, month, and semester, respectively.

4.6.1 Consumption by shock

Table 4.8 provides means of food consumption by poverty status, village of residence, and shock. Consider food consumption, for example. While the means are not different in November 1998, consumption is higher in treatment villages irrespective of shock status in the two 1999 semesters. The comparison of consumption levels for non-poor households with and without the shock in the two villages types reveals the following: first, households hit by a shock in control villages tend to have the lowest average consumption levels. Second, consumption for households hit by a shock is higher in treatment than in control villages, but almost always lower than consumption for households in treatment villages who have not been hit by a shock. However, the standard deviations are very large. We repeat this comparison by regressing non-poor log-consumption on these four categories, as in equation 4.7, adding the usual set of conditioning variables to take account of possible differences in observable characteristics by group, and to improve the precision of the estimates.

We now test the hypotheses that indirect program effects are positive irrespective of shock status, and that the indirect program effect on consumption is higher for households who suffered a shock. The first hypothesis requires only the standard randomization assumption, the second one also the other assumptions discussed in the previous section. The results are shown in Table 4.9. The estimates of the indirect program effects are positive and significant irrespective of shock status in both 1999 semesters. Food consumption in May and November 1999 increases significantly by 9.1% and 5.4% among households not hit by an idiosyncratic shock, and by 6.6% and 14.4% for households who were hit by a shock. Furthermore, while in May 1999 there is no statistical difference between their magnitude, in November 1999 the effect is 9 percentage points higher for households hit by a shock. The value of these changes in monthly *pesos* is 18.3 and 13.3 per adult equivalent (90 and 68 per household) for households without ($S = 0$) and with a shock ($S = 1$) in May 1999. The respective values for November 1999 are 10.7 and 27.7 per adult equivalent (or 55 and 152 per household). The program effects for poor households are positive irrespective of shocks, they increase over time, as we saw in the previous consumption table, and never differ by shock status. The absolute value of these consumption changes is 20.3 (for $S = 0$ households) and 24.3 monthly *pesos* (for $S = 1$ households) per adult equivalent in May 1999, and 27.2 and 25.7 *pesos* in November 1999. There is a similar pattern for non-food consumption, although the positive trend is less marked. However, once we split non-poor households by shock, we do not find any significant indirect program effect for non-food consumption.[30]

[30]Results available upon request.

4.6.2 Credit market by shock

We now proceed to estimate separate *ITEs* for households who have and have not been hit by a shock. Further, we test whether Progresa enables non-poor households hit by an idiosyncratic shock to borrow more (or receive more transfers) than households not hit by a shock. We perform such tests by estimating equation 4.7 for both the likelihood of receiving loans and transfers, and the amount received.

Table 4.10 reports means and proportion of credit resources by household and village type, and by shock occurrence. As we saw before, in general the indirect program effects on credit resources are larger over time. In addition, both the magnitude of these effects and the likelihood of receiving loans and transfers are larger for non-poor households hit by a shock, consistent with our expectations. A notable exception are monetary transfers in November 1999, when the effect for households hit by a shock is negative, and certainly lower than the effect for households not hit by a shock.

Table 4.11 provides estimates of the *ITEs*. In November 1998, households hit by an idiosyncratic shock are 2.6 percentage points more likely to receive loans or transfers if they live in treatment villages. Moreover, they borrow and receive about 22 pesos more per month. Instead, there is no significant effect for families not hit by a shock. It is striking that we detect these effects when Progresa had been implemented for a few months only. We interpret this fact as strong evidence that the program has an indirect effect on the credit market. We find the same type of effect for loans only, both in November 1998 and in May 1999. Once again, non-poor families who have been hit by a shock are about 2.5 percentage points more likely of borrowing money, and borrow 16.5 and 7.5 pesos more than their counterfactual in 1998 and 1999. These results are consistent with our hypothesis that Progresa may enable households to insure against risk by borrowing more: the estimated $ITEs$ are larger for households hit by a shock. Interestingly, the effects for monetary transfers vary over time. In November 1998 the average transfer size is positive for both groups, and larger for households hit by a shock, but these effects are imprecisely estimated. In November 1999, instead, we find that the effect of the program on transfers is positive and significant, and averages about 12 pesos per month, only for households not hit by a shock. One possible explanation for this counterintuitive finding may be that higher loans are crowding transfers out. Alternatively, this finding is consistent with the hypothesis that transfers occur between "closer" individuals than the parties involved in loans. The former group's geographic, as well as social higher proximity, results into shocks being positively correlated. Unfortunately, we do not have data on loans for the same semester, nor complete data on social networks, hence we cannot draw any conclusion regarding either effects. However, note that the finding that risk is shared through informal loans, rather than through transfers, is consistent with the evidence provided by Fafchamps and Lund (2003) for the Philippines. Lastly, note that there is no clear pattern for the likelihood of receiving in-kind transfers: the difference in $ITEs$ is negative in November 1998, and positive in November 1999. The magnitude of this difference is between 1.2 and 1.6 percentage points.[31], [32]

In sum, the results presented in this section are consistent with our hypothesis of better risk-sharing through a more liquid credit market. Non-poor households in Progresa villages indirectly benefit from the program by receiving more credit resources when hit by a shock.

[31]One potential limitation of this exercise is the small sample size: in each semester, there are no more than 500 non-poor families receiving transfers or borrowing money, which we compare after dividing them into four groups based on village type and shock occurrence. This number is even smaller when we consider loans and transfers separately.

[32]In the Appendix we report the effect of the program on the poor for comparative purposes.

4.6.3 Savings

We finally test for indirect program effects on savings by shock. Our prediction is that the ITE for households not hit by a shock is negative, while the effect on savings for $S = 1$ households is not clear. We test for changes in the stock of corn and beans, staples of the Mexican diet and easily storable commodities. The prediction is confirmed by the data: while there is no change in the stock of beans, $S = 0$ households reduce their stock of corn through a significant increase in the value of home consumption of about 10 *pesos* per month in May 1999. In the same semester, their likelihood of depleting agricultural stock is significantly higher by 8 percentage points in treated communities. Instead, there is no clear pattern for households hit by a shock: we observe a shift in production from sugar cane to corn, and a significant increase in home consumption of the same magnitude as $S = 0$ households, but their total grain stock appears to be unchanged. As regards livestock, we found evidence of a significant depletion for both types of households, especially of chickens and pigs. The change in stock does not differ by shock.[33]

Overall, the evidence is consistent with the conjecture that non-poor households in treatment areas may indirectly benefit from Progresa by being able to consume more because of changes in the credit and insurance markets.

4.7 Conclusions

Using the unique design of the experimental trial and the available data for the evaluation of Progresa, we show that non-eligible households who live in treatment villages benefit indirectly from the program by increasing their consumption level. We further show that the consumption increase occurs through changes in the credit

[33] However, the effect on the value of net sales is negative and significant only for $S = 1$ households, a reduction of 32 *pesos* in May 1999, and does not change for $S = 0$ households. It is possible that this latter group consumes part of its livestock, rather than selling it.

and insurance markets, which enable households to borrow more and to receive more transfers, permitting them to reduce their savings. We conclude that this class of aid programs improves consumption smoothing for non-treated households living in treatment areas, consistent with the findings that, though the program increases consumption and loans, and decreases savings for all non-poor households in treatment villages, the effects on consumption and loans are larger for families hit by a negative idiosyncratic shock. These results are consistent with our knowledge of the credit and insurance markets in developing countries, which operate through social networks. A positive income shock to any member is likely to benefit the whole network.

Our findings are interesting from several perspectives. They show how households in developing countries deal with credit market imperfections, and how a liquidity increase may have beneficial indirect effects on the local community at large, including households whose current income does not change. Thus, this class of aid policies has important spillover effects which should not be neglected when evaluating the impact of the program.

In addition, this exercise is a striking example of circumstances in which the SUTVA, a key identifying assumption normally used in the program evaluation literature, fails. This assumption is often non-testable, when the experimental design consists of one treatment and one control group. However, its likelihood of being violated may be a function of observable characteristics of the programs and of the local economies where it operates. We suspect that the characteristics of Progresa which cause a significant indirect program effect are that this program targets a large proportion of the population of a local economy; has generously-sized transfers; relaxes some existing binding constraints (lending constraints in this case); operates in areas where the treated and non-treated subjects are sufficiently "close" (from an economic, geographic, or social perspective); and has been ongoing for a sufficient amount of time. Analyzing similar features of different programs may provide guidelines on the robustness of the SUTVA.

4.8 Tables and Figures

Table 4.1: Average monthly food and non-food consumption levels per adult equivalent

	Food consumption			Non-food consumption		
	1998 Nov.	1999 May	1999 Nov.	1998 Nov.	1999 May	1999 Nov.
NP control	203.25	201.50	195.99	76.55	72.47	74.22
	(168.13)	(197.00)	(216.60)	(87.92)	(80.29)	(80.86)
NP treatment	198.41	220.87	212.27	74.54	77.51	71.08
	(160.43)	(281.17)	(263.19)	(85.53)	(90.13)	(75.76)
P control	137.47	144.59	139.39	34.58	36.91	37.24
	(103.10)	(141.25)	(116.05)	(43.39)	(41.99)	(40.80)
P treatment	151.16	168.52	168.59	36.09	42.59	41.97
	(122.17)	(178.94)	(191.25)	(41.88)	(46.02)	(41.86)

Note: the amounts are in pesos; the exchange rate was roughly 10 pesos per USD.

Table 4.2: Average effect of Progresa on log-food and non-food consumption

	Food consumption			Non-food consumption		
	1998 Nov.	1999 May	1999 Nov.	1998 Nov.	1999 May	1999 Nov.
$\hat{ITE}$	-0.0213	0.0514	0.0669	0.0629	0.1384	-0.0602
	[0.0256]	[0.0257]**	[0.0211]***	[0.0672]	[0.0655]**	[0.0623]
Obs.	4602	3824	4257	4771	4259	4443
$\hat{ATT}$	0.1033	0.1699	0.1892	0.0959	0.1675	0.1128
	[0.0236]***	[0.0219]***	[0.0211]***	[0.0651]	[0.0613]***	[0.0540]**
Obs.	10879	9605	10508	11484	10630	10856

Note: Standard errors in [brackets] clustered at the village level. ***, **, * indicates significance at the 1, 5, 10 % level respectively.

Table 4.3: Tobit estimates of program effect on per capita monthly household labor earnings

	1997 Sept.	1998 Nov.	1999 May	1999 Nov.
$\hat{ITE}$	-9.58	14.21	-7.72	64.82
	[46.99]	[39.57]	[44.63]	[49.49]
Obs.	5095	4539	3806	4160
$\hat{ATT}$	-16.20	44.29	-38.83	72.09
	[21.83]	[18.87]**	[20.20]*	[20.90]***
Obs.	12370	10818	9590	10426

Note: Standard errors in [brackets]. ***, **, * indicates significance at the 1, 5, 10 % level respectively.

Table 4.4: Differences in monthly sales of agricultural products

	1998 November			1999 May		
	Agriculture					
	Net sales	Costs	Gross sales	Net sales	Costs	Gross sales
$\widehat{ITE}$	-18.21	-8.87	-19.12	-35.95	-17.86	-41.25
	[15.02]	[3.15]***	[5.10]***	[21.58]*	[6.24]***	[11.06]***
Obs.	4017	4100	4479	3469	3509	3763
$\widehat{ATT}$	-3.72	3.19	-1.11	-3.90	1.67	-3.51
	[1.84]**	[0.75]***	[0.53]**	[3.97]	[1.27]	[1.40]**
Obs.	9458	9595	10696	8666	8759	8666
	Animals					
	Net sales	Purchases	Gross sales	Net sales	Purchases	Gross sales
$\widehat{ITE}$	1.99	-0.07	0.72	0.17	0.02	-0.59
	[1.47]	[0.41]***	[1.28]	[0.77]	[0.28]	[0.53]
Obs.	4491	4541	4557	3736	3777	3779
$\widehat{ATT}$	-0.17	-0.03	-0.03	-0.09	0.18	0.15
	[0.27]	[0.11]	[0.13]	[0.17]	[0.08]**	0.10
Obs.	10640	9595	10760	9390	9493	9501

Note: differences in net sales estimated by OLS, with standard errors clustered at the village level. Differences in gross sales and costs (purchases) estimated by tobit MLE.
***, **, * indicates significance at 1, 5, 10 % levels.

Table 4.5: Estimates of differences in participation to at least one alternative aid programs, in average number of programs benefits (for households who participate to at least one program), and in monetary receipt

	Non-poor				Poor			
	97	1998 Nov.	1999 May	1999 Nov.	97	1998 Nov.	1999 May	1999 Nov.
	Overall:							
At least one	0.033	0.013	0.043	0.013	-0.001	-0.089	-0.132	-0.078
	[0.024]	[0.029]	[0.032]	[0.031]	[0.026]	[0.025]***	[0.026]***	[0.032]**
	5260	4615	3822	4259	12482	10911	9606	10516
How many if >0	0.011	-0.023	0	-0.001	0.084	-0.114	-0.126	-0.115
	[0.038]	[0.028]	[0.028]	[0.030]	[0.058]	[0.019]***	[0.021]***	[0.022]***
	1293	2177	1661	1827	4356	4582	3396	3613
Monetary transfer	n.a.	10.928	5.975	-3.336	n.a.	-26.036	-19.178	1.133
		[16.633]	[4.583]	[5.354]		[6.078]***	[2.321]***	[1.747]
		4569	3784	4218		10803	9513	10411

Note: We estimate the difference in the likelihood of participating to at least one program by probit, the difference in the number of programs participating to (if at least one) by OLS, and the total monetary receipt by tobit. Standard errors in [brackets] clustered at the village level (apart from the tobit specification). ***, **, * indicates significance at 1, 5, 10 % levels.

Table 4.6: Credit resources: mean, recipient proportion, and average amount obtained per recipient, by household type and semester

	1998 November			1999 May			1999 November		
	Mean	%	Avg. receipt	Mean	%	Avg. receipt	Mean	%	Avg. receipt
				Total credit resources:					
NP control	47.795	11.91	404.681						
	[235.436]		[571.121]						
NP treatment	54.073	12.94	423.903						
	[249.015]		[574.522]						
P control	21.651	8.44	259.939						
	[125.693]		[357.858]						
P treatment	22.201	8.59	260.398						
	[125.184]		[349.228]						
				Loans:					
NP control	8.218	2.84	289.264	11.165	3.71	301.057			
	[79.402]		[378.541]	[105.302]		[464.083]			
NP treatment	11.509	3.28	358.194	16.403	5.30	314.247			
	[116.847]		[551.285]	[123.282]		[446.186]			
P control	3.955	2.73	147.627	7.653	4.62	167.729			
	[37.238]		[175.535]	[57.777]		[215.808]			
P treatment	4.012	2.75	147.324	7.257	4.30	170.541			
	[33.911]		[145.677]	[61.744]		[248.944]			
				Monetary transfers from family and friends:					
NP control	29.203	3.89	751.666				18.064	2.43	741.975
	[246.713]		[1018.716]				[246.090]		[1413.340]
NP treatment	30.212	4.57	671.928				26.991	3.74	737.183
	[223.084]		[825.028]				[201.382]		[768.112]
P control	11.480	2.57	446.990				3.596	1.33	274.826
	[110.974]		[536.127]				[43.101]		[262.137]
P treatment	10.198	2.61	397.255				7.081	1.69	421.680
	[94.221]		[439.463]				[80.080]		[457.063]
				In-kind transfers from family and friends:					
NP control		1.120						1.623	
		[1.054]						[1.261]	
NP treatment		1.921						0.964	
		[1.372]						[9.743]	
P control		1.492						1.607	
		[1.214]						[1.255]	
P treatment		1.398						1.033	
		[1.171]						[1.011]	

Note: Amounts are in pesos per month; the exchange rate was roughly 10 pesos per USD. The last percentile of positive values has been trimmed in the computation of the quantities but not for the proportions.

Table 4.7: Program effects on credit resources

	1998 Oct.		1999 May		1999 Nov.	
	Probit	Tobit	Probit	Tobit	Probit	Tobit
			Non-poor			
			Total credit resources:			
$\widehat{ITE}$	0.0037	4.9357				
	[0.0138]	[5.9083]				
Obs.	4598	4595				
			Loans:			
$\widehat{ITE}$	0.0073	4.0375	0.0151	3.7937		
	[0.0064]	[3.5252]	[0.0076]*	[2.0264]*		
Obs.	4598	4595	3671	3802		
			Monetary transfers from family and friends:			
$\widehat{ITE}$	0.0007	1.561			0.0074	6.562
	[0.0057]	[3.5437]			[0.0048]	[3.720]*
Obs.	4600	4525			4246	4194
			In-kind transfers from family and friends:			
$\widehat{ITE}$	0.0038				-0.0059	
	[0.0029]				[0.0036]*	
Obs.	4479				3973	
			Poor			
			Total credit resources:			
$\widehat{ATT}$	0.0024	0.5373				
	[0.0080]	[1.685]				
Obs.	10893	10885				
			Loans:			
$\widehat{ATT}$	0.0017	0.1068	-0.0084	-1.4556		
	[0.0042]	[0.5072]	[0.0061]	[0.6294]***		
Obs.	10893	10889	9478	9569		
			Monetary transfers from family and friends:			
$\widehat{ATT}$	-0.0007	-0.5267			0.0015	0.9958
	[0.0027]	[1.011]			[0.0015]	[0.8763]
Obs.	10894	10741			10500	10361
			In-kind transfers from family and friends:			
$\widehat{ATT}$	-0.0009				-0.0049	
	[0.0018]				[0.0019]**	
Obs.	10894				10500	

Note: Top 1% of positive values is trimmed in the Tobit.
Standard errors in [brackets] clustered at the village level
in the Probit regressions. ***, **, * indicates significance at the 1, 5, 10 % level respectively.

Table 4.8: Average (std. dev.) monthly food consumption levels per adult equivalent, by shock

	Food consumption		
	1998 Nov.	1999 May	1999 Nov.
S = 0			
NP control	203.87	200.89	198.13
	(165.61)	(179.71)	(175.01)
NP treatment	197.60	238.05	209.44
	(160.32)	(342.09)	(211.19)
S = 1			
NP control	202.39	201.94	191.68
	(171.63)	(208.55)	(282.50)
NP treatment	199.62	208.42	217.77
	(160.67)	(226.36)	(342.58)
S = 0			
P control	138.15	151.32	138.91
	(92.80)	(123.51)	(113.18)
P treatment	152.02	170.27	168.48
	(131.40)	(163.07)	(207.15)
S = 1			
P control	136.59	140.13	140.54
	(115.05)	(151.73)	(122.74)
P treatment	149.94	167.21	168.87
	(107.84)	(189.97)	(143.66)

Note: Amounts are in pesos, the exchange rate was roughly 10 pesos per USD.

Table 4.9: Effect of progresa on poor and non-poor food consumption, by shock

	Non-poor			Poor		
	1998 Nov.	1999 May	1999 Nov.	1998 Oct.	1999 May	1999 Nov.
$I\hat{TE}^{S0}$	-0.0235	0.0911	0.0544	0.0972	0.1347	0.1958
	[0.0303]	[0.0421]**	[0.0277]*	[0.0243]***	[0.0272]***	[0.0228]***
$I\hat{TE}^{S1}$	0.0168	0.066	0.1444	0.0991	0.1735	0.1832
	[0.0360]	[0.0350]*	[0.0364]***	[0.0368]***	[0.0285]***	[0.0376]***
$I\hat{TE}^{S1} - I\hat{TE}^{S0}$	0.0403	-0.0251	0.0900	-0.0019	0.0388	-0.0126
	[0.0422]	[0.0366]	[0.0431]**	[0.0140]	[0.0317]	[0.0309]
Obs.	4615	3825	4264	10911	9608	10517

Note: The standard errors in [brackets] are clustered at the village level. The parameters estimate should read *ATT* for poor. ***, **, * indicates significance 1, 5, 10 % respectively.

Table 4.10: Mean [std. dev.], proportion, and average receipt of the available credit resources, by shock

	1998 Oct.			1999 May			1999 Nov.		
	Mean	%	Avg. receipt	Mean	%	Avg. receipt	Mean	%	Avg. receipt
				Total credit resources:					
S=0									
NP control	59.691	12.61	554.707						
	[278.566]		[885.931]						
NP treatment	52.273	12.19	538.312						
	[236.336]		[917.846]						
S=1									
NP control	31.52	10.95	287.861						
	[157.179]		[391.637]						
NP treatment	56.791	14.07	551.742						
	[267.094]		[1413.64]						
S=0									
P control	23.769	8.17	318.507						
	[133.549]		[446.005]						
P treatment	25.532	8.73	316.756						
	[140.043]		[449.479]						
S=1									
P control	18.914	8.78	277.192						
	[114.728]		[540.304]						
P treatment	17.509	8.4	235.682						
	[100.426]		[404.897]						
				Loans:					
S=0									
NP control	10.613	3.08	344.271	16.295	4.63	352.318			
	[97.103]		[443.67]	[140.126]		[562.472]			
NP treatment	7.994	2.34	340.95	11.912	4.53	386.82			
	[84.879]		[445.567]	[84.185]		[839.32]			
S=1									
NP control	4.935	2.51	196.623	7.522	3.06	245.998			
	[44.821]		[211.166]	[70.734]		[329.858]			
NP treatment	16.836	4.70	804.233	19.636	5.86	392.704			
	[152.945]		[2289.049]	[144.983]		[692.395]			
S=0									
P control	2.919	2.35	166.348	8.299	4.55	285.46			
	[29.285]		[244.262]	[57.572]		[625.651]			
P treatment	4.848	2.74	176.758	9.772	4.58	249.653			
	[40.215]		[169.78]	[73.982]		[371.46]			
S=1									
P control	5.294	3.21	165.014	7.227	4.66	155.239			
	[45.471]		[196.799]	[57.921]		[222.573]			
P treatment	2.834	2.76	145.15	5.376	4.10	168.005			
	[22.12]		[264.066]	[50.628]		[476.358]			
				Monetary transfers from family and friends:					
S=0									
NP control	40.373	4.46	905.761				12.179	2.14	569.854
	[307.972]		[1171.238]				[117.691]		[586.588]
NP treatment	33.441	4.97	816.626				32.36	3.97	1176.285
	[254.179]		[1240.642]				[229.979]		[2134.937]
S=1									
NP control	13.71	3.09	443.478				29.848	2.43	984.971
	[115.694]		[502.789]				[392.487]		[2095.718]
NP treatment	25.37	3.96	641.159				16.428	3.74	502.241
	[165.848]		[553.774]				[127.094]		[508.405]
S=0									
P control	12.603	2.95	426.823				2.870	1.17	504.882
	[113.466]		[512.747]				[38.251]		[1476.105]
P treatment	12.986	2.73	531.477				7.614	1.74	546.007
	[114.76]		[614.372]				[86.942]		[1038.744]
S=1									
P control	10.034	2.07	484.054				5.345	1.72	311.628
	[107.69]		[582.13]				[52.969]		[267.511]
P treatment	6.274	2.44	317.059				5.737	1.59	360.816
	[53.035]		[517.004]				[59.312]		[310.396]
				In-kind transfers from family and friends (%):					
S=0									
NP control		0.77						1.75	
NP treatment		2.28						0.70	
S=1									
NP control		1.45						1.06	
NP treatment		1.68						1.49	
S=0									
P control		1.45						1.78	
P treatment		1.48						0.96	
S=1									
P control		1.27						1.17	
P treatment		1.27						1.29	

Note: Amounts are in pesos per month, the exchange rate was roughly 10 pesos per USD. Top 1% of positive values is trimmed in the computation of the quantities but not for the proportions.

Table 4.11: Effect of Progresa on credit for the non-poor, by shock

	1998 Oct.		1999 May		1999 Nov.	
	Probit	Tobit	Probit	Tobit	Probit	Tobit
			Total credit resources:			
$IT\hat{E}^{S0}$	-0.0117	-5.679				
	[0.0173]	[7.3795]				
$IT\hat{E}^{S1}$	0.0263	22.039				
	[0.0195]	[8.9567]**				
$IT\hat{E}^{S1} - IT\hat{E}^{S0}$	0.038	27.718				
	[0.0233]*	[11.9991]**				
Obs.	4598	4595				
			Loans:			
$IT\hat{E}^{S0}$	-0.0057	-4.2357	0.0063	-0.2219		
	[0.0079]	[4.3411]	[0.0127]	[2.7923]		
$IT\hat{E}^{S1}$	0.0264	16.6777	0.0249	7.5166		
	[0.011]***	[4.976]***	[0.0139]**	[2.6014]***		
$IT\hat{E}^{S1} - IT\hat{E}^{S0}$	0.0321	20.9134	0.0186	7.7385		
	[0.0109]***	[7.5773]***	[0.0163]	[4.1162]*		
Obs.	4598	4595	3671	3802		
			Monetary transfers from family and friends:			
$IT\hat{E}^{S0}$	0.0013	0.9502			0.0125	11.8295
	[0.0069]	[20.8995]			[0.0072]*	[4.3498]***
$IT\hat{E}^{S1}$	-0.0004	4.3990			-0.0009	-1.9762
	[0.0076]	[5.7945]			[0.0074]	[5.8670]
$IT\hat{E}^{S1} - IT\hat{E}^{S0}$	-0.0017	3.4488			-0.0134	-13.8057
	[0.0089]	[6.0505]			[0.0098]	[7.5441]*
Obs.	4600	4525			4246	4194
			In-kind transfers from family and friends:			
$IT\hat{E}^{S0}$	0.0102				-0.0078	
	[0.0052]**				[0.0033]**	
$IT\hat{E}^{S1}$	-0.0027				0.0083	
	[0.0044]				[0.0070]	
$IT\hat{E}^{S1} - IT\hat{E}^{S0}$	-0.0129				0.0161	
	[0.0063]**				[0.09]*	
Obs.	4479				3973	

Note: Top 1% of positive values is trimmed in the Tobit.
The standard errors in [brackets] are clustered at the village level in the Probit regressions.
***, **, * indicates significance at the 1, 5, 10 % level respectively.

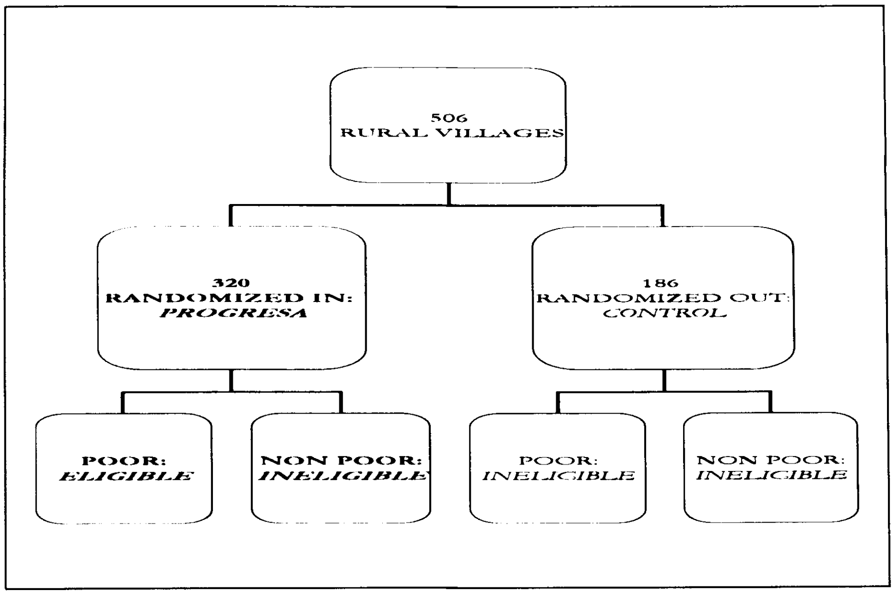

Figure 4.1: Design of the data for the evaluation of Progresa

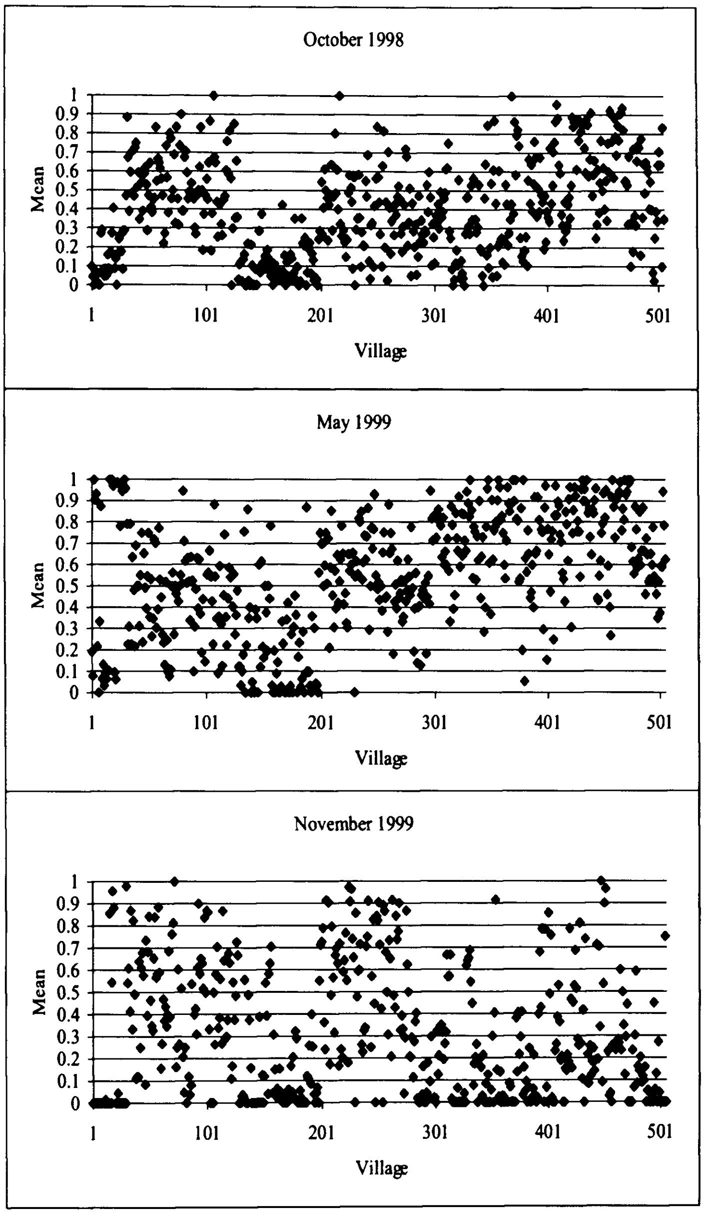

Figure 4.2: Fraction of households hit by a shock at the village level

4.9 APPENDIX

In this Appendix we describe how we created some of the relevant variables for our analysis: consumption, transfers and loans, school enrollment, hours of work, earnings, and prices.

4.9.1 Food consumption

We consider the three data waves collected after the program begins, in November 1998, May 1999, and November 1999. Considering food consumption first, households report both the expenditure in food purchased during the last week and the quantity bought. If expenditure on a particular item is missing, but we know the amount purchased, we consider the village median price. We compute the village price in the following way: we create household-specific prices by dividing the expenditure in food purchased during the last week by the quantity bought. If we have at least 20 household-specific prices per village, we use this information to compute median prices at the village level. Otherwise, we use either median municipality or state price (we use the lowest level of aggregation with at least 20 price observations). Once we have household-specific prices, we multiply them by quantity consumed. We do this because households produce part of the consumed food. Considering only food expenditure would underestimate the amounts actually consumed.

We use November 1998 prices to compute consumption values in May and November 1999 also. Unlike in 1998, in 1999 we know both how much food is purchased and how much is consumed, but we have no direct information on home-produced food. Hence, in order to be consistent between the three waves, we assume that in 1999 all food purchased is consumed if total consumption is smaller or equal than food purchased. If total consumption is greater than purchased goods we apply median prices to the difference, this means that either home-produced food, or food given as a present is evaluated at market prices. Since we could not convert different measurement units in a single one, we only consider those who have bought and consumed food in the same unit (Kilo, Liter or Units). We believe that the absence of measurement conversion does not pose any major problem, since only about 1% of the sample has different measurements for the same food. Lastly, we compute adult equivalents for both food and non-food data. For this purpose, we use the adult equivalence conversion estimated by Di Maro (2004) using Progresa data. According to Di Maro, children consume on average 73% of adults. For example, to estimate individual consumption per adult equivalent for a household with one child and one adult, we divide household consumption by 1.73.

An additional issue is how to treat missing observations. We noted that some aliments, which are not staples for rural Mexicans, have a large number of missing observations. Thus, we create three different food expenditure variables, each time dropping all households with missing observations. The first variable is aggregate expenditure in food consumption for all available categories (hence the one with the highest number of missing observations). In this way, we drop about 5% of the sample. The second one excludes industrially produced food (*pastelillos en bolsa*, soft drinks, coffee, sugar, vegetable oil). The third food consumption variable excludes industrially produced food, sliced bread (*pan de caja*), breakfast cereals, fish, and seafood. The results we show in the paper use the first consumption variable. However, they are robust to the use of these alternative variables.

The food consumption variable we use in the paper has the following number of non-missing observations for non-poor: 5004 in November 1998, 3857 in May 1999, and 4286 in November 1999. 371 (i.e. about 7%) households have zero food consumption in November 1998. Only 14 and 3 households have zero food expenditure in the May and November 1999 data, respectively. 10 households have a food consumption level larger than 10000 *pesos* per adult equivalent per month in the 1999 waves. We consider these extreme values to be due to measurement error, and we omit the corresponding observations from our sample. After these trimmings, we are left with the following sample sizes for non-poor households: 4633, 3839, and 4277 in November 1998, May 1999, and November 1999, respectively. We do the same for poor households, whose final samples have 10943, 9631, and 10547 households for the three waves.[34] There is a drop in the valid household size in May 1999, supposedly due to a higher proportion of non-responses (this drop is not limited to the consumption variables). However, the proportion of households in treated and control areas is roughly constant over time (for non-poor, this proportion ranges from 38.8% living in control areas in November 1998 to 39.9% in November 1999). Because of this, we believe that the smaller sample size in May 1999 does not pose attrition problems.

4.9.2 Non-food consumption

For non-food consumption, we also consider the three waves used above. The variable on non-food consumption is only available as expenditure on particular categories of non-food items. Our measure of monthly non-food consumption is the sum of expenditures in: transport both for adults and children; tobacco; personal and house hygiene; drugs and prescriptions; doctor visits; heating (ie. wood, gas, oil); electricity; clothing and shoes; school items (ie. pencils, books).[35] As for food

[34]Note that the sample sizes from the regressions may differ from the aforementioned ones because some conditioning variables may have missing observations.

[35]Information on expenditure on durable goods such as household appliances or home improvement is not included in our measure of non-food consumption because we observe these data only in the 1998 wave. For that year, we compute measures of both durable goods consumption and agricultural expenditure. We consider durable goods all home appliances purchased (ranging from blenders to vans), and all expenditures to improve the house. Our measure of agricultural-related expenditures refer to cost of seeds, fertilizer, pesticides, machinery, and labor, excluding land rental costs. Both measures refer to the previous 12 months.

consumption, we trim the extreme values because of possible measurement error. The value of the expenditure is then converted in real terms by applying the monthly CPI (Bank of Mexico, 2005).

When estimating the ITE on consumption, we perform a variety of robustness checks. First, we test for pre-program differences in consumption between non-poor households in treatment and control villages using March 1998 data.[36]. These differences are never significant. Second, since about 2.9% and 1.5% of non-poor households report zero food and non-food consumption, respectively, we re-estimated equation (4.3) using consumption levels (including zeros), rather than logs, and interpreting zero consumption as measurement error or infrequent purchases, in which case OLS estimates are consistent.[37] We find a significant increase in both food and non-food consumption for non-poor households in treated areas in both 1999 waves also when we include these additional households. As further robustness checks, we test whether these results are robust to a different type of trimming: for example, rather than dropping households who report a monthly food consumption larger than 10000 *pesos* per adult equivalent, we omit the largest percentile: the significance of the results does not change (although the point estimates are slightly different, of course). Moreover, we also estimate ITEs using the alternative measures of food consumption (described in the Appendix). Again, the magnitude and the significance of the estimates do not vary considerably.[38]

Table 4.12 provides estimates of average durable expenditure differences for the whole sample, the probability of having positive expenditures, and the difference in

[36]March 1998 consumption data are not directly comparable with the other data waves, given that only aggregate expenditure categories are available, and that aggregate expenditure information often results in under-reporting (Deaton (1997)). For example, there are only four food categories: vegetables and fruit; aliments of animal origin; processed foods; grains and cereals.

[37]Note that the 2.9% zero food consumption comes mainly from November 1998, where about 7% of non-poor households have zero food consumption.

[38]Results available upon request.

average expenditures for households with positive expenditure levels.

Table 4.12: Durable expenditure: house appliances/improvements and agricultural production/investment

	House appliances and improvements			Production costs/investments		
	OLS	Probit	OLS if >0	OLS	Probit	OLS if >0
$\hat{ITE}$	27.923	0.005	204.142	-128.4	-0.007	-127.54
	[37.405]	[0.015]	[234.422]	[106.924]	[0.006]	[109.436]
$\hat{ATT}$	-3.614	0.004	-48.861	23.872	-0.006	25.09
	[12.056]	[0.010]	[90.343]	[20.904]	[0.004]*	[21.164]

Note: The standard errors in [brackets] are clustered at the village level. ***, **, * indicates significance 1, 5, 10 % respectively.

4.9.3 Hours of work and labor earnings

The 1998 and 1999 surveys report hours of work for the sole sub-set of individuals who have a paid job, unlike the 1997 pre-program one, which collects working hour information for all individuals. In 1997 there is no explicit distinction between paid and unpaid jobs. Thus, in order to create a consistent measure, we excluded self-employed, business owners and *ejidatarios* from the computation of hours of work. We considered as unemployed all individuals who reported not having a job in the previous week (unlike those who said that they have a job but could not work). In case of disagreement (i.e. individuals reporting they do not have a job, but having a positive number of hours worked) we included the reported work time.

These variables are very noisy measures of work time and earnings, as at times we have to impute monthly earnings from daily, weekly or annual wages. To reduce measurement error, we trim the top and bottom percentile. However, we do not expect the type of error in measurement to differ systematically between households in treatment and control villages. Moreover, when used as dependent variable, classic measurement error affects only the precision, but not the consistency of the coefficient estimates. We try to offset this lack of precision by including a large set of conditioning variables.

4.9.4 Prices

Prices refer to the food and non-food goods used to compute the value of consumption. There are 57 different goods, but only food prices are available before the program begins, in March 1998. Thus, we use the 36 food prices available both before and during the program implementation to provide DD estimates of the effect of Progresa. In November 1998 and May 1999 we have up to two prices for each good. When two different prices for the same good are available, we compute the mean village price. Table 4.13 provides a list of the goods used in the DD analysis of the program effect.

We find a small positive effect on some food prices in November 1998. Prices of onions (p2), lemons (p8), eggs (p26), and coffee (p34) are significantly higher in treatment than in control areas. At the same time, though, the price of fish (p23) is significantly lower. Despite the fact that onions, eggs, and coffee are commonly consumed foods (Hoddinott *et al.*, (2000)), we do not expect these price changes to increase the cost of the food basket substantially, because prices of staples such as rice, beans, corn, and chicken do not change. Second, there is no price change in the later waves. Third, if we consider the pooled waves, the prices of 6 items increase, while the prices of 3 goods decrease in the observed time, out of a total of 36 items by 3 waves. This amounts to roughly a change in 8% of good prices. We believe that, perhaps with the exception of a minor price increase for some goods in the end of 1998, Progresa does not significantly change prices of treatment areas.[39]

As a further robustness check, we considered all 57 different (food and non-food) goods available in the 3 waves collected after the beginning of the program. We

[39]There is a large number of missing observations. Since there are 506 villages observed in 4 different points in time, each price should have about 2000 observations. Instead, the non-missing observations range between 313 and 1375.

Table 4.13: Food prices used to compute DD estimates of program effect on prices

p1	tomatoes (kilo)
p2	onions (kilo)
p3	potatoes (kilo)
p4	carrots (kilo)
p5	oranges (kilo)
p6	bananas (kilo)
p7	apples (kilo)
p8	lemons (kilo)
p9	lettuce (unit)
p10	nixtamal masa (kilo)
p11	corn grains (kilo)
p12	Bread (unit)
p13	Bread "de caja" (unit)
p14	wheat flour (kilo)
p15	soup (200 grs.)
p16	rice (kilo)
p17	Tortillas (kilo)
p18	corn "hojuelas" (unit)
p19	chicken (kilo)
p20	pork (kilo)
p21	beef (kilo)
p22	goat (kilo)
p23	fish (kilo)
p24	biscuits (kilo)
p25	beans (kilo)
p26	eggs (kilo)
p27	milk (liter)
p28	lard (kilo)
p29	pastry (bag)
p30	soft drink (bottle)
p31	Sardines (150 grs. in 98m, 400grs. after)
p32	Tuna can (175 grs.)
p33	aguardiente (liter)
p34	coffee (small pack)
p35	sugar (kilo)
p36	vegetable oil (liter)

pooled prices, creating a price basket that gives equal weight of one to each good. We then regressed this synthetic price indicator on a dummy for treatment and control villages, obtaining cross sectional estimates of the effect of Progresa on prices. Also in this case we reject the hypothesis that prices differ significantly between the two village groups.[40]

Table 4.14: Difference in differences estimates of the effect of Progresa on village prices

	p1	p2	p3	p4	p5	p6	p7	p8	p9
T	0.0826	-1.1782	-0.0435	-0.5462	0.1246	-0.0443	-0.0414	-1.3424	-1.6077
	[0.2582]	[0.4387]***	[0.2357]	[0.4391]	[0.2012]	[0.1677]	[0.3895]	[0.9675]	[1.2532]
T*1998 Nov.	0.095	1.2498	0.0609	0.877	-0.5583	0.2174	-0.2832	1.8117	1.6452
	[0.3840]	[0.4985]**	[0.3974]	[0.6473]	[0.3953]	[0.2851]	[0.6732]	[1.0401]*	[1.2807]
T*1999 May	0.0804	0.781	-0.3818	0.1531	0.0866	-0.3466	-0.1483	1.1914	1.2465
	[0.7453]	[0.5417]	[0.3099]	[0.5599]	[0.2987]	[0.2831]	[0.7589]	[1.0111]	[1.2876]
T*1999 Nov.	-0.8173	1.3505	-1.3779	1.5488	-1.6462	-1.1732	-0.325	3.2171	2.0641
	[0.3489]**	[0.8422]	[0.8913]	[2.5669]	[0.6641]**	[1.4749]	[0.5746]	[2.2543]	[1.3830]
Obs.	1034	990	948	369	678	698	426	548	413

	p10	p11	p12	p13	p14	p15	p16	p17	p18
T	-0.5039	0.0057	-0.3148	-1.69	0.0501	-0.3291	-0.1483	0.0265	0.8105
	[0.3412]	[0.2665]	[0.3250]	[1.3105]	[0.1785]	[0.4101]	[0.1409]	[0.1299]	[0.9233]
T*1998 Nov.	0.4998	0.2034	0.3913	1.4979	-0.0945	0.2555	-0.1047	0.0531	0.1171
	[0.4262]	[0.3482]	[0.3511]	[1.3912]	[0.2497]	[0.4202]	[0.1859]	[0.1758]	[1.3075]
T*1999 May	0.2958	-0.428	0.0864	2.0122	-0.2671	0.264	0.1945	0.3573	0.694
	[0.4109]	[0.3291]	[0.6539]	[1.3717]	[0.3055]	[0.4162]	[0.2806]	[0.3769]	[1.5655]
T*1999 Nov.	0.8789	2.5907	4.0232	0.4852	9.1043	-0.3298	-0.569	-0.4817	-1.6231
	[4.4228]	[1.7468]	[9.9232]	[1.4638]	[9.2109]	[0.5394]	[0.5809]	[0.4022]	[1.1575]
Obs.	365	640	750	390	678	1233	1375	424	565

	p19	p20	p21	p22	p23	p24	p25	p26	p27
T	-0.2255	-0.5765	-0.1634	-15.4505	-1.6939	-0.15	-0.3617	-0.9393	-0.34
	[0.6422]	[0.9607]	[1.2435]	[9.0361]*	[2.0772]	[0.2455]	[0.1621]**	[0.3405]***	[0.2410]
T*1998 Nov.	-1.8291	-1.8317	-1.3589	13.7839	-6.8775	0.1938	0.0172	1.1282	0.3435
	[1.3035]	[1.6762]	[3.2008]	[11.1087]	[3.9947]*	[0.2641]	[0.3141]	[0.4336]***	[0.3052]
T*1999 May	-0.5113	0.9343	1.1377	12.0755	5.714	0.129	0.3623	0.5862	0.3434
	[0.8389]	[1.2960]	[1.7734]	[11.5156]	[5.6679]	[0.2637]	[0.2650]	[0.4429]	[0.4402]
T*1999 Nov.	-1.2303	0.9294	-0.0991	15.3242	3.4506	0.1216	-0.2787	0.4151	-2.3614
	[2.2299]	[1.1187]	[1.2644]	[9.0350]*	[2.7004]	[0.4785]	[0.4876]	[0.7810]	[3.3976]
Obs.	486	566	313	334	344	1375	1194	1206	833

	p28	p29	p30	p31	p32	p33	p34	p35	p36
T	-0.0689	-0.1552	0.0916	-0.0645	-0.0312	-0.3793	-1.5319	-0.1483	-0.1349
	[0.3556]	[0.1894]	[0.2142]	[0.1266]	[0.0939]	[0.7805]	[0.5373]***	[0.0950]	[0.1196]
T*1998 Nov.	-0.0634	0.1684	-0.151	0.1084	-0.0026	1.2379	2.2448	0.1084	0.1579
	[0.5199]	[0.2163]	[0.2647]	[0.1734]	[0.1331]	[1.1330]	[0.7657]***	[0.1324]	[0.1828]
T*1999 May	-0.3067	-0.068	-0.3152	0.542	0.2061	1.5668	0.3668	0.1521	0.1635
	[0.5517]	[0.3063]	[0.3013]	[0.2638]**	[0.1485]	[1.5627]	[0.7374]	[0.2031]	[0.1800]
T*1999 Nov.	0.5788	0.6931	6.3978	-0.4729	-0.3143	0.3309	-0.0825	0.6451	2.5376
	[0.7721]	[0.9239]	[10.4329]	[0.5949]	[10.0995]	[0.8828]	[6.4283]	[0.7407]	[2.2574]
Obs.	634	488	922	1272	1021	757	636	1431	1219

Note: Standard errors clustered at the village level. ***, **, * indicates significance at the 1, 5, 10 % level respectively.

4.9.5 Alternative program receipt

Table 4.15 shows estimates of differences in participation to alternative aid programs. The first 4 rows of the Table present estimates of the differences in the

[40] Results available upon request.

Table 4.15: Differences in participation to alternative aid programs

	Non-poor				Poor			
	97 Sept.	1998 Nov.	1999 May	1999 Nov.	97 Sept.	1998 Nov.	1999 May	1999 Nov.
				Overall:				
At least one	0.033	0.013	0.043	0.013	-0.001	-0.089	-0.132	-0.078
	[0.024]	[0.029]	[0.032]	[0.031]	[0.026]	[0.025]***	[0.026]***	[0.032]**
	5260	4615	3822	4259	12482	10911	9606	10516
How many if >0	0.011	-0.023	0	-0.001	0.084	-0.114	-0.126	-0.115
	[0.038]	[0.028]	[0.028]	[0.030]	[0.058]	[0.019]***	[0.021]***	[0.022]***
	1293	2177	1661	1827	4356	4582	3396	3613
Monetary	n.a.	-26.005	5.827	-8.186	n.a.	-27.328	-24.914	-2.291
transfer		[33.063]	[7.757]	[12.531]		[13.978]*	[5.789]***	[4.794]
OLS		4569	3784	4218		10803	9513	10411
Monetary	n.a.	10.928	5.975	-3.336	n.a.	-26.036	-19.178	1.133
transfer		[16.633]	[4.583]	[5.354]		[6.078]***	[2.321]***	[1.747]
tobit		4569	3784	4218		10803	9513	10411
				Cash transfers (participation):				
Solidaridad	0.005	0.009	0.018	0.011	0.013	-0.096	-0.087	-0.04
	[0.008]	[0.007]	[0.006]***	[0.006]	[0.013]	[0.010]***	[0.010]***	[0.007]***
	5232	4426	3812	4246	12469	10911	9606	9193
INI	0	-0.001	—	—	-0.001	0	—	—
	[0.001]	[0.001]			[0.002]	[0.001]		
	5046	2696			10593	9214		
Probecat	0.001	-0.002	—	—	0	-0.002	—	—
	[0.001]	[0.002]			[0.001]	[0.001]**		
	5048	2621			12175	9867		
Empleo temporal	0.003	0.002	—	—	0.001	0	—	—
	[0.001]**	[0.001]			[0.000]*	[0.001]		
	5046	2860			12175	10521		
Procampo	n.a.	0.018	0.004	-0.001	n.a.	0.061	0.029	0.053
		[0.030]	[0.032]	[0.031]		[0.024]**	[0.026]	[0.026]**
		4615	3822	4259		10911	9606	10516
				Transfers in kind (participation):				
DIF food	0.032	-0.021	0	-0.005	-0.005	-0.067	-0.083	-0.093
	[0.021]	[0.016]	[0.018]	[0.020]	[0.022]	[0.014]***	[0.015]***	[0.018]***
	5258	4560	3822	4259	12469	10677	9606	10516
Desayuno DIF	0.001	n.a.	n.a.	n.a.	0.003	n.a.	n.a.	n.a.
	[0.008]				[0.013]			
	4459				12306			
Tortilla	0	0	0	0	0	0	0.001	0
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.001]**	[0.000]
	4185	3840	2854	1872	12030	8465	9297	8312
Milk	-0.005	-0.002	0.007	-0.003	0	-0.001	-0.005	-0.001
	[0.007]	[0.006]	[0.003]**	[0.003]	[0.006]	[0.006]	[0.003]	[0.002]
	5194	4600	3812	3438	12179	10911	9591	10500

Note: The likelihood of receiving individual program transfers is estimated by Probit. The total monetary transfer received is estimated both by OLS and by tobit. Standard errors in [brackets] are clustered at the village level (apart from tobit estimates). ***, **, * indicates significance at the 1, 5, 10 % level respectively.

aggregate measures. The following 8 rows show estimates of the differences in participation to the five cash transfer programs, and the three in kind programs. All these variables are at the household level, and refer to the month before the interviews took place. We find that there are no differences in alternative welfare receipt in 1997, with the exception of a slightly higher participation to the temporary employment program in treatment areas (for both poor and non-poor households). Thus, we rely on cross-sectional estimates of the differences for the remaining three data waves.

It is possible that non-poor participation to alternative welfare programs may have increased for all control villages: for example, households in control villages know of the existence of Progresa and may feel deprived, hence participate more to alternative programs. Alternatively, program officials may become more "generous" in communities excluded from Progresa. This increase in welfare receipts in control villages would be compatible with the results presented in Table 4.15. For instance, the absence of significant differences for non-poor may be consistent with equally increased participation rates in both types of villages. Similarly, part of the apparent decrease in receipt for treatment poor may actually be due to increased participation in control villages. In any case, note that if non-poor alternative program take-up increased equally in both treatment and control areas, it still would not explain why non-poor consumption is higher in treatment than in control villages.

One way to observe whether alternative programs take-up surged in both types of villages is to compare trends in these programs' pre- and during-Progresa participation rates. If we observed a substantial peak in 1998 and 1999 non-poor participation, compared to the 1997 levels, we may suspect that this is partly an indirect effect of Progresa in both treatment and control villages. Of course the supposed increase in welfare participation may as well be due to exogenous reasons. Table 4.16 compares average participation rates over time for Solidaridad, DIF food, Liconsa milk, and Tortilla. These are the only programs we observe for all 4 data waves. The comparison of program take-up rates over time clearly shows that there is no surge in participation in control areas after the beginning of Progresa. In fact, participation rates seem to slowly decrease over time.

Table 4.16: Average (std. dev.) monthly participation rates to selected alternative aid programs

	1997 Sept.	1998 Nov.	1999 May	1999 Nov.
	Solidaridad			
NP control	0.081	0.074	0.053	0.040
	(0.273)	(0.262)	(0.224)	(0.197)
NP treatment	0.079	0.077	0.079	0.052
	(0.270)	(0.266)	(0.269)	(0.222)
P control	0.141	0.136	0.121	0.049
	(0.349)	(0.343)	(0.326)	(0.216)
P treatment	0.153	0.026	0.005	0.003
	(0.360)	(0.160)	(0.071)	(0.057)
	DIF food			
NP control	0.112	0.099	0.100	0.094
	(0.316)	(0.299)	(0.300)	(0.292)
NP treatment	0.136	0.072	0.092	0.090
	(0.343)	(0.259)	(0.289)	(0.286)
P control	0.153	0.116	0.112	0.144
	(0.360)	(0.321)	(0.315)	(0.351)
P treatment	0.140	0.044	0.025	0.024
	(0.347)	(0.206)	(0.156)	(0.154)
	Tortilla			
NP control	0.000	0.001	0.001	0.001
	(0.022)	(0.024)	(0.026)	(0.034)
NP treatment	0.002	0.006	0.009	0.004
	(0.039)	(0.077)	(0.095)	(0.059)
P control	0.001	0.001	0.001	0.001
	(0.033)	(0.031)	(0.029)	(0.027)
P treatment	0.001	0.002	0.003	0.002
	(0.030)	(0.045)	(0.050)	(0.045)
	Liconsa milk			
NP control	0.037	0.027	0.012	0.012
	(0.189)	(0.162)	(0.109)	(0.108)
NP treatment	0.029	0.028	0.017	0.009
	(0.166)	(0.164)	(0.130)	(0.096)
P control	0.034	0.034	0.016	0.009
	(0.181)	(0.182)	(0.124)	(0.097)
P treatment	0.032	0.034	0.010	0.008
	(0.176)	(0.180)	(0.099)	(0.088)

4.9.6 Shocks

We observe two different measures of shock in our sample. The data record whether, in the six months preceding the interview, the household has been hit by any of the following types of natural disasters: drought, flood, frost, fire, plague, earthquake, and hurricane. We also know whether any of these natural disasters caused a damage to the household, such as: loss of land, harvest, housing, property, tools, animals, and household members; casualties, and members migrating to find jobs elsewhere. We create two different dummy variables. One records whether the household has been hit by any natural disaster. The second one, instead, considers whether the household suffered from any of the above losses. The natural disaster dummy may include households who did not incur any loss because of the natural disaster. However, we suspect it is more likely to be unrelated to household characteristics than the second variable. For instance, loss of land and animals are conditional on their ownership by the households, while the decision to migrate is constrained by the availability of savings to finance the trip. The results showed in the paper use our preferred shock variable, the natural disaster dummy. However, our estimates do not vary substantially when we use the loss dummy.[41] One potential shortcoming of this variable may be shocks hitting entire villages, leaving no within-village variation. For example, it is unlikely that an earthquake may hit only half a village. However, Figure 4.2 shows that this is not the case: when we plot within-village proportions of households hit by shocks, we notice that only in very few cases the entire village is hit.

In order to compare indirect program effects for households hit and not hit by a shock, we require that shocks are random both between and within villages. To

[41] Results available upon request.

Table 4.17: Tests of the Randomness of Idiosyncratic Shocks

	1998 Nov.	1999 May	1999 Nov.
	Difference in average shock level by village type		
ALL	-0.0029	-0.0254	0.0138
	[0.0239]	[0.0255]	[0.0311]
Obs.	14953	12979	14264
NP	0.0159	0.0028	0.0483
	[0.0294]	[0.0308]	[0.0366]
Obs.	4407	3667	4082
P	-0.0126	-0.0353	0.005
	[0.0267]	[0.0282]	[0.0331]
Obs.	10546	9312	10182

Difference in pre-program consumption by shock status

	Food	Non-Food	Food	Non-Food	Food	Non-Food
ALL	-0.0008	0.017	-0.007	-0.0189	0.004	0.030
	[0.010]	[0.018]	[0.120]	[0.021]	[0.012]	[0.022]
Obs.	13981	13794	12130	11974	13297	13135
NP	-0.011	0.005	-0.002	-0.035	-0.008	-0.025
	[0.019]	[0.037]	[0.021]	[0.044]	[0.023]	[0.038]
Obs.	4209	4145	3502	3448	3890	3838
P	0.001	0.021	-0.012	-0.015	0.009	0.051*
	[0.012]	[0.023]	[0.014]	[0.024]	[0.014]	[0.027]
Obs.	9772	9649	8628	8526	9407	9297

Difference in pre-program consumption by shock status and village type

	Food	Non-Food	Food	Non-Food	Food	Non-Food
NP	0.56	0.10	2.56*	2.41*	0.15	0.53
Obs.	4196	4132	3493	3493	3879	3827
P	0.12	1.00	0.25	1.31	0.14	1.67
Obs.	9757	9634	8614	8512	9393	9283

Note: The usual set of pre-program controls, including average shock intensity at the village level, is added to all regressions (but the upper panel). Standard errors clustered at the village level.
***, **, * are 1, 5, 10 % significance levels.

insure that this is the case, we perform three tests. First, we check whether shocks hit treated and control areas differentially, for all households, and for non-poor and poor families separately. The top panel of Table 4.17 presents the partial effects from a probit on the probability of being hit by a shock on village area dummies, and shows that the estimates are never statistically significant. Second, to test whether more vulnerable households are more likely to suffer from adverse shocks, we check whether households with lower pre-program consumption are more likely to suffer future shocks. We regress March 1998 log food and non-food consumption on the November 1998, May 1999, and November 1999 shock dummies, alternatively. We repeat this exercise for both the whole sample and for poor and non-poor separately. The middle panel in Table 4.17 presents the estimates of the shock dummy coefficients, which are never significant. Thus, we cannot reject the hypothesis that the average pre-program consumption levels are the same for households hit and not hit by adverse shocks. Lastly, we want to test whether average pre-program consumption levels are the same for the four groups defined by village of residence (treatment or control, i.e. $T = 1$ or $T = 0$) and shock status (hit or not hit, i.e. $S = 1$ or $S = 0$). We consider non-poor and poor households separately, and regress log consumption on dummies that group households accordingly (one dummy for $T = 1$ and $S = 0$, one for $T = 1$ and $S = 1$, a third one for $T = 0$ and $S = 0$) and test the hypotheses that the coefficients of these three dummies are jointly equal to zero, interpreting the null as evidence of the randomness of the idiosyncratic shocks. As before, we repeat this exercise three times, since we have shock data for each of the three semesters we consider. We report the values of the F-tests in the lower panel of Table 4.17. We add the usual set of pre-program controls, including average shock intensity at the village level, to all regressions in the Table. Note that the weak significance of the F test for May 1999 shocks disappears as we change the set of conditioning variables. Thus, also this third test confirms that these adverse shocks hit the households in our sample in a random way.

Table 4.18: Effect of Progresa on credit for the poor, by shock

	1998 Nov.		1999 May		1999 Nov.	
	Probit	Tobit	Probit	Tobit	Probit	Tobit
			Total credit resources:			
$AT\hat{T}^{S_0}$	0.0066	1.5004				
	[0.0104]	[2.182]				
$AT\hat{T}^{S_1}$	-0.0029	-0.7207				
	[0.0104]	[2.504]				
$AT\hat{T}^{S_1} - AT\hat{T}^{S_0}$	-0.0095	-2.2211				
	[0.133]	[3.2663]				
Obs.	10893	10885				
			Loans:			
$AT\hat{T}^{S_0}$	0.0048	0.7434	-0.0058	-0.7343		
	[0.0059]	[0.6643]	[0.0062]	[0.9255]		
$AT\hat{T}^{S_1}$	-0.0018	-0.6332	-0.0095	-1.7928		
	[0.0053]	[0.7313]	[0.0066]	[0.7903]**		
$AT\hat{T}^{S_1} - AT\hat{T}^{S_0}$	-0.0066	-1.3766	-0.0037	-1.0585		
	[0.0073]	[0.9627]	[0.0059]	[1.1142]		
Obs.	10893	10889	9478	9569		
			Monetary transfers from family and friends:			
$AT\hat{T}^{S_0}$	-0.0020	-0.7036			0.0031	1.839
	[0.0031]	[1.2475]			[0.0018]*	[1.031]*
$AT\hat{T}^{S_1}$	0.0015	-0.147			-0.0018	-0.7363
	[0.0041]	[1.574]			[0.0022]	[1.5037]
$AT\hat{T}^{S_1} - AT\hat{T}^{S_0}$	0.0035	0.5566			-0.0049	-2.5753
	[0.0047]	[1.9193]			[0.0027]*	[1.7884]
Obs.	10894	10741			10500	10361
			In-kind transfers from family and friends:			
$AT\hat{T}^{S_0}$	-0.0007				-0.0062	
	[0.0021]				[0.0019]***	
$AT\hat{T}^{S_1}$	-0.0012				0.0002	
	[0.0026]				[0.003]	
$AT\hat{T}^{S_1} - AT\hat{T}^{S_0}$	0.0005				0.0064	
	[0.0033]				[0.0033]*	
Obs.	10894				10500	

Note: Top 1% of positive values is trimmed in the Tobit. The standard errors in [brackets] are clustered at the village level in the Probit regressions. ***, **, * indicates significance at 1, 5, 10 % respectively.

Table 4.19: Difference in differences estimates of the effect of Progresa on the stock of animals

	Chickens		Goats		Pigs		Rabbits		Horses		Donkeys		Cows		Oxen	
	Probit	OLS	Probit	OLS	Probit	OLS	Probit	OLS	Probit	OLS	Probit	OLS	Probit	OLS	Probit	OLS
Non-poor																
T	0.007	0.367	0.039	0.072	0.037	0.059	0.001	-0.097	-0.014	-0.025	0.02	0.008	0.004	-0.176	0.012	0.009
	[0.022]	[0.327]	[0.022]*	[0.459]	[0.025]	[0.101]	[0.006]	[0.148]	[0.025]	[0.048]	[0.021]	[0.048]	[0.026]	[0.200]	[0.006]*	[0.063]
T*98N	-0.006	-0.619	-0.011	0.019	-0.029	0	0	0.413	0.004	-0.026	-0.005	0.034	-0.02	-0.158	-0.006	0.144
	[0.025]	[0.319]*	[0.010]	[0.482]	[0.018]	[0.112]	[0.005]	[0.211]*	[0.015]	[0.049]	[0.014]	[0.046]	[0.016]	[0.229]	[0.005]	[0.083]*
T*99M	-0.013	-0.33	0	-0.782	-0.064	-0.007	-0.004	-0.002	0.006	0.051	-0.003	-0.044	-0.028	-0.206	-0.005	-0.01
	[0.026]	[0.354]	[0.013]	[0.555]	[0.021]***	[0.109]	[0.005]	[0.230]	[0.017]	[0.050]	[0.017]	[0.049]	[0.018]	[0.217]	[0.006]	[0.117]
T*99N	0.014	-0.289	-0.007	0.878	-0.043	-0.046	0	0.083	0.023	0.018	0.006	0.003	0.007	0.148	-0.005	-0.143
	[0.023]	[0.354]	[0.015]	[0.556]	[0.021]**	[0.105]	[0.006]	[0.228]	[0.017]	[0.043]	[0.018]	[0.059]	[0.018]	[0.207]	[0.005]	[0.086]*
Obs.	19062	11985	19063	2760	19069	7345	19066	355	19062	4412	19065	4122	19065	4924	19060	662
Poor																
T	-0.007	-0.461	0.021	-0.216	-0.027	-0.115	-0.005	-0.15	0	-0.043	0.012	-0.026	0.002	-0.273	0.003	0.063
	[0.020]	[0.220]**	[0.018]	[0.310]	[0.026]	[0.112]	[0.003]*	[0.112]	[0.013]	[0.039]	[0.017]	[0.036]	[0.013]	[0.142]*	[0.002]	[0.057]
T*98N	0.05	0.501	-0.01	0.228	0.028	0.033	0.006	0.265	0.018	0.028	0.009	-0.01	0.01	0.101	0.004	0.036
	[0.020]**	[0.228]**	[0.008]	[0.308]	[0.020]	[0.114]	[0.004]	[0.152]*	[0.010]*	[0.040]	[0.013]	[0.030]	[0.010]	[0.139]	[0.003]	[0.102]
T*99M	0.056	0.723	-0.002	0.481	0.025	0.234	0.001	0.363	0.017	0.037	0.007	-0.036	0.017	0.122	0.003	-0.085
	[0.021]***	[0.243]***	[0.008]	[0.347]	[0.021]	[0.122]*	[0.004]	[0.183]**	[0.011]	[0.042]	[0.013]	[0.035]	[0.010]*	[0.186]	[0.003]	[0.1[illegible]3]
T*99N	0.063	0.611	0.001	0.89	0.035	0.053	0.003	0.045	0.026	0.01	0	-0.006	0.021	0.558	-0.002	-0.1[illegible]2
	[0.022]***	[0.246]**	[0.011]	[0.320]***	[0.021]*	[0.106]	[0.004]	[0.166]	[0.011]**	[0.038]	[0.014]	[0.033]	[0.012]*	[0.177]***	[0.002]	[0.106]
Obs.	46199	29900	46196	5495	46197	16724	46198	594	46196	6406	46193	8487	46195	6347	46189	814

Note: Odd columns are estimates from probits on likelihood of holding any animal. Even columns are estimates from OLS regressions for households holding at least one animal. Standard errors in [brackets] clustered at the village level. ***, **, * indicates significance at the 1, 5, 10 % level respectively.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] Abadie, A., Angrist, J. and imbens, G., (2002), "Instrumental Variables Estimates of the Effect of Subsidized Training on the Quantiles of Trainee Earnings", Econometrica, 70(1): 91-117.

[2] Albarran, Pedro and Orazio Attanasio. 2002. "Do Public Transfers Crowd out Private Transfers? Evidence from a Randomized Experiment in Mexico". WIDER discussion paper 2002/6.

[3] Angrist, J. and Imbens, G., (1994), "Identification and Estimation of Local Average Treatment Effects", Notes and Comments, Econometrica, 62(2): 467-475.

[4] Angrist, J. and Imbens, G., (1999), "Comments on James J. Heckman, Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations", Journal of Human Resources, 34(4): 823-827.

[5] Angrist, J., Grady, K. and Imbens, G., (2000), "The Interpretation of Instrumental Variables Estimators in Simultaneous Equations Models with an Application to the Demand for Fish", Review of Economic Studies, 67(3): 499-527.

[6] Angrist, J. and Krueger, A., (1991), "Does Compulsory School Attendance Affect Schooling and Earnings", Quarterly Journal of Economics, 106(4): 9791014.

[7] Angrist, J. and Krueger, A., (2001), "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments", Journal of Economic Perspectives, 15(4): 69-85.

[8] Attanasio, O., (2000), "Consumption", Handbook of Macroeconomics, Volume 1B, Taylor, J. and M. Woodford , eds., Amsterdam: Elsevier Science.

[9] Banerjee, A., (2004), "Inequality and Investment", mimeo, MIT.

[10] Banerjee, A., Duflo, E. and Munshi, K., (2003), "The (Mis)Allocation of Capital", Journal of the European Economic Association, 1(2-3): 484-494.

[11] Berhman, J., and Todd, P. (1999), "Randomness in the Experimental Sample of Progresa (Education, Health, and Nutrition Program)", International Food Policy Research Institute, Washington, D.C.

[12] Blundell, R., (2002), "Welfare-to-Work: Which Policies Work and Why?", Keynes Lecture in Economics, Proceedings of The British Academy, 117: 477-524

[13] Blundell, R. and Costa-Dias, M., (2000), "Evaluation Methods for Non-Experimental Data", Fiscal Studies, 21(4): 427-468

[14] Blundell, R., Costa Dias, M., and Meghir, C., (2003), "Impact of Wage Subsidies: A General Equilibrium Approach", Unpublished manuscript, Institute for Fiscal Studies.

[15] Blundell, R., Costa-Dias, M., Meghir, C. and van Reenen, J., (2004), "Evaluating the Employment Impact of a Mandatory Job Search Program", Journal of the European Economic Association, 2(4): 569-606

[16] Bobonis, G. and Finan, F. (2005), "Endogenous Social Interaction Effects in School Participation in Rural Mexico", mimeo, University of California Berkeley.

[17] Boeri, T., Layard, R. and Nickell, S., (2000), "Welfare-to-Work and the Fight Against Long-term Unemployment", Department for Education and Employment, Research Report, n. 206

[18] Buse, A., (1992), "The Bias of Instrumental Variable Estimators", Econometrica, Notes and Comments, 60(1): 173-180.

[19] Calmfors, l., (1994), "Active Labor Market Policy and Unemployment - A Framework for the Analysis of Crucial Design Features", OECD Economic Studies, 22(1): 7-47.

[20] Deaton, A. (1991), "Saving and Liquidity Constraints", Econometrica, 59(5): 1221-48.

[21] Deaton, A. (1997), "The Analysis of Household Survey Data", Baltimore: Johns Hopkins University Press.

[22] De Giorgi, (2005), "Long Term Effects of a Mandatory Multistage Program: The New Deal for Young People in the UK", IFS Working Papers, W05/08.

[23] Dorsett, R., (2006), "The new deal for young people: effect on the labour market status of young men", Labour Economics, 13(3): 405-422.

[24] DWP, (2004), "Building on New Deal: Local solutions meeting individual needs", DWP paper

[25] Frölich, M., (2004), "Programme Evaluation with Multiple Treatments", Journal of Economic Surveys, 18(2): 181-224.

[26] Di Maro, V. (2004), "Evaluation of the Impact of Progresa on Nutrition: Theory, Econometric Methods and an Approach to Deriving Individual Welfare Findings from Household Data", mimeo, University College London.

[27] DiNardo, J. and Lee, D., (2004), "Economic Impacts of New Unionization on Private Sector Employers: 1984-2001", Quarterly Journal of Economics, 119(4); 1383-1441.

[28] Duflo, E. and Saez, E. (2002), "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment," unpublished working paper, MIT and University of California, Berkeley.

[29] DWP, (2004), "Building on New Deal: Local solutions meeting individual needs", DWP paper

[30] Fafchamps, M. and Lund, S., (2003), "Risk-sharing Networks in Rural Philippines", Journal of Development Economics, 71(2): 261-87.

[31] Fan, J., (1992),"Design-adaptive Nonparametric Regression", Journal of the American Statistical Association, 87(420): 998-1004

[32] Fan, J. and Gjibels, I., (1996), "Local polynomial modelling and its applications", London: Chapman and Hall

[33] Hahn, J. and Hausman, J., (2003), "Weak Instruments: Diagnosis and Cures in Empirical Econometrics", in Papers and Proceedings, American Economic Review, 93(2): 118-125.

[34] Hahn, J., Todd, P. and van der Klaauw, W., (1999), "Evaluating the Effect of an Antidiscrimination Law Using a Regression-Discontinuity Design", National Bureau of Economic Research, Inc, NBER Working Papers: 7131

[35] Hahn J. Todd P. and van der Klaauw W., (2001), "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design", Econometrica, 69(1): 201-09

[36] Hansen, C, Hausman, J. and Newey, W., (2006), "Estimation with Many Instrumental Variables", MIT-mimeo.

[37] Heckman, J., 1979, "Sample Selection Bias as a Specification Error", Econometrica, 47(1): 153-162.

[38] Heckman, J., (1997), "Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations", Journal of Human Resources, 32(3): 441-462.

[39] Heckman, J., (1999), "Instrumental Variables: Response to Angrist and Imbens", in Comments, The Journal of Human Resources, 34(4): 828-837.

[40] Heckman, J., Ichimura, H., Smith, J. and Todd, P., (1998), "Characterizing Selection Bias Using Experimental Data", Econometrica, 66(5): 1017-1098.

[41] Heckman, J., Lalonde, R. and Smith, J., (1999), "The Economics and Econometrics of Active Labor Market Programs", Handbook of Labor Economics, Volume 3, Ashenfelter, A. and D. Card, eds., Amsterdam: Elsevier Science

[42] Heckman, J., Lochner, L. and Taber, C., (1998), "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents", National Bureau of Economic Research, Inc, NBER Working Papers: 6384.

[43] Heckman, J. and Smith, J., (2004), "The Determinants of Participation in a Social Program: Evidence from a Prototypical Job Training Program", Journal of Labor Economics, 22(4): 243-298.

[44] Hoddinott, J., Skoufias, E., and Washburn, R. (2000) "The Impact of Progresa on Consumption: a Final Report". International Food Policy Research Institute, Washington, D.C.

[45] Hotz, J., Imbens, G. and Klerman, J., (2000), "The Long-Term Gains from GAIN: A Re-Analysis of the Impacts of the California GAIN Program", National Bureau of Economic Research, Inc, NBER Working Papers: 8007.

[46] Imbens, G. and Angrist, J., (1994), "Identification and Estimation of Local Average Treatment Effects", Econometrica, 62(2): 467-475

[47] Katz, L., (1998), "Wage Subsidies for the Disadvantaged", in R. Freeman and P. Gottschalk, eds., Generating Jobs: How to Increase Demand for Less-Skilled Workers, Russell Sage, 21-53

[48] Katz, L., J. Kling, and J. Liebman (2001): "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment," Quarterly Journal of Economics, 116: 607-54.

[49] Kleibergen, F., (2002), "Pivotal Statistics for Testing Structural Parameters in Instrumental Variables Regression", Econometrica, 70(5): 1781-1803.

[50] Lalive, R. and Cattaneo, A., (2005), "Social Interactions and Schooling Decisions", Mimeo, University of Zurich.

[51] Lechner, M., (1999), "Identification and Estimation of Causal Effects of Multiple Treatments Under the Conditional Independence Assumption", IZA Discussion Papers 91, Institute for the Study of Labor (IZA).

[52] Lim, Y. and Townsend, R. (1998), "General Equilibrium Models of Financial Systems: Theory and Measurement in Village Economies", Review of Economic Dynamics, 1(1): 59-118.

[53] Lise, J., Seitz, S. and Smith, J., (2005a), "Evaluating search and matching models using experimental data", mimeo, University of Michigan.

[54] Lise, J., Seitz, S. and Smith, J., (2005b), "Equilibrium Policy Experiments and the Evaluation of Social Programs", mimeo, University of Michigan.

[55] Miguel, E. and Kremer, M. (2004),"Worms: identifying impacts on education and health in the presence of treatment externalities," Econometrica, 72(1): 159-217.

[56] Moreira, M., (2003), "A Conditional Likelihood Ratio Test for Structural Models", Econometrica, 71(4): 1027-1048.

[57] Mundell, R., (1968), "International Economics", New York: Macmillan, 1968, pp. 85-99.

[58] Munshi, K. and Rosenzweig, M., (2005),"Why is Mobility in India so Low? Social Insurance, Inequality, and Growth", CID Working Paper No. 121, Harvard.

[59] Parker, S. and Skoufias, E., (2000), "Final Report: The Impact of PROGRESA on Work, Leisure, and Time Allocation", International Food Policy Research Institute, Washington.

[60] Philipson, T, (2000), "External Treatment Effects and Program Implementation Bias", National Bureau of Economic Research, Inc, NBER Technical Working Papers: 250.

[61] Porter, J., (2003), "Estimation in the Regression Discontinuity Model", Mimeo

[62] Rosenzweig, M. (1988a), "Risk, Private Information, and the Family," American Economic Review, 78(2): 245-50.

[63] Rosenzweig, M. (1988b), "Risk, Implicit Contracts and the Family in Rural Areas of Low-Income Countries", The Economic Journal, 98(393): 1148-70.

[64] Rubin, D., (1980), "Discussion of Randomization Analysis of Experimental Data: The Fisher Randomization Test by D.Basu", Journal of the American Statistical Association, 75: 591-93.

[65] Rubin, D, (1986), "Which ifs have causal answers? Discussion of Hollands Statistics and causal inference", Journal of the American Statistical Association, 81: 961-62.

[66] Rubin, D., (2005), "Causal Inference Using Potential Outcomes: Design, Modeling, Decisions", Journal of the American Statistical Association, 100(469): 322-31.

[67] Ruppert, D., Sheather, S. J. and Wand, M. P., (1995), "An Effective Bandwidth Selector for Local Least Squares Regression", Journal of the American Statistical Association, 90(432): 1257–1270

[68] Sacerdote, B. (2001), "Peer Effects with Random Assignment: Results for Dartmouth Roommates," Quarterly Journal of Economics, 116: 681-704.

[69] Schultz, P., (2001), "School Subsidies for the Poor: Evaluating the Mexican PROGRESA Poverty Program", forthcoming in Journal of Development Economics.

[70] Shea, J., (1997), "Instrument Relevance in Multivariate Linear Models: A Simple Measure", in Notes, The Review of Economics and Statistics, 79(2): 348-352.

[71] Sianesi, B., (2001), "Swedish Active Labour Market Programmes in the 1990s: Overall Effectiveness and Differential Performance", Swedish Economic Policy Review, 8(2): 133-169.

[72] Sianesi, B., (2004), "An evaluation of the Swedish system of active labour market programmes in the 1990s", Review of Economics and Statistics, 86(1): 133-155.

[73] Skoufias, E., Davis, B. and Behrman, J., (1999a), "Final Report: An Evaluation of the Selection of Beneficiary Households in the Education, Health, and Nutrition Program (PROGRESA) of Mexico", International Food Policy Research Institute, Washington, D.C.

[74] Skoufias, E., Davis, B. and de la Vega, S., (1999b), "An Addendum to the Final Report: An Evaluation of the Selection of Beneficiary Households in the Education, Health, and Nutrition Program (PROGRESA) of Mexico. Targeting the Poor in Mexico: Evaluation of the Selection of Beneficiary Households into PROGRESA", International Food Policy Research Institute, Washington, D.C.

[75] Smith, J., Whalley, A. and Wilcox, N., (2006), "Are Program Participants Good Evaluators?", mimeo, University of California-Merced.

[76] Staiger, D. and Stock, J., (1997), "Instrumental Variables Regression with Weak Instruments", Econometrica, 65(3): 557-586.

[77] Stock, J. and Yogo, M., (2003), Testing for Weak Instruments in Linear IV Regression", mimeo, Harvard University.

[78] Thistlethwaite, D. and Campbell, D.(1960), "Regressiondiscontinuity analysis: an alternative to the ex post facto experiment", Journal of Educational Psychology, 51: 309-17

[79] Townsend, R., (1994), "Risk and Insurance in Village India", Econometrica, 62(3): 539-591.

[80] Townsend, R., (1995a), "Consumption Insurance: An Evaluation of Risk-Bearing Systems in Low-Income Economies", The Journal of Economic Perspectives, 9(3): 83-102.

[81] Townsend, R., (1995b), "Financial Systems in Northern Thai Villages," The Quarterly Journal of Economics, 110(4): 1011-46.

[82] Udry, C. (1994), "Risk and Insurance in a Rural Credit Market: An Empirical Investigation in Northern Nigeria," Review of Economic Studies, 61(3): 495-526.

[83] Udry, C. (1995), "Risk and Saving in Northern Nigeria", American Economic Review, 85(5): 1287-1300.

[84] van den Berg, G., van der Klaauw, B. and van Ours, J., (2004), "Punitive Sanctions and the Transition Rate from Welfare to Work", Journal of Labor Economics, 22(1): 211-41

[85] Van Reenen, J., (2003), "Active Labor Market Policies and the British New Deal for Unemployed Youth in Context", in R. Blundell, D. Card and R. Freeman, eds., Seeking a Premier League Economy, University of Chicago Press, 461-96

[86] Woodfield, K., Bruce, S. and Ritchie, J., (2000), "New Deal for Young People: the National Options. Findings from a Qualitative Study Amongst Individuals", Employment Service, Research and Development Report, ESR37