Dynamic Models for Policy Evaluation

Costas Meghir*
University College London and IFS

March 16, 2006

Abstract

The evaluation of interventions has become a commonly used policy tool, which is frequently adopted to improve the transparency and effectiveness of public policy. However, evaluation methods based on comparing treatment and control groups in small scale trials are not capable of providing a complete picture of the likely effects of a policy and do not provide a framework which allows issues relating to the design of the programme to be addressed. The longer term effects relate to decisions by individuals to change aspects of their life-cycle behavior not directly targeted by the intervention, so as to best take into account of its presence. They also relate to possible changes in prices that may change or even reverse the incentives designed by the programme. In this paper we show how experimental data from field trials can be used to enhance the evaluation of interventions and we also illustrate the potential importance of allowing for longer term incentive and General Equilibrium effects.

1 Introduction

The evaluation of interventions has become a commonly used policy tool. This has led to a better understanding of the overall benefits and costs of interventions, and of their distributional impact. It has also improved transparency in policy making and lead to choices that are founded in fact rather than political prejudice. Recent examples of policy-influential evaluations are the Job Training

*Acknowledgements I thank my co-authors, Orazio Attanasio, Richard Blundell, Monica Costa-Dias, Giovanni Gallipoli, Ana Santiago, and Gianluca Violante for allowing me to use material from our joint work as well as for stimulating discussions. I also thank Joe Altonji for his comments at the World Congress in London as well as Ken Judd and seminar participants at the Cowles lunchtime seminar in Yale.

Partnership act in the USA (see Heckman, Ichimura, and Todd (1997)) or the evaluation of the Education Maintenance allowance in the UK - a conditional cash transfer, incentivising 16 year old children to stay on at school in the UK (see Dearden, Emmerson, Frayne, and Meghir (2005)), the PROGRESA programme in Mexico, discussed below and many others. A large body of research has further stimulated its use and has addressed important intellectual challenges[1]. The standard basis for evaluations has been field trials where a policy is implemented in a small scale and its effects measured by comparing to a suitably chosen comparison group. The gold standard in this approach is a randomized trial where the unit of randomization may by an individual or a community, however defined. Other than that, a large number of alternative methods have been developed, or adapted from standard econometric techniques, based on observational data. These include matching and quasi-experimental methods such as Instrumental Variables and Difference in Differences[2]. However there are many reasons why this set of approaches can only offer limited answers to the question of policy design and need to be complemented with further methods of analysis. First, long term implementation of a programme may induce changes in behavior by individuals that cannot be measured by a simple comparison of the treated and control group in the short run. Thus a programme that subsidizes wages may indeed increase employment of the client group relative to a comparison grouping in some isolated labor market; however it may also change the incentives to invest in human capital (see Heckman, Lochner, and Cossa (2003) and Adda, Dustmann, Meghir, and Robin (2005)) an effect

[1] Rosenbaum and Rubin (1983), LaLonde (1986), Heckman, Ichimura, and Todd, (1997, 1998), Heckman and Robb (1985), Imbens and Angrist (1994) to name but a few.

[2] Effectively an instrumental variables method in itself.

that will occur a few years down the road. Second, the programme may have more wide-ranging impacts affecting individuals who are not explicitly targeted by the policy, including changing behavior to become a member of the targeted group and thus obtain the benefits of the programme. In certain cases, field trials can be designed to measure such effects, as distinct from the main effects of the programme; this is however difficult in practice and requires a large amount of resources to be expended. Thirdly, while a small field trial may not affect prices because the changes in supply and demand of say human capital that it induces are negligible, the same cannot be said of a broad implementation of a successful programme (see Heckman, Lochner, and Taber (1998a)). In this case the effects of the programme can be seriously mitigated to the extent that they are almost neutralized.

There are two main themes that we address in this paper. First, we discuss the use of randomized field experiments to estimate structural economic models that can then be used to improve on the design of interventions. We argue that this approach reinforces the usefulness of experiments in that it allows, given the model assumptions, to go beyond the simple conclusions that can be drawn by a comparison of means. Moreover, economic models can be used to guide the way we design experiments in the first place, so as to get most out of them and be able to identify important aspects of behavior. In this sense we pursue the ideal set forth byOrcutt and Orcutt (1968) where experiments on incentives are advocates are used to learn about structural parameters in an economic model. Thus, our first example draws from the paper by Attanasio, Meghir, and Santiago (2005) where data from a field trial in Mexico (PROGRESA) was combined with a structural model of education choice to produce a model that at the same time challenges the results obtained by conventional observational data and provides a framework for thinking of better ways to redesign the programme or compare with alternative policies.

The next theme of the paper addresses longer term evaluation questions that cannot be dealt with based on field experiments. Two issues arise requiring structural economic models. First, interventions may have longer run incentive effects along dimensions not intended. For example a programme designed to boost employment by wage subsidies may reduce human capital accumulation. This issue has been addressed empirically by Heckman, Lochner, and Cossa (2003), Blundell, Costa-Dias, and Meghir (2003) and more recently by Adda, Dustmann, Meghir, and Robin (2005). These papers consider different forms of wage subsidy interventions and show that these can affect substantially Human Capital accumulation. Second, as discussed in Heckman, LaLonde, and Smith (1999) and illustrated empirically in Heckman, Lochner, and Taber, (1998b), large scale interventions may have important consequences for prices and hence for outcomes, at least in some settings where factor price equalization internationally is somehow prevented. In this case these General Equilibrium (GE) effects can be very large and often neutralize or sometimes reinforce the effects of policy.[3] More recently the notion of using empirical GE models to evaluate policy interventions and to understand trends in the labor market is becoming more prevalent. Examples of such work include Alonso-Borrego, Fernndez-Villaverde, and Galdon-Sanchez (2004) who look at the impact of short term employment contracts and find that they decrease employment and Lee and Wolpin (2005) who use a GE model to understand the evolution of wages and employment in

[3]For an example of the latter see Gallipoli and Fella (2005).

the services sector.

Here we examine the potential importance of dynamic incentive and General Equilibrium effects by summarizing the results of two papers. The first is the paper by Gallipoli, Meghir, and Violante (2005) and the second by Blundell, Costa-Dias, and Meghir (2003). In these models risk-averse life-cycle consumers choose education, labor supply and consumption in a world with uninsurable idiosyncratic risk. The economy is stationary and consists of overlapping generations. Output is produced by a production function which is constant returns to scale in capital and three types of labor characterized by the level of formal education – less than high school, high school and some College. The models are capable of addressing the issues of dynamic incentives and endogenous prices. They thus allow for interactions between individuals due to a number of mechanisms: The government budget constraint, used to fund the intervention, the dynamic incentives created by the programme and the endogeneity of prices.[4] Using these models we consider the impact of tuition subsidies and of temporary wage subsidies, similar to those offered by the UK New Deal - an active labor market programme introduced in 1998.

2 School subsidies and Education choices in Mexico

Our first example is based on the paper by Attanasio, Meghir, and Santiago (2005) and illustrates the use of experimental data for fitting an economic model, with which it is then possible to address questions of design in a partial equi-

[4]There is a real issue here relating to the possibility of factor price equalization. This is a controversial topic we have not addressed. In our illustrations we take prices as endogenous. However, even if prices were exogenous the other issues relating to dynamic incentives over the life-cycle would remain.

librium framework.

2.1 The PROGRESA program and the evaluation data.

In 1997, the Mexican government started a large program to reduce poverty in rural Mexico, by focussing on increasing education using subsidies ("conditional cash transfers") and improving health and nutrition. The programme is known as PROGRESA[5] was first implemented in a randomly chosen set of village communities out of a population of eligible ones; for the purposes of evaluation. Data was collected both from the communities that were randomized into the programme and those randomized out. The programme was implemented in all communities with a delay of 1.5 years. Greater details of the programme can be found in Schultz (2003) and Attanasio, Meghir, and Santiago (2005).

Once a locality qualifies for the programme, individual households within it could qualify or not, depending on a single indicator of poverty. The largest component of the program is the education one. Beneficiary households with school age children receive grants conditional on school attendance. The size of the grant increases with the grade and, for secondary education, is slightly higher for girls than for boys. In Table 1, we report the grant structure as it was at the start of the programme. All the figures are in current pesos, and can be converted in US dollars at approximately an exchange rate of 10 pesos per dollar. To keep the grant, children have to attend at least 85% of classes. Upon not passing a grade, a child is still entitled to the grant for the same grade. However, if the child fails the grade again, it looses eligibility.[6] In addition to the bimonthly payments, beneficiaries with children in school age receive a small

[5] The Spanish acronym for 'Health, Nutrition and Education'

[6] All the figures are in current pesos, and can be converted in US dollars at approximately an exchange rate of 10 pesos per dollar.

PROGRESA bimonthly monetary benefits	
Type of benefit	1998
Nutrition support	190
Primary school	
3	130
4	150
5	190
6	260
secondary school	
1st year	
boys	380
girls	400
2nd year	
boys	400
girls	440
3rd year	
boys	420
girls	480
maximum support	1,170

Table 1: Table Caption

annual grant for school supplies.

The randomized trial, offers the opportunity to obtain unbiased estimates of the average impact of the programme on any desired outcome non-parametrically, simply by comparing the estimated mean outcomes between treatment and control villages. There results can also be broken down by characteristics that are exogenous or determined pre-programme. However, there are a number of intricate aspects to the design of the programme or deeper questions that cannot be answered directly by the randomized trial. For example, the policymakers in increasing the grant by grade are responding to the fact that the school drop out rates increase quite dramatically during the teenage years. But whether the best use of funds has been made in this direction cannot be addressed by the randomized trial. While it would be possible and scientifically desirable to design a field trial with many directions of variation by age this is not practical.

To proceed further an economic model can be fitted to the observed responses by combining the exogenous source of variation induced by the experiment by other sources of variation that can be considered conditionally exogenous. This combines the experiment with reasonable assumptions to further enhance our understanding and the scope of the data. Within this context it will also be possible to test whether the experimental data would lead to the same conclusions as those obtained by a model estimated solely on observational data.

The analysis is based on the evaluation data whose collection began in 1997 and following the start of the in 1998. This is a household based panel, which collects information on schooling, work patterns, earnings and income as well as on consumption and household assets.

Table 2 provides the difference in differences estimates of the effects of the policy for the eligible individuals drawn from Attanasio, Meghir, and Santiago (2005). These are obtained by comparing the growth of school attendance between baseline and after the programme started in treatment and control villages. While a simple post-programme comparison between treatment and control would have provided unbiased estimate, in this context the difference in differences improves precision. The key conclusion from this table, which refers to boys is that the programme has little or no impact on young children, since they nearly all attend in the firs place. However, it has a substantial impact on the 14-17 age group. To go beyond this result to questions relating to the design of the programme we need a model. or alternatively a much more elaborate experiment.

Post-program differences in educational attendance between treatment and control communities	
age group	eligible
6-17	0.034
	(0.012)
6-9	0.003
	(0.005)
10-13	0.032
	(0.011)
14-17	0.084
	(0.031)
Standard errors in parentheses are clustered at the locality level. Source: Attanasio et al (2005)	

Table 2: Experimental Results October 1998

2.2 The model and its estimation

We consider the choice of attending school versus work for boys from age 11 to 16. The trade-off is between current earnings in the labor market and future increased earnings. These are set against other schooling costs, such as travel, equipment etc. The choice takes place in an environment of idiosyncratic risk: First, the child may not pass the grade and will have to repeat; second the costs of schooling may change - we model this latter aspect as a shock to preferences. The monetary opportunity cost of schooling is the local wage. We assume that children have the possibility of attending school up to age 17. All formal schooling ends by that time. In our model schooling is taken as an investment. The parents are assumed to act in the best interest of the individual child and hence do not trade-off welfare between children.

When the grant is offered conditional on attending school it should have the same effect as a reduction in schooling costs or alternatively a reduction in the wage that the individual would earn if she worked. The implication is that it should be possible to estimate the effect of the grant on school participation by learning about the sensitivity of individual schooling choices to wages. This is the traditional approach to public policy analysis and dates back at least to Marschak and Andrews (1944). For example, learning about the implication of a tax rate on demand can be achieved by looking for a price increase under similar conditions at some other historical moment or point in space. For the PROGRESA case Todd and Wolpin (2003) use this idea to examine whether a model estimated on non-experimental data and relying on wage variation can predict the impact of the policy; the latter can be estimated in an unbiased way because of the experimental design.

We take a different tack in Attanasio, Meghir, and Santiago (2005). We specify a dynamic model of education choice and include the grant as one of the current benefits of schooling. Subject to suitable scaling we can then test whether the wage has the same effect (but in opposite direction) to the grant. However, our model is more than just a testing vehicle. It provides a framework to simulate changes in the design that would improve targeting and further increase school participation, at the same overall cost.

The heart of the model is the schooling flow utility function which we specify to be a function of school availability in the village, distance from school and direct costs of schooling. These variables either determine the cost of attending school or are direct components of the cost, such as fees, shoes etc. In addition we condition on parental education and ethnic background, to capture indirect or "psychic" costs of schooling that may be driven by the support that the child gets at home. Finally we condition on whether the household is programme eligible in the sense of being below the poverty threshold defined by the intervention, as well as living in an eligible village. These variables account for permanent differences between eligible and ineligible households allowing for the human capital differences that could occur because of different income levels. They also control for pre-programme differences that were identified among the ineligible households. Going to school may build a habit and a taste for this, as pupils learn to learn and create social networks. We thus include the past level of attainment as an additional component in the flow utility. Beneficiary households receive the grant, which is also included among the variables - its value for boys varies by grade. The final component on the flow utility of schooling is a an unobserved component. This consists of a permanent effect and an iid shock. The former reflects unobserved tastes for schooling or factors such as risk aversion that are not explicitly modelled here, since we do not include consumption in a nonlinear way. The latter is an iid stochastic shock to schooling costs that can be interpreted as transitory factors affecting attendance, such as health for instance. Finally, when in work the flow utility is proportional to the wage he can earn.

The decision to attend school depends on comparing the expected discounted flows under the two alternatives, allowing for optimal decisions in the future. In both cases we assume that beyond 17 the individual will be working and earning wages commensurate to their educational attainment. Before that, individuals either obtain the (possibly negative) utility cost of schooling when they attend or earnings from the labor market, depending on their age and qualification. In simple within village regressions the returns to schooling are as low as 1% for a year of education. However the relevant return is the one obtained when working in the urban centres. This is of the order of 10%. In other words, individuals who obtain education are probably committing themselves to moving out to seek better work opportunities.

To model schooling choice we need a forward looking dynamic model. The model we consider is dynamic for three reasons. First, the fact that one cannot attend regular school past age 17 means that going to school now provides the option of completing some grades in the future. This source of dynamics becomes particularly important when we consider the impact of the PROGRESA grants, because to take full advantage of the subsidy one has to attend continuously and pass the grade. Second, we allow for state dependence: The number of years of schooling affects the utility of attending in this period. State dependence is important because it may be a mechanism that reinforces the effect of the grant. Third education choices when young affect future earnings and probably mobility choices.

The model contains two sources of uncertainty. We have already mentioned the iid shock to the cost of schooling. This implies that the costs of schooling are not known in future periods and consequently the individual has to take into account that in future these may be higher or lower. The other source of uncertainty originates from the that one may not pass a grade. This may affect the value of attending. In our model we do not allow for an endogenous supply of effort to affect the probability of passing. While this is a potentially important source of programme impact we need to leave this for future work; this will have to allow for the fact that the composition of those attending and hence whose grade performance is observed are different in the treatment and control group.

Solving the model involves computing the value of the two alternatives -

school and work, given the state variables, namely the current level of education, and the wage and given the set of observed and unobserved characteristics (at each point of the support of their distribution). This solution assumes future optimal decisions and accounts for the impact of current choices on future choices. Solving this model is straightforward because of the logistic assumption on the iid school cost shock.

The randomized availability of the programme provides an important source of exogenous variation for the identification of the structural economic model and should increase our confidence in using it. While this is an important source of variation is not sufficient to identify all key parameters.[7] First, the model includes the actual amount of the grant, a quantity that does not vary randomly. Second the model also includes the wage for the child which reflects the opportunity cost of schooling.

To identify the impact of the grant we use the way it is designed to vary with the grade. The lack of perfect correlation of grade and age generates the required variation. Wages do vary from locality to locality and this variation is used to identify the wage effect on schooling. This requires one to assume that the variability is due to different labor demand conditions, rather than changes in labor supply induced by unobserved factors. In addition we use the adult wage as an instrument for predicting wages. This controls for measurement error and for missing wages caused by the choice to attend school.[8]

The experiment plays an important role in identifying the model. However, it is also true that identification would benefit from a richer experimental field

[7] See Attanasio, Meghir, and Santiago (2005) for a full discussion of identification issues

[8] There is also an important initial conditions problem which causes an identification problem. This issue and its solution is discussed in Attanasio, Meghir, and Santiago (2005)

trial. For example randomizing the amounts received across villages as well as randomizing the age gradient of the grant would have been particularly useful. In this sense thinking more broadly about the information that we would wish to extract from an experiment, other than the simple quantification of a particular design is very important when coming up with the design of an evaluation. It is thus possible to come up both with a richer set of experimental results and create the conditions for estimating economic theory based models with fewer statistical and functional form assumptions, which tend to have no foundation in economic theory.

The model we described can be estimated quite straightforwardly using maximum likelihood. The likelihood function is based on the distribution of ε which is assumed to be logistic. Unobserved heterogeneity is then integrated out using a discrete mixture as in Heckman and Singer (1984).

2.3 Results and policy design

The estimated model fits very well the experimental results, which is quite remarkable for such a parsimonious specification. However, its value lies in its ability to guide policy choices. In fact it provides a tool both for redesigning the grant structure to obtain stronger attendance results and possibly to compare the relative merits of such a conditional cash transfer to other infrastructure-type policies; this would exploit the school availability variables and distance to school, both of which are included.

A further potentially important contribution of such a is in its ability to test whether the evaluation results of the experiment were similar to those that would be predicted using observational data and relying on wage variation alone. It turns out that the effect of the grant is three times higher than the wage, when this is scaled up to represent lost earnings from school attendance. Although this points to the importance of running the experiment and collecting the evaluation data base, one should not dismiss the underlying economic model on the basis of this finding: First wages may not be measured well enough leading to an attenuating effect. Beyond this simple statistical explanation the effect of the grant and of earnings may have different effects because of who receives each of them. Thus intrahousehold allocations may play an important role here, implying that the source of income and to whom it is paid may have important effects in itself. This point emphasizes the usefulness of the field trial; even within the context of a structural model the information obtained from the trial can be of critical importance in understanding how the policy works.

Using these estimates we can also address the question of redesigning the programme. The grant is provided to all children over the age of 6. Our model fits behavior from 10 onwards. The impact of the grant is very small at earlier ages. In this policy experiment we consider redesigning the grant so as to make it zero up to and including age 11 and redistribute the grant equally beyond that date, keeping the overall cost of the programme the same. We show that this simple redesign can almost double the effect of the policy on participation at 15 and even more than double them at 16 - all this at no extra financial cost (see Attanasio, Meghir, and Santiago (2005)). This simulation takes as sole aim of the policy to get children into school. It may well be that the grant before 12 was considered a pure transfer, conveniently handed out via the school, since most children attend anyway. However, if school participation is the sole intention, and because the programme has no effect on younger children, an properly targeted unconditional transfer for this group may be more efficient at alleviating poverty.

Our aim in this section has been to illustrate the power of combining experimental data with economic modelling. We illustrated how this can help understand potential failings when using observational data and how the model can enhance the use of the experiment in a policy context. We gave one example, but with the possibility of comparing this policy to alternative options, such as the use of infrastructure the scope of such an approach is very broad. We now proceed to consider wider issues with General Equilibrium models.

3 Policy Interventions, Long Term Incentives and General Equilibrium

The example we detailed in the previous section did not consider longer term incentives of the programme and was set in a partial equilibrium context where prices remain constant. In the longer run a number of changes can occur that will alter the impact of the programme in important ways.

First, individuals may change their work and/or consumption behavior to become eligible for the benefits. This point is well understood in economics and there is a vast literature on the work-disincentive effects of welfare benefits. These disincentives may not be important enough for one to want to avoid implementing a programme. After all, the programme is in place partly to alleviate a perceived market imperfection that prevents (in the PROGRESA case) children from obtaining valuable education. However their magnitude need to be measured if the programme is to be designed to achieve effectively its aims.

Second, beyond these longer term incentive effects large scale programmes that successfully alter the supply of human capital may well affect prices. Thus an increase in educated individuals may compress educational differentials, or an increase in re-employment rates may put a downward pressure on wages. This in itself may partly overturn the effects of the programme as estimated from some limited field trial. This point has been eloquently made by Heckman, LaLonde, and Smith (1999); Heckman, Lochner, and Taber (1998a) and Lee (2001) have provided a quantification of the magnitude of the effects in quite different contexts. These papers show quite clearly that results can differ quite substantively, depending on the particular set up of the model and the choices faced by individuals. They also illustrate the potential importance for policy of allowing for such effects. In many cases the GE effects are lower than the partial equilibrium ones. However this need not be the case. Thus, Gallipoli and Fella (2005) show that the impact of an education subsidy on crime is reinforced in General Equilibrium with substantially larger effects.

Considering the general equilibrium effects of a policy is necessarily much more complicated because one has to take a stand on a number of important issues, such as market structure for which the evidence may be scant. In GE models the trade-off between the richness of the model and its tractability can be even more stark than usual. First, comes the problem of having available data on all the required dimensions in a compatible way. Second, comes the issue of defining the way agents interact, which broadly speaking relates to defining the market structure as well as the role of the government in funding policy interventions. Thirdly comes the issue of potentially dealing with macroeconomic shocks. Inevitably compromises have to be made.

We illustrate the issues involved in evaluating policy using two similar General Equilibrium models chosen to fit US and UK data respectively. In both cases individuals choose labor supply, consumption and education. They face uncertainty on the returns to education. The models are of the overlapping generations type and only steady state equilibria are considered. In other words questions relating to transitional dynamics are not addressed. In the first case the focus of the model is on analyzing human capital policies such as tuition subsidies. The model and results are drawn from Gallipoli, Meghir, and Violante (2005) and we use this model to illustrate the potential issues of introducing tuition subsidies. The main question is how effective can a policy be, funded by proportional taxation of earnings, in encouraging College education and hence increasing overall education levels and earnings.. In the second example we consider a policy affecting individuals later in life, namely a wage subsidy available to individuals at an age following the age when most people would have obtained College education. The results we present are taken from Blundell, Costa-Dias, and Meghir (2003) and illustrate not only the importance of allowing for price changes but also the importance of dynamic incentive effects: individuals are shown to change their early education choices in response to the future availability of the temporary subsidy. This illustrates, how a programme designed to reduce unemployment can in the long run have opposite effects once dynamic incentive and price changes combine. The motivation is to consider issues relating to the New Deal - an active labor market programme in the UK designed to help the long term unemployed back to work. To economize in space we base our description on the model developed by Gallipoli, Meghir, and Violante (2005); when required we highlight some key differences between the models..

3.1 The model

We consider a closed economy where a unique good is produced, and it can be either consumed or used as physical capital. Our specification effectively precludes factor price equalization and hence both the interest rate and human capital prices are taken as endogenous to the economy. This is a contentious and important issue that can affect policy in dramatic ways because when prices are exogenous the GE effects we discuss will not take place. However the individual dynamic incentives will still be altered by the policy.

We specify an overlapping generations general equilibrium model. Consumers maximize an intertemporal utility function over their finite life-cycle, with respect to education, labor supply and consumption/savings. Agents can accumulate assets representing ownership of shares on physical capital. They have a maximum lifetime and they plan to consume their entire assets. However they may die before that, leaving accidental bequests. The maximum possible lifetime is 99 years. Individuals can work up to 65 but not beyond. They can however decide not to work before that. Retirement is financed by the accumulated assets. The population consists of 99 overlapping generations ex-ante heterogeneous agents, each with an ex-ante identical distribution of heterogeneity.

Young and old households are not linked in any direct way. Bequests are pooled together and redistributed to all newly born individuals according to the steady state equilibrium wealth distribution. This reflects both inter-vivo transfers for education and actual bequests. We now provide a brief description of the components of the model.

The Individual problem Individuals can live up to a maximum age of 99, but can die earlier with probability given by the US life tables. They maximize an intertemporal utility function of consumption and leisure, additive over time but nonadditive between consumption and leisure. They discount the future at a personal discount factor β. Time can be used for education, work or leisure up to age 22 and only for work or leisure after that age. the individual can choose between education and work and leisure. When in education individuals have to pay an annual fee; subsidizing this fee will be the policy instrument which we will examine in this paper. Individuals differ by an ability factor θ. This affects the utility cost of education, by changing the time input required for obtaining a particular qualification (high school or College). This same factor also affects the wage the individual can earn in the labor market. Individuals also choose consumption in each period and can invest in a riskless asset at a rate r, which is endogenous to the model. Assets have to be zero at the maximal age of 99. However, individuals may die accidentally, leaving positive or negative assets. As explained above these are redistributed to those entering the economy. The individual faces uncertainty because of shocks to human capital that are not insurable; these are described below.

The within period utility function is assumed to take the isoelastic weakly separable form

$$u\left(c_j, l_j \mid \mathrm{d}_j = 0\right) = \frac{\left[c_j^{\nu} l_j^{1-\nu}\right]^{(1-\lambda)}}{1-\lambda}$$

$$u\left(c_j \mid \mathrm{d}_j = 1,\ l_j = f^e\left(\theta\right)\right) = \frac{\left[c_j^{\nu} f^e(\theta)^{1-\nu}\right]^{(1-\lambda)}}{1-\lambda}$$

The individual problem is solved recursively by backwards induction. This results in education, consumption and labor supply choices all as a function of unobserved heterogeneity, human capital prices, the interest rate and the state variables - which include the individual amount of human capital and wealth. as well as consumption/savings decisions. In the model simulations we present here we have set the intertemporal elasticity of consumption to 0.75 as in Blundell, Browning, and Meghir (1994) and Attanasio and Weber (1993). Given this, a value of $\nu = 0.33$ and hence $\lambda = 2$ leads to a solution that matches the labor supply data very well.

3.2 The wage process

A key element of the model is the wage process because it drives both the incentives to obtain education and the uncertainty faced by individuals. The model we use is in effect the Roy model where the sector is identified with level of education and is similar in nature to the model of Heckman and Sedlacec (1985). The wage equations are stochastic and the shock to wages is the main source of uncertainty in the model. For this we use the well established specification as in MaCurdy (1981), Abowd and Card (1989) and Meghir and Pistaferri (2004). Thus, the empirical specification we estimate on the PSID data has the form

$$\ln w_{eit} = w_{et} + g_e\left(age_{eit}\right) + u_{eit} \tag{1}$$

where w_{et} represents the log of the aggregate price of human capital for education group e and where $g_e\left(age_{eit}\right)$ is the education specific profile of wages. The unobservable component u_{eit} is specified to be

$$\begin{array}{c} u_{eit} = z_{eit} + m_{it} \\ z_{eit} = z_{eit-1} + \varepsilon_{eit} \\ z_{ei0} = \theta_i \end{array} \tag{2}$$

where θ_i represents the effect of ability on initial wages. Consistent with the literature z_{eit} is the (persistent) component of unobserved human capital, which evolves as a random walk. The shock to wages ε_{eit} is drawn from a normal distribution with education specific variance. Finally m_{it} represents measurement error, is assumed *iid* and does not affect behavior. Thus the model allows for one more important channel for education choice, namely that of insurance. Since individuals are risk averse, they will take into account the different variance of the shocks when making education decisions.

The wage process is estimated using the PSID. The variance of wage shocks and of the measurement error are estimated using the variances and autocovariances of residual wage growth after taking out time and age effects based on the PSID data. The approach followed is based on Meghir and Pistaferri (2004). The main difficulty that needs to be addressed however, is estimating the impact of ability on wages. This is important because it will allow education choice to depend on labor market ability. To deal with this we use an IQ test administered by the PSID in 1972. We thus regress wages on an age polynomial and the IQ score, separately for each education group to obtain the effect of the IQ test on wages. This is then used when solving the individual's problem.

3.3 Uncertainty or unobservability

An issue of central importance in such models is the real degree of uncertainty. Measured wages vary; however there is no reason why this variability should represent uncertainty, because it may well be that the individual anticipates much of this variability, or because it just reflects measurement error. The model described here strips out measurement error from uncertainty. It also ignores any transitory shocks. Finally, uncertainty is taken to be the set of future innovations to the permanent component of wages. There are two questions: First, what is the extent to which such uncertainty, due to the wage process, is insurable? Second, when does the "innovation" become known to the individual? These are very hard issues to resolve; papers that have gone in that direction are Blundell and Preston (1998), Blundell, Pistaferri, and Preston (2005) and in the context of education choices, Cunha, Heckman, and Navarro (2004). Identifying what is in the information set and what is not at the time of a decision is of course a very hard task which requires strong identifying restrictions. In the present paper we assume that the variance of the permanent shocks to wages reflect uncertainty. This shock explains just a fraction of the observed conditional cross sectional variance. This relates to the accumulation of past shocks, already in the information set, as well as measurement error.

3.4 Production Structure

As far as the effectiveness of human capital policy is concerned the structure of the production function is of central importance. In one extreme all types of education can be perfect substitutes in the production function and the only pertinent difference between individuals in this respect is amount of efficiency units of HC that they possess following their education and given their ability θ. In this case, increasing the supply of say College graduates increases overall human capital but does not affect the relative prices across types and hence there are no GE effects. The other extreme is a Leontieff type technology where GE effects will lead to full crowding out of any partial equilibrium effects since the increase in one input has no impact on production unless the other inputs increase by the right proportions. Thus reliable estimates of the production technology are central to understanding the policy impact of an intervention. We thus specify the production function to be

$$Y = K^{\alpha}\left[(\delta_1 H_1^{\rho} + \delta_2 H_2^{\rho} + (1-\delta_1-\delta_2)H_2^{\rho})\right]^{\frac{1-\alpha}{\rho}}$$

In principle one could allow different elasticities of substitution between the various skill groups. However, on US data we found these to be of very similar magnitude and we could not reject they were in fact equal. In the UK paper on wage subsidies we use a more general specification as described below. Details on estimation are provided in Gallipoli, Meghir, and Violante (2005) and draw from Heckman, Lochner, and Taber (1998a). The key issue is that we do not observed human capital aggregates corresponding to the types of education in the model. Thus we use the relative price series obtained from estimating the wage equation on the PSID to compute efficiency units of human capital for each individual in the CPS, based on their observed earnings. We then aggregate these quantities by education type and use NIPA data to fit the production function using GMM.

The resulting elasticity of substitution based on an estimated ρ of 0.45 (se 0.13) is 1.82. Thus in the aggregate economy there seems to be quite a lot of substitutability between the various types of human capital. This need not reflect the flexibility of technology within firms but also the shift of production between different sectors relying on different mixes of these inputs, as relative factors change. The hypothesis of perfect substitutability however ($\rho = 1$) is clearly rejected and thus changes in the supply of the different types of human capital can lead to changes in their relative prices. However, we do accept the hypothesis that the elasticity of substitution is the same across different pairs of human capital.

3.5 The effects of a tuition subsidy

Given the preference parameters, the estimated wage process and the production function, the remaining parameters including the time costs of education $f(\theta)$ are obtained by calibration. The baseline model is calibrated on the basis that direct costs of College education are 30% of median annual income. The model also includes a Government sector whose sole role is to raise proportional labor taxes to fund any interventions. In other work we also consider the relative merits of taxes on capital income.

When we solve the model we use the notion of a stationary recursive competitive equilibrium Lucas (1980). In equilibrium all individual decisions are optimal, input prices are set equal to the marginal product and the goods and asset markets clear.

The model we have set up can address a number of topical policy issues, including the impact of welfare benefits, minimum wages, active labor market programmes or of education policies. The key point of course is that such interventions have effects on the decisions in the entire life-cycle, altering both work incentives and incentives to accumulate human capital. The fact that individuals are *ex ante* different is a central characteristic of these models. It allows for the existence of a group of individuals who justify the existence of the programme as an attempt to compensate for low productivity and/or high costs of schooling. It also allows an examination of how a policy affects inequality driven be exogenous factors.

Here we illustrate some of the issues by considering a tuition subsidy for College. This is an interesting policy to consider because it is high on the agenda of many governments. In the UK for example there has been extensive

The Impact of Tuition Subsidy

GROUPS	Edu. Participation (aggr.shares)			Human Capital Aggregates		
Benchmark (30% of median income)						
All	Less than HS	HS	College	Less than HS	HS	College
	0.251	0.583	0.167	3.312	7.837	2.812
	Edu. Shares by ability			Marg. Products after Tax. and Depr.		
Ability 1 (lowest)	1.0	0.0	0.0	0.955	1.0	1.413
Ability 2	0.576	0.394	0.03			
Ability 3	0.305	0.595	0.10			
Ability 4 (highest)	0.121	0.615	0.263	% with zero wealth	0.162	
Partial Equilibrium results with a 50% tuition subsidy						
	Less than HS	HS	College	Less than HS	HS	College
	0.248	0.562	0.191	3.280	7.622	3.207
	Edu. Shares by ability			Marg. Products after Tax. and Depr.		
Ability 1 (lowest)	1.0	0.0	0.0	0.948	1.0	1.294
Ability 2	0.576	0.379	0.045			
Ability 3	0.303	0.577	0.120			
Ability 4 (highest)	0.116	0.590	0.295	% with zero wealth	0.157	
General Equilibrium results with a 50% tuition subsidy						
	Less than HS	HS	College	Less than HS	HS	College
	0.250	0.576	0.174	3.324	7.802	2.820
	Edu. Shares by ability			Marg. Products after Tax. and Depr.		
Ability 1 (lowest)	1.0	0	0	0.954	1.0	1.396
Ability 2	0.568	0.389	0.042			
Ability 3	0.307	0.580	0.113			
Ability 4 (highest)	0.119	617	0.265	% with zero wealth	0.162	

Elasticity of.Substitution=1.5. Source: Gallipoli et al. (2005)

Table 3: Tuition subsidy simulations with a CES Production function

debate on how much the subsidy should be culminating with a reduction in the tuition subsidy. In many other European countries no fees are charged and all is paid for through general taxation. In the US there is an extensive loan system as well as a number of scholarship programmes. The policy we consider is a very simple one, namely a blanket reduction of College fees by 50% paid for by general taxation. We will consider the effect of the subsidy on overall numbers attending College, how this is distributed by ability groups, what is the impact on between group inequality, and what is the effect on total human capital supplied in each category.

The top panel of Table 3 shows the results for the benchmark economy. The panel below shows what happens in partial equilibrium, when prices for human capital do not change. However taxes must change to fund the tuition subsidy and of course the underlying wealth distribution does change as well as the work behavior. Thus even when the prices are taken as exogenous, this dynamic model can give different results from a simple pilot/control group comparison because the need to fund the intervention internally will affect all individuals. The lowest panel shows the general equilibrium results where human capital prices and the interest rate are allowed to change.

In partial equilibrium this universal subsidy leads to an increase of college graduates from 16.7% to 19.1%. Note that in this economy there are no liquidity constraints and hence the tuition subsidy impact originates from the distortion in education prices alone. There is also a counteracting effect from increased taxation to fund the subsidy. The subsidy leads to large relative increases in College attendance among the second and third ability groups, but not among the lowest group who did not even complete high school before. It also increases College completion for the highest ability group. The total supply of human capital increases by 14%. The key to what is going to happen next however, is the large decline in the marginal product of College graduates, which declines from 1.413 to 1.294. If we now assume that prices are endogenous we obtain the results in the lower part of the Table. On aggregate the increase in the proportion of College graduates relative to the benchmark is less than 1%. (17.4% College graduates). Interestingly though, the ones who drop out are primarily the highest ability group. For them the decline in returns has the highest impact; next largest drop comes from the second highest ability group. Ability group 2 is hardly affected. Whether the policy is judged as effective or not depends very much on the original objectives and on the social welfare function: what has been achieved is an increase in College completion by lower ability individuals, although the aggregate effects are small. In other simulations we show that in the presence of liquidity constraints the policy impact is larger, and has a stronger impact, as one would expect, on higher ability individuals who wish to attend College because it offers funding they would otherwise not have. However, here the absence of formal liquidity constraints replaces the funding for College that would in many cases be offered by parents.

To take the policy analysis further one needs to link individuals to families. One important source of College finance comes from parents. Ability and parental income are correlated, if anything because wealthier parents invest in their children at an early age. Modelling this relationship and allowing for liquidity constraints will allow better to consider the targeting of the policy and analyze its effects in a possibly more realistic context. However these results, obtained using parameters obtained from US data show that standard pilots only contain part of the story. They are more useful as a tool for estimating the impact of incentives on individuals than on the ultimate design of policy. The latter requires us to put together knowledge obtained from a variety of sources and combined with theory and reasonable assumptions. The final outcomes can be dramatically different from those implied by a small scale pilot study. By the way this conclusion is also true, at least to an extent, even when prices are exogenous because of the impacts of changes in taxes and because of changes in incentives not directly targeted by the policy, such as say work incentives or incentives to train on the job, which have not been modelled here.

3.6 Wage subsidies education and employment

Using a model similar to the one described above Blundell, Costa-Dias, and Meghir (2003) analyze the effect of a wage subsidy on employment and education. There has been increased interest in such policies in a number of different countries. Examples include the British "New Deal" introduced in 1998 and the Canadian the "Self Sufficiency Programme". The impacts of such programmes have shown mixed results (see Card and Hyslop (2005) for Canada and Sianesi (2003) for Sweden). The effectiveness of such policies in terms of increasing wages and hence the longer term impact on employment remains an open question. Certainly in the results we present with the empirical dynamic model, where a time limited programme is used, the partial equilibrium results are modest.

However we wish to bring out both the importance of longer term incentives and the potential divergence between partial and general equilibrium policies. Thus, the key point being emphasized here is that policies that are designed to encourage employment may have substantially different effects in the long run, as incentives for education choice and the attractiveness of unemployment changes.

In the model used by Blundell, Costa-Dias, and Meghir (2003) a period is five years. However, the key difference from the model described in the section above is that education choices can be made throughout the life-cycle, rather than just at the start. Moreover work experience leads to the accumulation of human capital. Thus while an individual is working her/his human capital increases at a rate which depends on individual ability. This feature is important in this model because it allows us to examine the basic premise of active labor markets, namely that once individuals are placed in work their labor market attachment will improve because their earnings increase. However, experience is education type specific and is lost once the individual decides to obtain more education and change type. Heterogeneity in this model is of course crucial because it allows to define a margin of individuals who are potential clients of a programme and who could change their long term investments as a result of the programme's existence. In this model labor supply is endogenous but does not enter the utility function. It is determined as a function of a fixed cost of work. The utility function is isoelastic in consumption. Hence uncertainty matters and this originates in the stochastic process for wages. Thus the specification described above has been designed to allow the programme to have an effect on education choice and also to allow for the possibility that individuals substitute between formal education and on-the-job-training offered by the programme, when placed in work through a wage subsidy.

The production sector is a CES between the three types of Human Capital and overall it is a Cobb-Douglas between HC and physical capital - much as in the model before with the difference that this model allows for a different level of substitutability between the three types of human capital. It was found in fact that the elasticity of substitution between College graduates and High school graduates was about 5. However between the less-than-high school group and the aggregate of the other two the elasticity of substitution is 1.0. The different levels of substitutability will imply that GE effects for policies will differ depending on the skill group at which they are targeted.

With this model we consider two policies. One where those unemployed are offered a wage subsidy of 25% the minimum wage and one of 50%. The wage subsidy is only offered to the young, which in this model relates to the first period of five years. We then examine the impact of the policy on those registered in the programme in the first year for the subsequent part of the life-cycle. In the baseline individuals have the choice to work, study or remain unemployed in each and every period. They must also choose their consumption/savings level. The policy is introduced by adding a new option in the first five-year period of life. Under this option the individual may choose to remain unemployed and then will obtain the subsidized wage as described above. Although our graph depicts both policies we only discuss the stronger one which has some discernible long term impact.

The partial equilibrium results reported in figures 1 and 2 are not comparable to the results one would obtain from a small scale short -run pilot study because it allows individuals to change their work and education decisions (as well as consumption) early on in the life-cycle. The results reported relate to the impact of the programme for those who register. The programme makes a spell of early unemployment more attractive and education less attractive, particularly for individuals on the margins of ability enough for them to work or continue with High School in the baseline economy. Such effects are potentially very important but cannot normally be picked up from a standard field trial. An intermediate approach is to run trials to learn about behavior and incorporate this knowledge within the structural model, along the lines we described in the PROGRESA experiment.

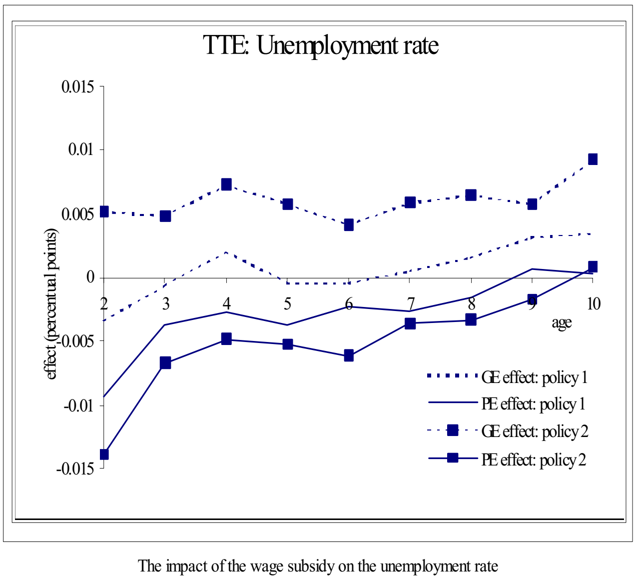
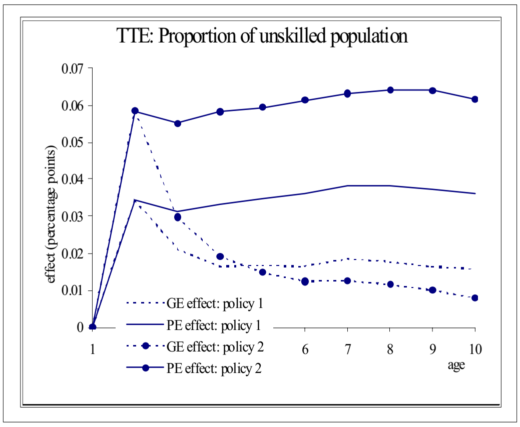


Figure 1: The impact of wage subsidies on life-cycle employment (Source: Blundell et al., 2005)

As seen in the figures, the partial equilibrium effect of the policy, i.e. where human capital prices are kept constant is to reduce unemployment by 1.5 percentage points. The effect is due to the increased labor market attachment induced by the improvement in productivity, obtained as a result of the programme. Five year later the effect declines to half the size. At the same time figure 2 demonstrates that education rates decline, with an increase in the pro-



Effect of the wage subsidy on th education rate for those enrolling in the programme

Figure 2: The Impact of the wage subsidies on education rates (Source: Blundell et al., 2003)

portion of those with less-than-high school of 6 percentage points. In General equilibrium the effect of the policy is actually reversed. The reason for this at first sight surprising result is precisely the fact that the composition and size of the unemployment pool changes as a result of the policy, initially drawing in individuals who would otherwise not be unemployed. This reduces the price of unskilled labor to such an extent that the policy now ends up increasing long-run unemployment, because of the number of people with low ability who prefer to go for the subsidy in period 1 of working life.

In terms of education the GE effect mitigates the adverse impact of the programme. Some individuals find it beneficial to invest in formal education once the programme period has ended and when they observe the decline in unskilled wages. However, there are still more unskilled workers now relative to the benchmark economy.

The key point to take away from this is that the longer term incentives structure introduced by an intervention designed to reduce unemployment may in fact achieve the opposite and may discourage the accumulation of human capital ultimately defeating the original aim of the programme. What our model shows is that with parameters obtained from data these concerns can be of practical relevance. More research needs to be carried out to establish the reliability of such models and the sensitivity of human capital prices to changes in domestic supply. Placed in an international context, factor prices may be equalized, in which case some of the effects we have documented will not be relevant. However, even with exogenous prices of human capital, the issue of dynamic incentives is likely to be important, with policies implemented over the long run having very different impacts from the short run returns.

4 Conclusions

Evaluation has progressed with leaps and bounds in the recent years. The progress has built upon the important advances in our understanding of methods as well by the increasing demands for high quality evaluation by the governments and international organizations. In some sense increasing emphasis has been given to non-parametric, "model free" methods and in particular to randomized experiments. These have been very valuable in establishing the success or otherwise of specific programmes, but by their nature they are quite limited in their ability to allow for redesign of programmes or for evaluating major interventions, that are likely to affect prices, thus violating a key assumption in the evaluation literature, namely that the control group is not affected.

In this paper we have explored two themes. In the first we consider a simple structural model that is estimated based on data from a randomized experiment but is then used to address issues of how the programme could be redesigned. The estimated model is richer than one that would be estimated by traditional observational data and one could argue that its identification rests on firmer ground than is often the case with such models. However, it uses assumptions from economic theory, that may or may not appeal to some; with this combination of data and assumptions one obtains a tool that can go far beyond the abilities of the simple field trail.

In the second theme we explore the use of General Equilibrium models for policy evaluation with heterogeneous agents. In many ways this is quite a new field which requires exploration. Almost inevitably, these models require crucial modelling choices and often heroic assumptions to be made. It is also often the case that the model structure one wishes to impose leads to specifications that are either very difficult to estimate in a rigorous way or for which suitable data is not available, requiring short cuts to be made in the way that parameters are chosen or estimated. Despite all the difficulties however, the results themselves justify the need to explore further and develop these models: GE outcomes can be very different from PE ones both overall and in detail. Thus small pilot studies are frequently incapable of providing even a remotely accurate picture of how a policy will operate when rolled out nationally. One may be justified in being deeply sceptical in certain circumstances about results from GE models; however these models illustrate that one should also be deeply skeptical about concluding from the results of small scale pilot studies on the effects of major policy interventions that alter long run incentives and that could affect prices.

References

Abowd, J. M., and D. Card (1989): "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 57(2), 411–45.

Adda, J., C. Dustmann, C. Meghir, and J.-M. Robin (2005): "Career Progression and Formal versus on the Job Training," mimeo, UCL.

Alonso-Borrego, C., J. Fernndez-Villaverde, and J. E. Galdon-Sanchez (2004): "Evaluating Labor Market Reforms: A General Equilibrium Approach," IZA Discussion Papers 1129, Institute for the Study of Labor (IZA).

Attanasio, O., C. Meghir, and A. Santiago (2005): "Education choices in Mexico: using a structural model and a randomised experiment to evaluate Progresa," IFS working paper EWP05/01.

Attanasio, O., and G. Weber (1993): "Consumption Growth, the Interest Rate and Aggregation," *Review of Economic Studies*, 60(3), 631–649.

Blundell, R., M. Browning, and C. Meghir (1994): "Consumer Demand and the Life-Cycle Allocation of Household Expenditure," *Review of Economic Studies*, 61(1), 57–80.

Blundell, R., M. Costa-Dias, and C. Meghir (2003): "The Impact of Wage Subsidies on Education and Employment: A General Equilibrium Approach," mimeo, Institute for Fiscal Studies.

Blundell, R., L. Pistaferri, and I. Preston (2005): "Consumption inequality and partial insurance," IFS Working Paper W04/28, November , 2004, revised May 2005.

Blundell, R., and I. Preston (1998): "Consumption Inequality and Income Uncertainty," *Quarterly Journal of Economics*, 113(2), 603–640.

Card, D., and D. R. Hyslop (2005): "Estimating the Effects of a Time-Limited Earnings Subsidy for Welfare Leavers," *Econometrica*.

Cunha, F., J. Heckman, and S. Navarro (2004): "Counterfactual Analysis of Inequality and Social Mobility," Mimeo.

Dearden, L., C. Emmerson, C. Frayne, and C. Meghir (2005): "Education subsidies and school drop-out rates," Institute for Fiscal Studies W05/11, June.

Gallipoli, G., and G. Fella (2005): "Education and Crime over the Life Cycle," Mimeo.

Gallipoli, G., C. Meghir, and G. Violante (2005): "Education Decisions, Equilibrium Policies and Wages Dispersion," Mimeo IFS.

Heckman, J., H. Ichimura, and P. Todd (1997): "Matching as an Econometric Evaluation Estimator," *Review of Economic Studies*, 65(2), 261–294.

——— (1998): "Characterizing Selection Bias Using Experimental Data," *Econometrica*, 66(5), 261–294.

Heckman, J., L. Lochner, and R. Cossa (2003): "Learning-By-Doing Versus On-the-Job Training: Using Variation Induced by the EITC to Distinguish Between Models of Skill Formation," in *Designing Inclusion: Tools to Raise Low-end Pay and Employment in Private Enterprise*, ed. by E.Phelps. Cambridge: Cambridge University Press.

Heckman, J., L. Lochner, and C. Taber (1998a): "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents," *Review of Economic Dynamics*, 1(1), 1–58.

——— (1998b): "General Equilibrium Treatment Effects: A Study of Tuition Policy," *American Economic Review, Papers and Proceedings*, 88(2), 381–386.

Heckman, J., and R. Robb (1985): "Alternative Methods for Evaluating the Impact of Interventions," in *Longitudinal Analysis of Labor Market Data*, ed. by J. Heckman, and B. Singer. Cambridge: Cambridge University Press.

Heckman, J., and G. Sedlacec (1985): "Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market," *Journal of Political Economy*, 93(6), 1077–1125.

Heckman, J., and B. Singer (1984): "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, 52(2), 271–320.

Heckman, J. J., R. LaLonde, and J. Smith (1999): "The economics and econometrics of active labor market," in *Handbook of Labor Economics, vol. 3A*, ed. by O. Ashenfelter, and D. Card. North-Holland.

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

LaLonde, R. (1986): "Evaluating the Econometric Evaluations of Training Programs with Experimental Data," *American Economic Review.*

Lee, D. (2001): "An Estimable Dynamic General Equilibrium Model of Work, Schooling and Occupational Choice," Ph.D. thesis, University of Pennsylvania.

Lee, D., and K. I. Wolpin (2005): "Intersectoral Labor Mobility and the Growth of the Service Sector," forthcoming Econometrica.

Lucas, R. E. (1980): "Equilibria in a Pure Currency Economy," in *Models of Monetary Economics*, ed. by J. Kareken, and N. Wallace. Federal Reserve Bank of Minneapolis, Minneapolis.

MaCurdy, T. E. (1981): "An Empirical Model of Labor Supply in a Life-Cycle Setting," *Journal of Political Economy*, 89(6), 1059–85.

Marschak, J., and W. H. Andrews (1944): "Random Simultaneous Equations and the Theory of Production," *Econometrica*, 12(3/4), 143–205.

Meghir, C., and L. Pistaferri (2004): "Income Variance Dynamics and Heterogeneity," *Econometrica*, 72(1), 1–32.

Orcutt, G. H., and A. G. Orcutt (1968): "Incentive and Disincentive Experimentation for Income Maintenance Policy Purposes," *American Economic Review*, 58, 754–772.

Rosenbaum, P. R., and D. B. Rubin (1983): "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrica*, 70(1), 41–55.

Schultz, T. (2003): "School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program," *Journal of Development Economics*.

Sianesi, B. (2003): "Swedish Active Labour Market Programmes in the 1990s: Overall Effectiveness and Differential Performance," *Swedish Economic Policy Review*, 8(2), 133–169.

Todd, P., and K. Wolpin (2003): "Using Experimental Data to Validate a Dynamic Behavioral Model of Child Schooling and Fertility: Assessing the Impact of a School Subsidy Program in Mexico," Mimeo the University of Pennsylvania.