DISCUSSION PAPER SERIES

IZA DP No. 3292

Gone for Good?
Determinants of School Dropout in Southern Italy

Niall O'Higgins
Marcello D'Amato
Floro Ernesto Caroleo
Adriana Barone

January 2008

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Gone for Good? Determinants of School Dropout in Southern Italy

Niall O'Higgins
Università di Salerno, Celpe and IZA

Marcello D'Amato
Università di Salerno and Celpe

Floro Ernesto Caroleo
Università di Napoli Parthenope and IZA

Adriana Barone
Università di Salerno and Celpe

Discussion Paper No. 3292
January 2008

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 3292
January 2008

ABSTRACT

Gone for Good?
Determinants of School Dropout in Southern Italy*

The aim of the present paper is to gain some insight into the causes of dropping out of school and, more generally, of the factors that induce parents to review their choices about their child's schooling careers. To this end we apply to data from a school dropout survey insights from a model of sequential decision making by parents, where the initial decision can be reviewed in the light of new information emerging about the ability and opportunities of the child in benefitting from education relative to her outside (in the unskilled market). Analysis of the data confirms the role of both economic capacity (opportunity costs) and cultural capacity (ability to disentangle signals about future opportunities) of the family of origin shape observed choices about drop-out and return to school by individuals in our sample. Dropping out behaviour also appears to be strongly influenced by mismatches between school and student, however, and many of those who leave are not "gone for good".

JEL Classification: I21, J13, J24

Keywords: young people, school dropout, human capital

Corresponding author:

Floro Ernesto Caroleo
Department of Economics
University of Naples "Parthenope"
Via Medina 40
80133 Naples
Italy
E-mail: caroleo@uniparthenope.it

* We acknowledge financial support from the Province of Salerno in undertaking the survey on which this paper is based. Thanks are also due to participants in the workshop on Education Supply and Demand in Italy, Turin, 18 May 2007 and at the AIEL conference in Naples, 13-14 September 2007 and in particular to Daniele Checchi and Giuseppe Bertola, who also organised the Turin meeting, along with an anonymous referee for their insightful comments on previous versions and which we very much hope lead to the paper's substantial improvement. Unfortunately none of these can be blamed for any remaining errors of omission or commission which remain firmly with the authors.

1. Introduction

In this paper, data from the "school dropout survey" undertaken in Salerno Province by the Centre for Labour Economics and Economic Policy (CELPE) are analysed. A key feature of our sample is that a substantial proportion of school "dropouts" subsequently return to some form of education. Abandoning school for a period and then possibly returning to education in a different school is obviously an indication of distress and suggests that education choices are subject to several kind of errors and mistakes: mistakes parents make about the kids ability to pursue a specific education career, changes in the child's attitudes towards school in a period of their life (adolescence) when her character, social skills and specific capacity to adapt to learning and education evolve and are subject to shocks somewhat not perfectly controlled or forecasted by families. These social and psychological aspects make the study of these problems from the economic point of view quite entangled. From a purely choice theoretic perspective, parental aspirations and their choices about their child's educational career involve subtle psychological and sociological aspects not easily squared with a simple investment approach to education choice. Such choices may be reviewed in the light of the new information about opportunities and abilities that unravels when the child grows as an adolescent. This new information may arise from parental experience about the cultural evolution of the young man as a student, about the quality of the match between a specific child and a specific school or teacher, or simply about the opportunity cost of education given what parents see as the return to have their child educated. However, even if it is a hard task to square decisions made by adolescents and their parent during this period of the child's life with a choice theoretical framework, studying the economic implications of these decisions is important from an economic perspective. Choices made during childhood have important effects later in life, shaping the destiny of the young person both in terms of social status and of economic returns of investments in human capital made during adolescence. More generally, the presence of mistakes and errors may be an indication of an inefficient allocation of talents to school tracks during adolescence and therefore, given the large involvement of public resources in the education systems, the study of these problems is very important also from the perspective of policy design as well.

The aim of our work is to gain some understanding of the causes of dropping out of school and, more generally, of the reasons that induce parents to review their choices about their child's schooling careers. The paper proposes a model of sequential decision making by parents where the decision can be reviewed in the light of new information emerging about the ability and opportunities of the child in profiting from education relative to her outside (in the unskilled market). The model allows interpretation of such dropout and return behaviour and emphasises the separate role of both economic and cultural capacity (as regards the ability to disentangle signals about future opportunities) for equilibrium decision making. On the other hand, being a partial equilibrium model, we do not tackle efficiency issues that obviously arise, and use the model as a guide for the implementation of the analysis of the data and as a theoretical framework to interpret empirical results. Specifically, the nature of the theoretical model leads fairly naturally to the implementation of a censored bivariate probit model of initial dropout and return to school by young people.

The rest of the paper is organized as follows: in section 2 we present some stylized facts and some related literature, in section 3 we present the model, in section 4 we present the empirical results; section 5 concludes.

2. Some Stylised facts and a glance at the literature

The Lisbon agenda includes several quantitative targets for education and training systems in Europe. One specific goal is reducing the share of early school leavers – i.e. 18-24 years-old with at most a lower secondary education qualification and not in further education – to less than 10% by 2010. (Commission of the European Communities, 2002). The reason is that rising skills demands imply the necessity of the completion of at least upper secondary education for successful labour market entry and for further participation in lifelong learning (see OECD, 2000). Even though present trends show, in general, clearly decreasing levels of early school-leavers in the Member States - especially in Spain, Italy, Greece, France, Belgium and Finland - (Quintini and Martin, 2006), further efforts are required to enable the EU to reach the target, as the loss of stock of human capital is still unacceptably high.

A recent analysis carried out by the Ministry of Public Education (Ministero della Pubblica Istruzione, 2006, pp. 12-13, tab.7) points out some peculiarities of the drop out phenomenon in Italy, underlining that:

- both failure at school and dropout are so widespread that about a third young men and more than one-in-five young women enrolled in upper secondary school do not obtain the "diploma";
- female educational performance is better than males, particularly in technical and vocational schools;
- students attending classical and scientific schools have the highest probability of completing their studies in all the geographical areas of Italy while those attending technical schools in the North of Italy have a lower probability of completing their studies than in other areas. This may plausibly be explained by the, at times opposing, motivations for dropping out related to the opportunity cost of, and expected (employment and income) returns to education. Thus, dropping out is a characteristic of both depressed and developed areas. The early entry into the labour market in Lombardia and other areas in the North-East where the labour demand is high can cause a crowding effect of the schooling and training system, raising the risk of lower qualified occupations in the future (table 1).

Table 1: Probability of obtaining final diploma by gender, geographical area and type of school (graduates from the school per 100 enrolled at the first year 5 years before) – Upper secondary school public and private - year 2004-05

	Total	Scientific and classical	Psycho-pedagogical	Technical	Vocational	Arts
TOTAL	72,3	85,6	82,3	79,0	47,8	62,9
North	71,1	85,3	84,7	73,4	51,3	68,5
Centre	74,4	86,7	78,5	82,3	48,4	61,2
South and Islands	72,4	85,4	82,1	82,6	44,8	59,5
MALES	67,1	84,7	87,5	74,5	41,5	56,7
North	64,7	83,5	90,7	67,9	43,8	59,5
Centre	69,7	88,7	69,9	75,9	42,2	55,8
South and Islands	67,9	83,8	91,5	79,6	39,7	55,2
FEMALES	78,1	86,3	81,5	87,6	56,5	66,0
North	78,1	86,8	83,8	83,2	60,3	72,6
Centre	79,5	85,3	80,1	95,5	56,5	63,8
South and Islands	77,5	86,5	80,6	88,6	52,9	61,8

Source: Ministero della Pubblica Istruzione (2006b, table 7, p. 13)

Concerning the case under study, table 2 reports residents at the beginning of the years 2003-2005 by age in the Province of Salerno and students enrolled in the schooling year 2004-05. Each age group comprises 13-14.000 individuals. The gap between the numbers of those enrolled and age-specific age population is around 10% of the population which in itself provides a very rough indicator of the extent of dropping out before completion of secondary school. Moreover, both the numbers enrolled and the proportion in State schools decline from 15 to 18 years. The first phenomenon is an aspect of what is known "school dispersion" and the second signals that students shifts from public to private schools in order to become secondary school graduates.

Table 2: Number of the residents and enrolled in the Province of Salerno by age

Age	Total residents at 01/01/03	Men	Total residents at 01/01/04	Men	Total residents at 01/01/05	Men	Enrolled to schools (A)	Enrolled to public schools (B)	Enrolled to public school/ enrolled to schools (A/B) (%)
15	13390	6825	13954	7056	13474	6971	12885	12616	97.9
16	13723	7002	13447	6845	14030	7101	13025	12645	97.1
17	13937	7177	13769	7009	13482	6879	12211	11749	96.2
18	14329	7317	13980	7209	13801	7031	10370	9844	94.3
Total	55379	28321	55150	28119	54787	27982	48491	46854	96.6

Source: Istat (2003-2005) and Ministry of Public Education (2006a)

Data taken from the 2001 Census of Population and reported in Table 3 confirm that the number of residents aged 15-18 in the Province of Salerno not enrolled to a regular course of study increases as the age increases.

In the last decade the Italian authorities, aiming at reducing drop out rates, have opted to change the minimum school-leaving age introducing compulsory schooling and training up to age 18 (Law n.144/1999 "obbligo formativo" and Law n.53/2003 "diritto dovere all'istruzione e alla formazione"). The institutional actors, public or private, involved in the realization of such obligation are manifold (Schools, Training Centres, Employment Agencies) and at different level of governance: central (Ministry of Education) and local (Regions, Provinces, Local Employment Agencies). The duty of all of them is to follow the students in pursuing their studies and to make easier the school-to-work transitions, developing broad-based community partnerships aimed at helping at-risk youth.

Table 3. Number of the residents and enrolled/not enrolled in the Province of Salerno

Age	Residents (A)	Enrolled in a regular course of study (B)	Not enrolled in a regular course of study (C)	(C/A)%
15	13867	11914	1953	14
16	13881	11131	2750	19.81
17	14482	10858	3624	25
18	14383	9528	4855	33.75
Totale	56613	43431	13182	23.28

Source : ISTAT(2001)

Students who drop out face a lot of economic and social difficulties. Several studies find that adult earnings are higher when students are compelled to take an extra year of school (Oreopoulos 2007; Angrist and Krueger 1991; Acemoglu and Angrist 2001; Blundell, Sianesi and Dearden, 2003). The positive social externalities are also considerable. Drop outs are more likely to be unemployed, reporting poor health, being depressed, being in a low skilled manual occupation, drawing on social assistance and other welfare programs, and to end up in jail (Lochner and Moretti 2004).

Many factors have been identified as influencing dropping out. Some are school-related: a drop out does not like school in general or the school he/she is attending, he is failing, getting poor grades, or cannot keep up with school work, he does not get along with teachers and/or students. Other factors are student-related: a drop out has disciplinary problems, is suspended, or expelled, he does not feel safe in school, he has different traits than those who graduate (Eckstein and Wolpin, 1999) for example: low ability and/or motivation, low expected returns to graduation, better market opportunities for the jobs that don't require graduation, lower consumption value of school attendance. Reasons for dropping out may be related to personal problems as well: drug and alcohol abuse, obesity or health problems (Celpe 2006). Other factors are family-related: stressful/unstable home life, lack of family support, socioeconomic status, single-parent households, poor education of parents, desires to get married and/or getting pregnant (Cardoso and Verner, 2006). There are also some categories of youth that are at more at risk of dropping out: certain ethnic groups, students living in large cities or in a poor areas, or attending schools whose structure or academic and social organization may not favour the holding in of students at risk (Lee and Burkam 2003).

The standard human capital model assumes that students are rational and time consistent, so the choice of the optimal education attainment level may be treated as an investment decision (Becker 1964; Card 1999). That level is defined by the point where the opportunity costs from additional schooling outweigh the benefits. Students who decide to drop out because of psychological or motivational problems, are simply evaluating that forgone earnings and effort costs from attending school are higher than the estimated expected benefits.

However, it is widely recognized that the standard model cannot completely explain drop out behaviour. Policymakers want to update compulsory school laws and to introduce restrictions on the choice of students because they are persuaded that leaving secondary school before graduation is a sub-optimal outcome. Oreopoulos (2007) finds "significant lifetime rewards to wealth, health, and overall happiness from having to take another year of school" as well as substantial income gains so that disutility from not attending school is very high.

There are other reasons that may account to some extent for the decision to leave the school. Credit constraints may mean that students, especially those from low income families, cannot borrow against the expected higher outcomes to finance the ex ante optimal level education (Carneiro and Heckman 2002). Myopic behaviour of drop outs may also lead to a similar outcome. Students may give greater emphasis on the present, because they make erroneous predictions of future returns or underestimate the real gains from school or have negative expectations about the future, so they attach more weight to the current non pecuniary or monetary costs in calculating the school attainment decisions. Sociological and psychological research points to the importance of a student's social group in determining their active involvement in school. This literature (for a review see Akerlof and Kranton 2002) considers schools as institutions, with social goals besides imparting skill, and highlights that educational outcomes of students depend on their identification with the *school's social category* and its *ideal* student. A further aspect that has been investigated is the role that schools –their structures, their academic organisation and their social organisation- play in students' decisions to stay in school or leave before graduating (Lee and Burkam 2003). Factors related to school quality are also important in determining whether students who leave their high schools either transfer to another school (and thus stay in school) or leave school altogether (Rumberger and Thomas 2000; Hanushek, Lavy and Hitomi 2006).

Different dropping out explanations, obviously, call for different policy designs. Compulsory school legislation may be effective when staying in school has significant benefits in term of future rewards to wealth and well being. On the contrary, policies aimed at forcing students to remain in school until they

graduate may have little impact if the reasons for leaving school are related to the individual characteristics: the lack of motivation, low ability, or the greater preference for low skilled jobs. In this case, it becomes necessary to alter the traits with which youths come to high school. Policy options, aimed at reducing the costs of schooling (both direct and indirect) or offsetting immediate costs with immediate benefits, may improve the student's skill upgrading when drop outs do not correctly evaluate the future returns to schooling because of a myopic behaviour or because of the family background, credit constraints, extremely poor economic and social environment. As noted above, psychologists and ethnographers often point at the importance of a student's attitude towards school, rooted in their social and cultural background, in influencing their school choice decision, so it becomes important that schools reinforce the social inclusion and the social identity of their students. Finally, sociological theory stresses also the relevance of organizational and structural characteristics of school -i.e. size and sector, curricula offered, the character of relationship between students and teachers, etc.- in the decision of students, implying that policymakers have to pay a special attention to strengthen the quality of schools.

3. A framework for the analysis. A simple model of parental sequential choices about their children's schooling careers.

In this section we build up a framework for the analysis of the main determinants of the decision to quit school, and possibly return to it, after having spent a period out. The model is based on a simple idea of repeated learning by parents about their children's abilities and (future) job market opportunities[2]. Sending the child to school allows parents to assess their ability in the labour market and formulate expectations about the child's wage on the skilled labour market relative to the unskilled one. The main building blocks of the model are as follows: parents use school signals to extract information about their child's ability as a skilled worker based on their prior beliefs and on the evidence provided by school assessment. If the signal at school is bad, parents re-assess the value of investment in human capital depending on their private signal on the child's ability. After balancing the private with the public signal (the relative precision of these being crucial for the assessment) parents make a decision about school enrolment in subsequent grades. If the child leaves school, she spends a period out of education and further information accrues to her parents about her specific ability in the unskilled job market. Conditional on this information, parents may decide to leave the child outside school or to give her a second chance to proceed in education. This latter hypothesis, i.e. that some information about future alternative opportunities to the child is collected when he spends a period out of school, is the key element to understanding the process of re-entry.

Parental capacity to assess their child's quality and use the public signal at school, along with the opportunity cost of education and some form of irreversibility of education choices will determine the equilibrium choices in this model: parental culture (precision of the private signal of the child's ability), school quality (the precision of the public signal provided in grading the child's performance at school), and economic status (the opportunity costs of educational investment) matter in our framework. As for irreversibility we will assume that high school can only be completed when the child is adolescent and stays with her parents. We will also assume that after completing high school the decision about the young person's future career (university versus unskilled labour market) is irreversible.

---

[2] The identification of the parent(s) as the decision-maker rather than the individual directly concerned aligns this paper with the convention usually adopted in overlapping generations models, rather than the standard human capital model which tends to identify the (potential) student as the decision-maker. Indeed this may look strange to readers more familiar with the latter type of model. However, the crucial simplifying assumption here is that there is a single decision-maker rather than who that decision-maker is. For a model with altruistic parents choosing their children's human capital investment see, for example, Mookherjee & Ray (2003).

Specifically, we consider two basic paths leading the child to being prepared, as an adult, either for the skilled labour force (say university enrolment) or for the unskilled labour force. We assume that decisions are made by altruistic parents. These choices are based on the parental assessment of the benefits of the child's schooling alternatives, given their expectation about wages in the skilled and unskilled labour market. These assessments regard the child's ability in the two labour markets. School is assumed not to add skills valuable on the job market, this is a strong simplification but a richer role for school could easily be considered within the model[3].

Choices are made sequentially, conditional on the signals parents get about their children's ability and their market opportunities. In making their sequential choices, parents trade-off these benefits with their opportunity costs of having their children educated. Sequentiality of choices reviewed in the light of the new information is modelled as follows. We divide the child's time horizon (call it adolescence) for the family decisions into three periods. The existence of compulsory basic education means that parents have to send children to school for at least one period at the end of which a signal is collected about the child's ability. At the beginning of period two, parents make a choice about the child's future career: either they withdraw him from school or they have him complete his education. This choice is made conditional on the signal being collected at school in the first period (educational score at school in period one) balancing it with their priors about the child's ability. If parents choose to leave the child at school he completes his schooling career in period two and has to decide whether to enrol at a higher level (university, say) or get back to the unskilled labour market.

Conditional on the parental decision to withdraw the child from school at the beginning of the second period, we assume the child spends a period out learning his alternative opportunities on the unskilled labour market[4]. Again, after having spent a period out of school and having obtained a signal about her

[3] For a richer view of child's identity and education see Akerlof and Kranton (2002). Assuming education is productive would increase the incentive to stay in school and, where relevant, to return after dropping out but would not change the implications of the model.

[4] The description of a period out of school during adolescence has to be interpreted in quite broad terms. It may vary depending on the actual socio-economic contexts in which the family investment in educating their children is cast. For example in less developed countries or areas, where agriculture or small scale family business play important role, failing at school will likely involve for the kid a period of actual training in the traditional occupation. In more developed countries or regions kids' failing at school may only involve further information acquisition by parents about alternative opportunities (this information acquisition by parents may range from a more careful investment in the parent-child relationship entrusting a psychologist or a private tutor, discussing the issue with friends and relatives to better assess the child's talents for alternative occupations. We do not enter into such details- which can be quite important- since our working hypothesis is meant to deal with several specific socioeconomic environments: after the kid is perceived to have failed at school parents have incentives to review their prior about the kids talents in different occupations and condition their education investment upon this new information.

child's opportunities on the (unskilled) labour market they make their decision again: they can leave the child in the unskilled labour market or have her return to school.

At equilibrium, in each family, choices about a child's career will depend on the skill premium expected given the signals (i.e. the relative expected benefits from the two alternative paths), the opportunity cost of education, and on the precision of the signal about the child's potential ability as a skilled worker (at school) and her opportunities on the unskilled job market.

To analyze the model we make a bunch of simplifying assumptions none of which will affect our main results whose economics will turn out to be quite intuitive. We assume that there is one parent and one child, the decision is made by the parent, leaving the child with no meaningful decision. Our model is a partial equilibrium model, i.e. we do not derive the equilibrium level of the skill premium as a function of family choices. Learning occurs optimally given the information accruing to parents as time unravels. We will assume risk neutrality in parental utility and that all the signals used for parental decisions and the parameters on which the decision making is made contingent on are normally distributed[5]. We also assume no discounting by parents and, just for simplicity, no tuition fees for high school enrolment with anticipated costs of university enrolment being equal to T_s[6].

As already discussed, once children complete high school, the choice between going to university and working in the unskilled labour market becomes irreversible. Analogously, after two periods in the unskilled labour market children are not allowed to return to school. This is equivalent to making the cost of switching back to education after a certain age sufficiently large. A less drastic assumption would have been to provide parents and children with the option to switch across educational careers for a longer horizon at a smoothly increasing costs. This assumption is only made to avoid recursive formulation of the more general dynamic programming problem that would arise in the absence of complete irreversibility. This richer formulation would not change the main results, however. Finally we assume that the children's abilities are relevant on the skilled labour market but are not on the unskilled labour market[7]. This will simplify the derivation of the formula for the information updating process by the parent quite a bit, but again, most of the results would go unchanged with a more general role of ability in both markets. More specific comments about these assumptions will be provided in the following.

[5] Both these assumptions can be dispensed with the results going almost unchanged if we assume risk aversion on the child's benefits and risk neutrality with respect the parents' income and i.i.d properties of the shocks accruing to parents' information set.

[6] Which may include perceived consumption costs as well as tuition fees and the like.

[7] Equivalently the assumption is that child's ability index valuable in the skilled labour market is independent of the ability index that defines the child's market opportunities in the unskilled labour market.

The details of the decision making process and the time line faced by parents are as follows.

At time *t=0* parents enrol their child at school. The child's ability is denoted α, which is equal to the child's value (productivity) expected on the skilled labour market (denoted αw_s), and is related to parental beliefs by the relation:

(1) $\alpha = \theta + \varepsilon$,

where θ is the parent's initial assessment, ε is the noise affecting parental judgement. We assume the signal θ is unbiased[8] i.e., ε follows a normal distribution with mean 0 and variance σ_ε^2. This latter can be family specific and reflect cultural heterogeneity among families in their capacity to assess child's ability valuable for the skilled sector. Having assumed it is unbiased, the prior assessment about the child's value on the skilled labour market is θ. A Parent's unconditional expectation about his child's abilities is therefore:

(2) $E[\alpha] = \theta$

Sending the child to school provides a signal on his ability level

(3) $z_t = \alpha + s_t$

Where s_t follows a normal distribution with mean 0[9] and variance σ_s^2. The precision $1/\sigma_s^2$ can be interpreted as a quality index for the assessment process at school.

At time *t=1* the school assessment of the child's ability is revealed as $z_1 = \alpha + s_1$. The parent observes it and makes his first decision about the child's career: either she is withdrawn from school or

[8] An unbiased signal is not necessarily a very realistic assumption ("even a beetle is a beautiful to his mother" as the Neapolitan saying goes) however it removes issues to do with inefficient schooling decisions and the dynastic segregation of jobs based upon social confidence by the upper classes. As we will see, the model already allows for the role of cultural bias in education decisions even with an unbiased signal: if the parent is a skilled worker, he is likely to have a low σ_ε^2 and therefore attribute relatively more weight to his own assessment than to school grading than a (less educated) parent with high σ_ε^2 leading to the result that skilled parents are less likely to withdraw their children from school conditional on a bad grade.

[9] Again, we do not consider another important source of segregation into occupations, i.e. the impact of cultural homogeneity among skilled families and school teachers on education choices.

is allowed to continue in education. The perceived benefits from continuing in education are measured by the conditional expectation of the child's ability :

(4) $\hat{\alpha}_{t+1} w_s$ where $\hat{\alpha}_{t+1} = E[\alpha | I_t]$

This is a measure of the expected wage in the skilled labour market as a function of the estimated ability of the child conditional on the parent's information set. Specifically I_t is the parent's information set at time *t* including all the signals received until then. In the first period, $I_1 = \{\theta, z_1\}$.

The benefits of withdrawing the child from the school are given by the expected wage on the labour market plus the option value of returning to school next period.

At time *t=2* further information is released to parents. If the child was thusfar in school, a second signal about her ability in the skilled sector $z_2 = \alpha + s_2$ is revealed and a decision about her future is made: either she continues her path towards the skilled labour market (by entering university) or she returns to the unskilled labour market. Benefits from continuation are therefore measured by

(5) $\hat{\alpha}_3 = E[\alpha | I_2]$

The information set after two periods in school is $I_2 = \{\theta, z_1, z_2\}$.

If the child is withdrawn from school after the first period, she spends a period out of school (say working), a signal is collected about the child's opportunities in the unskilled labour market. Define

(6) $w_t^u = w^0 + u_t$

the child's expected wage in the unskilled labour market. Where w^0 is the parent's prior belief on the child's potential wage in that market and u is the specific talent of the child in this job (unobserved by the parent). u follows a normal distribution with mean 0 and variance σ_u^2. The parent does not observe w_t^u, but an unbiased signal of child's opportunities in the unskilled labour market (say the child's current wage provided she is paid one), defined as

(7) $\phi_t = w^u + \delta_t$

where δ_t follows a normal distribution with mean 0 and variance σ_δ^2. The information set after one period in school and one period out of school is then $I_2^D = \{\theta, z_1, \phi_2\}$.

Summarizing, we model the period of adolescence as a learning process by parents about their childrens' abilities and opportunities in the presence of some degree of irreversibility of educational choices. Parents make sequential decisions about their child's educational career based on their attainment and grades z_t. We will show that a good signal at school (z_1) on the child's ability α increases parental expectations about what the child's utility will be as a skilled worker and will increase the chances that the parent will keep financing the "child's" education. If this signal is relatively bad the parent will try the unskilled labour market where a signal about the child's opportunities will be collected (ϕ_t). If this latter signal also turns out to be bad, the parent may be willing to reconsider his choice and enrol the child again at school.

All of this is quite intuitive, the advantage of making the model explicit is to derive precise predictions about educational choices as a function of the parameters of the problem, i.e. the priors, the average skill premium in the market, the opportunity costs of children's education for the families and the precision of different signals about abilities and market opportunities. Notice that, as simple as the model is, it can allow for economic and cultural heterogeneity at the family level. Economic heterogeneity is captured by the opportunity cost faced by parents who bear the full direct cost of education, T_s. Cultural heterogeneity is caught by the precision of the signal the parent has on their child's ability in the skilled sector σ_ε^2 and in the unskilled sector σ_δ^2. The time line, the evolution of information sets and the decision tree is summarized in the following figure:

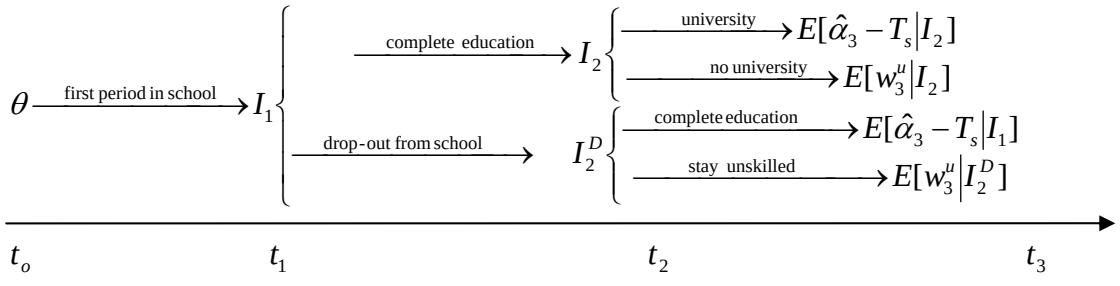

Figure 1: Parental choice

3.1 Updating information about the child's abilities and opportunities

The updating formulas are iterations on the conditional expectation operator where conditioning occurs on the signal recovered (this is a version of the Kalman Filter, see Sargent 1987, pp.230-231). Define the optimal weight parents attribute to signals from school as

(8) $$K_t = \frac{\Sigma_{t-1}}{\Sigma_{t-1} + \sigma_s^2}$$

Where $\Sigma_t = E(\alpha - \hat{\alpha}_t)^2$, therefore the updating formula on the child's ability during his school career under the assumption of the present model is given by

(9) $$\hat{\alpha}_t = (1 - K_t)\hat{\alpha}_{t-1} + K_t z_t$$

Notice that *K* is a measure for the relative precision of the signal at school relative to the precision of the signal that parents collect on their own.

Define the weight parents give to the signal about their child's opportunity in the unskilled labour market as

(10) $$H = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\delta^2}$$

Notice that *H* is a measure for the relative precision of the signal in the unskilled labour market relative to the precision of the signal that parents have on their child's opportunities on the unskilled labour market. Notice that, since the child can only re-enter after one period out of school, we do not need iterative formulas for updating signals on the unskilled labour market opportunities. Therefore the formula for the expected wage in the unskilled labour market is given by

(11) $$\hat{w} = E[w|\phi_1] = (1 - H)w_0 + H\phi_1 .$$

We are able now to compute expected benefits in both sectors conditional on the relevant information set.

Consider the expected productivity in the skilled sector after one year at school when the score z_1 is obtained :

$$\begin{aligned}(12)\qquad \hat{\alpha}_1 &= (1-K_1)\hat{\alpha}_0 + K_1 z_1 \\ &= (1-K_1)\theta + K_1 z_1\end{aligned}$$

Where, $\hat{\alpha}_0 = \theta$, $\Sigma_0 = \sigma_\varepsilon^2$ and $K_1 = \dfrac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_s^2}$.

After two years of schooling, expected productivity in the skilled sector will be given by

$$(13)\qquad \hat{\alpha}_2 = (1-K_2)\hat{\alpha}_1 + K_2 z_2$$

Where $\Sigma_1 = E[\alpha - \hat{\alpha}_1]^2 = (1-K_1)^2\sigma_\varepsilon^2 + K_1^2\sigma_s^2$ and $K_2 = \dfrac{\Sigma_1}{\Sigma_1 + \sigma_s}$.

After one (compulsory) period at school the expected wage in the unskilled sector will be given the unconditional expectation w_0. After one period at school and one period out of school, expected productivity in the unskilled wage is given by

$$(14)\qquad \hat{w} = E[w|\phi_1] = (1-H)w_0 + H\phi_1$$

Having defined the information updating process for the decision maker we can solve now for the equilibrium choices.

3.2 Equilibrium decision making about children's careers

This is a sequential parental choice model about educational careers of children. Choice at each point in time is made optimally, in the light of the current information set and given choices made in the past. An equilibrium of the decision making problem will be a set of thresholds for the scores obtained at school and for the signals obtained conditional on dropping out defining, in period 1, the choice to withdraw the child from school and, in period 2, to enrol or not in the university conditional on being at

school and to re-enter or not education provided one period was spent out of school. To obtain equilibrium decisions we work backwards from the final period.

After two periods at school the child, by now a young woman, has produced two signals z_1 and z_2 about her productivity in the skilled labour market and no signals on the unskilled labour market. Therefore[10] a university career will be chosen if

$$(15) \qquad \underbrace{\hat{\alpha}_3 w_s - T_s}_{\substack{\text{conditional expected productivity in the} \\ \text{skilled labor market net of tuition fees}}} \geq \underbrace{w_0}_{\substack{\text{unconditional expected productivity} \\ \text{in the unskilled labor market}}}$$

After some trivial algebra this can be written as

$$(16) \qquad (1-K_2)(1-K_1)\theta + K_1(1-K_2)z_1 + K_2 z_2 \geq \frac{w_o + T_s}{w_s}$$

Which gives the cut off value of $z_2, \tilde{z}_2(z_1)$ such that, for any value of z_1 consistent with continuation, if $z_2 > \tilde{z}_2(z_1)$ a university career leading to a skilled job will be chosen i.e.

$$(17) \qquad z_2 \geq \tilde{z}_2(.) = \frac{w_0 + T_s - (1-K_2)(1-K_1)\theta - K_1(1-K_2)z_1}{K_2 w_s}.$$

This characterizes family decisions in the third period as a function of the history of scores at school, their relative precision and other parameters of the model.

After the first period of schooling, i.e. conditioning on z_1 alone, with no discounting, the net benefits from continuation at school have to be traded off against benefits from switching to the alternative of dropping out. In the Appendix, using the iterated expectation law, we show that the choice of dropping out reduces to the following inequality

$$(18) \qquad \underbrace{\mathrm{E}[\hat{\alpha}_2 | I_1] w_s - T_s}_{\substack{\text{conditional expected productivity in the} \\ \text{skilled labor market net of tuition fees}}} \geq \underbrace{w_0}_{\substack{\text{unconditional expected productivity} \\ \text{in the unskilled labor market}}}$$

That is:

$$(19) \qquad [1-K_1(1-K_2)]\theta + K_1(1-K_2)z_1 \geq \frac{w_0 + T_s}{w_s}$$

[10] Remember that after completing high school we assume that the choice between university and unskilled labour market becomes irreversible. This hypothesis is made mainly to simplify the model and is equivalent to make the cost of switching back to education after a certain age sufficiently large. However, the idea that irreversibility of choices about occupational careers is a realistic feature of the model. A less drastic assumption would give a longer horizon to agents maintaining the option to go back and forth different education choices for a larger number of periods at increasing costs, until it is no more worthwhile to gather information about different alternatives.

Therefore a parent will choose to have the child to continue in school for the second period if the signal in the first period is good enough

$$z_1 \geq \tilde{z}_1(.) = \frac{w_0 + T_s}{K_1(1-K_2)w_s} - \frac{[1-(1-K_2)K_1]\theta}{K_1(1-K_2)} \tag{20}$$

and drop out otherwise. This characterize family decisions as a function of first period signal at school.

Finally we need to characterize the decision about re-entering school after one period out. To this end we define the expected benefits for continuing in the unskilled labour market conditional on two signals: ϕ_2 (z_1 is assumed to be irrelevant to review priors about child's ability in the unskilled labour markets) and the expected benefits from re-entry conditioning on z_1 alone (since ϕ_2 is assumed to be irrelevant to assess opportunities in the skilled labour market). The latter is equal to the left hand side of equation (18). The former is given by:

$$V_2^R = (1-H)w_0 + H\phi_2 \tag{21}$$

Therefore re-entry will occur if and only if

$$(1-H)w_0 + H\phi_2 \leq \{[1-K_1(1-K_2)]\theta + K_1(1-K_2)z_1\}w_s - T_s \tag{22}$$

i.e.

$$\phi_2 \leq \tilde{\phi}_2(.) = \frac{\{[1-K_1(1-K_2)]\theta + K_1(1-K_2)z_1\}w_s - T_s - (1-H)w_0}{H} \tag{23}$$

At equilibrium, conditional on having dropped out of school, the probability of re-entry is negatively related to the signal parents obtain about the child's opportunities in the unskilled labour market. The intuition for the result is quite straightforward: the child is given a second chance at school if the signal about alternative opportunities is even worse than the signal obtained at school. What is of interest for the empirical part is the analysis of the determinants of the thresholds in equation (20) and (23).

3.3 Comparative statics and predictions for the empirical results

The main conclusion we draw from the theoretical framework described in the present section is that during school families *learn* about their child's talent in alternative occupations and *react* to this information through the choice of educational investment. We identify a set of parameters that characterize the family's socio-economic status as an important determinant of drop-out behaviour. The model also shows that there is the possibility that a parent may rationally decide to have his child return to school after spending a period outside as a reaction to new information about his kid's talent. The next section will investigate whether this is empirically relevant and whether the determinants of dropping out and re-entry reflect the determinants highlighted in the model we presented. Importantly for the discussion of our empirical results we consider the comparative statics for the equilibrium threshold for dropping out of school and the equilibrium threshold for re-entry. Results can be stated about the effects of the relative precision of the signals in the parent's information sets, providing some characterization of the effect of cultural heterogeneity on educational choices.

Dropping out after the first period signal was characterized by the threshold $\tilde{z}_1(.)$ defined in eq. (20). By using the definition for K_1 and K_2, this latter expression can be rewritten as:

$$\tilde{z}_1(.) = \frac{w_0 + T_s}{w_s} + (1 + \frac{\sigma_s^2}{\sigma_\varepsilon^2})(\frac{w_0 + T_s}{w_s} - \theta) \tag{24}$$

We immediately obtain the following comparative statics results:

$$\frac{\partial \tilde{z}_1(.)}{\partial w_s} < 0, \quad \frac{\partial \tilde{z}_1(.)}{\partial w_0} > 0, \quad \frac{\partial \tilde{z}_1(.)}{\partial T_s} > 0, \quad \frac{\partial \tilde{z}_1(.)}{\partial \theta} < 0. \tag{25}$$

A larger perceived skill premium (larger w_s or lower w_0) intuitively makes families' choices less selective, reducing the threshold and allowing their kids to continue even in the face of a bad school signal[11]. Larger educational costs T_s make families more selective and increase the probability of

[11] Though intuitive, this results is interesting: raising trends (as recently observed in Oecd countries) in the skill premiums induce families to be more pushy in the career choice of their kids diluting talent allocation in the skilled sector, see D'Amato and Mookherjee, (2007).

school abandonment. A larger θ i.e. a better prior about the child's talents in the skilled sector makes families less selective and reduces the probability of dropping out.

Those results are quite intuitive and will be used extensively in the interpretation of our empirical results. More subtle and interesting issues arise when it comes to the results about σ_ε and σ_s. Remember that σ_ε is the (inverse) precision of the family in assessing the kid's talents and perspectives in the skilled labour market, whereas σ_s is the (inverse) precision parents attach on the school grading process in the assessment of the same index.

Define $\rho = \frac{\sigma_s^2}{\sigma_\varepsilon^2}$ as a measure for the relative weight parents assign to the first period signal at school, given the option to recover a second signal at school relative to the prior. It is a measure of the cultural confidence parents have in making their decision based on the public assessment at school relative to their private assessment. We obtain

$$\frac{\partial \tilde{z}_1(.)}{\partial \rho} = \frac{w_0 + T_s}{w_s} - \theta \tag{25}$$

Therefore the derivative above is signed according to the sign of $w_0 + T_s - \theta w_s$, i.e. the effect of the relative precision of the signals depends on parameter values. To interpret the result above notice that $w_0 + T_s < \theta w_s$ implies that from the point of view of the family it is *ex-ante efficient* [12]to send the child to school (remember that θ is an unbiased estimate of α). As a consequence the model predicts that, ceteris paribus (i.e. with given perceived skill premiums and opportunity costs of education), families with larger cultural capacity (larger ρ) will be less selective with respect to kids for whom education is ex ante efficient (assessment based on θ). In other words dropping-out from school is less likely to occur in more educated families, conditional on the kid's talent being ex-ante assessed valuable for a skilled occupation[13]. Also notice that for given θ and perceived skill premium, the larger the costs of education, the more selective is the family.[14]

[12] With risk neutrality sending the child to school is ex-ante efficient when, conditioning on the prior information θ, the expected return of education is above its opportunity costs (equal to the sum of direct education costs and foregone returns in the alternative occupation).

[13] One implication of the model is that less precise school signals (larger σ_s) make families less strict and dilutes the average ability level of kids continuing in education.

[14] It is important to note that in a more imperfect (less intergenerationally mobile) world than the one considered in this model, children in educated families would have better perspectives in the skilled sector, ex ante and independently of their school attainment, due for example to inheritance of (skilled) jobs. In terms of our model we can capture this effect by

Therefore both the quality of the public signal (lower quality of the grading process at school *z*) and the private one (quality of the parental assessment θ) affect drop-out rates. Summarizing, the implication of the model for our empirical specification of the probability of dropping out of school is that both the impact of school quality and the impact of parental cultural capacity has a non linear effect on drop-out rates, interestingly the model captures different channels through which the family cultural capacity and the family's economic capacity influence the education investment decision. This feature allows us to introduce parental education and measures of families permanent income as separate determinants of the investment choice.

The decision to re-enter after a period out of school was characterized by the threshold in eq. (23). Straightforward algebra shows that

$$(26) \qquad \frac{\partial \tilde{\phi}_2(.)}{\partial w_0} < 0, \quad \frac{\partial \tilde{\phi}_2(.)}{\partial T_s} < 0, \quad \frac{\partial \tilde{\phi}_2(.)}{\partial \theta} > 0, \quad \frac{\partial \tilde{\phi}_2(.)}{\partial z_1} > 0.$$

In words, re-entry is less likely when the family expects the child to face better opportunities in the unskilled labour market (larger w_0) and/or larger costs of education (larger T_s). A larger prior about the child's opportunities on the skilled market (θ) and a better grade z_1 in the first period at school increase the probability of re-entry. Moreover, children from richer families are more likely to re-enter conditional on drop-out.

This conclusion is reinforced when we consider the comparative statics with respect to the precision of the signals. To compute $\frac{\partial \tilde{\phi}_2(.)}{\partial H}$ rewrite equation (23) using the definition of $\tilde{z}_1$ from equation (20), we get:

$$(26) \qquad \tilde{\phi}_2(.) = w_0 + (z_1 - \tilde{z}_1) w_s \frac{K_1(1-K_2)}{H}$$

evaluated at $z_1 \leq \tilde{z}_1(.)$. Since *H* does not enter into (20) we get

$$(27) \qquad \frac{\partial \tilde{\phi}_2(.)}{\partial H} > 0$$

---

noticing that in more educated families, children would start their school career with a larger θ: the effect of cultural capacity would reinforce the standard economic argument based on wealth constraint: the sign of equation (25) is more likely to be negative in more educated families, i.e. given z_1 children in more educated families would be less likely to drop out.

Remember that from equation (10), *H* measures the relative weight placed on the signal received about opportunities on the unskilled labour market relative to the weight placed on the prior by each family. To derive empirical predictions from this latter comparative statics result consider that in the case of opportunities in the unskilled market, the weight put on the prior (relative to a signal) is likely to be larger for a parent working in the unskilled sector: there is less to learn about the unskilled opportunities facing his child for an unskilled worker than for a skilled one. We can conclude that, conditional on school abandonment, the probability of re-entry is larger for skilled families than for unskilled ones. In a specific sense the link between parental status and children's educational choice exhibits persistence driven not only by wealth constraints and financial market imperfections (larger costs of education in less well to do families) but also by the cultural and social ability of different families to process signals about their children's abilities in two alternative occupations[15].

Summarizing, in this section we have laid out a simple model of parental decision making about educational choices when the process involves learning about the child's abilities and job opportunities in two alternative occupations (skilled and unskilled). The model shows that cultural capacity and economic status of the family affect the probability of being a drop out, along with traditional determinants such as the skill premium as perceived by the decision maker. In adherence with the aim of the empirical analysis the model also shows that conditional on being a drop out the probability of returning back to school is affected by the same variables in an intuitive way. With this theoretical framework in mind we now move to the empirical analysis.

[15] As already noticed this does not mean that persistence is rooted on efficiency grounds in our model. Of course, along with financial market imperfections, cultural bias and other forms of intergenerational persistence of job allocation may have important social costs when we consider the problem of allocating talent to jobs. We do not address this point here.

4. Empirical Analysis: data and descriptive statistics

On the basis of the theoretical model and general considerations offered on the likely causes of dropping out behaviour offered above, in this section we report the results of estimating an empirical model of school-leaving and permanent dropping out behaviour amongst young people from Salerno and province.

Data used in our analysis are drawn from a survey carried out in 2004 by CELPE (Centro Interdipartimentale di Economia del Lavoro e di Politica Economica, University of Salerno) on behalf of the Province of Salerno, and supplemented in 2005. The aim was to better understand the magnitude and determinants of the upper secondary school dropout phenomenon in the province of Salerno of those aged 15-18.

Table 4 reports some information, provided by the local CSA (Centro per i Servizi Amministrativi di Salerno), on the distribution of enrollment in the high secondary public school in the Province during the schooling year 2003-2004.

Table 4: Number of enrolled in the school and students interviewed in the survey

Type of school	Number of enrolled	Enrolled(%)	Number of schools in the sample	Number of students interviewed
Scientific	15000	23.2	5	135
Vocational	13100	20.3	4	125
Technical- commercial	11000	17.1	4	104
Technical- industrial	8.000	12.4	3	60
Classical	6000	9.3	2	56
Psycho-pedagogical	6000	9.3	2	56
Others (ITG,LL,IA)	5400	8.4	1	57
Total	64500	100	21	593

Source: Data provided by CSA of Salerno

We constructed our sample on the basis of this distribution. As regards students regularly attending school, we employed a two stage procedure in which first we extracted 21 of the total of 85 schools in the province taking into account the distribution of schools across geographical areas and school type. Secondly around 600 students were randomly drawn from the 21 schools' records.

As regards the dropouts (and returnees) in our sample, a list of names of persons identified as "dropouts" by schools was provided by the local CSA and contained information on around 900

individuals born between 1987 and 1989 inclusive. From these, 178 individuals were interviewed in June/July 2004 (126 "permanent" drop-outs and 52 "returnees")[16]. In early 2005, a refreshment sample of dropouts was added through renewed requests for dropout lists from all the schools and further attempts were made to contact the members of this enlarged group of dropouts (1326 as opposed to 900) including also dropouts born in 1990.

Thus, our empirical sampling base consisted of a sample of students along with the entire population of dropouts identified by the CSA. In other words, relatively speaking, dropouts (including returnees) are deliberately overrepresented in our sample. This was necessary in order ensure a group of dropouts sufficiently large to compare with students[17]. Of this group of 'officially identified' dropouts a further one-fifth (or 274) 'dropouts' were excluded because at the time of the interview, there were discovered to not have interrupted their education (i.e. they were misclassified). Once those who were addresses were incorrect and so on were excluded , a group of 380 dropouts were identified, amongst whom 51 (or 13.4%) refused to be interviewed.

To summarize, the final full sample consists of 918 individuals, 593 of which are students, 206 were permanent dropouts and 123 returnees. Returnees are defined as those who at the time of the interview, were actually attending full time educational institutions, having spent some time out of school. Given that it is likely that 'returnees' were more likely to be interviewed – for fairly obvious reasons it is natural that they would have been more likely than dropouts to be contacted at the sampling stage – the substantial numbers of returnees in our sample (around one-third of all 'dropouts') is likely to be an overestimate of the numbers of returnees in the dropout population. There is no reliable way of verifying this, however, for our purposes, the main concerns are:

a) that those identified as dropouts had actually interrupted their education; and,
b) those identified as returnees had actually interrupted and then returned to education – as opposed simply to have transferred form one school to another.

The procedures adopted leave us confident that this was largely achieved. To further ensure that this was the case, a further barrier for inclusion in the empirical analysis was adopted. In Italy, if one fails a scholastic year twice, one must in any event change educational institution (D.L 297/94)[18]. Thus, young people in the sample who reported having repeated a specific school year twice were excluded from the sample. Individuals were also excluded from the sample if they never entered upper

[16] The questionnaires for students and dropouts can be found at www.unisa.it/CELPE

[17] On the basis of the data provided to us by the CSA and table 2 above, "dropouts" thus comprised around 2.7% of the registered school student population in Salerno Province

[18] The law includes some exceptions and modifications of this basic rule, however, effectively this rule is

secondary education. In addition to the exclusion of individuals who had missing information on specific explanatory variables this lead to an empirical sample of 839 observations on 589 students, 151 dropouts and 99 returnees. Descriptive statistics on the final sample are reported in Table 5.

Table 5: Descriptive Statistics for variable used in the analysis, by final status.

	Students n = 593		Returnees n = 99		Dropouts n = 151	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Ownhome	0.812	0.391	0.713	0.454	0.797	0.403
no. Of family members (nat. log.)	1.466	0.231	1.460	0.232	1.530	0.267
no. Of durable goods in HH	2.713	0.229	2.707	0.219	2.500	0.330
father working	0.919	0.274	0.870	0.337	0.776	0.418
mother working	0.407	0.492	0.463	0.501	0.343	0.476
both parents working	0.379	0.485	0.380	0.488	0.238	0.427
Both parents living at home	0.917	0.276	0.889	0.316	0.853	0.355
Mother with incomplete secondary or vocational qualification	0.440	0.497	0.463	0.501	0.476	0.501
Mother acadmeic secondary or tertiary education	0.452	0.498	0.306	0.463	0.154	0.362
Male	0.560	0.497	0.602	0.492	0.594	0.493
Firstborn	0.418	0.494	0.324	0.470	0.315	0.466
Parents read to child	0.105	0.307	0.056	0.230	0.021	0.144
Parents helped with homework	0.244	0.430	0.176	0.383	0.140	0.348
Obese	0.083	0.276	0.093	0.291	0.245	0.431
health problems	0.027	0.163	0.019	0.135	0.056	0.231
Private lessons	0.385	0.487	0.315	0.467	0.210	0.409
Parents attended school meeting	0.603	0.490	0.537	0.501	0.483	0.501
Child interviewed at school	0.331	0.471	0.370	0.485	0.231	0.423
work experience at school	0.056	0.230	0.046	0.211	0.021	0.144
Failed a year during middle school	0.015	0.123	0.019	0.135	0.105	0.307
Failed a year during secondary school	0.088	0.284	0.426	0.497	0.462	0.500
mother sec/tertiary education interacted with failing a year	0.024	0.152	0.139	0.347	0.070	0.256
suspended during school	0.022	0.147	0.028	0.165	0.084	0.278
regular truant	0.051	0.220	0.102	0.304	0.371	0.485
Liceo	0.323	0.468	0.111	0.316	0.021	0.144
Population density	7.099	6.526	5.172	5.078	6.349	6.015
Unemployment rate	11.970	2.014	11.325	2.034	11.447	1.456
large local labour market	0.730	0.444	0.694	0.463	0.671	0.471

Source: Salerno Province Dropout Survey

5. Empirical Analysis: Model and Results

The theoretical framework outlined above leads fairly naturally to the implementation of a two equation probit model. It also raises some quite interesting identification issues. In this section these issues will be discussed and the results of the empirical implementation presented. The discussion above thus provides a framework within which to look at the factors influencing school dropout and subsequent return behaviour. It also provides some testable predictions which are discussed below. However, it is evident that the framework outlined above is very simple and is unlikely to capture, as it stands, all the factors influencing dropout behaviour. Moreover, the link between the framework and its empirical implementation requires further assumptions on the link between observed variables and their unobserved counterparts in the framework. The extent to which the framework captures essential elements of the empirical results is the focus of the discussion of the results. This allows us to suggest missing elements which would be needed to move towards a more complete model of dropout behaviour.

In any event, on the basis of the simple framework outlined above, a parent will decide to take its child out of school at time t_1 iff:

$$L^* \equiv \frac{w_0 + T_s}{w_s} + \left(1 + \frac{\sigma_s^2}{\sigma_\varepsilon^2}\right)\left(\frac{w_0 + T_s}{w_s} - \theta\right) - z_1 > 0 \qquad (28)$$

Where L* may be thought of the (unobserved) tendency to initially leave school. Similarly, rearranging (22), the child will permanently drop out at time t_2 iff:

$$D^* \equiv \frac{w_0 + T_s}{w_s} + \frac{H(\phi - w_0)}{w_s} + \left(1 + \frac{\sigma_s^2}{\sigma_\varepsilon^2}\right)\left(\frac{w_0 + T_s}{w_s} + \frac{H(\phi - w_0)}{w_s} - \theta\right) - z_1 > 0 \qquad (29)$$

Where D* represents the (unobserved) tendency to permanently drop out.

The empirical counterparts of the conditions for school leaving and permanent dropout are, assuming standard normally distributed error terms:

$$L^* = X_1\beta_1 + e_1 \qquad e_1 \sim N(0,1) \qquad (28')$$

and

$$D^* = X_2\beta_2 + e_2 \qquad e_2 \sim N(0,1) \qquad (29')$$

Of course we only observe whether or not the person permanently drops out and/or leaves school initially, so we need the additional relations:

(30) L = 1 iff L*>0

and

(31) D = 1 iff D*>0

Furthermore, D (=0,1) is only observed if L*>0.

Very clearly the factors influencing L and D are very similar. The only difference between (28) and (29) lies in the additional terms involving $H(\phi-w_0)$. This has several implications:

i) in practice, the error terms, e_1 and e_2 are likely to be correlated which implies the estimation of a censored bivariate probit model;

ii) The only way a person can return to school at t_2, having left at t_1 is if $\phi<w_0$ – the signal from the labour market is worse than the parent's *a priori* expectation;

iii) a censored bivariate probit model of this form requires at least one exclusion restriction in order to identify β1, β2 and the correlation between the error terms, ρ_{12}, either through functional form or, more usually, by excluding at least one variable from (28') included in (29'). Since (29') is simply (28') with some additional terms, it is not at all obvious how one would exclude variables from (28') included in (29') thus creating an issue of identification.

We reports results using alternative approaches to identification. In order to illustrate the restrictions imposed, it helps to rewrite the model slightly:

(28'') $$L^* = X\beta_1 + Z\alpha_1 + e_1$$

(29'') $$D^* = X\beta_2 + Z(\alpha_1 + \alpha_2) + e_2$$

(32) $$\text{Cov}(e1, e2) = \begin{pmatrix} 1 & \rho_{12} \\ & 1 \end{pmatrix}$$

Note that X and Z are the same in both equations. Notionally, the distinction between X and Z is that the variables in X only have an impact on D* through their impact on L*, whereas the Z variables may

affect D* both through their impact on L*, but also in their impact on $H(\phi\text{-}w_0)$. Clearly as it stands the model is not identified, nor is it consistent with the underlying theoretical framework.

The results reported here involve in turn the following identifying restrictions:

(i) $\rho_{12} = 0$ i.e. two univariate probit equations.

(ii) $\beta 1 = \beta 2$ i.e. The X variables have the same impact on the tendency to leave school as they do on the tendency to permanently drop out.

(iii) $\alpha_1 + \alpha_2 = 0$ for some elements of Z. i.e. For at least some of the variables affecting both L* and $H(\phi\text{-}w_0)$, the overall effect cancels out.

Our preferred approach is the adoption of restriction (ii)[19]. It is closest in spirit to the theoretical framework although it does involve the exploitation of the non-linearity of the model to obtain identification. Restrictions of the type (iii) are more conventional in empirical work, however, in the current context, actually appear more arbitrary than (ii).

Finally, in order to arrive at the equations actually estimated, one needs to posit a relation between the 'explanatory' variables in (28) and (29) and their empirical counterparts, X and Z. We make no attempt to formally estimate the parameters of (28) and (29), rather we identify factors likely to influence elements of the underlying theoretical model. Above-all it is important to identify variables likely to affect $H(\phi\text{-}w_0)$.

Specifically, the variables included are:

- living in family owned accommodation, the number of components of the family[20] and the number of durable goods[21] in the household are included to represent permanent income. These variables are posited to enter only through their influence on the school leaving decision – specifically through their impact on financial constraints of further education facing the family – here represented by T_s - and/or, à la Heckman, through their impact on cognitive ability, on z_1;
- The mother's level of education, a fairly crucial variable in as much as it is expected to affect both K and H[22] ;

[19] Derivation of the formal identification of the model in this instance goes beyond the scope of this paper, however, that the model is identified is fairly obvious if one thinks in terms of the likelihood function to estimate. See Maddala (1983) and particularly Cameron & Trivedi (2005) for discussions of the issues involved.

[20] Obviously inversely related to family permanent income.

[21] Individuals were asked whether they had access to a variety of durable goods (e.g. TV, computer, car and so on). These were simply summed to provide an indicator of family wealth.

[22] Note however that this by no means necessarily implies correlation between K and H.

- whether one or other (or both) of the parents are working, intended to reflect both family income (affecting the decision at t_1), and, conceivably also H;
- Whether both parents are living at home, the amount of attention and help given by parents to their children during the early scholastic period and whether the child is firstborn, all of which have been identified in the literature as factors influencing scholastic performance, presumably through the impact on cognitive ability and so on z_1;
- Obesity and the possession of permanent health problems are both likely to affect scholastic performance, z_1;
- Experiences at school – whether the child and/or parents attended regular meeting with teachers at school, whether the child undertook work experience at school, whether the child had private lessons outside school;
- Whether the child failed a year – clearly related to z_1 but also feasibly reflecting a series of underlying factors returned to below – and whether the child was suspended from school and whether they frequently played truant;
- The type of school, although not formally part of the theoretical framework, is also, given its obvious importance is also included, although in a limited way i.e. whether the child attended a Liceo, the more academic stream;
- Finally, three variables are included to reflect local conditions – population density, the local unemployment rate and the size of the local labour market are all included and likely to affect the direct and opportunity costs of education as well as influence the labour market experiences of school leavers[23].

The results are presented in table 6. One might observe that the estimation has quite a lot of explanatory power. In particular, the two pseudo-r^2 statistics in the uncorrelated probits, each of .32, are rather high for this type of estimation. One might also note that both bivariate probits reject restriction (i) of no correlation between the equations. In order to arrive at the equality restrictions (restriction (ii)), a priori expectation regarding variables which were thought not to affect elements of $H(\phi\text{-}w_0)$, in one case modified by observation of the independent probit equations, were used to arrive at the reported form. Although a Likelihood ratio test of the restrictions as a whole is not possible since the unrestricted model is not identified, LR tests of the single restrictions are never rejected at a .05 significance level. The model estimated with exclusion restrictions produces rather similar results to the

[23] These data are defined according to the national statistical agency's (ISTAT) definition of local labour markets. On this criteria, Salerno province is divided into 20 local labour markets. Data for four other bordering local labour markets were also included due to some students not actually living in Salerno Province itself.

bivariate probit model with equality restrictions. The restrictions themselves were largely determined by empirical observation as opposed to a priori consideration. Variables were excluded which, on the basis of the uncorrelated probit model appeared to have little or no impact on the decision to drop out permanently.

The first group of explanatory variables in the table relate to family permanent income. Curious amongst these is home ownership which very clearly has a different effect on permanent dropping out than it does on initial school-leaving[24]. One possible (albeit not very convincing) interpretation might take the form of some sort of 'Oswald' home ownership effect on the likelihood of finding employment. Otherwise, the permanent income variable perform as expected although they are not always statistically significant, they always have the expected sign. The positive coefficient on the 'mother working' dummy also invites comment. Note that this is measuring the effect of **only** the mother working since 'both parents working' is also included as a variable.

Of particular significance in terms of the theoretical framework, maternal education is also, net of the effect of family permanent income as measured here, important in explaining behaviour. Specifically, higher levels of mothers' education is negatively (and statistically significantly) associated with school leaving over and above the measures of family income, as predicted by the theoretical framework. Two possibly connected reasons can be offered for the inclusion of mother's, rather than father's, education. First, mother's education is more clearly associated with school-leaving[25] than is father's education, and, second, mother's education is likely to be less closely related to family permanent income. As noted above we interpret this in terms of the effects of family culture operating through the relative precision of the initial school signal. The effect seems to disappear with regard to dropping out (in the correlated probits) which is also reasonable within the framework outlined above.

As regards individual characteristics, it is curious that being male has a negative coefficient albeit not statistically significant, since a much higher proportion of young males leave school than young females. This implies that they tend to have more of characteristics associated with school leaving (and

[24] I both bivariate models the difference in coefficients is statistically significant at $p<0.01$.
[25] See the single (full) probit estimations reported in the appendix.

Table 6: Probit models of school leaving and permanent dropout

	(i) uncorrelated probit				**(ii) bivariate probit equality restrictions**				**(iii) bivariate probit exclusion restriction**			
	School leaving (t_1)		dropout (t_2)		School leaving (t_1)		dropout (t_2)		School leaving (t_1)		dropout (t_2)	
	Coeff.	Std. Error	Coeff.	Std. Error	Coeff.	Std. Error	Coeff.	Std. Error	Coeff.	Std. Error	Coeff.	Std. Error
Ownhome	-0.10	0.135	**0.58**	0.234	-0.08	0.133	**0.52**	0.207	-0.07	0.134	**0.51**	0.193
no. Of family members (nat. log.)	0.17	0.246	*0.88*	0.479	*0.34*	0.192	*0.34*	0.192	0.17	0.241	*0.68*	0.353
no. Of durable goods in HH	**-0.50**	0.223	**-1.34**	0.408	**-0.67**	0.174	**-0.67**	0.174	**-0.47**	0.220	**-0.90**	0.366
father working	-0.12	0.227	-0.25	0.379	-0.12	0.224	-0.16	0.335	-0.16	0.213	-	-
mother working	**0.94**	0.317	0.33	0.545	**0.94**	0.314	0.02	0.464	**0.85**	0.308	-	-
both parents working	**-0.81**	0.337	-0.64	0.590	**-0.80**	0.334	-0.37	0.508	**-0.79**	0.323	-	-
Both parents living at home	0.09	0.211	0.12	0.370	0.10	0.160	0.10	0.160	0.13	0.198	-	-
Mother with incomplete secondary or vocational qualificat	**-0.51**	0.145	**-0.50**	0.243	**-0.51**	0.144	-0.27	0.220	**-0.49**	0.144	-0.19	0.217
Mother acadmeic secondary or tertiary education	**-0.62**	0.180	-0.56	0.386	**-0.61**	0.178	-0.28	0.384	**-0.60**	0.179	-0.14	0.346
Male	-0.18	0.117	-0.17	0.221	-0.14	0.092	-0.14	0.092	-0.16	0.117	-0.05	0.175
Firstborn	-0.17	0.115	-0.09	0.215	-0.12	0.090	-0.12	0.090	-0.17	0.115	0.05	0.173
Parents read to child	*-0.38*	0.224	0.11	0.626	-0.28	0.192	-0.28	0.192	*-0.41*	0.224	0.31	0.476
Parents helped with homework	-0.16	0.142	0.15	0.286	-0.03	0.117	-0.03	0.117	-0.13	0.142	0.30	0.231
Obese	**0.46**	0.166	**0.94**	0.306	**0.53**	0.131	**0.53**	0.131	**0.46**	0.165	**0.59**	0.276
health problems	0.62	0.311	0.85	0.607	**0.57**	0.252	**0.57**	0.252	*0.59*	0.310	0.39	0.525
Private lessons	**-0.26**	0.128	-0.14	0.244	*-0.19*	0.101	*-0.19*	0.101	**-0.26**	0.128	-0.04	0.197
Parents attended school meeting	-0.03	0.113	-0.03	0.207	-0.03	0.087	-0.03	0.087	-0.02	0.113	-0.02	0.167
Child interviewed at school	-0.09	0.121	-0.35	0.228	-0.13	0.095	-0.13	0.095	-0.07	0.120	-0.14	0.186
work experience at school	-0.23	0.285	-0.63	0.588	-0.30	0.227	-0.30	0.227	-0.25	0.285	-0.55	0.455
Failed a year during middle school	0.47	0.298	0.48	0.505	0.46	0.297	0.31	0.470	0.49	0.297	0.16	0.448
Failed a year during secondary school	**0.96**	0.152	0.12	0.241	**0.97**	0.150	-0.21	0.251	**0.96**	0.151	-0.34	0.223
mother sec/tertiary education interacted with failing a year	**0.68**	0.290	-0.01	0.472	**0.65**	0.286	-0.26	0.442	**0.69**	0.289	-0.48	0.396
suspended during school	0.31	0.282	0.86	0.576	*0.38*	0.224	*0.38*	0.224	0.30	0.280	0.57	0.468
regular truant	**0.74**	0.170	**1.05**	0.264	**0.75**	0.137	**0.75**	0.137	**0.75**	0.168	**0.51**	0.256
Liceo	**-0.80**	0.171	*-0.95*	0.494	**-0.74**	0.156	**-0.74**	0.156	**-0.82**	0.164	-	-
Population density	*-0.02*	0.011	*0.04*	0.022	*-0.02*	0.011	**0.04**	0.019	*-0.02*	0.011	**0.04**	0.016
Unemployment rate	**-0.13**	0.038	-0.05	0.075	**-0.12**	0.038	0.00	0.066	**-0.12**	0.037	-	-
large local labour market	**0.49**	0.160	0.02	0.271	**0.47**	0.159	-0.10	0.245	**0.48**	0.156	-0.16	0.195
Intercept	**2.30**	0.771	**2.63**	1.471	**2.34**	0.672	**1.84**	0.886	**2.09**	0.760	*1.84*	1.019
N	839		250		839		250		839		250	
Log-Likelihood	-348.93		-114.57		-467.84				-465.41			
Pseudo R-squared	0.32		0.32		-				-			
r	-		-		**-0.56**				**-0.82**			
Chi-square test of independence	-		-		**8.67**				**6.76**			

note: coefficients which are significant at $p < .05$ are reported in **bold**. Coefficients which are significant at $.10 > p > .05$ are reported in *italics*

dropout) than do their female counterparts. In line with findings in the literature, being the first born is negatively correlated with leaving school. Youthful obesity is clearly positively correlated with school leaving[26]. Parental attention to children is negatively associated with school leaving although this is not statistically significant, as is the variable indicating whether the person had private lessons while at school – this may be explained by the likely duality in the motivation. On the one hand, children are more likely to have private lessons if their parents care more about their performance at school. On the other, private lessons are also likely to be associated with poor performance at school.

Turning to experiences at school, it will be observed that failing a year, at secondary level[27], is an important determinant of initial school leaving. Obviously this is directly related to the school signal, z_1. The coefficient is positive and strongly statistically significant. The effect on permanent drop out is rather different however. In this case the coefficient is close to zero and, for secondary failure, is actually negative. Although the coefficient is not statistically different from zero, the difference in the effect of failure at secondary level on initial school-leaving and permanent dropping out very clearly is[28]. The implication is that the kind of person who is likely to fail at school may also tend to have other characteristics which cause him or her to fail, and/or have expectations which are unreasonable and/or inaccurate concerning, the unskilled labour market. Moving beyond the theoretical framework, an alternative plausible explanation is that failure at school is an indicator of mismatch between schools and students. The inclusion of an interaction term between high maternal education and failure at school is also positive and statistically significant. Again, this takes us beyond the simple model outlined above and points towards a slightly more complex relationship between parental education and the school leaving signal than is suggested by the underlying model. A plausible story here would be that better educated parents are more likely to interpret the school signal as an issue of mismatch between school (or teacher) and pupil (their child). If better educated parents are more likely to interpret their child being failed at school as a problem with the teacher or the school, that is if z_1 is interpreted as an indicator of school quality rather than the students' intrinsic potential, they are more likely to react to a poor school signal by removing the child from the school and seeking an alternative. This also fits with the non significant interaction coefficient in the permanent drop out equation.

[26] The objection might be raised here concerning possibly endogeneity of obesity. Young people leaving school may become depressed etc.. and so obese. We would argue that this is unlikely to be the case – one does not become obese overnight and secondly tests of the issue in a companion paper suggest this not to be the case.

[27] It will be recalled that those reporting their last school as a middle school were excluded form the analysis, which means that those dropping out before arriving at secondary school are excluded, with the obvious consequences on the effect of middle school failure compared to secondary failure.

[28] Statistically significant at $p < .01$.

Implying ceteris paribus, a stronger negative reaction to bad labour market signals. There are various ways that this type of phenomenon might be captured in a more sophisticated theoretical model. However, in this paper we limit ourselves to identifying the issue.

Discontinuous attendance is also important in determining initial school leaving. This may indicate in part relatively low ability and/or low utility value of school (a big T_s). The effect of truancy on school leaving and permanent dropout are very similar (and indeed in the second model, the coefficient is constrained to be equal across equations). Suggesting that the effect is only operating through its effect on performance and the consumption value (cost) of education. The estimation also includes a variable representing attendance at a more academically oriented school. Again, in order to properly account for this type of effect one would need a more articulated model, specifically with more than one type of school.

The final group of variables reflect local conditions. Based on indicators for local labour markets, these are intended to capture geographical variations in factors affecting school attendance. The local unemployment rate is interpreted as an (inverse) indicator of the opportunity cost of education. Indeed the coefficient in the school-leaving equation is always negative and statistically significant. In the dropout equation the effect disappears suggesting a counterbalancing impact of the unemployment rate through its effect on the reaction to local labour market experiences. A not unreasonable possibility, remaining within the framework, is that the unemployment rate is positively related to H. The higher the unemployment rate the more likely that unexpectedly poor labour market experiences are likely to be interpreted as reflecting real possibilities rather than just bad luck. This is stretching the framework somewhat, and is certainly not the only possible explanation, however, it does suggests that the finding in itself does not invalidate the framework. A larger local labour market has a positive impact on school leaving which again disappears when permanent dropout behaviour is estimated. Here again one would expect ceteris paribus that the larger the local labour market the greater the opportunities for the unskilled and so the higher the opportunity cost of education. The disappearance of the effect in the permanent dropout equation might once again be interpreted in terms of its effect on H, however as before one would not want to push this argument too far. The population density which has a negative impact on school leaving and a positive impact on permanently dropping out is probably more plausibly explained by a mismatch argument than by the framework outlined here. Specifically, a higher population density is likely to be associated with a wider variety (and number) of schools nearby. A greater variety of schools is likely to increase the probability of a ‘good’ initial match between school and student, thereby reducing the likelihood of initial school-leaving. At the same time,

it will reduce the cost (or increase the benefits) of returning to education in the event that one does leave. This does go very much beyond the simple model outlined above.

Thus, the results presented here suggest that the theoretical framework goes some way to providing a plausible model of school-leaving and return. In particular, it provides an explanation for the importance of parental education in encouraging educational participation which is independent of income and which is verified in the results reported above. However, the framework does not account for all observed features of dropping out considered here. The most notable gap in the model relates to issues of educational mismatch. The empirical results strongly suggest that there is a role for educational mismatch of one type or another which cannot be accounted for by the theoretical framework, given the existence of only one type of school. This would suggest that the incorporation of different types of school in the framework would be desirable.

5. Conclusions

Dropping out of school is at the centre of current debates in the EU concerning long growth and competitiveness. In line with the Lisbon Strategy, in the last decade the Italian authorities have built up a complex institutional networking system aimed at reducing school drop out and increasing the duration of educational participation.

This paper proposes a simple theoretical framework and an empirical analysis aimed at throwing greater light on this phenomena in Southern Italy. In the theoretical model and its empirical implementation, a two-stage decision process is hypothesised in line with observed behaviour which shows that many young people 'dropout' of school only to return subsequently. Such behaviour is not compatible with a standard full information human capital model.

The results confirm the role of both economic and cultural capacity of the family of origin in separately shaping observed choices about drop-out and return to school by individuals in our sample. Interestingly we find that whilst poor performance at, and low attachment to, school – measured by repetition of the school year through end of year failure and attendance records - is a key determinant of initial dropping out, the former does not seem to affect subsequent return to education.

The results reported here also provide substantial evidence of mismatch between school and student as determining behaviour. The answer then to the question in the title of this paper, interpreted in its normative sense, is no: the process of allocation of talents to school tracks is subject to many trial errors and revisions by families and many of those who leave school return to it. The point is whether this process of allocation of talents to school tracks is efficient and whether we should see review of family decisions as the natural outcome of the learning process. Specifically is return to school an indicator of allocative inefficiency? It is argued there that much depends on the determinants of school re-entry behaviour. We do not address normative aspects in the model (indeed every family equilibrium decision is individually rational, conditional on the information set they act upon so that the allocation of the child to the school track is interim Pareto-efficient) and can only base our judgment on the suggestive empirical evidence. This latter however allows us to shape a judgment on this issue: a policy maker should not necessarily be happy after observing a large fraction of students return to school after a period out. Since cultural and economic capacity matter a lot for the actual review of the decision, the comparison of students' irregular careers among different education systems (across regions or across countries) should be used as an indication of the social cost of the mistakes associated to irregular careers. Moreover since there is evidence of a role for the mismatch between a child and the school in

influencing the decision, indicators of irregular careers should be adopted by decision-makers to assess the quality of education provision by public agencies and to allocate public resources in this area. Clearly further work is needed on this point to assess public school intervention in education and to a more effective policy design.

Our findings, suggest that students' decisions as whether to leave education definitively or to return back after a period in the 'real world' are also due to an information gathering process by parents about children's attitudes, expected wages in the unskilled market and school quality. In other words, to maximize the investment on education parents need to evaluate their options, in terms of the child's educational choices, on the basis of economic criteria. However, given that information is imperfect and costly to improve, mistakes are made and additional information collected may cause parents to revise their previous decisions. Consequently, one policy implication of our analysis is that efforts should be made to reduce these information costs. To help children make their best educational choice, educational guidance and counselling services from primary school onward should be implemented. Programmes of career guidance –career orientation, academic and occupational orientation with high and post-secondary schooling, work-based learning and skills development- should enhance the capacity to evaluate work opportunities. In association with the reorganization of the vocational education and training system, this is likely to reduce the gap between formal education and training and the world of work. Ensuring school quality calls for a broader education system reform encompassing school structure, academic organization and social organization.

References

Acemoglu D. and J. Angrist (2001), "How Large Are Human Capital Externalities? Evidence from Compulsory Schooling Laws," *NBER Macroannual*, pp. 9-59.

Akerlof G. and R. Kranton (2002), "Identity and Schooling: Some Lessons for the Economics of Education" *Journal of Economic Literature*, No. 40, Vol. 4 (December), pp. 1167-1201.

Angrist J. D. and A. Krueger (1991), "Does Compulsory School Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics*, Vol. 106, No. 4, pp. 979, 1014.

Becker G. S. (1964), *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. New York: Columbia University Press.

Blundell R., B. Sianesi and L. Dearden (2003), "Evaluating the impact of education on earnings in the UK: Models, methods and results from the NCDS", *IFS Working Paper* W03/20, London.

Cameron, A.C. and P.K. Trivedi (2005), *Microeconometrics*, Cambridge University Press, Cambridge.

Card D. (2001), "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems," *Econometrica*, vol. 69(5), pp. 1127-60.

Cardoso A. R. and D. Verner (2006), "School drop-out and push-out factors in Brazil : the role of early parenthood, child labour, and poverty", *IZA DP* No. 2515, December.

Carneiro P. and J. J. Heckman (2002), "The Evidence on Credit Constraints in Post-Secondary Schooling," *NBER Working Paper* 90955, July.

Carneiro P and J. J. Heckman (2005), "The Evidence on Credit Constraints in Post-Secondary Schooling", *Economic Journal*, 112, pp. 705-734.

CELPE (2006), *L'analisi del fenomeno dei dropout nella provincia di Salerno*, Celpe, Salerno

Checchi D. and L. Flabbi (2006), "Intergenerational mobility and schooling decisions in Italy and Germany", mimeo.

Checchi D. and F. Zollino (2001), "Sistema scolastico e selezione sociale in Italia", *Rivista di Politica Economica*, n. 7/8, pp43-84;

COMMISSION OF THE EUROPEAN COMMUNITIES (2002), "*European benchmarks in education and training: follow-up to the Lisbon European Council"* COM(2002) 629 final, Brussels.

Eckstein Z. and K.I. Wolpin (1999), "Why Youths Drop Out of High School: the Impact of Preferences, Opportunities, and Abilities", *Econometrica*, vol. 67, n. 6, November, pp. 1295-1339.

Hanushek E. A., V. Lavy and V. L. K. Hitomi (2006), "Do students care about school quality? determinants of dropout behaviour in developing countries", *NBER Working Paper* 12737.

Heckman J. J. and P. Carneiro (2003), "Human Capital Policy", *NBER working paper* 9495.

Istat (2001), 14° Censimento Generale della Popolazione e delle Abitazioni, data on line at www.istat.it.

Istat (2003-2005), Popolazione residente nella Provincia di Salerno per età e sesso, anni 2003-2005, data on line at demo.istat.it.

Lee V.E and D. T. Burkam (2003), "Dropping Out of High School: The Role of School Organization and Structure", *American Educational Research Journal*, Summer, Vol. 40, No. 2, pp. 353–393.

Lochner L. and E. Moretti (2004), "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports", *American Economic Review*, Vol. 94(1), pp. 155-89.

Maddala G.S. (1983), *Limited Dependent and Qualitative Variables in Econometrics*, Cambridge University Press, Cambridge.

Ministero della Pubblica Istruzione (2006a), *La dispersione scolastica*, Direzione Generale Studi e Programmazione, Ufficio di Statistica, Rome.

Ministero della Pubblica Istruzione (2006b), *La scuola in cifre*, Direzione Generale Studi e Programmazione, Rome.

Mookherjee D. and D. Ray (2003), "Persistent Inequality", Review of Economic Studies 2003, 70, (2), pp. 369-393.

OECD (2000), *From initial education to working life: make transitions work*, Paris.

Oreopoulos P. (2007), "Do Dropouts Drop Out Too Soon? Wealth, Health, and Happiness from Compulsory Schooling," *Journal of Public Economics*, forthcoming.

Quintini G. and S. Martin (2006), "Starting Well or Losing their Way? The Position of Youth in the Labour Market in OECD Countries", *OECD Social Employment and Migration Working Papers*, n. 39, Paris.

Rumberger R. W. and S. L. Thomas (2000), "The distribution of dropout and turnover rates among urban and suburban high schools", *Sociology of Education*, 73(1), 39–67.

Sargent T.J., (1987), *Macroeconomic Theory*, New York: Academic Press.

Appendix 1

Here we show that after the first period of schooling, i.e. conditioning on z_1 alone and given that parents rationally anticipate $\tilde{z}_2$ and $\tilde{\phi}_2$, with no discounting, the choice between going on and dropping out is as given in the text.

The expected benefits from continuation at school is given by

$$V_2^s = \underbrace{[1 - F(\tilde{z}_2)]}_{\text{expected given success}} \underbrace{\{E[\hat{\alpha}_2|\hat{\alpha}_1]w_s - T_s\}}_{\text{benefits from continuing in the second period at school}} + \underbrace{F(\tilde{z}_2)E[w^u|z_1]}_{\text{expected benefits from continuing in the second period at school given failure in the}}$$

$$= [1 - F(\tilde{z}_2)]\{E[\hat{\alpha}_2|\hat{\alpha}_1]w_s - T_s\} + F(\tilde{z}_2)w_0$$

The expected benefits from dropping out is given by

$$V_2^d = F(\tilde{\phi}_2)\underbrace{\{[1-F(\tilde{z}_2)]\{E[\hat{\alpha}_2|\hat{\alpha}_1]w_s - T_s\} + F(\tilde{z}_2)E[w^u|z_1]\}}_{\text{expected benefits from dropping out given re-entry in the future}} + \underbrace{[1-F(\tilde{\phi}_2)]E[w^u|z_1]}_{\text{expected benefits from dropping out given no re-entry in the future}}$$

$$= F(\tilde{\phi}_2)\{[1-F(\tilde{z}_2)]\{E[\hat{\alpha}_2|\hat{\alpha}_1]w_s - T_s\} + F(\tilde{z}_2)w_0\} + [1-F(\tilde{\phi}_2)]w_0$$

Simple algebra shows that the equilibrium choice satisfies

$$V_2^s \geq V_2^d$$

whenever

$$E[\hat{\alpha}_2|\hat{\alpha}_1]w_s - T_s \geq w_0$$

That is:

$$[1-K_1(1-K_2)]\theta + K_1(1-K_2)z_1 \geq \frac{w_0 + T_s}{w_s}$$

equivalently

$$z_1 \geq \tilde{z}_1(.)$$

Characterizes equilibrium parental choice conditional on z_1.