Estimating human capital's returns in a segmented labour market through mixture regression model: an application for Italy

Michele Battisti[♣]

(Draft version, do not quote)[♦]

Abstract

The aim of this paper is to verify the presence of dualism in a wage regression for the Italian labour market.

In conformity with Labour market segmentation theory, in an economy there is a clear division between primary and secondary workers, that is not given by the characteristics of workers, but rather by the jobs. One standard way to assess this situation is by looking at the human capital's returns of comparable persons working in different segments. Tipically we expect respectively a flat and a steep profile of the experience/education – earnings relationship.

In the attempt to avoid arbitrary choices about the determination of the segments, we use the tool of mixture regression models for an endogenous determination of the segments. Our results for Italy are generally similar to those obtained by Cipollone (2001) with the switching regression framework of Dickens and Lang.
According to the results there is a strong demarcation between homogeneous workers, so that policy strategies have to consider these characteristics in the implementation of some labour market measures, like an improvement in the supply of human capital, or some ad hoc measures favouring some sectors of economy against other ones.

[♣] University Luiss Guido Carli, Rome. E – mail address for correspondence: mbattisti@luiss.it

1. Introduction

Between economists there has always been a debate about the presence of dualism either in theoretical[1], or in empirical field.

During the second half of this century a current of thinking used dualistic concepts to challenge the standard view of the labour market described by the mainstream economists[2]. The idea of a dual structure regards not only the labour market but the whole economy of a given country or region, for example characterized by the presence of both traditional and high – tech industries. Italy has often been seen as an interesting case study for this arguments but there are few econometric evidences to support this ides with some interesting exceptions like the study of Cipollone (2001) that is our benchmark in this field. We use the results for comparison but after this we depart from his analysis in several ways, like for example, the tool and the number of segments used.

This work tryes to measure the dualism between upper and lower segments of the labour market using the econometric tool of mixture regression. This allows us to avoid ad hoc definitions of the segments (a priori allocation of workers or sectors) that could crucially influence the results. The paper is organized as follows: in the second section there is a review about the concept of the segmentation in the economic theory and about the empirical methods to assess it, while in the third one there is the analysis about the presence of more than one segment in the wage regression for a group of workers and then the study ends with some conclusive considerations.

2. Labour market segmentation: theoretical survey and methods of investigation

According to Ryan (1984) we may talk of segmentation when we have the formation of "…*different groups of participants in the labour market which is evoked by the concepts of non competing groups and balkanisation*".

---

[1] Some considerations about this question are done by Pigou (1944), and for example Dunlop (1957) finds that these arguments dates back in Mill and Cairnes.
[2] See the surveys of Leontaridi (1998) and Guidetti (2001).

It is important to stress that differences between workers, either suscettible of economic evaluation like in the human capital theory[3], (that focus the attention on economic reasons, like training and returns of schooling and formation) or given by discrimination[4] are not a necessary element in Labour market segmentation so we could think that the Institutionalist approach explains different work rewards (wage, career and so on) with the difference in job characteristics. The institutionalist approach borns by the works of the American School during the '50s (Kerr (1954) and Dunlop (1957)): Dunlop emphasizes the differences between the labour market seen by the economists and the reality.

This difference is clearly described by Kerr (1954): the *wage market* is the place seen by the economists, where there is a single wage, fixed by the market; the *job market* is a zone geographically and industrially defined where the workers can move themselves freely, but there is scarce or absent mobility between job markets and a wage can be the same for two job markets but not for the entire labour market. After two decades, there is the fundamental work of Doeringer and Piore (1971) about internal labour markets ILMs: the market is a mix among two different realities, the internal and the external labour market.

The first is *"..An administrative unit, such as a manufacturing plant, within which the pricing and allocation of labor is governed by a set of administrative rules and procedures.*", while the second is the place where the conventional economic theory holds. The main factors reinforcing the formation of an ILM are peculiarity of labour, technology and custom. The specificity of labour and of technology causes an increase in the cost of turnover, while the custom is an "environmental" factor. According to the ILM structure, in the same plant or industry can live together several groups of workers. The cost of turnover causes a difference between employed and unemployed and it is also the basis of the neoclassic approach. The access to primary sector is rationated (there is a queue) through "ports of entry" and then there is a ladder system of promotions.

If the ILM describes a microeconomic environment, the macroeconomics of this way of thinking is highlighted by Piore (1980 nn.1 and 2).

---

[3] The pioneers are Becker (1962) and Mincer (1962).
[4] The classic reference for discrimination like a "taste" is the book of Becker (1971) while a survey containing statistical causes is contained in Altonji and Blanck (1999).

The idea under the dualism is that in the primary sector there is less possibility to change the employment for cyclical fluctuations, because relations are formal and there are more guarantees for workers, but also because in the primary segment there are high formation's costs, so turnover is very expensive and the stock of labor is rigid[5]. Then, the secondary sector has the function to allow cyclical fluctuations.

By these premises it is clear that the secondary segment is less stable, remunerative and there is less return of human capital, then at the extreme it is thought like a sector that gives negative skill accumulation (more properly a depauperation of skills[6]), because the required labour is not expert and the duties are usually only menial and unskilled.

Another way to look this situation is that there are economic barriers given by the insider theory in the primary segment and a perfect competition in the secondary one (so that the wage is at the competitive level).

A *traît d'union* of this way of thinking, with more standard formalized models in which agents maximize their utility has been investigated by Bulow and Summers (1986)[7]. They set an efficiency – wage bargaining system a là Shapiro – Stiglitz in the primary sector of the economy while there is perfect competition in the residual sector so that there is wage equilibrium into the sectors and identical workers are differently payed.

Ideas of dualism could be useful to describe the Italian labour market ? Piore (1980 n.1) looks at the Italian situation like the clearest example of dualism in the industrialised countries. According to this author the secondary sector developed after the '70s with the conquer of new rights for unions and workers, because it became an outlet for the impossibility to adapt the size of the primary sector. This argument is reinforced by the structural change of the Italian economy in the '70s (broadly and briefly speaking, lower growth, adverse international shocks and more participation to work). It is even more important in the light of the presence of a large quantity of black

---

[5] Oi (1962) supposed that a production function could have three factors: capital, rigid labor and flexible labor and that in the design of the production process, the first two factors were substantially unchangeable. At the extreme, the short run production function has only the secondary labor like flexible factor.

[6] Taubman and Wachter (1986).

[7] A number of scenarios and policy implication of the segmentation is showed by Saint – Paul (1996).

economy[8], especially in the Southern Italy, that is traditionally poorer and structurally weaker so that the problem may not be the unemployment, but the quality of the employment.

There are several ways that could be used to test the presence of dual markets in an economy; if we knew the segments it was very simple to test human capital's returns for people with same characteristics working in different segments, through the use of simple mincerian regressions.

It is important to stress that problems may emerge from this type of test for endogeneity and unobserved ability. The problems of endogeneity between some of the regressors and the error term causes the well known problem of identification, so we need to establish conditions to identify the equations for the estimators. This problem has been treated with instrumental variables and the question is to find a good instrument. Tipical instruments used are proximity to college and parents' education. The unobserved ability is the classical problem of omitted variables in a regression, so we could think that a constant in an earning equation is an average of a lot of different constants that are proxies of different abilities not captured from data. For example if schooling is positively correlated with ability (a clever worker needs less year to complete study qualifications) the real cause of more earnings is ability but the result will be an higher return of schooling[9]. So it is important to stress that the bias deriving by individual ability is very important in the estimation of the effect of education about earnings and the effectiveness of the estimates of schooling returns is strongly influenced by the goodness of the instruments and by unobserved ability, otherwise the risk is an overvaluation of the coefficients of education. The solution is to find a variable strongly related with schooling and unrelated with earnings, so to "clean" the effect of endogeneity with a first regression about the instrument.

---

[8] Bovi and Castellucci (2001) show evidences about a structural underground sector, proportionally much more important in the South of country.

[9] Possible solutions are proxies for unobserved ability like intelligency score or family background, but for example studies about twins give not unambiguous evidence and the general question is if twins are really equal why they should make different choices about human capital 'investments ? More promising seems to be the use of natural experiments. Angrist and Krueger (2001) tried to use birth dates of children after and before an exogenous event like a law for a greater lenght of compulsory education. Another test of Angrist and Krueger (1998) used the draft lottery to leave for Vietnam that depended by the birth date and the only alternative was the college. For example persons born with one day of distance (8th and 9th July) had respectively numbers 13 and 277 and there was a call till 195 (first date was 14th September) but there were no other reasons to think that their choice about schooling must be sistematically different.

After a preliminary treatment of these unconvenients, an index of dualism is given by the evidence of different wage equations with dissimilar human capital – earnings profiles, while an intermediate line is that we find when we estimate the overall market. Several papers estimate schooling returns' in Italy, like for example Flabbi (1999) or Brunello, Comi and Lucifora (2000). These authors estimate schooling's returns about the surveys of the Bank of Italy for women and men, using different approaches. Flabbi estimates schooling's returns about the survey 1991 of the Bank of Italy for women and men, using different approaches and a slightly larger specification of the wage equation. The results obtained with instrumental variables' techniques show an average annual return of about 4% with gender differences so that for men there returns of 5.3%[10]. Similar data are in Brunello, Comi and Lucifora. According Labour market segmentation theory we expect very different returns between segments.

Unfortunately, when our aim is to assess the presence of dualism, it is too arbitrary to choose segments[11] looking at sectorial or earnings demarcations[12] and it is important to consider the role of voluntary choices so that is important to choose homogenous groups. Methods that do not require a priori choices to verify the presence of dualism either in labour market or in the whole economic structure of the economy are:

1. Cluster analysis

    The objective is finding very homogeneous subgroups with respect to a number of variables, so there is the advantage to not constrain individuals to enter into a fixed number of groups.

2. Factor analysis

    In this way a researcher could reduce the dimension of a population through the identification of a smaller number of factor used to represent more complex relations for a set of interrelated variable. To confirm the hypothesis of a dual structure one has to find a common factor to divide individuals into the subgroups.

3. Switching regressions

---

[10] It is very close to the returns obtained in the OLS regression in table n.1, but the sample used is not the same, mainly because Flabbi uses people working in all sectors.
[11] Also the number of segments could be difficult: for example Rumberger and Carnoy used three segments, while Osberg, Apostle and Clairmont (1987) determine an higher number.
[12] Taubman and Wachter (1986) stress for the risks of bias with these a priori separations.

This tool makes separated inferences for the segments without the predetermination of the membership, so that persons are treated like unknown in relation to the segment they belong to.

The structure of a switching regression model can be described by the following set of equations (in compact structure):

$$Y_1^* = \beta_1 X_1 + \varepsilon_1$$

$$Y_2^* = \beta_2 X_2 + \varepsilon_2$$

$$Y_3^* = \beta_3 X_3 + \varepsilon_3$$

First two equations are segments' equations (for example earnings equations of a worker if is a member of a union) and third equation is participation's equation with Y_3^* that is the latent variable that determine the threshold of participation (for example the enrolment to an union).

By using Y_3^* as a threshold we will observe:

$$Y_1 = Y_1^* \text{ if } Y_3^* > 0 \quad \text{and}$$

$$Y_2 = Y_2^* \text{ and } Y_3^* \leq 0$$

With a similar procedure to models *a là Heckman* we can estimate the model, starting with a probit for the threshold and then by regressing the equations, respectively for a value lower or greater than the threshold for the individual – i.

This methodology has been applied to test dual labour markets in a series of works by the 1985 by Dickens and Lang. Their test (Dickens and Lang 1985) has two objectives: the first is to show that two wage equations could explain the real world better than one and that in the secondary equation there are almost nil returns of schooling and training on the job. Secondly, to verify that non economic barriers could prevent secondary workers by the achievement of a primary job. Works that apply this technique found different evidences of segmentation for specific countries (see for example Roig (1999) and Sousa-Poza (2004)). The method used here is a mixture of regressions that is not

7

popular in applied economics works as in other disciplines like for example medicine or marketing research[13].

The idea of the general class of mixture densities is that the true density function of a phenomenon is given by the mixture of more functions one for each segment, weighted by the probability of belonging to different parts of the population. If we have a mixture of linear normal regression with s segments, j regressors and i individuals [14], the normal density function of a worker's wage conditional to belong to the group s is:

$$f(w_i | s = k; \theta_i) = \left(2\pi\sigma_s^2\right)^{\frac{1}{2}} e^{-\frac{\left(w_i - \sum_j \beta_j x_{ij}\right)^2}{2\sigma^2}}$$

Where w are the individual's wages, and the j regressors of the mincerian equation are schooling, experience, experience squared, firm size, city dimension and a dummy equal to one for people working in the South of country. The mean is substituted with a linear predictor and the link between these two terms is the identity in the case of a normal density.

Given Bayes 'rule we may extract the probability of y for s equal to the a given segment like a joint probability, that is the ratio of conditioned probability and the probability to membership in a segment and if we sum all values of s, the result is the unconditional density of w_i.

$$f(w_i, \theta_i) = \sum_s \mu_s \left(2\pi\sigma_s^2\right)^{\frac{1}{2}} e - \frac{\left(w_i - \sum_j \beta_{js} x_{ij}\right)^2}{2\sigma_s^2}$$

The unknown is the vector θ (that contains also the weights μ_S): a solution is to find an initial value of the parameters, compute the density for these parameters and recompute the final θ, by the maximization of loglikelihood (EM algorithm) or alternatively through routines of numerical optimization. In the case of two components the loglikelihood takes the form:

---

[13] General references are Mc Lachlan and Peel (2000), Wedel and Kamakura (1999), Hamilton (1995), and Titterington, Makov and Smith (1985).
[14] See De Sarbo and Cron (1988).

$$\text{Ln L} = \sum_{i=1}^{n} \ln \left(\mu_1 \frac{1}{\sqrt{2\pi}\sigma_1} e^{\left(\frac{-(w_i - \sum \beta_{1j}x_{ij})^2}{2\sigma_1^2} \right)} + \mu_2 \frac{1}{\sqrt{2\pi}\sigma_2} e^{\left(\frac{-(w_i - \sum \beta_{2j}x_{ij})^2}{2\sigma_2^2} \right)} \right)$$

After estimated f we compute the posterior probability that observation – i comes from s, by the Bayes' theorem:

$$p_{is} = \frac{\hat{\mu}_s \, f_{is}\left(w_i \mid x_{ij}, \hat{\sigma}_s^2, \hat{\beta}_{js} \right)}{\sum_{s=1}^{S} \hat{\mu}_s \, f_{is}\left(w_i \mid x_{ij}, \hat{\sigma}_s^2, \hat{\beta}_{js} \right)}$$

And the observation –i is assigned to a segment s* when $p_{is}^* \geq p_{is}$ for all $s \neq s^*$.

By summarize the EM algorithm has two alternated steps: in the expectation step we compute the densities and in the maximization step we estimate the parameters of different regression (De Sarbo and Cron showed how this second step is equivalent to run least squares regression, weighted for the square roots of the probabilities to belong to a segment). A problem is that maximization of likelihood could converge in a local optimum so there is the need of many attempts with the choice of the highest value of likelihood. Thus, the algorithm computes the probabilities through functions and parameters and we go back to weighted regressions, with the criterion of stop decided according the convergence of loglikelihood. Probably the most attractive feature of this metodology is the possibility to easily extend the number of segment without the constraint to enter entirely in a given segment (there is not a threshold but a worker could stay proportionally in different segments).

3. Testing dualism in Italy through mixture regression

In this test we used data about households' budgets of Bank of Italy. This survey has a biennial frequency beginning from 1989 to 2004 (from 1977 there is an historical

survey with a lower number of information). The analisys uses hourly earnings (annual earnings divided an annual number of hours given by hours worked in a week, months in a year and number of weeks worked).

The choice of the sample is an answer to two different and contrasting esigencies: there is a trade – off between the possibility to find diversification between workers and the possibility that the results will be affected by voluntary choices[15]. This second question is probably much more serious so that the other one is left behind.

Initial sample extracted from the survey of 2002 uses males between 20 and 65, heads of family, dependant workers of private non – agricoltural sectors, working at least 20 hours for week[16]. The criteria chosen are a number of iterations between 20 and 400 and exit from maximization when the increase in loglikelihood is smaller than 0.005. We compute this routine with 200 random starting values probabilities.

Now in the table n. 1 we look at results obtained with the mixture regression model, compared with the OLS regression for the whole sample (1156 observations)[17].

Different starting probabilities randomly generated are used and results are very similar with each other, with the choice of the maximum level of likelihood (that has a monotonic pattern).

Looking at the OLS results, we may appreciate that the regression has a good fit and the explanatory variables (R^2 equal to 0.28) all significant and have expected signs of the coefficients (expect for the city dimension that is not significant)[18].

The mixture model describes a completely different situation among groups of workers because human capital's returns are more than 7% in the first segment, where also other variables have a greater weight in the coefficient except that for the

---

[15] For example it could be more plausible that women work part time more than men, to stay more time with children and in this way they renounce to higher possibility of career. In this way a lower return of education could be only a voluntary choice.

[16] This sample has same criteria of the study of Cipollone (2001), but there the survey used was 1995 and here there are eliminated observations with logarithm of hourly wage lower than 0, and greater than 5 to have a well blended set and to avoid outliers.

[17] Experience is proxied with actual age less the age of first job, while schooling is computed as the number of years necessary to have a given qualification (unfortunately we cannot know the real time necessary for each individual to complete each step). Similar results are obtained if we use the years of payroll like proxy for experience, but in this case we lost some observations.

[18] For the firm size we can think there are either economic or legislative reasons. Economic reasons could be compatible with dual market theory, because large, innovative firms are in the primary sector and the institutional argument could be highlighted for example by the fact that over a certain number of workers is easier organize an union so to bargain with the firm.

constant[19]. The size of this group of workers is about 37% of workers according the richest specifications of wage equations. Second columns shows also the importance of non linearities in experience. It is fairly usual that experience has not a linear profile compared with wage, as we know by the simple Mincerian function because an optimal investment has to give returns until the marginal contribution is close to zero. We observe as schooling and experience for primary workers are very higher than those for secondary ones.

It seems also that Southern secondary workers are less penalized with respect to the national average, while for primary workers the residence it is not so unimportant so that the simple dummy South is a good indicator of local market conditions. It is interesting to simulate profiles of earnings with respect to the educational qualifications and potential experience for all three models.

The graph n.1 shows as the relation is very close to theoretical predictions: there is a pronounced difference between workers with a ripid and a flat profile (OLS regression is closer to the second one according the proportions). In the simulation we considered 38 years of working life (necessary time to arrive to the retirement age) and we use absolute value and squared value to capture decreasing marginal returns. In this way we see that in general, decisions of investment given by primary workers are optimal in this situation because there are negative marginal returns close to the end of the working life (at the 34th year, looking at the higher line that is the primary segment, while the intermediate is the simulation given by the results of the OLS regression).

One obvious question could regard the presence of particular requirements for people staying in the primary sector, for example degree compared to diploma or work in transports respect to financial intermediation, and so on. The data contained in the tables nn.2 and 3 try to compare general characteristics and sectorial membership of the population chosen, with those of the people having more than 50% of probabilities to stay in the favoured sector.

Table n.2 is explanatory about the fact there are not sistematic differences between individuals belonging to different groups also if we consider for example the number of

---

[19] There are at least two potentially complementary explanations: first, maybe there are omitted regressors explaining better secondary segment's equation; second, that the intercept is lower for primary workers because they paid a part of their training at the beginning of job period of with lower wage, while secondary workers have a flat profile, as we expected.

earners divided components of households (this measure could affect labour – leisure trade – off of an individual). It corroborates predictions of dual theory. Table n.3 displays one slight but interesting difference, so that favoured persons work proportionally more in the industry. This fact is expected and it helps to explain the role of variable "firm size" because broadly speaking industry has a greater size with respect to other sectors and it is generally a more innovative sector.

In general we are in the position to sustain that a single wage equation is a too restrictive assumption also for this homogeneous group. If we use different surveys for the period 1991 – 2002 for two segments, the average weight of the favoured one is about 35% and also the proportions between coefficients of human capital's returns shows a clear divarication.

Now the next step is the problem of endogeneity.

We have already discussed as the main problem of this procedure could be in the coefficients of schooling because this variable may not be exogenous with respect to the error term and in this case we are not able to identify the parameters. We use a classical solution to deal with endogeneity, through an instrument to describe the variable schooling. The classical two – stage estimation is implemented with a first stage regression with schooling like dependant variable and parents' education as independent one (in addition to a constant).

In this way I obtain a fitted variables (the R^2 of this regression is about 23%) for schooling and I use this variable in the mixture regression framework. The full sample is almost unchanged while the main result is a growing spread between segments and the table n.4 compares results into the segments of the regressions respectively with schooling and fitted schooling (now the sample is slightly smaller because there are not data of parents 'education for all the workers): 9.6 against 2.1% for schooling respect to 8.0 and 2.2% (figure n. 2).

The I.V. coefficient for the entire sample is a good investment globally considered, but if we believe that markets are segmented and that the return depends by the job place and by the segment one person belongs, then the investments may not be always good.

A return of 2.2% in the secondary segment of the market could not justify the investment in schooling. The final point is whether other facts (in addition to theoretical predictions) could justify the choice of two segments instead of a greater number.

If we use three segments with schooling determined as before from a first – stage regression with parent's education (table n. 5) we obtain that a third segment had a size of about 36% (25% for the primary and 39% for the secondary) but it only consists of a further division of the secondary sector. In facts the third segment has almost nil human capital's returns (and not significant coefficient for education and experience) while the wage is more influenced from city size. Also the comparison between workers characteristics and sectors give not further insights because we have really close results to those of table nn.1 and 2, so the only effect is to reinforce the distance between favoured and unfavoured workers.

By these results schooling is a good investment ?

Brunello, Comi and Lucifora observe like schooling is often more expensive for Italian graduates with respect to the other countries. Using the data quoted by these authors, if the cost of one additional year of schooling is for example 600€ (it is a very conservative data, representing only average tuition fees in 1996) and the return is 2.5% of 12.000€ (300€), with a working life double than the schooling life, it is not clear if there is an incentive to study, also if we do not consider the opportunity cost of alternatives. Looking at the results, the crucial difference to judge the goodness of human capital's investments is in the choice between two and three segments.

If we use test based on the likelihood (Hawkins, Allen and Stromberg (2001) simulate power of 22 different tests in various situations of separation between segments) there is uncertainty between two and three components (AIC is less penalizing for greater number of coefficients) in the mixture but also the theories about dualism often tend to divide primary segment in an upper and a lower tier, so that three segments could be a good choice.

However, this evidence is still not sufficient because else if we try to choose workers with comparable characteristics it is possible that stay in the secondary segment is a voluntary choice. We need to justify that persons working in the secondary segment are locked in this part of the market, so that when they have the possibility they enter in the first segment. About this situation we may only observe indirect evidence, and this is another reason for choosing a small homogeneous group of workers.

Our expectation is that people working in primary segment remains in this part of the market and a proportion of people working in other pieces arrive in the primary

segment, if the queue process is not close (this is consistent with the hypothesis of ports of entry in primary segments). The strategy used here is to take the panel components of each survey and to measure changes of segments survey after survey, from 1995 to 2004 for the two segment mixture model[20].

Table n. 6 highlights the results of this strategy respectively for short period and long period transition. The most relevant result is given by the first part of the table (in a longer period we could have a lot of factors influence the transition and we have very few observation). There is always the expected result with the expection of the survey 2000 where the solution of two segment seems not to be the best specification (loglikelihood criteria prefer choices with 3 or 4 segments) and about one half of workers stay in the primary sector but the human capital return are almost the same between the segments. Generally speaking the primary segment seems to be the preferred situation because people tends to remain in this segment and there is a net transfer of workers towards this segment.

4. Conclusions

This work tried to verify the presence of more segments in the wage regression for Italian labour market starting from an homogeneous group of dependant workers. We found some robust evidences to sustain the idea that the use of only one regression is a strong limitation not supported by the data so that it is better to use different equations for groups of workers. The use of mixture regression models is an useful tool to avoid too arbitrary assumptions about the allocation of workers.

The first result seems to be the presence of a favoured part of the market that has a weight of about 37% of the market. The coefficients of human capital variables' returns for the primary workers in the survey 2002 of Bank of Italy are about four times greater than those of the secondary one.

There is a substantial confirmation of the prediction of Labour market segmentation theory with the unfavoured part of the market that has a flat profile in terms of human capital 's variables. The problem to have endougenos' s schooling decisions is corrected

---

[20] This allows us simpler computation of the transition between the segments.

with the use of an instrument and in this case the distance between returns into the segments becomes also higher.

Against the alternative of human capital' s theory there seem not to be specific characteristics that could justify competitive reasons for this phenomenon. More precisely workers with a probability to stay in the best part of the market higher than 50% have not characteristics much different with respect to the others, in the requirements of experience or education or other, while jobs in industry are proportionally more represented.

There are not elements to prefer one or more than three components in the mixture regression framework according to the loglikelihood, but three segments are apparently only a further division with a group of workers that have nil returns of human capital and another one closer to primary segment.

Southern workers are more disadvantaged in the primary market and also the firm size is one important factor of explanation of earnings. This result is not so surprising either in the vision of a dualism between core and pheriferic industries, but also because for example Italian legislation about firings is more restrictive for big firms (over a threshold of 15 dependant workers) and it is confirmed by the sectorial comparison.

Finally, the results show as the primary segment is the preferred location for the workers because there is always a positive net transition towards this group for the individuals followed for many years.

Policy implications of segmented labour market are very relevant because looking for example at the schooling we have to note that only a small part of the market has good returns from education, but for the majority of workers educational requirements are scarcely important to explain wage level.

An annual average return of 2% is probably lower than the expected, when parents decide to invest to finance human capital improvement for their sons, then a possible situation is that there is a net profit only for workers in primary groups, so globally the investment is not positive and above all there is a redistribution between households. In this way a reason for justifying the investment could be only a strong propensity to risk for the households but this situation is usually not seen as the common one.

Possible future extensions regard for example the use of more powerful instruments such as source of schooling differences given by natural experiments

References

Altieri G. (1999) The transformation of the Italian labour market: a process of segmentation, pagg.175 – 184 in Global Money, Capital Restructuring and the Changing Patterns of Labour, *Edward Elgar edition* by R. Bellofiore

Altonji, J.G. and R.M. Blank (1999) Race and Gender in the Labor Market, in: O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, vol.3 pp. 3143 – 3259

Angrist J.D. and A.B.Krueger (2001) Instrumental variables and the search for identification: from supply and demand to natural experiments, *NBER Working papers* no. 8456

Angrist J.D. and A.B.Krueger (1998) Empirical strategies in labor economics, Princeton University, Industrial Relations Sector Working Paper no. 401

Becker G. (1962) Investment in Human Capital: A theoretical analysis, *The Journal of Political Economy*, vol.70 no.5, part.2 pp. 9 – 49

Becker G.(1971) The economics of discrimination 2nd edition. *The University of Chicago Press*

Bovi M. and L.Castellucci (2001) Cosa sappiamo dell'economia sommersa in Italia al di là dei luoghi comuni ? *Economia Pubblica* n.6, pp. 77 – 119

Bulow J.I. and L.H.Summers (1986) A theory of dual labor markets with application to industrial policy, discrimination and Keynesian unemployment, *Journal of Labor Economics* vol.4, no.3 pp. 376 – 414

Card D.E. (1998) The casual effect of education on earnings, *Handbook of labor economics*, chapter n.30 pp.1802 – 1860

Cipollone P. (2001) Is the Italian Labour Market Segmented ?, *Tema di discussione del Servizio Studi* n. 400 della Banca d'Italia

Dell'Aringa C. (2003) The Italian Labour Market: Problems and Prospects, *Quaderni dell'istituto di Economia dell'impresa e del lavoro* n. 33

De Sarbo W.S. and W.L. Cron (1988) A Maximum Likelihood Methodology for Clusterwise Linear Regression, *Journal of Classification* 5, pp. 249 – 282

Dickens W.T. and K.Lang (1985) A test of Dual Labor Market Theory, *The American Economic Review,* vol. 75 no.4 pp. 792 – 805

Dickens W.T. and K.Lang (1988) The Reemergence of Segmented Labor Market Theory, *American Economic Review, Papers and Proceedings*, Vol.78 no.2 pp. 129 – 134

Dickens W.T. and K.Lang (1988) Labor Market Segmentation and the Union Wage premium, *The Review of Economics and Statistics* n.LXX 3 pp. 527–530

Doeringer P.B. and M.J. Piore (1971) Internal Labor Market and Manpower Analysis, *Health Lexington Books*

Dunlop J.T. (1957) The task of contemporary wage theory, in "The Theory of Wage Determination: Proceedings of a Conference Held by the International Economic Association" by John Dunlop

Flatau P.R. and P.E.T.Lewis (1993) Segmented Labor Markets in Australia, *Applied Economics* n. 25 pp. 285 – 294

Guidetti G. (2001) Teoria della Segmentazione e dei mercati del lavoro. Attualità, limiti ed evoluzione dell'analisi istituzionalista, *Economia Politica*, Anno XVIII, n.2, pp. 249– 306

Hamilton J.D. (1994) Time Series Econometrics, trad. it. 1995 ed. Monduzzi

Hawkins D.S., D.M.Allen and A.J.Stromberg (2001) Determining the number of components in mixtures of linear models, ***Computational Statistics and Data Analysis*** n.38, pp. 15 − 48

Kerr C. (1957) Wage Relationship − The Comparative Impact of Market and Power Forces, "The Theory of Wage Determination" ***Proceedings of a Conference Held by the International Economic Association*** by John Dunlop

Kerr C. (1955) The Balkanisation of Labor Market, reprinted in "The Balkanisation of the Labor Market,(1975) pp. 21 − 37, ***University of California Press***

Kerr C.(1950) Labor Market: Their Character and Consequences, reprinted in "The Balkanisation of the LaborMarket, (1975) pp. 38 − 52, ***University of California Press***

Kerr C. and L.H.Fisher (1950) Effect of Environment and Administration on Job Evaluation, reprinted in "The Balkanisation of the Labor Market, (1975) pp. 53 − 88, ***University of California Press***

Lang K. and W.T.Dickens (1992) Labor Market Segmentation Theory: Reconsidering The Evidence, ***NBER Working Papers***, no. 4087

Lang K. and W.T.Dickens(1993) Bilateral Search as an Explanation for Labor Market Segmentation and other Anomalies, ***NBER Working Papers*** no.4461

Leontaridi R.M.(1998) Segmented Labour Markets: Theory and Evidence, ***Journal of Economic Surveys***, vol.12 No.1

Lindbeck A. and D.J. Snower (1988) The Insider − Outsider Theory of Employment and Unemployment, ***The MIT Press***, Cambridge

Mc Lachlan G. and D. Peel (2000) Finite mixture models. ***John Wiley and Sons***, NY

Mincer J. (1962) On-the-job training: costs, returns and some implications, ***The Journal of Political Economy***, vol.70 no.5 part.2, pp. 50 − 79

Mincer J. (1970) The Distribution of labor Incomes: A survey with special reference to the Human Capital approach, ***Journal of Economic Literature*** vol.8, no.1 pp.1 − 26

Oi W.F. (1962) Labor as a quasi-fixed factor, ***The Journal of Political Economy***, pp. 538 − 555

Osberg L., R.Apostle and D.Clairmont (1987) Segmented labour markets and the estimation of wage functions, ***Applied Economics*** no.19, pp.1603 − 1624

Oster G. (1979) A Factor Analytic Test of the Theory of the Dual Economy, ***The Review of Economics and Statistics*** n.61, pp. 33 − 39

Osterman P. (1975) An Empirical Study of Labor Market Segmentation, ***Industrial and Labor Relations Review***, n.28 pp. 508 − 523

Pigou A. (1944) Lapses from full employment, ***ed. McMillan*** London

Piore M.J. (1980 n.1) Dualism as a response to flux and uncertainty, in Berger S. And Piore M.J., Dualism and Discontinuity in Industrial Societes, ***Cambridge University Press***

Piore M.J. (1980 n.2) The technological foundations of dualism and discontinuity, in Berger S. and Piore M.J., Dualism and Discontinuity in Industrial Societes, ***Cambridge University Press***

Reich M., Gordon D. M. and Edwards R.C. (1973) A theory of labour market segmentation, ***The American Economic Review***, vol. 63, num. 2, pp. 359 − 365

Roig A.H. (1999) *Testing Spanish labour market segmentation: an unknown-regime approach*, ***Applied Economics*** n. 31, pp. 293 − 305

Rumberger R.W. and M.Carnoy (1980) Segmentation in the U.S. Labour Market: its Effects on the mobility and earnings of whites and blacks, ***Cambridge Journal of Economics*** no.4 pp. 117 – 132

Ryan P.(1984) Segmentation, Duality and the Internal Labour Market, Wilkinson F. "The Dynamics of Labour Market Segmentation", London Academic Press

Saint – Paul G.(1996) The Dual Labor Market, ***The Mit Press***

Sloane P., P.D.Murphy, I.Theodossiou and M.White (1993) Labour market segmentation: a local labour market analysis using alternative approaches, ***Applied Economics*** n. 25 pp. 569 – 581

Sousa-Poza A. (2004) Is the Swiss Labor Market Segmented ? An Analysis Using Alternative Approaches, ***Labour*** n.18(1) pp. 131 – 161

Taubman P. and M.L.Wachter (1986) Segmented Labor Markets, ***Handbook of Labor Economics*** chapter n.21, pp.1183 – 1217

Titterington D.M., A.F.M.Smith and U.E.Makov (1985) Statistical analysis of finite mixture distributions, ed. ***John Wiley & Sons***

Wedel M. and Kamakura W.A. (2000) Market Segmentation. Conceptual and Methodological Foundations ed. ***Kluwer***
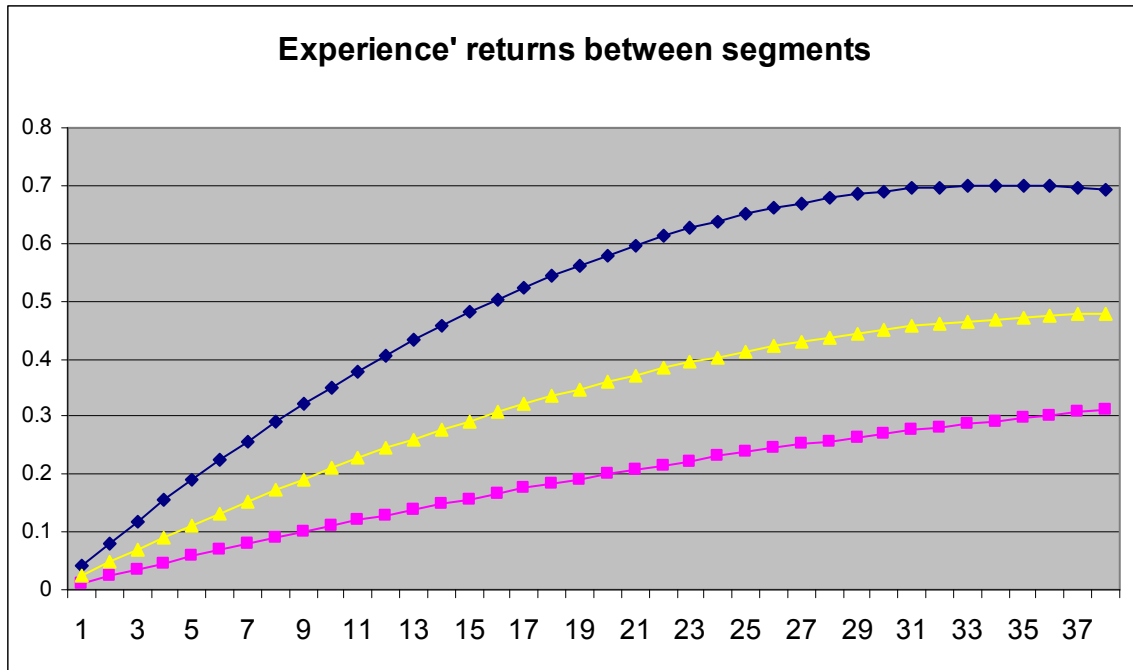
APPENDIX : RESULTS

Experience' returns between segments

Table n. 1

Results mixture model year 2002

	OLS	Segment 1	Segment 2
Constant	1.164	0.632	1.558
Schooling	0.044	0.076	0.018
Potential Experience	0.022	0.045	0.008
Potential experience ^{2}	-0.0003	-0.0007	-0.00001*
Dummy South	-0.079	-0.223	0.020*
City size	-0.003*	-0.001*	-0.0112
Firm size	0.044	0.045	0.045
Weight %	37.6	62.4	

19

Table n.2 Comparative characteristics of primary segment		
	mean sample	mean segment 1
Schooling	10.82	12.00
Experience	24.68	25.20
N° earners	0.61	0.59
%South w.	28.20	34.30
Firm size	3.66	3.76

Table n.3 Sectorial membership		
	weigth% sample	weight% seg.1
Industry	42.56	49.65
Building	12.40	14.27
Commerce	11.16	12.89
Transport	9.09	9.86
Intermediation	19.01	7.01
Estate agencies and others	3.31	3.20
Domestic services	2.47	3.12

Table n. 4

Dealing with schooling decisions' endogeneity

	Segment 1	Segment 2	Segment 1	Segment 2
Schooling	0.080	0.022		
Schooling fitted			0.096	0.021
Potential Experience	0.048	0.010	0.050	0.010
Potential experience [2]	−0.0007	−0.0001	−0.0008	−0.0002
Dummy South	−0.234	0.012*	−0.309	0.006*
City size	−0.006*	−0.015	0.002*	−0.015
Firm size	0.036	0.051	0.066	0.061
Weight %	34.3	65.7	30.1	69.9

Schooling returns when we deal with endogeneity
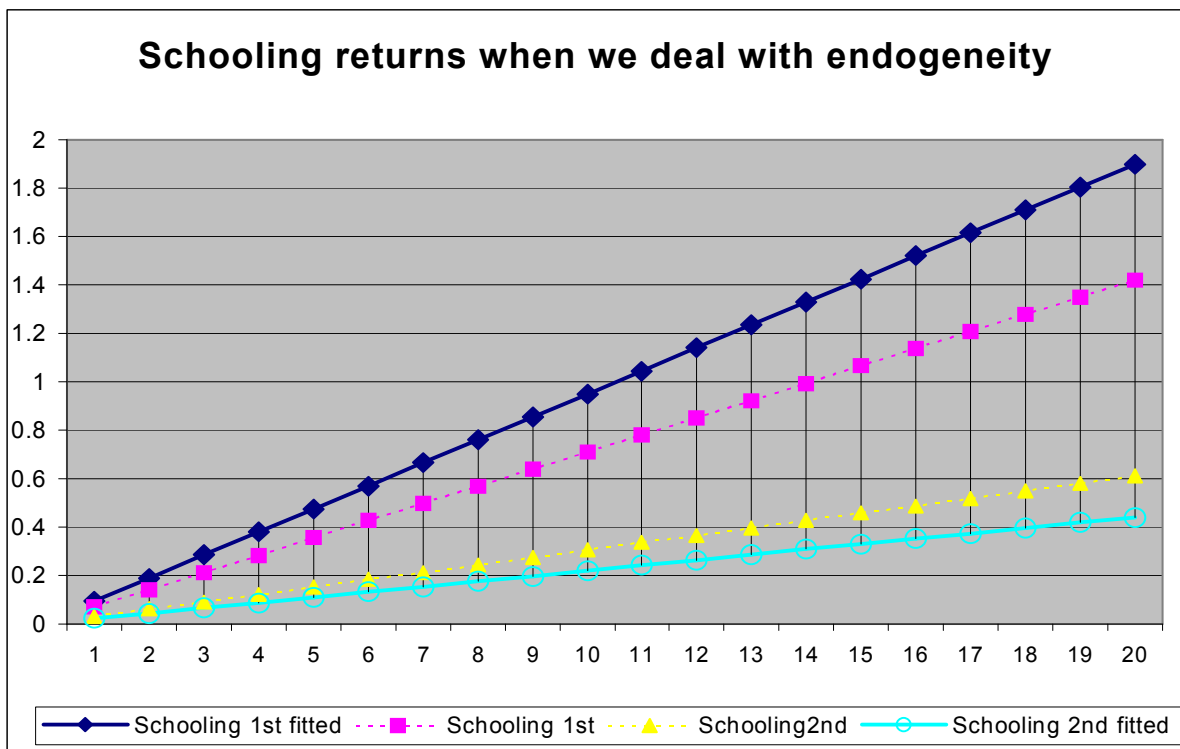
Table n. 5

Three segments

	Segment 1	Segment 2	Segment 3
Constant	1.108	1.031	1.802
Schooling	0.081	0.051	0.0001*
Potential Experience	0.046	0.029	0.002*
Potential experience ^{2}	–0.007	–0.0004	0.0000001*
Dummy South	–0.249	–0.194	0.171
City size	–0.0003*	0.0008*	–0.002
Firm size	0.043	0.057	0.034
Weight %	24.7	39.1	36.2

Table n. 6

Net transition towards 1ˢᵗ segment

	1995 – 1998	1998 – 2000	2000 – 2002	2002 – 2004
Short run transition				
Net change of segment	**+7.3%**	**+17.8%**	**– 9.9%**	**+3.5%**
N° panel obs	300	208	283	255
	1995 – 2000		1995 – 2002	1995 – 2004
Long run transition				
Net change of segment	**+26%**		**+10.3%**	**+13.2%**
N° panel obs	169		107	76