University Studies and Employment. An Application of the Principal Strata Approach to Causal Analysis

Leonardo Grilli, Fabrizia Mealli

Statistics Department "G. Parenti", University of Florence, Italy

Summary. In this paper, we propose a methodology, based on the principal strata approach to causal inference, for assessing the relative effectiveness of two university study programmes with respect to the employment status of their graduates. The analysis relies on a parametric model fitted by maximum likelihood. In that context, we discuss some relevant modelling issues and the implications of the results for policy.
Keywords: Causal effects; Effectiveness; Employment; Principal strata.

1. Students' propensity for employment

Traditional analyses of the effect of study programmes on employment (also called *external effectiveness analyses*) are performed on the sole basis of graduated students, neglecting the fact that the students who are able to graduate in a programme are in principle different from those who are able to graduate in another one. In other words, two programmes could select different kinds of students with specific propensity for employment.

The analysis of the employment of graduates mixes the "direct" effect of the study programme on the employment status with the "indirect" effect through the graduation status. The possibility to disentangle the two effects could be important from a political point of view. For example, if there is a direct effect on employment, then the programme with smaller effectiveness should try to adapt its contents in order to match labour market requirements.

If, instead, the occupational success of a programme is merely due to different selection criteria of the university career (e.g., one programme is more difficult than the other one and thus selects better students), the problem becomes an issue of educational policy. We should evaluate whether it is desirable for the society to graduate students with low ability or to allow the existence of study programmes with different difficulty levels.

To study the direct effect of programmes on employment, avoiding the possible bias caused by different graduation processes, it is necessary to envisage a joint study of graduation and employment. In this respect, a convenient framework is that of principal stratification (Frankgakis & Rubin, 2002), a development of the potential outcomes approach to causal inference (Rubin, 1974). Barnard *et al*. (2003) recently used the framework of principal stratification for the analysis of a complex randomized experiment in the educational context.

In the following, the *treatment variable* is the degree programme, while the *intermediate (post-treatment) variable* defining the principal strata is the graduation status (graduated or not). The key point is that if a student does not graduate, the *outcome variable*, that is the employment status, cannot be defined for assessing the external effectiveness of a given programme. This is an example of the so-called *censoring by death* (Zhang & Rubin, 2003).

The present analysis is limited to the comparison of two study programmes. The extension to three or more degree programmes entails some technical difficulties, but the conceptual framework would remain unaltered.

We compare the programmes of Economics and Political Science of the University of Florence, which are supposed to be similar with respect to the contents of the courses and to the job market opportunities. In the light of this similarity, the choice of a student to enrol in a programme should be weakly related to unobserved characteristics that potentially affect also graduation and employment status, so the ignorability assumption discussed later seems reasonable.

The paper is organised as follows. Section 2 describes the data, while Section 3 outlines the principal strata framework and the probabilistic structure used to model the data at hand. Section 4 describes the model fitting and shows the main results. Section 5 concludes the paper with some remarks.

2. The data

A joint analysis of the academic careers and job opportunities of university students requires the merging of two data sources: an administrative database of a cohort of freshmen and survey data on employment of the graduates belonging to that cohort. The two sources are: (*i*) the administrative database of the 1992 cohort of freshmen enrolled in the two programmes to be compared; and (*ii*) three surveys on the occupational status of the graduates of the years 1998, 1999 and 2000, respectively. The matriculation number is used to merge the datasets.

Overall, 1941 freshmen belong to the examined 1992 cohort: 1068 enrolled in Economics and 873 in Political Science.

Table 1. Percent distribution of curriculum status of the 1992 cohort by the end of the year 2000

Status	*Economics*	*Political Science*
Dropped	51.0	60.9
Graduated	25.3	20.2
Still enrolled	23.7	18.9
(n)	(1068)	(873)

Table 2. Employment status at the interview of the 1992 cohort

Status	*Economics*	*Political Science*
Graduates (n)	(270)	(176)
% graduates interviewed	69.3	56.2
% graduates with permanent job	51.9	36.4

The choice of the 1992 cohort is motivated by the availability of survey data for the graduates of the years from 1998 to 2000. Only 21 students of the 1992 cohort graduated before 1998, while among the students who did not drop out, the majority of them graduated in the triennium 1998-2000.

The status of the students by the end of the year 2000 is summarised in Table 1. For the students still enrolled at the end of 2000 there are no data on their employment status. Therefore, for the purpose of the present analysis, *graduation* means "graduation within nine years from enrolment". This definition is not particularly harmful for the analysis since graduation after nine years typically concerns students who already have a regular job during their studies.

All the interviewees responded to the question on the employment status. Apart from 21 students who graduated before 1998 and are out of target, almost all missing interviews are due to failed contact. The outcome variable of the analysis is a binary indicator of permanent job at the time of the interview, i.e. from one to two years after graduation. The employment status at the interview for the subset of graduated students is reported in Table 2.

The administrative database includes some additional information on every student of the 1992 cohort, which is used to define five dichotomous covariates: gender, residence (Florence *vs*. others), high-school degree (Gymnasium *vs*. others), high school grade (high grade, i.e. 50-60, *vs*. low grade, i.e. 36-49), late enrolment (i.e. the student did not enrol soon after high school). Table 3 reports the sample means of the covariates.

The covariates have different distributions in the two programmes, i.e. in the two treatment groups, since the assignment mechanism is not random. In particular, the high school grade is higher for the Economics students. The most striking difference concerns late enrolment, which is relatively rare in Economics, but reaches 22% in Political Science.

Table 3. Percent estimates of the covariates for the two study programmes

Covariate	*Economics (n=1068)*	*Political Science (n=873)*
Female	0.41	0.54
Residence in Florence	0.23	0.31
High school: gymnasium	0.34	0.45
High grade at high school	0.37	0.25
Late enrolment	0.06	0.22

3. The principal strata representation

Let n denote the total number of individuals under study, i.e. the dimension of the 1992 cohort of freshmen enrolled in Economics or Political Science. The *treatment variable* Z_i is thus defined as $Z_i = 1$ if student i was enrolled in Economics, and $Z_i = 0$ if student i was enrolled in Political Science.

Now let z_i denote the realized value of Z_i and let $\mathbf{z}$ denote the vector of z_i for all n individuals. In the potential outcomes framework every post-treatment variable, i.e. any relevant variable that takes its value after treatment assignment, depends on the vector of treatment assignments $\mathbf{z}$. However, in the present application it is reasonable to make the standard SUTVA (*Stable Unit Treatment Value Assumption*): for any individual i every post-treatment variable depends on $\mathbf{z}$ only through z_i. This excludes possible interactions between individuals.

Given SUTVA, every post-treatment variable has many potential versions as the number of possible treatments (two in the present application). Therefore, the post-treatment variables can be defined as follows.

The first potential post-treatment variables are the intermediate variables $S_i(z)$, z=0,1: $S_i(z) = 1$ if student i graduated by the end of 2000 (i.e. within 9 years) when enrolled in study programme z, and $S_i(z) = 0$ otherwise.

The availability of survey data for the graduates up to the year 2000 suggested to identify S as to represent the event "the student graduated by the end of 2000"; in this way the groups "Dropped" and "Still in course" collapsed in the same category.

Other potential post-treatment variables are the response indicators $R_i(z)$, z=0,1: $R_i(z) = 1$ if student i responded to the question on the employment status when enrolled in programme z and graduated, and $R_i(z) = 0$ otherwise.

The last potential post-treatment variables are the outcome variables $Y_i(z)$, z=0, 1: $Y_i(z) = 1$ if student i had a permanent job at the time of the interview, i.e. from one to two years after graduation, when enrolled in programme z and graduated, and $Y_i(z) = 0$ otherwise.

For each individual, the treatment variable assumes a single value; for every post-treatment variable, only one of the two versions can be observed. Therefore, we introduce the notation $S_i^{\text{obs}} = S_i(Z_i)$, $R_i^{\text{obs}} = R_i(Z_i)$, $Y_i^{\text{obs}} = Y_i(Z_i)$.

Since both the treatment variable and the intermediate variable are dichotomous, four principal strata can be defined through the latent variable L_i:

- L_i = '*GG*' (Graduated, Graduated) if $S_i(1)$=1 and $S_i(0)$=1: students who are able to graduate in both study programmes;
- L_i = '*GN*' (Graduated, Not graduated) if $S_i(1)$=1 and $S_i(0)$=0: students who are able to graduate in the first programme (Economics), but are *not* in the second (Political science);
- L_i = '*NG*' (Not graduated, Graduated) if $S_i(1)$=0 and $S_i(0)$=1: students who are *not* able to graduate in the first programme (Economics), but are able to graduate in the second (Political science);
- L_i = '*NN*' (Not graduated, Not graduated) if $S_i(1)$=0 and $S_i(0)$=0: students who are *not* able to graduate in either programme.

Note that each student belongs to a single stratum, although the data cannot in general reveal which stratum a person belongs to. In other words, the principal strata are latent classes and the data only allow estimating the probability that a given student belongs to a certain latent class. Also, note that the couple of potential values of the intermediate variable define the principal strata, so the strata are not affected by the treatment and thus can be viewed as an unobserved pre-treatment covariate.

The relationship between the observed groups, defined by Z_i and S_i^{obs}, and the principal strata is described in Table 4, along with the corresponding supports of R_i^{obs} and Y_i^{obs}.

For the post-treatment variables the sample proportions in the two treatment groups are: $P_{S,1}$ = 0.253 that is the sample proportion of graduates among students enrolled in Economics (Z_i=1), and $P_{S,0}$ = 0.202 is the analogue for Political science (Z_i =0); $P_{Y,1}$ = 0.516 is the sample proportion of Economics students with a permanent job (Z_i=1), who graduated (S_i^{obs} = 1) and responded to the interview (R_i^{obs} = 1), while $P_{Y,0}$ = 0.364 is the analogue for Political Science (Z_i =0).

Therefore, Economics has a higher graduation rate and a higher employment rate among the graduates. The analysis should assess if the better performance of Economics can be attributed to a positive causal effect.

Table 4. Observed groups, principal strata and variables of the study

Observed group $O(Z, S^{obs})$	Z_i	S_i^{obs}	R_i^{obs}	Y_i^{obs}	*Latent group* L_i *(principal stratum)*
$O(1,1)$	1	1	$\in \{0,1\}$	$\in \{0,1\}$	*GG* or *GN*
$O(1,0)$	1	0	not defined	not defined	*NG* or *NN*
$O(0,1)$	0	1	$\in \{0,1\}$	$\in \{0,1\}$	*GG* or *NG*
$O(0,0)$	0	0	not defined	not defined	*GN* or *NN*

Since the purpose of the study is to evaluate the effectiveness of graduation in a given programme with respect to the job market, the outcome variable *Y* is defined only for the graduates. Therefore the causal effect $Y_i(1)$-$Y_i(0)$ on the employment status is properly defined only for the *GG* stratum, i.e. the students who are able to graduate in both programmes.

In principle, if data were available, the outcome variable *Y* could be defined for all the enrolled students, allowing comparisons within the other strata. Anyway, such comparisons would not address the issue of relative effectiveness of graduation in different programmes.

The parameter of main interest is thus the *average* causal effect (ACE) for the *GG* stratum. When interest rests only on the population at hand, this estimand is simply the difference of the means of the two potential outcomes of *Y* for the individuals belonging to the *GG* stratum: $\overline{Y}_{GG}(1) - \overline{Y}_{GG}(0)$.

However, we are interested on the more general data generation mechanism, so the results will be implicitly referred to a superpopulation and expressed in probability terms. The estimand is thus the difference between the probabilities of having a permanent job under the two treatments, again for the *GG* stratum: $E(Y_{GG}(1)) - E(Y_{GG}(0)) = P(Y_{GG}(1) = 1) - P(Y_{GG}(0) = 1)$.

The probabilities of the principal strata are also of interest since they allow deepening the analysis of the effectiveness of the programmes with respect to graduation, as explained in Section 4.

Since *Z* is not randomised, there may be some confounders that influence both Z and S or both *Z* and *Y*: in such a case, the simple association between *Z* and *Y* cannot be interpreted as a causal effect. The available covariates $\mathbf{x}_i$, described in Table 3, may alleviate this problem and this is what underlies the following assumption (*Unconfoundedness of treatment assignment*): $Z_i \perp \{S_i(0), S_i(1), Y_i(0), Y_i(1)\} \mid \mathbf{x}_i$.

In the present application, this assumption would be violated if students with the same observed covariates would base their enrolment decision on reliable predictions on graduation and employment that depend on unobserved covariates. However, this behaviour seems unlikely, since the two competing degree programmes have many common features.

The data on the graduates' outcomes also suffer from a problem of missing data: in fact, even if the outcome variable *Y* is defined only for the graduates, it is available only for the interviewed ones. We assume that the information about *Y* is *Missing at Random*: $R_i(z) \perp Y_i(z) | \{\mathbf{x}_i, S_i(z)=1\}$ for each z=0,1. Under this assumption, the response mechanism is ignorable, so the analysis can be safely based on the available responses (conditional on observed covariates). Overall, the assumption of missing at random seems reasonable in the data at hand[1] because almost all missing interviews are due to failed contact.

---

[1] Mealli *et al*. (2004) discuss alternative assumptions on the response mechanism.

Under the assumptions of SUTVA, treatment unconfoundedness and missing at random, the data generating process can be defined in terms of the following two sets of probabilities:

A. *Probabilities of the principal strata*: $\pi_{GG:i}$, $\pi_{NG:i}$, $\pi_{GN:i}$, $\pi_{NN:i}$, where, for example, $\pi_{GN:i} = \Pr(L_i = 'GN' \mid \mathbf{x}_i)$ is the probability that student i belongs to principal stratum GN, i.e. he or she is able to graduate in Economics but not in Political Science;

B. *Probabilities of outcome conditional on the principal stratum*: $\gamma_{1,GG:i}$, $\gamma_{0,GG:i}$, $\gamma_{1,GN:i}$, $\gamma_{0,NG:i}$, where, for instance, $\gamma_{0,GG:i} = \Pr(Y_i^{obs} = 1 \mid Z_i = 0, L_i = 'GG', \mathbf{x}_i)$ is the probability that student i has a permanent job when he/she belongs to the principal stratum GG and graduated in Political science ($Z_i = 0$). From the unconfoundedness assumption, $\gamma_{0,GG:i} = \Pr(Y_i(0) = 1 \mid L_i = 'GG', \mathbf{x}_i)$.

The probabilities corresponding to combinations of the programme and principal stratum, other than the four listed, are not defined in the present application. The probabilistic structure is analogous to that of latent class models, except that in the present case belonging to a certain latent class determines not only the values of the probabilities of Y, but also whether they are defined or not. The estimand of main interest is the *average causal effect (ACE) on employment in the GG stratum*, i.e. the difference between the probabilities of the outcome Y under the two treatments for individuals belonging to the GG stratum:

$$\Pr(Y_i(1) = 1 \mid L_i = 'GG', \mathbf{x}_i) - \Pr(Y_i(0) = 1 \mid L_i = 'GG', \mathbf{x}_i) = \gamma_{1,GG:i} - \gamma_{0,GG:i} \, . \quad (1)$$

Also the probabilities of the principal strata ($\pi_{GG:i}$, $\pi_{NG:i}$, $\pi_{GN:i}$, $\pi_{NN:i}$) are interesting in itself, as they throw light on the dynamics of the graduation process in the two programmes. In fact, the *average causal effect (ACE) on graduation* is

$$\Pr(S_i(1)=1 \mid \mathbf{x}_i) - \Pr(S_i(0)=1 \mid \mathbf{x}_i) = (\pi_{GG:i} + \pi_{GN:i}) - (\pi_{GG:i} + \pi_{NG:i}) = \pi_{GN:i} - \pi_{NG:i} \, . \quad (2)$$

Therefore, the probability of the GG stratum, $\pi_{GG:i}$, is irrelevant for the ACE on graduation, but it can still describe different scenarios. In particular, as $\pi_{GG:i}$ diminishes, the graduates of the two degree programmes tend to be more heterogeneous. Even in the case of a homogeneous population, the probabilities π's and γ's are not directly estimable from the data without further assumptions. In fact, there are three non redundant π's and four γ's, compared with only four sample proportions ($P_{S,1}$, $P_{S,0}$, $P_{Y,1}$, $P_{Y,0}$). In particular, $P_{S,1}$ and $P_{S,0}$ allow us to get a point estimate of the π's only after fixing one of them, as long as the π's are the same in both treatment arms. Moreover, the γ's cannot be directly estimated, since they are defined conditional on the principal stratum. Rather, the data allow to estimate (through $P_{Y,1}$ and $P_{Y,0}$) the following mixtures of probabilities, so that estimation requires some mixture deconvolution:

$$\gamma_{1,GG:i}\frac{\pi_{GG:i}}{\pi_{GG:i}+\pi_{GN:i}}+\gamma_{1,GN:i}\frac{\pi_{GN:i}}{\pi_{GG:i}+\pi_{GN:i}};$$
$$\gamma_{0,GG:i}\frac{\pi_{GG:i}}{\pi_{GG:i}+\pi_{NG:i}}+\gamma_{0,NG:i}\frac{\pi_{NG:i}}{\pi_{GG:i}+\pi_{NG:i}}. \qquad (3)$$

4. Model specification and fitting

Model specification and estimation is a difficult task, since in the principal strata framework the latent groups lead to mixtures of distributions that are difficult to disentangle. The covariates are extremely useful to identify the model: identification can be achieved by several alternative restrictions whose plausibility should be judged case by case, as illustrated by Jo (2002) in the special instance of non-compliance with two latent groups. However, the likelihood function is usually flat, so its maximization is not trivial. A Bayesian analysis (Imbens & Rubin, 1997) may circumvent these difficulties, but, apart from the computational complexity, the choice of suitable prior distributions is tricky.

Here we perform a maximum likelihood analysis, which turns out to be effective for the problem at hand. As noted in the previous section, the data generating process can be defined in terms of two sets of probabilities, the π's, leading to the principal strata sub-model, and the γ's, leading to the outcome sub-model. The variables available for each individual are Z_i, S_i^{obs}, R_i^{obs}, Y_i^{obs} (if $R_i^{obs}=1$) and the vector of covariates $\mathbf{x}_i$.

In the present application, the 19 individuals with missing values on the covariates are simply deleted, so the covariates can be treated as fully observed. Extensions to handle missing values in the covariates are shown in Barnard *et al.* (2003).

Now let us collect the parameters in the vector $\boldsymbol{\theta}$ and the variables for the n individuals in the vectors $\mathbf{Z}$, $\mathbf{S}^{obs}$, $\mathbf{R}^{obs}$ and $\mathbf{Y}^{obs}$ and in the matrix $\mathbf{X}$. Then the likelihood can be written as a product over the four observable groups defined by Z_i and S_i^{obs}, where $i\in O(k,h)$ stands for $Z_i=k$ and $S_i^{obs}=h$:

$$\begin{aligned}
&L(\boldsymbol{\theta} \mid \mathbf{Z}, \mathbf{S}^{obs}, \mathbf{R}^{obs}, \mathbf{Y}^{obs}, \mathbf{X}) = \\
&\prod_{i\in O(1,1)}\left\{\pi_{GG:i}\left[(\gamma_{1,GG:i})^{Y_i^{obs}}(1-\gamma_{1,GG:i})^{1-Y_i^{obs}}\right]^{R_i^{obs}}+\pi_{GN:i}\left[(\gamma_{1,GN:i})^{Y_i^{obs}}(1-\gamma_{1,GN:i})^{1-Y_i^{obs}}\right]^{R_i^{obs}}\right\}\times \\
&\times\prod_{i\in O(1,0)}\{\pi_{NG:i}+\pi_{NN:i}\}\times \qquad (4)\\
&\times\prod_{i\in O(0,1)}\left\{\pi_{GG:i}\left[(\gamma_{0,GG:i})^{Y_i^{obs}}(1-\gamma_{0,GG:i})^{1-Y_i^{obs}}\right]^{R_i^{obs}}+\pi_{GN:i}\left[(\gamma_{0,NG:i})^{Y_i^{obs}}(1-\gamma_{0,NG:i})^{1-Y_i^{obs}}\right]^{R_i^{obs}}\right\}\times \\
&\times\prod_{i\in O(0,0)}\{\pi_{GN:i}+\pi_{NN:i}\}
\end{aligned}$$

The model is based on the assumptions of SUTVA, treatment unconfoundedness and missing at random. In the likelihood (4) the individuals who did not respond to the interview (R_i^{obs} =0) do not contribute to the estimation of the γ's, but do contribute to the estimation of the π's. Therefore, the π's are estimated from all the individuals in the sample, while information about the γ's is given only by the individuals who graduated and were interviewed (overall, 15% of the sample), so estimation of the γ's relies on scarce data.

As in the majority of current applications of the principal strata approach, the treatment and the intermediate variable are binary, leading to four principal strata. While in many settings it is sensible to assume that certain strata are empty (e.g. the assumption of no defiers in an experiment without compliance), in the present context such assumptions are not plausible in the light of the symmetry of the two treatments, so all the strata exist. This level of generality implies a considerable increase in model complexity since, as it is clear from the likelihood (4) that every observed group $O(k,h)$ is generated by a mixture of two distributions to be disentangled.

The probabilities of the principal strata (π's) are subject to some restrictions since they must lie in the [0, 1] interval and must sum to one. Therefore, in order to model the dependence of these probabilities on the covariates it is useful to operate a transformation to a set of unbounded parameters, using a multinomial logit specification (where *NN* is the reference category). For example, the model probability for GG is

$$\pi_{GG:i} = \frac{\exp(\eta^{\pi}_{GG:i})}{1+\exp(\eta^{\pi}_{GG:i})+\exp(\eta^{\pi}_{GN:i})+\exp(\eta^{\pi}_{NG:i})}.$$

For the probabilities of the outcome (γ's) the transformation to unbounded parameters can be obtained through separate logit specifications. For example,

$$\gamma_{1,GG:i} = \frac{1}{1+\exp(-\eta^{\gamma}_{1,GG:i})},$$

and four linear predictors are defined: $\eta^{\gamma}_{1,GG:i}$, $\eta^{\gamma}_{0,GG:i}$, $\eta^{\gamma}_{1,GN:i}$ and $\eta^{\gamma}_{0,NG:i}$.

Then, the η^{π}'s and the η^{γ}'s are assumed to depend linearly on the covariates. In the most general version of the model, each of these parameters has its own set of regression coefficients. In the current application the most general model we consider entails an unconstrained linear specification of the η^{π}'s, e.g. $\eta^{\pi}_{GG:i} = \alpha^{\pi}_{GG} + \boldsymbol{\beta}^{\pi}_{GG}{}'\mathbf{x}_i$, and a linear specification of the η^{γ}'s with a specific intercept, but a common vector of slopes, e.g. $\eta^{\gamma}_{1,GG:i} = \alpha^{\gamma}_{1,GG} + \boldsymbol{\beta}^{\gamma}{}'\mathbf{x}_i$. In this way it is assumed that each covariate has the same effect in each principal stratum and that the ACE on employment in the *GG* stratum is additive on the logit scale, $\alpha^{\gamma}_{1,GG} - \alpha^{\gamma}_{0,GG}$, i.e. is the same for all levels of the covariates. In our application, this specification seems a reasonable one. Other specifications could be devised (Jo, 2002), but in the present case their use is hindered by the limited sample information.

Model identification is possible only with a suitable number of covariates. The model specification just outlined has 27 parameters and 5 covariates, so that conditions for identification are met (Grilli & Mealli 2005). However, empirical underidentification problems are likely with models of this kind, and in fact in the model selection process we had to put some constraints on some parameters entering the η^{γ}'s.

Maximum likelihood estimation was performed by means of the NLMIXED procedure of the SAS system (SAS Institute, 1999). The procedure has several maximizing algorithms, the default being quasi-Newton with a BFGS update of the Cholesky factor of the approximate Hessian.

For certain values of the covariates, some principal strata have very low predicted probabilities, meaning that they are nearly empty. In particular, for the baseline individual, which was chosen to be the most frequent pattern in the sample and characterised by all the covariates being zero, the *NG* stratum seems to be empty, since the corresponding value on the multinomial logit scale is -7.826 (s.e. 14.763).

Therefore, in order to follow a simple and clear model selection strategy, we redefined the coding of the covariates in order to obtain a new baseline individual with non-negligible probabilities for all the four strata. The goal was achieved simply by switching the coding of the covariate *Late enrolment*, from now on labelled *Regular enrolment*.

In the unrestricted model, six of the estimated β^{π}'s are below –5 with huge or not available standard errors, meaning that when the covariate switches from zero to one the corresponding principal stratum disappears. In particular, with the exception of some students who enrolled late, the *NG* stratum appears to be empty. This is not surprising, since the overall proportion of graduates is modest and is lower for Z_i =0, so that the *NG* stratum (Not graduated if Z_i =1 and Graduated if Z_i =0) is necessarily very limited. In addition, the opposite *GN* stratum seems to be empty in some cases.

Therefore, model selection goes on by fixing to $-\infty$ the aforementioned β^{π}'s, leading to the results shown in Table 5. The reduction from 27 to 21 parameters entails a negligible reduction in the deviance, while the other parameters and standard errors are essentially unchanged. Some of the β^{π}'s are not significant at conventional levels, so the principal strata sub-model could be further refined. However, model selection was stopped since a parsimonious principal stratum sub-model is not an objective of substantive interest and has little effect on the precision of the estimates of the outcome sub-model.

In the outcome sub-model the β^{γ}'s are not significant at conventional levels, though two of them (*Gymnasium* and *Regular enrolment*) are high in magnitude: more data would be needed to assess the influence of the covariates on the outcome. Nonetheless, the primary target of the analysis, i.e. the causal effect on the logit scale $\alpha^{\gamma}_{1,GG} - \alpha^{\gamma}_{0,GG}$, is estimated as 0.666 with a s.e. of 0.301, so it is significantly different from zero at the 5% level.

Table 5. Parameter estimates (and standard errors) of the model

Principal strata submodel (π 's)			*Outcome submodel (γ 's)*		
α^{π}_{GG}	-4.402	(0.448)	$\alpha^{\gamma}_{1,GG}$	1.262	(1.241)
α^{π}_{GN}	-2.647	(0.752)	$\alpha^{\gamma}_{0,NG}$	-1.365	(1.568)
α^{π}_{NG}	-3.207	(0.835)	$\alpha^{\gamma}_{0,GG}$	0.596	(1.185)
$\beta^{\pi}_{GG,gymnasium}$	1.275	(0.157)	$\alpha^{\gamma}_{1,GN}$	0.484	(1.058)
$\beta^{\pi}_{GN,gymnasium}$	$-\infty$		$\beta^{\gamma}_{gymnasium}$	-0.410	(0.374)
$\beta^{\pi}_{NG,gymnasium}$	$-\infty$		$\beta^{\gamma}_{high\_grade}$	-0.036	(0.263)
$\beta^{\pi}_{GG,high\_grade}$	1.205	(0.146)	$\beta^{\gamma}_{regular\_enrolment}$	-0.932	(0.979)
$\beta^{\pi}_{GN,high\_grade}$	1.113	(0.652)	β^{γ}_{female}	0.070	(0.272)
$\beta^{\pi}_{NG,high\_grade}$	$-\infty$		$\beta^{\gamma}_{Florence}$	0.104	(0.333)
$\beta^{\pi}_{GG,regular\_enrolment}$	2.023	(0.425)	ACE on empl. *GG* $\alpha^{\gamma}_{1,GG}-\alpha^{\gamma}_{0,GG}$	0.666	(0.301)
$\beta^{\pi}_{GN,regular\_enrolment}$	-0.009	(0.792)			
$\beta^{\pi}_{NG,regular\_enrolment}$	$-\infty$				
$\beta^{\pi}_{GG,female}$	0.117	(0.137)			
$\beta^{\pi}_{GN,female}$	-0.622	(0.755)			
$\beta^{\pi}_{NG,female}$	0.991	(1.111)			
$\beta^{\pi}_{GG,Florence}$	0.280	(0.144)			
$\beta^{\pi}_{GN,Florence}$	$-\infty$				
$\beta^{\pi}_{NG,Florence}$	$-\infty$				

Table 6. Estimated percent probabilities for some covariates' patterns

Probability	**00000**	**00100**	**00110**	**00101**	**01100**	**10100**	**11100**	**11111**
$\pi_{GG:i}$	1.1	8.0	9.1	10.9	20.3	24.9	52.5	62.2
$\pi_{GN:i}$	6.3	6.0	3.3	0.0	14.0	0.0	0.0	0.0
$\pi_{NG:i}$	3.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0
$\pi_{NN:i}$	89.0	86.0	87.6	89.1	65.7	75.1	47.5	37.8
ACE on graduation $\pi_{GN:i}-\pi_{NG:i}$	2.7	6.0	3.3	0.0	14.0	0.0	0.0	0.0
$\gamma_{1,GG:i}$	77.9	58.2	59.9	60.7	57.3	48.0	47.1	51.5
$\gamma_{0,GG:i}$	64.5	41.7	43.4	44.2	40.8	32.2	31.4	35.3
$\gamma_{1,GN:i}$	61.9	39.0	40.7	41.5	38.1	29.8	29.0	32.8
$\gamma_{0,NG:i}$	20.3	9.1	9.7	10.0	8.9	6.3	6.1	7.1
ACE on empl. for *GG* $\gamma_{1,GG:i}-\gamma_{0,GG:i}$	13.5	16.5	16.5	16.4	16.5	15.8	15.7	16.2
$\gamma_{0:i}$	30.6	41.7	43.4	44.2	40.8	32.2	31.4	35.3
$\gamma_{1:i}$	64.2	49.9	54.8	60.7	49.5	48.0	47.1	51.5
$\gamma_{1:i}-\gamma_{0:i}$	33.6	8.2	11.4	16.4	8.6	15.8	15.7	16.2

The pattern (x_1,x_2,x_3,x_4,x_5) stands for *Gymnasium* = x_1, *High grade* = x_2, *Regular enrolment* = x_3, *Female* = x_4, *Florence* = x_5

To aid the interpretation of the results, Table 6 reports the estimated probabilities for some covariates' patterns. The patterns are in increasing order of $\pi_{GG:i}$. Note that $\pi_{GN:i}-\pi_{NG:i}$ is the ACE on graduation (2) and $\gamma_{1,GG:i}-\gamma_{0,GG:i}$ is the ACE on employment in the *GG* stratum (1), while $\gamma_{0:i}$ and $\gamma_{1:i}$ are the probabilities of employment for the two programmes obtained through the mixtures defined in (3).

The estimated proportion of students belonging to the *GG* group varies a lot with the covariates, from a minimum of 1.1% to a maximum of 62.2%. Moreover, the proportion of students belonging to the *GN* and *NG* groups (i.e. the students able to graduate in only one programme) tends to diminish as the *GG* stratum grows even if the *NN* stratum goes down.

On one extreme, the individual with all the covariates equal to one (a female with residence in Florence coming from a gymnasium with high grade and regular enrolment) has the highest probability of graduation (62.2%) which is entirely attributed to the *GG* group.

On the other extreme, the baseline individual (a male with residence outside Florence coming from a school other than a gymnasium with a low grade and late enrolment) has a low probability of graduation in at least one of the programmes (11.0%), mainly attributed to the *GN* and *NG* groups.

Since the ACE on graduation, $\pi_{GN:i}-\pi_{NG:i}$, stems from the *GN* and *NG* groups, it follows that the two degree programmes have a differential effect on the probability of graduation only for students having a weak background. Special guidance policies should be designed for this kind of students.

In general, knowing the sizes of the principal strata is important for guidance: for example, the students who benefit from enrolling in Economics are the ones belonging to the *GN* stratum, as they graduate only in that degree programme.

Looking at the effect on employment, some results are worth stressing. The level of the probability of employment varies a lot with the covariates, ranging from 47.1% to 77.9% for the graduates in Economics, and from 31.4% to 64.5% for the graduates in Political science. However, the ACE on employment in the *GG* stratum, which we assumed constant on the logit scale to avoid identification problems, generates a quite stable differential of about 15% in the employment probability.

Of course, the reliability and also the substantive importance of such an effect depends on the size of the *GG* stratum: for example, the causal effect in the *GG* stratum has little relevance for our baseline individual, which has a probability of only 1.1% to be a *GG*.

5. Concluding remarks

In the paper, two programmes of the University of Florence have been compared in order to evaluate their effectiveness with respect to employment rate.

The principal strata approach to causal inference was used to set up a general framework for the analysis of the problem, with a precise definition of the quantities of interest. Inference was drawn through a model fitted with maximum likelihood. Some care was needed in the model selection strategy to account for the possibility that some principal strata were empty.

The causal effect in the *GG* stratum (i.e. the students who are able to graduate in both programmes) is positive (namely in favour of Economics graduates). Moreover, the model gives some additional insight into the phenomenon, as it shows how the principal strata structure changes with the covariates: this information is crucial to understand the enrolment process and to interpret consciously the estimated causal effect.

Unfortunately, the limited sample information on the employment status led to a lack of statistical significance in most parameters of the outcome submodel and prevented us from exploring structures that are more complex.

The model could be developed also in a Bayesian framework, which entails several difficulties (specification of the priors, computational complexity), but offers some advantages which become crucial as the complexity of the model increases (Barnard *et al.*, 2003). Alternatively, if the assumptions of the parametric model are judged too restrictive, large-sample non-parametric bounds can be calculated (Grilli & Mealli, 2005).

References

BARNARD J., FRANGAKIS C.E., HILL J.L., RUBIN D.B. (2003) Principal stratification approach to broken randomized experiments: a case study of school choice vouchers in New York City, *Journal of the American Statistical Association*, **98**: 299-323.

FRANKGAKIS C.E., RUBIN D.B. (2002) Principal stratification in causal inference, *Biometrics*, **58**: 21-29.

GRILLI L., MEALLI F. (2005) Causal inference through principal stratification to assess the effect of university studies on job opportunities (*Submitted manuscript*).

IMBENS G.W., RUBIN D.B. (1997) Bayesian inference for causal effects in randomized experiments with noncompliance, *Annals of Statistics*, **25**: 305-327.

JO B. (2002) Estimation of intervention effects with noncompliance: alternative model specifications, *Journal of Educational and Behavioral Statistics*, **27**: 385-409.

MEALLI F., IMBENS G.W., FERRO S., BIGGERI A. (2004). Analyzing a randomized trial on breast self-examination with noncompliance and missing outcomes, *Biostatistics*, **5**, 2, 207-222.

RUBIN D.B. (1974) Estimating causal effects of treatments in randomized and nonrandomized studies, *Journal of Educational Psychology*, **66**: 668-701.

SAS INSTITUTE (1999) *SAS/STAT User's Guide Version 8*. Cary: SAS Institute Inc.

ZHANG J.L., RUBIN D.B. (2003) Estimation of causal effects via principal stratification when some outcomes are truncated by 'death', *Journal of Educational and Behavioral Statistics*, **28**: 353-368.