# XI ENCUENTRO DE ECONOMÍA PÚBLICA <br> Barcelona, 5 y 6 de Febrero de 2004 

# EXPLICANDO LA AVERSIÓN SOCIAL A LAS DESIGUALDADES EN SALUD 

Ignacio Abasolo ${ }^{1(*)}$ y Aki Tsuchiya ${ }^{2}$<br>${ }^{1}$ Departamento de Economía de las Instituciones, Estadística Económica y Econometría. Universidad de La Laguna, Tenerife. e-mail: iabasolo@ull.es<br>${ }^{2}$ Sheffield Health Economics Group, University of Sheffield, UK

TRABAJO EN CURSO
(POR FAVOR, NO CITAR SIN EL PERMISO DE LOS AUTORES)


#### Abstract

Resumen El principal objetivo de este trabajo consiste en explorar los factores que explican la aversión de los ciudadanos a las desigualdades socioeconómicas en salud. Se llevó a cabo una encuesta sobre una muestra representativa de la población española ( $\mathrm{n}=1209$ ). Después de informar que los individuos de las clases sociales altas tienen una esperanza de vida al nacer mayor que los individuos de clases sociales bajas, se pide a los entrevistados que elijan entre dos políticas de salud: una política que aumentaría en la misma cantidad la esperanza de vida de la clase alta y de la clase baja (política neutral), y otra política que tendría como objetivo únicamente aumentar la esperanza de vida de la clase social baja, reduciéndose así las desigualdades en salud (política igualitaria). Además, se administraron dos variantes de la pregunta, siendo la principal diferencia que una variante tenía un apoyo visual mientras que la otra no lo tenía. Se diseña un modelo que explica la preferencia por la política igualitaria, de acuerdo con diferentes variables demográficas, socioeconómicas e ideológicas. Las estimaciones probit muestran que el $70 \%$ de los entrevistados manifestaron una preferencia por la política igualitaria. Individuos de derechas o aquéllos viviendo en regiones con mayor renta per capita tienen una menor probabilidad de preferir la política igualitaria. Sin embargo, sorprendentemente, ni la educación ni la renta de la familia está asociada con una mayor propensión individual a elegir la política igualitaria. Además, los jóvenes y los más viejos tienen una menor probabilidad de elegir la política igualitaria que aquellos otros de edad mediana. Finalmente, la forma en que se administra la pregunta también importa: los entrevistados que tienen un apoyo visual tienen un $20 \%$ menor de probabilidad de preferir la política igualitaria.


(287 palabras)

## Agradecimientos:

Agradecemos la colaboración y sugerencias de Juan Diez-Nicolas, Paul Dolan, Roberto Gonzalez, Andrew Jones, Jaime Pinilla, Rebecca Shaw y Alan Williams. También estamos agradecidos a todos los entrevistados que tomaron parte en el estudio y al Instituto de Estudios Fiscales por el apoyo financiero. Cualquier error es de nuestra responsabilidad.
(*) autor para correspondencia.

# EXPLAINING SOCIAL AVERSION TO INEQUALITIES IN HEALTH 


#### Abstract

The main objective of this paper is to explore the factors that can explain peoples' aversion to inequality in health across socioeconomic groups. A representative sample of the Spanish population was surveyed ( $\mathrm{n}=1,209$ ). After being informed that those from the higher social class have longer life expectancy at birth than those from the lower social class, respondents were required to choose between two health policies: one would increase life expectancy of the high and low social groups by the same amount (the "distribution neutral" policy), and the other option would target the lowest social class group, thereby reducing current health inequalities (the "targeting" or "egalitarian" policy). Two variants of the question were administered, the main difference being that one had a visual aid whilst the other did not. A model is developed that explains the preference for targeting according to different demographic, socioeconomic and ideological variables. Probit estimations show that $70 \%$ of respondents report a preference for the egalitarian policy. As expected, right wingers or/and individuals living in a high per capita income region are less likely to choose the egalitarian policy. However, surprisingly, neither individual's education nor household income is associated with individual's preference to target. Age is also a determinant of individual choice: younger and older individuals are less likely to target the egalitarian policy than those in the middle age. Finally, results also suggest that the way in which the question is asked matters: respondents that have a visual aid are less likely to target by $20 \%$.


Key words: inequality aversion, background characteristics, mode of presentation.

## I. INTRODUCTION

Maximising average population health is a major objective of public health care systems. However, modern health care systems have another objective: equality in the distribution of health. Various studies have explored how members of the public see the balance between these two objectives and found that neither health maximisation or health equalisation alone is supported, and that these two objectives are traded off against each other (see for example Johannesson \& Gerdtham, 1996; Cuadras-Morato et al, 2001; Anderson \& Lyttkens, 1998; Dolan \& Robinson, 2001; Dolan et al, 2002).

The motivation of this paper is to explore what factors can explain inequality aversion in health. Based on a survey data from Spain, the relationship between the respondents' backgrounds characteristics and their preferences to reduce inequalities in health is analysed. In addition, whether different modes of asking people about this issue lead to different answers is investigated.

This is done by analysing interview data obtained in Spain (see Abasolo et al. 2001). First, respondents were informed of a current socioeconomic health inequality, namely, that men from the highest social class have a longer life expectancy at birth ( 75 years) than men from the lowest social class ( 70 years $)^{1}$. Then, respondents were asked to choose between two health programmes with the same cost: programme A, that extends the life expectancy of both these population groups by two years each, and programme B, that targets the disadvantaged group and increases their life expectancy by four years. Since, by targeting all the benefits to the disadvantaged group, programme B will reduce the inequality by the same amount, we will refer to this as the "targeting policy" or the "egalitarian policy". The aim of the questionnaire was to tap into the preferences of the respondent as a citizen, comparing public policies addressing issues of inequality and distribution, as opposed to a consumer, purchasing insurance policies for their own benefit. The survey question continued to identify the strength of inequality aversion by subjecting those respondents who choose the egalitarian policy to explicit trade offs between efficiency and equality (for details of the questionnaire, see Shaw et al); however, the present paper concentrates on the first exercise alone, so choices are not conditioned by efficiency considerations as both programmes have the same total health gain. At the end of the survey, socioeconomic, demographic and ideological information were gathered to serve as the explanatory variables. Two different ways of presenting the options were used (a Pictorial variant and a Verbal variant), with the aim of testing whether or not these two different devices to present the same information lead to the same results.

## II. THE MODEL AND ESTIMATION METHODS

We specify a model that explains the probability of a given respondent to target the disadvantaged group, as defined above. An underlying (or latent) variable ( $T^{*}$ ) represents an individual's propensity to chose this policy. We anticipate that demographic, socioeconomic and ideological characteristics are associated with people's attitudes towards egalitarianism. With respect to the first, age (A) and gender (G) are considered. Secondly, since we are dealing with attitudes regarding socioeconomic inequalities, we may expect there to be some pattern in the responses by respondents' socioeconomic status; proxies used to explore this possibility are household income

[^0]$(Y)$, education $(E)$, per capita income of the region of residence $(R)$, and being in the labour market $(L)$. In general, it is expected that all these variables are negatively related with the propensity to support the egalitarian health policy. Thirdly, political affiliation, or ideology ( $I$ ), is also assumed to have a role in the model. After all, people's attitude towards egalitarianism is an ideological issue, which might be favoured more by a left wing person. Lastly, we include in the model another variable $(Q)$ representing the questionnaire variant. The question could be presented to the respondents with the aid of pictures ( P ) or verbally with no visual aids (V). The aim of this regressor in the model is to determine whether there are significant differences in the estimated probabilities depending on the way the question is asked.

Thus, the model can be written as:

$$
\begin{equation*}
T_{i}^{*}=T(G i, A i, E i, Y i, L i, I i, R i, Q i)+\varepsilon i \tag{eq.1}
\end{equation*}
$$

In model [eq.1], the $i$ subscripts represent individual respondents, and $\varepsilon i$ captures influences unobserved by the analyst, which we assume to have a standard normal distribution with zero mean and constant variance. In practice, $T^{*}$ is unobserved. Rather, we observe $T i$, which is a dummy variable representing whether or not the individual actually chooses the targeting policy; therefore it is the realization of a binomial process defined by:

$$
T i=1 \text { if }\left[T_{i}^{*}>0\right]
$$

So, if the individual's propensity to target is positive ( $T_{i}^{*}>0$ ) $\mathrm{s} / \mathrm{he}$ will choose the egalitarian policy $(T i=1)$, and if otherwise $\left(T_{i}^{*} \leq 0\right) \mathrm{s} /$ he will not $(T i=0)$.

In order to select the functional form of the empirical model, socioeconomic and statistical criteria are used. This has allowed us to consider the definitions of the set of dummy variables and whether the only continuous variable (age) enters the model in natural units, as logarithms or as higherorder powers. Interactions between regressors are also included and tested in the model. In particular, in order to test whether the mode of presenting the question $(P$ or $V$ ) has a different effect on the propensity to target across the different background characteristics specified in the model, we will test the joint hypothesis that the interactions of $Q$ with the rest of regressors are not significantly different from zero.

The estimation process will be undertaken through non-linear probit regressions. Likelihood ratio (LR) tests and Reset specification tests will be carried out to appraise the appropriateness of the different functional forms. Throughout, a $5 \%$ significance level is used.

Estimations of model [eq.1] will allow us to empirically assess the relevance of the different hypothesised explanatory variables and appraise whether the way in which the question is formulated does not account or it does, and if it does in which way.

Model [eq.1] may be subject to selection bias due to incomplete survey data. In surveys of this sort respondents do not always provide answers to all the questions of the survey. It is the so-called "item non-response". If the pattern of non-response is not at random, conventional estimators may be biased and inconsistent. Tests for selection bias and correction, if necessary, are undertaken estimating a probit with sample selection (Greene 1997). The probit with sample selection works in
a manner very similar to the Heckman model except that the response variable is binary. This method requires additional exogenous variables (or identifying variables), which should explain the probability of participating but have no direct impact on the probability to target.

So, let us assume an underlying (unobserved) variable $Y_{i}^{*}$ that determines selection into participating groups, i.e. $Y i=1$ when $Y_{i}^{*}>$ threshold, and $Y i=0$ when $Y_{i}^{*} \leq$ threshold. $Y_{i}^{*}$ would represent the inclination for the individual to participate answering all the relevant questions being $Y_{i}^{*}$ a linear function of some of the exogenous variables in model [eq.1] as well as some identifying variables as follows:

$$
\begin{equation*}
Y_{i}^{*}=Y(G i, A i, L i, R i, Q i, M i, H i)+u i \tag{eq.2}
\end{equation*}
$$

The identifying variables include $M i$, which represents the marital status of the individual (married, single, divorced or widow), and $H i$, which represents whether the individual lives in a rural area. The main criteria used here for proposing both set of identifying variables are that the variables have an impact on the probability to participate but are unrelated to the individual's preference for egalitarian policies. $u i$ is a random error term normally distributed with zero mean and constant variance.

Selection bias occurs when there is correlation between $Y$ and $\varepsilon$ (and therefore between $\varepsilon$ and $\boldsymbol{u}$ ); that is, when unobservable factors that influence the potentially selection are also influencing the probability to target. If so, selection bias will be corrected. To check whether selection bias is absent we will test, firstly, whether $\rho$ (the correlation of residuals) is significantly different from zero: if the covariance between $\varepsilon$ and $\boldsymbol{u}$ is significantly different from zero, then we cannot reject that there is no selection bias. In addition, a comparison of the estimates of both the initial probit and the probit with selection is undertaken: a large change in the coefficients, a change of the sign of the coefficients or a change in the statistical significance of the coefficients between the initial probit and the probit with selection indicate the existence of selection bias.

## III. DATA AND VARIABLES DEFINITION

The data were collected during 1999 in Spain. A survey of 1,209 individuals over 18 years of age was undertaken. Face to face interviews were assigned across the 17 "Comunidades Autónomas" (Regions for short), reflecting the local resident population proportionally. Within each of the Regions, interviews were randomly allocated so that the achieved sample will be representative of the general Spanish population in terms of socio-demographic characteristics. Each of the two variants of the questionnaire was administered to approximately the same number of respondents: 602 for the P -variant and 607 for the V -variant. In terms of background characteristics, the distribution of these variables across the two variants is not statistically significantly different from each other. For descriptive statistics, see Table 1 of the Appendix.

The binary dependent variable, target, takes the value 1 if individual $i$ targets the worse off group, 0 if otherwise. Age is a continuous variable, which takes values between 18 and 94, and also enters the equation in quadratics. The binary variable female indicates whether the individual is female or male. Education is recorded by level of schooling and has been categorised in three dummy variables representing low education lowedu (those with primary school education or less), middle education midedu (those with secondary school education, the baseline category), and high
education highedu (those with higher and university education). Household income is also categorised in three dummies high income (more than 1,653 euros per month), middle income (between 600 and 1,653 euros per month, the baseline category) and low income (less than 600 euros per month ${ }^{2}$. Per capita income in the region of residence is captured by three dummies: high income regions highreg (those resident in Madrid, Navarra or País Vasco), low income regions lowreg (those who live in Andalucia or Extremadura) and middle income regions midreg (residents in the rest of Spain, the omitted category). The binary variable non-labour market (nonlabmkt) indicates whether the individual is not currently in the labour market (i.e. is retired, unemployed, homemaker or student) or whether $\mathrm{s} / \mathrm{he}$ is. Political affiliation is recorded by three categorical indicators, right (those who report as being centre-right, right or extreme right wing), left (those who report as being centre-left, left or extreme left wing), and centre (those who are in the political centre, the baseline category). Type is the dummy variable representing the type of question administered, taking value 1 if the V variant was provided, and 0 if the P variant was administered. Regarding the identifying variables, single, divorced and widow are three dummy variables representing whether the individual belongs to one of such marital status (as opposed to the omitted category married); and the binary variable rural, representing whether the individual is resident in a rural area (of less than 10,000 residents).

## IV. RESULTS

Descriptive statistics and estimation results are reported in table 1. Overall, 70\% of respondents chose to target the egalitarian policy. Model [eq.1] passes the Reset specification test, indicating that there is no evidence of functional form problems (table 2). In addition, when dealing with the issue of the relevance of the question variant in the model, likelihood ratio tests show that specification [eq.1] fits the data significantly better than other alternative specifications. The alternative equations considered include, on the one hand, a model that omits the type variable and, on the other hand, a model which includes the type variable together with interactions with the rest of independent variables. Therefore model [eq.1] shows that the way in which the question is presented counts, although there are no differential effects across the individual characteristics considered in the model.

Item non-response leads to 514 missing cases, which corresponds to $42.5 \%$ of the entire data, leaving 695 individuals as valid cases. Estimates for the probit with sample selection [eq.2] can be seen in table 4. The correlation coefficient (rho) is not statistically different from zero. It is also true that the confidence interval for the parameter estimate is wide. However, sign, magnitude and t-ratios of coefficients of the probit with selection are very similar to those of the initial probit estimation [eq.1]. Overall, the results suggest that we cannot reject the null hypothesis that there is no selection bias.

Probit average and marginal effects evaluated at sample means (table 3) show that: firstly, other things equal, age is a statistically significant explanatory variable of the probability to target the disadvantaged group. However, the marginal effect of age on targeting is not constant and changes with the age of the respondent. For young adults between 18 and 44 the probability of targeting increases with age at a diminishing rate. A maximum is reached, on average, at the age of 44 , after which the probability of targeting starts to decrease with age at an increasing rate.

[^1]Secondly, those who are politically right wing have a significantly lower probability of targeting the egalitarian policy, as compared to those who are in the political centre (and consequently, as compared also to those who are left wing). In particular, the probability of a right wing individual targeting is on average $12 \%$ less than the reference individual, other things being equal. Finally, living in one of the richest regions such as Madrid, Navarra or País Vasco is associated with a lower probability of targeting the egalitarian policy, other things being equal, by $17.4 \%$, compared to those living in average per capita income regions. There is no evidence that other demographic and socioeconomic factors included in the model have a significant influence in the attitudes towards health inequalities.

Regarding the effect of the question variants, our model suggests that mode of presentation makes a difference. When individuals are administered the V -variant (as opposed to the P -variant), there is a $7 \%$ higher probability of targeting the egalitarian policy (at sample means). A comparison of this average effect with those of other variables shows that, at sample means, the average effect of the question variants is higher than that the marginal effect of age but lower than the average effect of being right wing or resident in a rich region. However, there is no evidence that the question variant has differential effects on the probability of targeting across the different demographic, socioeconomic and ideological characteristics, when these are taken jointly.

## V. CONCLUSION

A survey that elicited peoples' aversion to inequality in health was carried out in Spain. The current study explored the factors associated with the respondents' propensity to choose a policy that reduced the inequality, as opposed to a policy that did not. These included the respondents socioeconomic and demographic characteristics, their political affiliations, and the mode of presentation used in the questionnaire. The results suggest that age, ideology and the affluence of the region where they live explain people's aversion to inequality in health: older individuals, right wingers and/or those living in affluent regions are significantly less likely to support egalitarian policies. Unexpectedly, other socio-economic background characteristics like individual's education or household's income do not seem to have a significant impact. The evidence also shows that the way in which the question is administered also counts: the regression results indicate that on average, there is a significantly higher probability to target when people deal with numbers rather than with pictures.

## VI. REFERENCES

Abasolo I., Shaw R., Tsuchiya A. (2001) ¿Hasta qué punto están más preocupados los españoles por la equidad que por la eficiencia en salud? (To what extent are the Spanish more concerned about equity than efficiency in health?) paper presented at the Asociación de la Economía de la Salud, 6-8 June, Ovideo.

Acheson D. (1998) Independent Inquiry into Inequalities in health. London, HMSO.
Andersson F, Lyttkens CH (1999). Preferences for equity in health behind a veil of ignorance. Health Economics;8(5):369-78.

Borrell C., Regidor E., Arias L. (1999) Inequalities in mortality according to educational level in two large southern European cities. International Epidemiological Association, 28: 58-63.

Cuadras-Morato X, Pinto-Prades JL, Abellán-Perpiñan JM (2001). Equity considerations in health care: The relevance of claims. Health Economics; 10(3):187-205.

Dolan P, Robinson A (2001). The measurement of preferences over the distribution of benefits: The importance of the reference point. European Economic Review;45(9):1697-709.

Dolan P, Tsuchiya A, Smith P, Shaw R, Williams A. (2002) Determining the parameters in a social welfare function using stated preference data: an application to health. Sheffield Health Economics Study Group, University of Sheffield. Discussion Paper Series 02/2.

Greene WH (1997) Econometric Analysis. Prentice Hall, third edition. New Jersey.
Johannesson M, Gerdtham UG (1996). A note on the estimation of the equity-efficiency trade-off for QALYs. Journal of Health Economics;15:359-368.

Wagstaff A (1994). QALYs and the equity-efficiency tradeoff. In: Layard A, Glaister S, eds. CostBenefit Analysis. second edition: Cambridge University Press (reprinted from Journal of Health Economics, 10:21-41, 1991 with corrections)

Williams A (1997). Intergenerational equity: an exploration of the 'fair innings' argument. Health Economics 6: 117-132

## APPENDIX

TABLE 1. MEAN VALUES AND STANDARD DEVIATIONS OF VARIABLES

| Variable | Obs | Weight | Mean | Std. Dev. | Min | Max |
| ---: | :---: | :---: | :---: | :---: | :---: | ---: |
| target | 695 | 694.516207 | .6904674 | .4626339 | 0 | 1 |
| female | 695 | 694.516207 | .485911 | .5001614 | 0 | 1 |
| age | 695 | 694.516207 | 45.26725 | 17.5467 | 18 | 91 |
| agesq | 695 | 694.516207 | 2356.567 | 1709.339 | 324 | 8281 |
| highedu | 695 | 694.516207 | .2362348 | .425074 | 0 | 1 |
| midedu | 695 | 694.516207 | .4139528 | .492895 | 0 | 1 |
| lowedu | 695 | 694.516207 | .3498124 | .477254 | 0 | 1 |
| highinc | 695 | 694.516207 | .2259842 | .4185301 | 0 | 1 |
| midinc | 695 | 694.516207 | .4759503 | .499781 | 0 | 1 |
| lowinc | 695 | 694.516207 | .2980655 | .4577378 | 0 | 1 |
| nonlabmkt | 695 | 694.516207 | .5352979 | .4991117 | 0 | 1 |
| left | 695 | 694.516207 | .4718532 | .4995667 | 0 | 1 |
| centre | 695 | 694.516207 | .3091312 | .4624682 | 0 | 1 |
| right | 695 | 694.516207 | .2190156 | .4138771 | 0 | 1 |
| reghigh | 695 | 694.516207 | .1453269 | .352684 | 0 | 1 |
| regmid | 695 | 694.516207 | .6314791 | .482751 | 0 | 1 |
| reglow | 695 | 694.516207 | .223194 | .4166872 | 0 | 0 |

TABLE 2. RESULTS PROBIT ESTIMATION

| Probit estimat Log likelihood | $=-410.96762$ |  |  | Number <br> LR chi2 <br> Prob > <br> Pseudo <br> RESET | f obs $=$ <br> 13) $=$ <br> hi2 $=$ <br> 2 $=$ <br> st $0.04 ;$  | $\begin{array}{r} 695 \\ 38.09 \\ 0.0003 \\ 0.0443 \\ (1)=3.84 \end{array}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| target | Coef. | Std. Err. | z | $\mathrm{P}>\|\mathrm{z}\|$ | [95\% Con | Interval] |
| female | . 0356847 | . 1081395 | 0.33 | 0.741 | -. 1762649 | . 2476343 |
| age | . 0439998 | . 0173643 | 2.53 | 0.011 | . 0099663 | . 0780333 |
| agesq | -. 0005051 | . 0001825 | -2.77 | 0.006 | -. 0008628 | -. 0001475 |
| highedu | . 2193948 | . 1342858 | 1.63 | 0.102 | -. 0438006 | . 4825902 |
| lowedu | . 2170204 | . 1467327 | 1.48 | 0.139 | -. 0705704 | . 5046111 |
| highinc | -. 1167251 | . 1338214 | -0.87 | 0.383 | -. 3790103 | . 14556 |
| lowinc | . 0217484 | . 136658 | 0.16 | 0.874 | -. 2460964 | . 2895933 |
| nonlabmkt | . 1568169 | . 1272843 | 1.23 | 0.218 | -. 0926558 | . 4062896 |
| left | -. 0148879 | . 1222525 | -0.12 | 0.903 | -. 2544983 | . 2247225 |
| right | -. 3359247 | . 1401156 | -2.40 | 0.017 | -. 6105463 | -. 0613031 |
| reglow | . 1271331 | . 1309688 | 0.97 | 0.332 | -. 1295612 | . 3838273 |
| reghigh | -. 4672769 | . 1422487 | -3.28 | 0.001 | -. 7460792 | -. 1884745 |
| type | . 2037721 | . 1021195 | 2.00 | 0.046 | . 0036215 | . 4039227 |
| _cons | -. 4742725 | . 4152314 | -1.14 | 0.253 | -1.288111 | . 3395662 |

## TABLE 3. PROBIT MARGINAL AND AVERAGE EFFECTS

Log likelihood $=-410.96762$
Pseudo R2 $=0.0443$

| target | $\mathrm{dF} / \mathrm{dx}$ | Std. Err. | z | $\mathrm{P}>\|\mathrm{z}\|$ | x-bar | [ 95\% | C.I. ] |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| female* | . 0124443 | . 0376975 | 0.33 | 0.741 | . 485911 | -. 061442 | . 08633 |
| age | . 0153486 | . 0060523 | 2.53 | 0.011 | 45.2672 | . 003486 | . 027211 |
| agesq | -. 0001762 | . 0000636 | -2.77 | 0.006 | 2356.57 | -. 000301 | $-.000052$ |
| highedu* | . 0740483 | . 0436885 | 1.63 | 0.102 | . 236235 | -. 01158 | . 159676 |
| lowedu* | . 0742953 | . 049191 | 1.48 | 0.139 | . 349812 | -. 022117 | . 170708 |
| highinc* | -. 041359 | . 0481081 | -0.87 | 0.383 | . 225984 | -. 135649 | . 052931 |
| lowinc* | . 0075692 | . 0474503 | 0.16 | 0.874 | . 298065 | -. 085432 | . 10057 |
| nonlab~t* | . 054818 | . 0445468 | 1.23 | 0.218 | . 535298 | -. 032492 | . 142128 |
| left* | -. 0051945 | . 0426634 | -0.12 | 0.903 | . 471853 | -. 088813 | . 078424 |
| right* | -. 1220372 | . 0524927 | -2.40 | 0.017 | . 219016 | -. 224921 | -. 019153 |
| reglow* | . 0434998 | . 0438869 | 0.97 | 0.332 | . 223194 | -. 042517 | . 129516 |
| reghigh* | -. 1738084 | . 0550714 | -3.28 | 0.001 | . 145327 | -. 281746 | -. 06587 |
| type* | . 0709977 | . 0354917 | 2.00 | 0.046 | . 50066 | . 001435 | . 14056 |
| obs. P | . 6904674 |  |  |  |  |  |  |
| pred. P | . 6978101 | (at x-bar) |  |  |  |  |  |

(*) $\mathrm{dF} / \mathrm{dx}$ is for discrete change of dummy variable from 0 to 1
$z$ and $P>|z|$ are the test of the underlying coefficient being 0

TABLE 4. PROBIT WITH SAMPLE SELECTION

|  | $=$ | 1209 |  |
| :--- | :--- | :--- | :--- |
|  | Number of obs | $=$ | 514 |


|  | Coef. | Std. Err. | z | $\mathrm{P}>\|\mathrm{z}\|$ | [95\% Conf. Interval] |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| target |  |  |  |  |  |  |
| female | . 0057774 | . 1480517 | 0.04 | 0.969 | -. 2843985 | . 2959534 |
| age | . 0489134 | . 0217279 | 2.25 | 0.024 | . 0063275 | . 0914993 |
| agesq | -. 0005496 | . 000214 | -2.57 | 0.010 | -. 0009691 | -. 0001302 |
| highedu | . 2097602 | . 1384156 | 1.52 | 0.130 | -. 0615294 | . 4810498 |
| lowedu | . 2111436 | . 1451737 | 1.45 | 0.146 | -. 0733917 | . 4956788 |
| highinc | -. 1141935 | . 1327207 | -0.86 | 0.390 | -. 3743213 | . 1459343 |
| lowinc | . 0226788 | . 1338767 | 0.17 | 0.865 | -. 2397148 | . 2850724 |
| nonlabmkt | . 1352117 | . 1500073 | 0.90 | 0.367 | -. 1587972 | . 4292207 |
| left | -. 0149317 | . 1195964 | -0.12 | 0.901 | -. 2493362 | . 2194729 |
| right | -. 3305266 | . 1424466 | -2.32 | 0.020 | -. 6097169 | -. 0513364 |
| reglow | . 1492482 | . 1469055 | 1.02 | 0.310 | -. 1386813 | . 4371778 |
| reghigh | -. 534301 | . 2321363 | -2.30 | 0.021 | -. 9892797 | -. 0793222 |
| type | . 2003553 | . 1029275 | 1.95 | 0.052 | -. 0013789 | . 4020894 |
| _cons | -. 7403945 | . 9407205 | -0.79 | 0.431 | -2.584173 | 1.103384 |


| Y |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| female | -. 2007844 | . 0792574 | -2.53 | 0.011 | -. 356126 | -. 0454427 |
| age | . 0364361 | . 0136711 | 2.67 | 0.008 | . 0096411 | . 063231 |
| agesq | -. 0003817 | . 0001372 | -2.78 | 0.005 | -. 0006505 | -. 0001129 |
| nonlabmkt | -. 0980364 | . 088269 | -1.11 | 0.267 | -. 2710405 | . 0749677 |
| reghigh | -. 4232526 | . 0963924 | -4.39 | 0.000 | -. 6121783 | -. 2343269 |
| reglow | . 1806764 | . 0976276 | 1.85 | 0.064 | -. 0106701 | . 372023 |
| single | -. 1022373 | . 1099509 | -0.93 | 0.352 | -. 317737 | . 1132624 |
| divorced | -. 1270649 | . 2394361 | -0.53 | 0.596 | -. 596351 | . 3422211 |
| widow | . 2583478 | . 152567 | 1.69 | 0.090 | -. 0406781 | . 5573737 |
| rural | . 1091552 | . 0928946 | 1.18 | 0.240 | -. 0729148 | . 2912253 |
| _cons | -. 3541731 | . 3366173 | -1.05 | 0.293 | -1.013931 | . 3055846 |
| / athrho | . 2713856 | . 9422371 | 0.29 | 0.773 | -1.575365 | 2.118136 |
| rho (p) | . 2649136 | . 8761116 |  |  | -. 917875 | . 9714895 |
| LR test of in | . eqns. | = 0) : | 2 (1) | 0. | Prob > ch | $=0.7742$ |


[^0]:    ${ }^{1}$ The data are based on those for England (Acheson, 1998). Available data on life expectancy at age 25 by education groups in Madrid and Barcelona show similar patterns (Borrel et al. 1999).

[^1]:    ${ }^{2}$ Although it would have been desirable to derive equivalent income, this has not been possible given the available information in the survey and also given that the variable household income was not available as a continuous variable.

