Response to the referee,

First of all, we would like to thank all your comments that have been of great help to improve this revised version of our paper.

As you said in your referee report, indeed, though the title could suggest that the issue to be addressed in the paper was a problem where dynamics were of interest, our model is a static model. In this paper we have chosen to focus our attention on the treatment of the question of sample selection bias in a static linear pseudo panel data model.

We also agree with your comments on Deaton (1985) and Moffit (1993) and we have tried to reflect their ideas in the Introduction. We have also tried to make a clear distinction in the text between matching instruments and other type of instruments as in Ridder and Moffit (2007) (RM).

Below, we describe the content of this revised version and respond to the issues raised in your report. We avoid answers to the specific comments on formulae of the old version because in general they are not in the revised version.

To make the revised version we have followed the final suggestion of your referee's report (last paragraph). We have analyzed in depth Semykina and Wooldridge (2010) (SW) article and the survey of RM and we have placed our proposal for the consistent estimation of pseudo panel parameters in the presence of selection bias within the framework of the generalized method of moments (GMM).

The second section of our revised paper introduces the main element of our approach. The estimation of a static linear pseudo panel data model in the presence of selection bias can be carried out in the traditional way if we remove a type of omitted-variables bias in the model. The omitted variable, the bias correction term, is proportional to the Heckman-Lee lambda (as you said) whose argument is a consistent estimation of the conditional probability that an individual is observed given that the individual belongs to a particular cohort. This proposal is in the line of Gronau (1974).

The main difference between our proposal and Gronau's, is that the latter leads to inconsistency in situations where the time-varying variables (tvc) that determine the selection process are relevant. In particular, this inconsistency occurs primarily when tvc are non-monotonous with respect to time. In the case of monotonous variables with respect to time Gronau's proposal produces good results because of the high correlation between Gronau's lambda and argument-consistent-estimation lambda.

The augmented estimating equation suggests a sample selection test through a test of H_0: $\rho=0$, where ρ is the lambda parameter, that can be performed in the usual way (for known lambda).

Our pseudo panel model consists of a main equation and a time-varying selection equation (that has to be estimated in the cross-section). It is a pseudo panel data version of

SW model. The third section of the revised version provides the identifying assumptions, discusses the two equations system and suggests the moments equations that are associated with the system. The estimation in the cross-section of the selection equation arises the procedure to treat the presence of fixed effects in the equation by means of Mundlak (1978) modeling device (used in SW too).

The fourth section contains the estimation of the cohort model in a GMM framework. In fact, as suggested by RM our method is a two-step GMMC. In the first stage we estimate T cross-section selection equations and construct lambda cohort-time estimates; in the second stage we estimate the augmented main equation (errors-in-variables). We review the well-known asymptotic properties of the estimators and provide a consistent estimation of the covariance matrix for known lambda. To take into account biases derived from estimated lambdas we follow and idea in Newey and McFadden (1993) and give an upper bound for the covariance matrix.

The fifth section is dedicated to a Monte Carlo experiment in which we analyze the power and size of the sample selection test in pseudo panel data models when using the lambdas derived from Gronau and our proposal. We show that the power of both tests is very similar and very high. In terms of size of the test, we find similar sizes in all situations whenever the selection mechanism is not determined by a tvc non-monotonous with respect to time. In the latter case the test derived from Gronau presents a very small size.

The sixth section is devoted to the empirical study.