Homework for Quantitative Economics for the Evaluation of the European Policy Homework for Period II

Irene Brunetti[1] Davide Fiaschi[2] Angela Parenti[3]

January 19, 2016

[1]Dipartimento di Economia e Management, Università di Pisa, Via Ridolfi 10, Pisa. *E-mail address:* ireneb@ec.unipi.it.

[2]Dipartimento di Economia e Management, Università di Pisa, Via Ridolfi 10, Pisa. *E-mail address:* davide.fiaschi@unipi.it.

[3]Dipartimento di Economia e Management, Università di Pisa, Via Ridolfi 10, Pisa. *E-mail address:* aparenti@ec.unipi.it.

Exams of Period II: First Part of the Course

The couples of student randomly matched and number of assignment for each couple **for the first part of the course**:

1. Dabrowski, Misior
2. Cattaruzzo, Gabellini
3. Scelzo, Saporito
4. Budiakivska, Carli
5. Asadova, Gasperin
6. Owusu, Lolini
7. Chen, Nkyi
8. Rolla, Palagi
9. Koyuncu, Tabarrani

The results must be commented and analysed according to the **theoretical models** discussed in the classes.

We expected a paper of **about 30 pages** and **a zip file with R codes** used to produce pictures, tables and estimates reported in the paper.

1. Analysis of GDP per worker together with manufacturing sector

The general question you should answer is about the relationship between the distribution of GDP per worker in European regions, the regional output composition and the European Regional Policy.

- download the dataset on the EU NUTS2 regions "datasetNUTS2_CE2010_EU27_forGRV_analysis" for the period 1991-2001 from the website of the course;
- estimate the density in the initial and final years of relative GDP per worker and discuss the resulting distribution dynamics in discrete (Markov matrix) and continuous time (stochastic kernel). Discuss the results in light of the possible reallocation of factors across EU regions;
- do the exploratory spatial analysis of the GDP per worker and the share of manufacturing in the initial and final years using the second-order contiguity matrix and map the distributions. Discuss the results in light of theoretical spatial models;
- estimate the impact of the change of the share of manufacturing between the initial and final year and the Objective 2 funds (relative to the periods 1989-1993 and 1994-1999) on the GDP per worker using both parametric and semiparametric methods. Discuss the results in the light of the purpose of the EU regional policy and of the theoretical model on the impact of EU funds on GDP per worker;
- test the endogeneity of both the change of the share of manufacturing and the Objective 2 funds using as instrument the initial level of share in manufacturing and the lagged funds (total funds of the periods 1975-1988) respectively;
- if the null hypothesis of exogeneity can be rejected at the usual level of significance (i.e., 5%) estimate the parametric model using TSLS and the semiparametric using the CF. Discuss the results;
- estimate the model as a Spatial Durbin model using the second-order contiguity matrix. Discuss the results in light of theoretical spatial models.

2. Analysis of GDP per worker together with construction sector

The general question you should answer is about the relationship between the distribution of GDP per worker in European regions, the regional output composition and the European Regional Policy.

- download the dataset on the EU NUTS2 regions "datasetNUTS2_CE2010_EU27_forGRV_analysis" for the period 1993-2008 from the website of the course;
- estimate the density in the initial and final years of relative GDP per worker and discuss the resulting distribution dynamics in discrete (Markov matrix) and continuous time (stochastic kernel). Discuss the results in light of the possible reallocation of factors across EU regions;
- do the exploratory spatial analysis of the GDP per worker and the share of construction in the initial and final years using the inverse distance (squared) spatial matrix and map the distributions. Discuss the results in light of theoretical spatial models;
- estimate the impact of the change of the share of construction between the initial and final year and the Objective 2 funds (relative to the periods 1994-1999 and 2000-2006) on the GDP per worker using both parametric and semiparametric methods.Discuss the results in the light of the purpose of the EU regional policy and of the theoretical model on the impact of EU funds on GDP per worker;
- test the endogeneity of both the change of the share of construction and the Objective 2 funds using as instrument the initial level of share in construction and the lagged funds (Objective 2 funds relative to the period 1989-1993) respectively;
- if the null hypothesis of exogeneity can be rejected at the usual level of significance (i.e., 5%) estimate the parametric model using TSLS and the semiparametric using the CF. Discuss the results;
- estimate the model as a Spatial Durbin model using the inverse distance (squared) spatial matrix. Discuss the results in light of theoretical spatial models.

3. Analysis of GDP per worker together with non-market services sector

The general question you should answer is about the relationship between the distribution of GDP per worker in European regions, the regional output composition and the European Regional Policy.

- download the dataset on the EU NUTS2 regions "datasetNUTS2_CE2010_EU27_forGRV_analysis" for the period 1993-2008 from the website of the course;
- estimate the density in the initial and final years of relative GDP per worker and discuss the resulting distribution dynamics in discrete (Markov matrix) and continuous time (stochastic kernel). Discuss the results in light of the possible reallocation of factors across EU regions;
- do the exploratory spatial analysis of the GDP per worker and the share of non-market services in the initial and final years using the inverse distance matrix and map the distributions. Discuss the results in light of theoretical spatial models;
- estimate the impact of the change of the share of non-market services between the initial and final year and the Objective 1 funds (relative to the periods 1994-1999 and 2000-2006) on the GDP per worker using both parametric and semiparametric methods. Discuss the results in the light of the purpose of the EU regional policy and of the theoretical model on the impact of EU funds on GDP per worker;
- test the endogeneity of both the change of the share of non-market services and the Objective 1 funds using as instrument the initial level of share in non-market services and the lagged funds (Objective 1 funds relative to the period 1989-1993) respectively;
- if the null hypothesis of exogeneity can be rejected at the usual level of significance (i.e., 5%) estimate the parametric model using TSLS and the semiparametric using the CF. Discuss the results;
- estimate the model as a Spatial Durbin model using the inverse distance matrix. Discuss the results in light of theoretical spatial models.

4. Analysis of GDP per worker together with financial intermediation sector

The general question you should answer is about the relationship between the distribution of GDP per worker in European regions, the regional output composition and the European Regional Policy.

- download the dataset on the EU NUTS2 regions "datasetNUTS2_CE2010_EU27_forGRV_analysis" for the period 1991-2008 from the website of the course;
- estimate the density in the initial and final years of relative GDP per worker and discuss the resulting distribution dynamics in discrete (Markov matrix) and continuous time (stochastic kernel). Discuss the results in light of the possible reallocation of factors across EU regions;
- do the exploratory spatial analysis of the GDP per worker and the share of financial intermediation sector in the initial and final years using the inverse distance matrix with cut-off (equal to the second quantile of the distance distribution) and map the distributions. Discuss the results in light of theoretical spatial models;
- estimate the impact of the change of the share of financial intermediation sector between the initial and final year and the total Structural and Cohesion funds (relative to the periods 1989-1993, 1994-1999 and 2000-2006) on the GDP per worker using both parametric and semiparametric methods. Discuss the results in the light of the purpose of the EU regional policy and of the theoretical model on the impact of EU funds on GDP per worker;
- test the endogeneity of both the change of the share of financial intermediation sector and total Structural and Cohesion funds using as instrument the initial level of share in financial intermediation sector and the lagged funds (total funds of the periods 1975-1988) respectively;
- if the null hypothesis of exogeneity can be rejected at the usual level of significance (i.e., 5%) estimate the parametric model using TSLS and the semiparametric using the CF. Discuss the results;
- estimate the model as a Spatial Durbin model using the inverse distance matrix with cut-off (equal to the second quantile of the distance distribution). Discuss the results in light of theoretical spatial models.

5. Analysis of GDP per capita together with manufacturing sector

The general question you should answer is about the relationship between the distribution of GDP per capita in European regions, the regional output composition and the European Regional Policy.

- download the dataset on the EU NUTS2 regions "datasetNUTS2_CE2010_EU27_forGRV_analysis" for the period 1991-2001 from the website of the course;
- estimate the density in the initial and final years of relative GDP per capita and discuss the resulting distribution dynamics in discrete (Markov matrix) and continuous time (stochastic kernel). Discuss the results in light of the possible reallocation of factors across EU regions;
- do the exploratory spatial analysis of the GDP per capita and the share of manufacturing in the initial and final years using the second-order contiguity matrix and map the distributions. Discuss the results in light of theoretical spatial models;
- estimate the impact of the change of the share of manufacturing between the initial and final year and the Objective 2 funds (relative to the periods 1989-1993 and 1994-1999) on the GDP per capita using both parametric and semiparametric methods. Discuss the results in the light of the purpose of the EU regional policy and of the theoretical model on the impact of EU funds on GDP per capita;
- test the endogeneity of both the change of the share of manufacturing and the Objective 2 funds using as instrument the initial level of share in manufacturing and the lagged funds (total funds of the periods 1975-1988) respectively;
- if the null hypothesis of exogeneity can be rejected at the usual level of significance (i.e., 5%) estimate the parametric model using TSLS and the semiparametric using the CF. Discuss the results;
- estimate the model as a Spatial Durbin model using the second-order contiguity matrix. Discuss the results in light of theoretical spatial models.

6. Analysis of GDP per capita together with construction sector

The general question you should answer is about the relationship between the distribution of GDP per capita in European regions, the regional output composition and the European Regional Policy.

- download the dataset on the EU NUTS2 regions "datasetNUTS2_CE2010_EU27_forGRV_analysis" for the period 1993-2008 from the website of the course;
- estimate the density in the initial and final years of relative GDP per capita and discuss the resulting distribution dynamics in discrete (Markov matrix) and continuous time (stochastic kernel). Discuss the results in light of the possible reallocation of factors across EU regions;
- do the exploratory spatial analysis of the GDP per capita and the share of construction in the initial and final years using the inverse distance (squared) spatial matrix and map the distributions. Discuss the results in light of theoretical spatial models;
- estimate the impact of the change of the share of construction between the initial and final year and the Objective 2 funds (relative to the periods 1994-1999 and 2000-2006) on the GDP per capita using both parametric and semiparametric methods.Discuss the results in the light of the purpose of the EU regional policy and of the theoretical model on the impact of EU funds on GDP per capita;
- test the endogeneity of both the change of the share of construction and the Objective 2 funds using as instrument the initial level of share in construction and the lagged funds (Objective 2 funds relative to the period 1989-1993) respectively;
- if the null hypothesis of exogeneity can be rejected at the usual level of significance (i.e., 5%) estimate the parametric model using TSLS and the semiparametric using the CF. Discuss the results;
- estimate the model as a Spatial Durbin model using the inverse distance (squared) spatial matrix. Discuss the results in light of theoretical spatial models.

7. Analysis of GDP per capita together with non-market services sector

The general question you should answer is about the relationship between the distribution of GDP per capita in European regions, the regional output composition and the European Regional Policy.

- download the dataset on the EU NUTS2 regions "datasetNUTS2_CE2010_EU27_forGRV_analysis" for the period 1993-2008 from the website of the course;
- estimate the density in the initial and final years of relative GDP per capita and discuss the resulting distribution dynamics in discrete (Markov matrix) and continuous time (stochastic kernel). Discuss the results in light of the possible reallocation of factors across EU regions;
- do the exploratory spatial analysis of the GDP per capita and the share of non-market services in the initial and final years using the inverse distance matrix and map the distributions. Discuss the results in light of theoretical spatial models;
- estimate the impact of the change of the share of non-market services between the initial and final year and the Objective 1 funds (relative to the periods 1994-1999 and 2000-2006) on the GDP per capita using both parametric and semiparametric methods. Discuss the results in the light of the purpose of the EU regional policy and of the theoretical model on the impact of EU funds on GDP per capita;
- test the endogeneity of both the change of the share of non-market services and the Objective 1 funds using as instrument the initial level of share in non-market services and the lagged funds (Objective 1 funds relative to the period 1989-1993) respectively;
- if the null hypothesis of exogeneity can be rejected at the usual level of significance (i.e., 5%) estimate the parametric model using TSLS and the semiparametric using the CF. Discuss the results;
- estimate the model as a Spatial Durbin model using the inverse distance matrix. Discuss the results in light of theoretical spatial models.

8. Analysis of GDP per capita together with financial intermediation sector

The general question you should answer is about the relationship between the distribution of GDP per capita in European regions, the regional output composition and the European Regional Policy.

- download the dataset on the EU NUTS2 regions "datasetNUTS2_CE2010_EU27_forGRV_analysis" for the period 1991-2008 from the website of the course;
- estimate the density in the initial and final years of relative GDP per capita and discuss the resulting distribution dynamics in discrete (Markov matrix) and continuous time (stochastic kernel). Discuss the results in light of the possible reallocation of factors across EU regions;
- do the exploratory spatial analysis of the GDP per capita and the share of financial intermediation sector in the initial and final years using the inverse distance matrix with cut-off (equal to the second quantile of the distance distribution) and map the distributions. Discuss the results in light of theoretical spatial models;
- estimate the impact of the change of the share of financial intermediation sector between the initial and final year and the total Structural and Cohesion funds (relative to the periods 1989-1993, 1994-1999 and 2000-2006) on the GDP per capita using both parametric and semiparametric methods. Discuss the results in the light of the purpose of the EU regional policy and of the theoretical model on the impact of EU funds on GDP per capita;
- test the endogeneity of both the change of the share of financial intermediation sector and total Structural and Cohesion funds using as instrument the initial level of share in financial intermediation sector and the lagged funds (total funds of the periods 1975-1988) respectively;
- if the null hypothesis of exogeneity can be rejected at the usual level of significance (i.e., 5%) estimate the parametric model using TSLS and the semiparametric using the CF. Discuss the results;
- estimate the model as a Spatial Durbin model using the inverse distance matrix with cut-off (equal to the second quantile of the distance distribution). Discuss the results in light of theoretical spatial models.

9. Analysis of GDP per worker together with Other Market Services

The general question you should answer is about the relationship between the distribution of GDP per worker in European regions, the regional output composition and the European Regional Policy.

- download the dataset on the EU NUTS2 regions "datasetNUTS2_CE2010_EU27_forGRV_analysis" for the period 1991-2008 from the website of the course;
- estimate the density in the initial and final years of relative GDP per worker and discuss the resulting distribution dynamics in discrete (Markov matrix) and continuous time (stochastic kernel). Discuss the results in light of the possible reallocation of factors across EU regions;
- do the exploratory spatial analysis of the GDP per worker and the share of Other Market Services sector in the initial and final years using the inverse distance matrix with cut-off (equal to the second quantile of the distance distribution) and map the distributions. Discuss the results in light of theoretical spatial models;
- estimate the impact of the change of the share of Other Market Services sector between the initial and final year and the total Structural and Cohesion funds (relative to the periods 1989-1993, 1994-1999 and 2000-2006) on the GDP per worker using both parametric and semiparametric methods. Discuss the results in the light of the purpose of the EU regional policy and of the theoretical model on the impact of EU funds on GDP per worker;
- test the endogeneity of both the change of the share of Other Market Services sector and total Structural and Cohesion funds using as instrument the initial level of share in Other Market Services sector and the lagged funds (total funds of the periods 1975-1988) respectively;
- if the null hypothesis of exogeneity can be rejected at the usual level of significance (i.e., 5%) estimate the parametric model using TSLS and the semiparametric using the CF. Discuss the results;
- estimate the model as a Spatial Durbin model using the inverse distance matrix with cut-off (equal to the second quantile of the distance distribution). Discuss the results in light of theoretical spatial models.

Homework of Period II: Second Part of the Course

The couples of student randomly matched and number of assignment for each couple **for the second part of the course** (the couples are the same of the first part of the course):

1. Dabrowski, Misior
2. Cattaruzzo, Gabellini
3. Scelzo, Saporito
4. Budiakivska, Carli
5. Asadova, Gasperin
6. Owusu, Lolini
7. Chen, Nkyi
8. Rolla, Palagi
9. Koyuncu, Tabarrani

The results must be commented and analysed according to the **theoretical models** discussed in the classes.

We expected a paper of **about 20 pages** and **a zip file with R codes** used to produce pictures, tables and estimates reported in the paper.

1. Assessing the impact of Job training provided under the Job Training Partnership Act

The U.S. Department of Labor began planning for an experimental evaluation of the training provided under the Job Training Partnership Act in 1986. The National JTPA study is the largest randomized training evaluation ever undertaken in the United States, collecting data on roughly 20,000 persons in 16 different sites around the country.

Between November 1987 and September 1989, eligible persons who applied for JTPA services were assigned to one of three different service groups classroom training, on-the-job training and some combination of the two. (Those with the weakest basic skills were assigned to classroom training.) During the application and assessment process, staff members explained to applicants that not all of them would be served and that a lottery would be conducted to determine who would participate. Therefore, it was only after assessment and assignment to service strategy group, that staff members telephoned a random assignment clerk at the central office to determine which applicants would be eligible to receive the treatment. One-third were assigned to a control group and prevented from receiving JTPA services.
Do the analysis only for MEN

- Download the dataset on the **JTPA.csv** and the codebook (**Code_variables_JTPA**) from the website of the course;

- To familiarize yourself with the setting, I strongly suggest you read the relevant parts of the paper "The benefits and costs of JTPA Title II-A programs: Key findings from the National Job Training Partnership Act study" (available on the course website).

- The assignment to treatment and control group were truly random. Discuss with the use of simple equations how in this setting it is possible to define and estimate the average treatment effect of trated.

- Test whether you see any differences between treatment and control in their baseline characteristics, such as race, previous work experience, marital status, age, etc.

- What is your estimate of the impact of the JTPA program on workers' earnings during 30 months after random assignment?

- Estimate the effect of the three different service strategies: classroom training, on-the-job training and other services.

- Generate an estimate of the program effect, after controlling for race, ethnicity, marital status, age and previous earnings using regression analysis. Does the estimated program effect change?

2. Assessing the impact of minimum wage laws on employment

In order to assess the impact of minimum wage laws on employment, two researchers (David Card and Alan Krueger) exploited the exogenous policy change in minimum wage that occurred in April 1992 in New Jersey (NJ): the hourly minimum wage was raised from 4.25 to 5.05 dollars in NJ but not in nearby States. Card and Krueger collected data at fast-food stores in NJ, before and after the policy change and also collected data at fast-food stores in a nearby State, Pennsylvania (PA), where the minimum wage was not changed.

- Download the dataset on the **card_data** and the codebook (**Code_variables_card**) from the website of the course.

- Define and discuss the Difference in Differences approach to estimating the average treatment effect on treated.

- How can you check in the data that the policy has indeed been implemented in NJ stores and not in PA stores? Implement the proposed check using the data.

- Write and estimate the regression model that give you the DiD estimatorof the effect of the policy change.

- How the results change when you include additional controls for: location within State, chain ownership, type of chain (i.e. KFC, Wendys, Roys or bk)? Comment the results.

- Discuss the validity of the DiD strategy implemented. Which problems can be arise with this approach?

3. Analysis of a natural experiment: same-sex siblings, fertility and women's labor supply

Suppose you are interested in estimating the effect of fertility on labor supply. In particular, you want to know how much a woman's labor supply falls when she has an additional kid. You have US census data from 1980 for a random sample of 30,000 married women aged 21-35 with two or more kids
Do the analysis only for WOMEN AGED BETWEEN 21-28

- Download the dataset on the **fertility.csv** and the codebook (**Code_variables_fertility**) from the website of the course.

- Think of weeks worked as the outcome variable (Y), and whether the woman has more than 2 kids as the treatment, (D). Assume a constant treatment effect, $A1 : \beta_{Di} = \beta_D$, and run the OLS regression of Y on D, W_1, W_2, W_3, W_4, W_5, W_6. Interpret the estimated coefficient on D.;

- U are all the other factors, besides D and the W_s above, that also affect Y. $E(U|D, W, ...W) = E(U|W, ...W)$ and $E(U|D, W, ...W) = 0$ are two alternative precise mathematical statements of the identifying assumption for the OLS estimate on D to represent the causal effect of fertility on labor supply. Discuss the plausibility of each assumption in this application;

- You decide to use Z as an instrument for D to estimate β_D. Interpret your estimate.

- Assess empirically whetheher *samesex* is a weak instrument.

4. Assessing the impact of Job training provided under the Job Training Partnership Act

The U.S. Department of Labor began planning for an experimental evaluation of the training provided under the Job Training Partnership Act in 1986. The National JTPA study is the largest randomized training evaluation ever undertaken in the United States, collecting data on roughly 20,000 persons in 16 different sites around the country.

Between November 1987 and September 1989, eligible persons who applied for JTPA services were assigned to one of three different service groups classroom training, on-the-job training and some combination of the two. (Those with the weakest basic skills were assigned to classroom training.) During the application and assessment process, staff members explained to applicants that not all of them would be served and that a lottery would be conducted to determine who would participate. Therefore, it was only after assessment and assignment to service strategy group, that staff members telephoned a random assignment clerk at the central office to determine which applicants would be eligible to receive the treatment. One-third were assigned to a control group and prevented from receiving JTPA services.
Do the analysis only for WOMEN

- Download the dataset on the **JTPA.csv** and the codebook (**Code_variables_JTPA**) from the website of the course;

- To familiarize yourself with the setting, I strongly suggest you read the relevant parts of the paper "The benefits and costs of JTPA Title II-A programs: Key findings from the National Job Training Partnership Act study" (available on the course website).

- The assignment to treatment and control group were truly random. Discuss with the use of simple equations how in this setting t is possible to define and estimate the average treatment effect of trated.

- Test whether you see any differences between treatment and control in their baseline characteristics, such as race, previous work experience, marital status, age, etc.

- What is your estimate of the impact of the JTPA program on workers earnings during 30 months after random assignment?

- Estimate the effect of the three different service strategies: classroom training, on-the-job training and other services.

- Generate an estimate of the program effect, after controlling for race, ethnicity, marital status, age and previous earnings using regression analysis. Does the estimated program effect change?

5. Analysis of a natural experiment: same-sex siblings, fertility and women's labor supply

Suppose you are interested in estimating the effect of fertility on labor supply. In particular, you want to know how much a woman's labor supply falls when she has an additional kid. You have US census data from 1980 for a random sample of 30,000 married women aged 21-35 with two or more kids
Do the analysis only for WOMEN AGED BETWEEN 29-35

- Download the dataset on the **fertility.csv** and the codebook (**Code_variables_fertility**) from the website of the course.

- Think of weeks worked as the outcome variable (Y), and whether the woman has more than 2 kids as the treatment, (D). Assume a constant treatment effect, $A1 : \beta_{Di} = \beta_D$, and run the OLS regression of Y on D, W_1, W_2, W_3, W_4, W_5, W_6. Interpret the estimated coefficient on D.;

- U are all the other factors, besides D and the W_s above, that also affect Y. $E(U|D, W, ...W) = E(U|W, ...W)$ and $E(U|D, W, ...W) = 0$ are two alternative precise mathematical statements of the identifying assumption for the OLS estimate on D to represent the causal effect of fertility on labor supply. Discuss the plausibility of each assumption in this application;

- You decide to use Z as an instrument for D to estimate β_D. Interpret your estimate.

- Assess empirically whetheher *samesex* is a weak instrument.

6. Assessing the impact of Workplace Training Programs on the re-employment probability of young unemployed workers

The workplace training program is a measure that aims to create a direct link between a job-seeker and a firm, and to allow the trainee to gain more experience to upgrade his curriculum and to facilitate a future possible business relationship with the host firm. The Province of Pisa adtopts this type of active labour program in order to accelerate the matching process between demand and supply in the labour market.
Do the analysis only for WOMEN

- Download the dataset on the **TrainingPisa.csv** and the codebook (**Code_variables_Training**) from the website of the course;
- Calculate means of covariates variables for treatment and control group;
- Perform matching estimations: first estimate the propensity score. Try varying the specification for the propensity score; compute histograms of the propensity scores for treatment and control groups; and check for balance;
- Estimate the Average Treatment Effect (ATE).

7. Assessing the impact of Job training provided under the Job Training Partnership Act

The U.S. Department of Labor began planning for an experimental evaluation of the training provided under the Job Training Partnership Act in 1986. The National JTPA study is the largest randomized training evaluation ever undertaken in the United States, collecting data on roughly 20,000 persons in 16 different sites around the country.

Between November 1987 and September 1989, eligible persons who applied for JTPA services were assigned to one of three different service groups classroom training, on-the-job training and some combination of the two. (Those with the weakest basic skills were assigned to classroom training.) During the application and assessment process, staff members explained to applicants that not all of them would be served and that a lottery would be conducted to determine who would participate. Therefore, it was only after assessment and assignment to service strategy group, that staff members telephoned a random assignment clerk at the central office to determine which applicants would be eligible to receive the treatment. One-third were assigned to a control group and prevented from receiving JTPA services.
Do the analysis distinguishing between:
ADULT (aged between 36 and 54) and YOUTH (aged between 22 and 35) MEN

- Download the dataset on the **JTPA.csv** and the codebook (**Code_variables_JTPA**) from the website of the course;

- To familiarize yourself with the setting, I strongly suggest you read the relevant parts of the paper "The benefits and costs of JTPA Title II-A programs: Key findings from the National Job Training Partnership Act study" (available on the course website).

- The assignment to treatment and control group were truly random. Discuss with the use of simple equations how in this setting t is possible to define and estimate the average treatment effect of trated.

- Test whether you see any differences between treatment and control in their baseline characteristics, such as race, previous work experience, marital status, age, etc.

- What is your estimate of the impact of the JTPA program on workers earnings during 30 months after random assignment?

- Estimate the effect of the three different service strategies: classroom training, on-the-job training and other services.

- Generate an estimate of the program effect, after controlling for race, ethnicity, marital status, age and previous earnings using regression analysis. Does the estimated program effect change?

8. Assessing the impact of the National Supported Work (NSW) Demonstration on post-intervantion earnings

Lalonde (1986) is the first to examine the impact of the randomized experiment, the National Supported Work Demonstration, NSW, a labor training program, on post-intervantion earnigns. He obtains an unbiased estimate of the training effect from the randomized experiment. Dehejia and Wahba (1999), using Lalonde's data, apply the propensity score methods to estimate the impact of the program and they demonstrate that, relative to the estimators that Lalonde evaluates, the propensity score estimates of the treatment impact are much closer to the experimental benchmark estimate.

- To familiarize yourself with the setting, I strongly suggest you read the relevant parts of the paper "Causal Effects in Nonexperimental Studies Reevaluating the Ealuation of Training Programs" (available on the website of the course).

- Install Matching package and download the dataset (lalonde) as explained on R'page: http://sekhon.berkeley.edu/matching/lalonde.html

- Calculate means of covariates variables for treatment and control group.

- Perform matching estimations: first estimate the propensity score. Try varying the specification for the propensity score; compute histograms of the propensity scores for treatment and control groups; and check for balance.

- Estimate the Average Treatment Effect on the Treated (ATT).

9. Assessing the impact of Workplace Training Programs on the re-employment probability of young unemployed workers

The workplace training program is a measure that aims to create a direct link between a job-seeker and a firm, and to allow the trainee to gain more experience to upgrade his curriculum and to facilitate a future possible business relationship with the host firm. The Province of Pisa adtopts this type of active labour program in order to accelerate the matching process between demand and supply in the labour market.
Do the analysis only for MEN

- Download the dataset on the **TrainingPisa.csv** and the codebook (**Code_variables_Training**) from the website of the course;
- Calculate means of covariates variables for treatment and control group;
- Perform matching estimations: first estimate the propensity score. Try varying the specification for the propensity score; compute histograms of the propensity scores for treatment and control groups; and check for balance;
- Estimate the Average Treatment Effect (ATE).