

**UNIVERSITÀ
DEGLI STUDI
DI PADOVA**

Sede Amministrativa: Università degli Studi di Padova

Dipartimento di Scienze Statistiche
SCUOLA DI DOTTORATO DI RICERCA IN SCIENZE STATISTICHE
CICLO XXVII

**Combining Jump and Kink ratio
estimators in Regression Discontinuity
Designs,
with an application to the causal effect of retirement on
well-being**

Direttore della Scuola: Prof.ssa Monica Chiogna

Supervisore: Prof. Enrico Rettore

Dottorando: Gianluca Mazzarella

Abstract

Regression Discontinuity Design (RDD) is one of the most popular designs in the field of causal inference in nonexperimental settings. It is based on the idea that the treatment is (totally or partially) determined by a threshold point of an observed continuous variable. When the treatment is just partially determined by that variable, it is usually defined *fuzzy* RDD. In this setting, given a certain outcome, the only effect that one is able to identify is the Average Treatment Effect (ATE) for the subpopulation of the *Compliers* at the threshold point. The ATE could be obtained by the ratio between the discontinuity at the threshold point in the average of the observed outcome divided by the discontinuity in the treatment probability.

This thesis explores, from a methodological and empirical perspective, how the change of slope at the threshold point is informative for the estimate of the parameter of interest. Starting from the changes of the eligibility criterion for retirement that took place in Italy in the '90s we propose an alternative estimator, based on *Instrumental Variables*, that is a combination of the discontinuity and the change of slope.

Furthermore we provide a simulation study to compare the efficiency of the different estimators. Then we analyze the effects of retirement on the subjective well-being. Finally we generalize the results using the *Two Sample Instrumental Variable* estimator, in order to improve the efficiency of estimates based on administrative data and to construct delayed outcomes for the same cohorts.

Sommario

Il Regression Discontinuity Design è una delle più diffuse tecniche nell'ambito dell'inferenza causale nei processi quasi-sperimentali. È basata sull'idea che l'esposizione ad un trattamento sia (parzialmente o totalmente) stabilita da un punto di soglia di una variabile continua e osservabile. Quando l'esposizione al trattamento è solo parzialmente stabilita da questa variabile, si è solito definirlo *fuzzy* Regression Discontinuity Design. In questo contesto, dato un determinato outcome di interesse, è possibile identificare soltanto l'effetto medio del trattamento per la sotto-popolazione dei *Compliers*. Tale effetto può essere ottenuto dal rapporto tra la discontinuità nel punto di soglia nella media dell'outcome divisa per la discontinuità nella probabilità di esposizione al trattamento.

La tesi esamina, da un punto di vista metodologico e empirico, come possano essere informativi per la stima del parametro di interesse i cambiamenti di pendenza nel punto di soglia. Partendo dalle modifiche nei criteri di ammissibilità al pensionamento avvenuti in Italia a partire dagli anni '90, abbiamo proposto uno stimatore, basato sulla logica delle *Variabili Strumentali*, che è una combinazione della discontinuità e del cambiamento di pendenza. In seguito abbiamo proposto uno studio di simulazione per confrontare l'efficienza dei diversi stimatori. Successivamente abbiamo analizzato gli effetti del pensionamento sulla soddisfazione personale percepita. Infine abbiamo generalizzato, usando lo stimatore a *Variabili Strumentali su Due Campioni* per migliorare l'efficienza delle stime con dati amministrativi o per costruire outcome successivi al pensionamento.

Acknowledgements

First of all I want to thank my advisor, prof. Enrico Rettore, for the opportunity to work with him, for directing me in these two years, for opening my mind with his useful suggestions, and, most of all, for his infinite patience with, probably, the most bipolar PhD student of the history.

To my family, for their support, for their encouragement, even if mostly of time they do not have understood my decisions.

A special thank is to the XXVII cycle: Mareg, Marco (the capitalist), Vera (DCF), Leo and Paola. *A PhD class is a mixture between a university class and a kindergarten* (cit.). These three years are characterized by a great sharing of time, ideas (few), moods (much more). I am sure that I would not be able to finish this PhD without you.

In my professional life a special place is forever reserved to Irvapp, the first (and until now the last) who dared to offer me a position.

Among the others a special thanks is reserved to Simone Schüller for her useful comments to my work.

I was lucky to have worked in places in which there are a lot of friends that have helped me with life and work suggestions that in most cases I have ignored. Among the others I want to thank: Giovanna, Nicola, Clovis, Roberta and all the younger PhD students (DipStat); Marco, Michele and Monica (endogeneity-land); Ilaria, Nadir, Lollo, Davide, Silvia, Simona, Zek (Irvapp).

To Damiano, Gianni, Giorgio, Pip and Vito. Offlaga say that friendship has a natural expiration date; for us it does not work even if 15 years start to be too much!

To Peppe and Fede of the Safarà for guaranteeing during my holidays the alcoholic level necessary to reduce my *work-related* stress.
Last but non least I want to thank prof. Adimari to be my job-motivator.

Contents

Contents	vii
List of Figures	ix
List of Tables	xi
1 Introduction	1
1.1 Overview	1
1.2 Main contributions of the thesis	3
2 Pension Reforms and Retirement Effects	5
2.1 Literature Review	5
2.2 <i>Aspetti della Vita Quotidiana</i> Survey	10
2.3 First Stage Estimation	13
3 Combining Jump and Kink ratio	17
3.1 Review of the literature	17
3.1.1 Regression Discontinuity Design	17
3.1.1.1 Estimation	20
3.1.2 Regression Kink Design	21
3.2 A combined version of the RDD & RKD estimators	24
3.3 Simulation Study	26
4 Results with <i>Aspetti della Vita Quotidiana</i> Survey	33
4.1 Subsample selection	36

CONTENTS

4.1.1	Robustness	36
4.1.2	Stratification	37
5	Two Sample Instrumental Variable	41
5.1	The Work Histories Italian Panel database	42
5.2	Delayed effects	45
6	Conclusions	49
	Appendix	52
A	Alternative specifications of the IV regression	53
B	Conditional Moments of the Running Variable in the IV regression	61
	References	67

List of Figures

2.1	Ratio of older dependents (older than 64) to the working-age population (age between 15 and 64).	6
2.2	Retirement Probability across cohorts – Age 52	14
3.1	Conditional average of the first stage equation for $K = 1$	28
5.1	First Stage Estimation using Whip for 52	44
5.2	Delayed effects 52 Male	47

List of Tables

2.1	Old age pension: evolution of eligibility rules	8
2.2	Seniority pension: evolution of eligibility rules	8
2.3	Sample size for age 52	12
3.1	Root Mean Square Error of the different estimators with $K = 1$ and $L' = 0$	30
3.2	Root Mean Square Error of the different estimators with $K = 1$ and $L' = 0.5$	31
4.1	Frequency distribution of the outcome variables about personal satisfaction	34
4.2	Results Age 52	35
4.3	Results with bandwidth reduction	37
4.4	Results excluding individuals with high educational attainment .	38
4.5	North-South Stratification	39
5.1	Sample size of Whip in comparison to AVQ survey	43
5.2	First Stage Comparison	44
5.3	Estimates for 52 y.o. population obtained combining AVQ and Whip data source	45

Chapter 1

Introduction

1.1 Overview

In the last three decades pension schemes have been often modified due to the increasing of life expectancy and the population aging. The aim of the changes is to maintain the financial sustainability of the social security system. The alternatives considered are typically the reduction of the amount of the pension or the change in the eligibility rules. Often policy-maker have decided to increase the requirements in order to preserve a good level of consumption in the older age (see for example Galasso, 2008).

In Italy the first reform of the pension scheme was put in place in 1992 with the so called Amato's law that modifies the eligibility criteria for the retirement age. This was followed by the Dini's law in 1995, the Prodi's law in 1997, the Maroni's law in 2004, which all changed the eligibility criteria for the seniority pension. These laws have brought an increase in the eligibility rules to retirement.

Retirement is a fundamental event in the life-cycle of a person. Many papers try to investigate the effects of retirement in order to understand which could be the reaction to an event that increases the leisure time and potentially reduce the wage of the retirees. Battistin et al. (2009) investigate the effects of retirement on consumption. Billari and Galasso (2014) analyze the fertility decision of the couples affected by the reforms and consequently have lower income perspectives. Many papers try also to investigate the effect of retirement on the subjective well-

being and depression. Charles (2004) analyze the causal effects of retirement in US, finding that people that are retired are less depressed and that there is an improvement in the subjective well-being. Bonsang and Klein (2012) study the retirement effects in Germany focusing on the differences between *voluntary retirements* and *involuntary retirements* (i.e. the person wants to continue working but is forced to retire). Kim and Moen (2002) underline the contrast between the so-called short-term “honeymoon” effects (i.e. the person is more happy in the first period after retirement) and the negative long-term effects in the US. Bertoni and Brunello (2014), instead, find that in Japan, for a wife the probability to be depressed increases when the husband retires.

The aim of the thesis is to estimate the effects of retirement on well-being. The analysis is based on the same logic as in the Regression Discontinuity Design (hereafter RDD), i.e. to identify the causal effects by comparing cohorts marginally unaffected by the reforms to those marginally affected. RDD, first introduced by Thistlethwaite and Campbell (1960), has found increasing success in the late '90s (see for example Hahn et al., 2001 for the identification assumption or Imbens and Lemieux, 2008; Lee and Lemieux, 2010 for a review and a guide to practice). The key idea is that the treatment status is (totally or partially) determined by a threshold point of a continuous observable variable (often defined *assignment*, *forcing* or *running* variable).

In the recent literature, some papers investigate the treatment effect estimation using the information given by the change of slope in the treatment probability or in treatment amount, the so called *Regression Kink Design* (hereafter RKD). Some papers use the RKD in order to estimate the effects of a *continuous treatment*, such as the price sensitivity of demand for prescription drug using the kinked reimbursement schemes provided by the Danish law (see Simonsen et al., 2015). Instead Card et al. (2012) investigate the effect of unemployment insurance benefits on the duration of joblessness in Austria, where the benefit schedule has kinks at the minimum and maximum benefit level. Dong (2014) investigates the RK estimator in a binary treatment setting (see Rubin, 1974). She shows that the kink ratio is a combination of the parameter of interest (the average treatment effect for the compliers at the threshold point, see Angrist et al., 1996

for the definition of compliance in a randomized experiment) and an unobserved bias term, but she analyzes in detail what could be the best combination between the standard jump-ratio and the kink-ratio just in the case in which the last one is not biased.

1.2 Main contributions of the thesis

The main contributions of the thesis are in the following directions:

- Using data from the survey called *Aspetti della Vita Quotidiana* (Daily Life Aspects) edited by the *Italian National Institute of Statistics* we have estimated the effect on a series of socio-economic outcomes in order to understand which could be the effects of retirement for the cohorts across the introduction of the reforms.
- The thesis proposes an alternative estimator based on *instrumental variables* that takes into account the information provided by the Jump and the Kink ratio. The aim is to find a combination of the two ratios that improves the efficiency of the standard Jump ratio when the Kink ratio is potentially biased and not only (as in the Dong's paper) that minimizes the variance.
- We have implemented a simulation study in order to test the performance of the different estimators.
- We have used administrative data provided by the *Italian National Social Security Institute* together with survey data. The two main advantages of the administrative data is a large sample size and the possibility to observe the exact moment in which a worker retires. Unfortunately these data do not provide interesting outcomes, for that reason it was necessary to combine administrative and survey data in order to implement a *Two Sample Instrumental Variable* estimator (see Angrist and Krueger, 1992).
- Finally we have used the next waves of the survey data in order to have delayed outcomes of the same cohorts. Unfortunately the survey does not

1. Introduction

provide any repeated measure of the same individuals, so the estimates are constructed using the same logic of the Two Sample Instrumental Variable.

Chapter 2

Pension Reforms and Retirement Effects

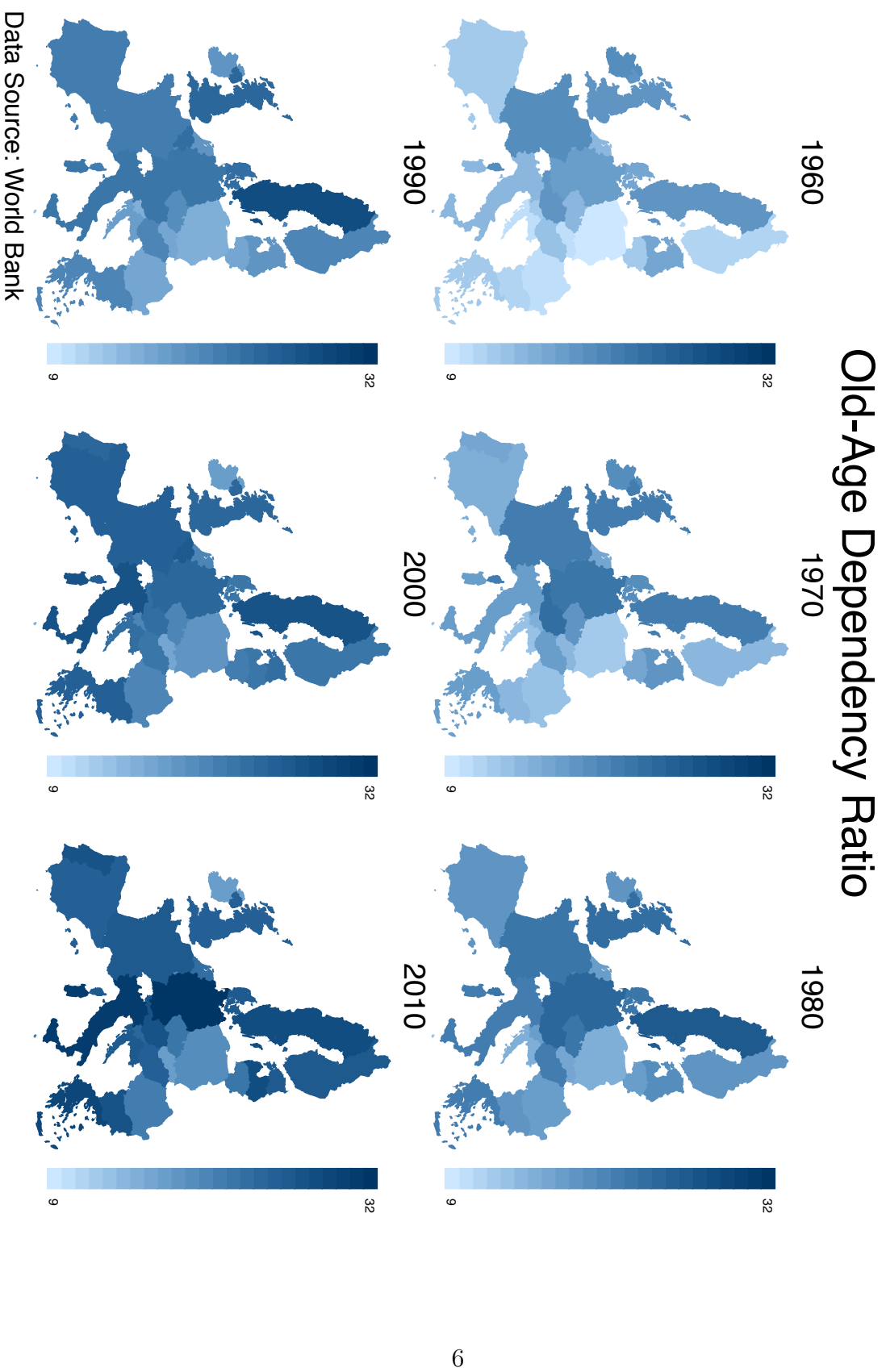
2.1 Literature Review

The increase in life expectancy and the birth reduction have caused, in the second part of the XX century, a relevant situation of population ageing. Figure 2.1 summarizes the phenomenon across the European Union. In addition to that, data suggest an increase in anticipated retirement and consequently a reduction of the participation at work of elderly people (see for example Costa Dora, 1998).

These two trends have progressively unbalanced the ratio between retired and working people, compromising the financial sustainability of the social security system (see Galasso and Profeta, 2004). In general the possible solutions of this problem could be the increase of the contribution rate for workers, the reduction of the amount of the pensions and the tightening of the eligibility criterion for the retirement. The first possibility was hardly ever considered in particular for inter-generational equity reasons. The second was often avoided in order to preserve a good level of standard of living in older age (see Galasso, 2008).

Therefore policy makers typically have decided to increase the age for retirement. In Italy the first change was put in place in 1992 with the so called Amato's law that modifies the eligibility criteria for the *old age pension*. This was followed by

Figure 2.1: Ratio of older dependents (older than 64) to the working-age population (age between 15 and 64).



2. Pension Reforms and Retirement Effects

the Dini's law in 1995, the Prodi's law in 1997, the Maroni's law in 2004, which all changed the eligibility criteria for the *seniority pension*.

The Amato's law, starting from 1993 to 2001, has brought a progressive increase in the years of contributions (from 15 to 20 years) and the retirement age (from 60 to 65 years old for men and from 55 to 60 years for women).

The Dini's law introduced two alternative rules regulating pension eligibility, stating that this can be obtained either at the age of 57 with 35 years of contributions, or with 40 years of contributions regardless of the age. As for the Amato's law, the introduction of the Dini's law was gradual, with age and contribution criteria increasing from 1996 to 2006 and further evolution of the contribution criteria from 1996 to 2008. The Dini's law also introduced the so called "*contribution system*" for the calculation of the pension amount. For people who started to work after of the 1st January 1996 the amount of pension is not calculated on the basis of wages in the last 5 year, but on the basis of the total amount of contributions paid.

The Prodi's law in 1997 only anticipated the role out of Dini's law from 2006 to 2002.

The Maroni's law in 2004 increased the minimum age for retirement from 57 to 65 for men and 60 for women, but it did not change the 40 years of contribution condition to obtain pension regardless of age.

A summary of the regulatory changes that characterize these intensive period of reforms is presented in Tables 2.1 and 2.2.

Many papers try to investigate the effects of retirement on consumption. In the economic theory consumption is supposed to be smooth across life cycle. The expenditure, seen as a function of age, on consumption is concave and gradually decreases in the older age, but it is not discontinuous at retirement (see Gourinchas and Parker, 2002, for an explanation on the role of precautionary saving in the older age). However empirical results underline a drop at retirement. This phenomenon has been called "*Retirement Consumption Puzzle*". There are several explanations of that discontinuity, such as the cessation of work-related expenses, the inadequate foresee of the decline in income associated with retirement, the increase of the leisure time etc. (see Blau, 2008; Hurd and Rohwedder,

2. Pension Reforms and Retirement Effects

Table 2.1: Old age pension: evolution of eligibility rules

Period	Age		Contribution Year
	Male	Female	
- 31/12/1992	60	55	15
01/01/1993 - 31/12/1993	60	55	16
01/01/1994 - 31/12/1994	61	56	16
01/01/1995 - 30/06/1995	61	56	17
01/07/1995 - 31/12/1996	62	57	17
01/01/1997 - 30/06/1998	63	58	18
01/07/1998 - 31/12/1998	64	59	18
01/01/1999 - 31/12/1999	64	59	19
01/01/2000 - 31/12/2000	65	60	19
01/01/2001 -	65	60	20

Table 2.2: Seniority pension: evolution of eligibility rules

Period	Age &	Only
	contribution	contribution
- 31/12/1995	-	35
01/01/1996 - 31/12/1997	52 & 35	36
01/01/1998 - 31/12/1998	54 & 35	36
01/01/1999 - 31/12/2000	55 & 35	37
01/01/2001 - 31/12/2001	56 & 35	37
01/01/2002 - 31/12/2003	57 & 35	37
01/01/2004 - 31/12/2005	57 & 35	38
01/01/2006 - 31/12/2007	57 & 35	39
01/01/2008 -	57 & 35	40

2005, for a detailed explanation).

Among the others, Aguila et al. (2011) estimate the effects on nondurable consumption for US retired between the 1980 and 2000. They find that there is not a significant reduction, instead they observe a reduction of food consumption at retirement. Battistin et al. (2009) analyze the Italian situation finding a reduction of about the 10% of the nondurable expenditure and a reduction of about the 14% of food expenditure for people at the threshold of the pension eligibility criteria. Others papers underline that the reduction in food expenditure is caused by the increasing of the home production caused by the increment of the leisure time (see among the others Luengo-Prado and Sevilla, 2013; Stancanelli and Van Soest, 2012).

Other studies try to investigate what could be the effects, in a sociological and psychological perspective, of retirement. For example is highly debate the role of retirement on subjective well-being. Charles (2004) analyzes the effects of retirement on well-being using as a source of exogenous variation the Social Security Amendment of 1983 that in US discouraged retirement before 65 years old and increased the benefits for people that decided to continue to work after 65. He focuses his attention on the voluntary retirement and shows that there is an improvement of subjective well-being if is taken into account the endogeneity of retirement (i.e. the people with lower well-being decide to withdraw from the labor force). Bonsang and Klein (2012); Hershey and Henkens (2013), respectively in Germany and in Netherlands, distinguish the role of *voluntary* and *involuntary* retirement. Bonsang and Klein (2012) find, for voluntary retirement, a positive effect on satisfaction with leisure time and a negative effect on satisfaction with household income. They also underline as for involuntary retirement the effect on satisfaction with income is more negative and the effect on satisfaction with leisure is less positive in comparison with voluntary retirement. Hershey and Henkens (2013) analyze the Satisfaction With Life (SWL) of Dutch retired. They show as involuntary retirement has a negative effect on SWL, in particular when it is caused by some health problems. Moreover the voluntary retirement seems to have a negligible positive effect. Kim and Moen (2002) starting from data of the employers of the New York State analyze the effect of retirement on morale and on

2. Pension Reforms and Retirement Effects

the probability to have depressive symptoms . They find a positive effect on well being for men that are just retired and this effects is bigger for those who have the lower level of pre-retirement well-being. However they show that there is also an increase in the probability to developing depressive symptoms. Finally they do not find any effects for retired women. Bertoni and Brunello (2014) proposed an analysis on the so called *Retired Husband Syndrom* in Japan. They estimate the effect of the husband's retirement on the wife's health using the exogenous variation provided by the 2006 revision of retirement eligibility criterion for the cohorts born after the 1945. They find a negative effect on wife's health and this effect is exacerbate for employed women.

Gall et al. (1997) try to give some empirical evidence to the theory proposed first of all by Atchley (1976) in which there is a positive short term effect of the retirement on well-being (often defined as an *honeymoon* period), and a negative effect in the middle-long term period. They show that 1 years after retirement in Canada there is a positive effect on some indicators such as psychological health, financial and interpersonal satisfaction and locus of control. However at 6-7 years after retirement they find a decrease of the interpersonal satisfaction and of the psychological health, while the locus of control continue to increase and the financial satisfaction remain stable.

2.2 *Aspetti della Vita Quotidiana* Survey

The aim is to evaluate the retirement effects using the exogenous variation caused by the change in the eligibility criterion provided by the Dini's Law. The first step of that Law was the 1st January 1996 and it imposed that private sector employers have reached at least 52 years of age and 35 years of contributions or 36 years of contributions regardless of age (in comparison with 35 years of contributions) to retire. However workers cannot retire in every moment of the year because the law introduced also the so called *retirement windows* that are fixed periods in which it is possible to stop to work (see Leombruni et al., 2012, for details). For that reason in private sector most of the retirements are in the 31th December and the first day of retirement is the 1st January of the next year.

So we expect that the first reduction in the number of retired are in 1997. It is reasonable to suppose that the reforms has involved a very selected part of the population, namely the individuals that have reached 35 years of contribution at age of 51. So we expect that they started to work very young, and, consequently, they have a low level of education.

In order to understand how the reforms have changed the retirement probability across different cohorts, we need a data source that allow us to compare, at a given age, people born in different years. Moreover these data should provide information about health, interpersonal, family and economic satisfaction, use of leisure time etc.

For all of these reasons we have decided to start from the survey named *Aspetti della Vita Quotidiana* (Aspects of Daily Life, hereafter AVQ) carried out by the Italian National Statistical Office. This survey each year involves about 50000 individuals belonging to about 20000 households and it is a part of an integrated system of social surveys called *Indagine Multiscopo sulle Famiglie* (The Multipurpose Surveys on Household). It covers the period from 1993 to 2012 (excluding 2004) for a total of 19 datasets. It includes information about which is the quality of individual life, the degree of satisfaction of their conditions, their economic situation, the area in which they live, the functioning of all public utility services.

As already said, we needed to compare across cohorts the retirement probability at a certain age. So we have fixed an age and we have extracted all the records with that age in 4 surveys before the 1997 (1993-1996) and 4 after 1997 (1997-2000). Furthermore we wanted to observe if their outcomes change across time but we do not have repeated measure for the same individuals, but we could observe individuals with one year more in the next year survey. Obviously they are not the same individuals, but they are both representative of the same population, for that reason it seems to be a reliable outcome to estimate possibles delayed effects of retirement. Table 2.3 summarizes the sample size of the obtained dataset for age 52.

2. Pension Reforms and Retirement Effects

		Year of survey														
		1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	
Age	52	287	309	310	331	289	473	343	355							
	53		322	347	292	312	346	335	370	298						
	54			346	275	321	373	242	325	305	341					
	55				341	323	376	279	268	290	301	281				
	56					283	384	278	287	244	292	278				
	57						361	303	308	224	249	302		248		
	58							293	283	237	237	246		256	266	

(a) Male

		Year of survey														
		1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	
Age	52	306	326	321	305	299	490	349	351							
	53		375	320	300	302	370	340	335	304						
	54			341	323	314	412	326	358	292	343					
	55				367	294	400	305	304	314	308	326				
	56					321	482	316	338	232	312	307				
	57						429	303	311	259	271	312		285		
	58							307	288	244	250	232		267	276	

(b) Female

Table 2.3: Sample size for age 52

2.3 First Stage Estimation

In this section we display how the retirement probability is changed due to the introduction of the Dini's law. Moreover will be discussed the endogenous nature of the retirement and because it is not possible just to compare the two groups (retired and not retired) in order to obtain a causal estimand. Finally will be presented some of the assumptions that are necessary to identify the causal effects. Starting from the data described in the previous section, we have selected just the records with the age of reference (in other words just the first row of the tables in 2.3) and we have estimated simple *linear probability models* (see section 15.2 of Wooldridge, 2010) in the form:

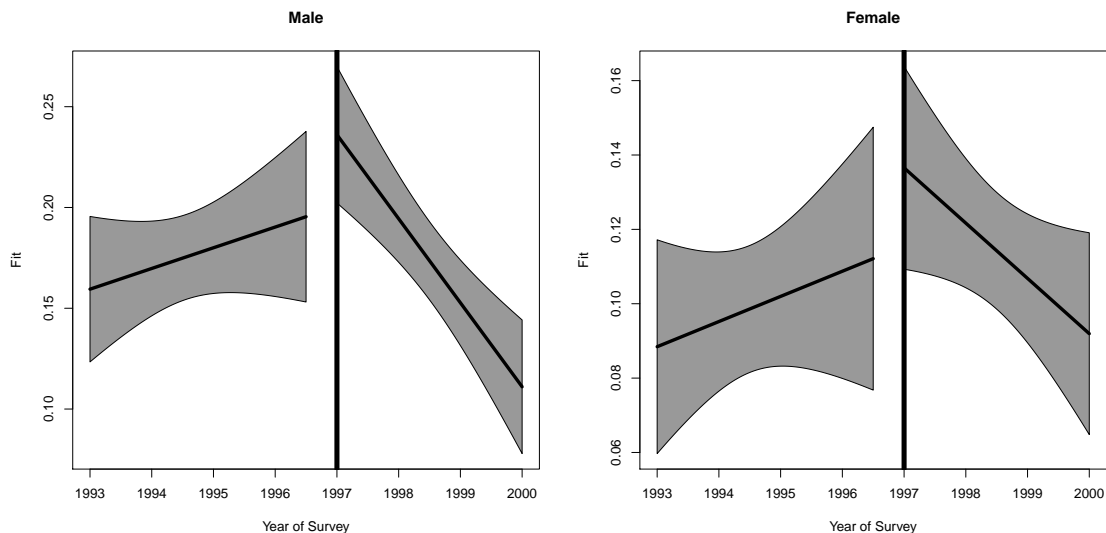
$$D = \alpha_0 + \alpha_1 Z + \alpha_2(X - 1997) + \alpha_3 Z(X - 1997) + \zeta, \quad (2.1)$$

where D is a variable that is equal to one if the person is retired, zero otherwise; X is the year of survey, $Z = \mathbb{1}_{\{X \geq 1997\}}$ and ζ is an error term with zero mean. The form chosen in equation (2.1) is due to the necessity to consider the conditional retirement probability as a function of the cohort membership. Moreover the term that depends from Z is included to capture the change at the introduction of the law. Finally the term related to $Z(X - 1997)$ (the interaction term) is included to have more flexibility in the conditional average and to underline how the eligibility criteria of retirement have been progressively tightened. Section 3 will present the identification strategy just based on this interaction term. Figures 2.2 present the results for individual at age 52.

Results show how, before the introduction of the law, the retirement probability is almost constant across cohorts. It is just possible to see a small increasing trend, probably due to some anticipatory effects of the law (i.e. someone that is eligible to retirement before the introduction of the law that in absence of reforms would decided to continue to work, anticipates his retirement decision. This because he is worried about not being eligible anymore after the introduction of the law). After the introduction of the law, is evident a decreasing trend that underlines how progressively less people are eligible to retirement. However does not appear any relevant reduction in the number of retirees at the first step of

2. Pension Reforms and Retirement Effects

Figure 2.2: Retirement Probability across cohorts – Age 52



the law sign that very few people are affected in that moment by the reform.

Retirement is complex decision in the life cycle of a person. It involved several aspects and it could be only partially explained by observed indicators. In econometric theory retirement is defined endogenous for outcomes such as subjective well-being, consumptions etc., in the sense that these outcomes influence the retirement decision, but the retirement status itself has an effect on these outcomes. More generally retirement is not randomly assigned so the comparisons between retired and workers is deeply affected by the so called *Selection Bias* (see Heckman, 1979) that is the differences between the two groups that there would be even if no one would be retired. Following Rubin (1974) we can define as $Y_i(1)$ the value of referring outcome if the unit i decides to retire and $Y_i(0)$ the value of the same outcome for the same unit if he decides not to retire (usually $Y(1)$ and $Y(0)$ are defined *Potential Outcomes*); and we could just observe the difference:

$$\begin{aligned} \mathbb{E}[Y(1)|D = 1] - \mathbb{E}[Y(0)|D = 0] &= \mathbb{E}[Y(1) - Y(0)|D = 1] \\ &+ \mathbb{E}[Y(0)|D = 1] - \mathbb{E}[Y(0)|D = 0], \end{aligned}$$

where the blue part represents the effect of retirement for the retired group and

the red part is the already defined selection bias.

In addition to that, it is not possible to assume the independence between the cohorts and the potential outcomes, so the independence between the X variable and $Y(1); Y(0)$.

In the next section will be discussed the assumptions and the identification strategy.

Chapter 3

Combining Jump and Kink ratio

As seen in chapter 2 there is not a relevant discontinuity in the retirement probability at the introduction of the law. However the progressive tightening provided by the law creates a significative change of slope at the threshold point. In this section we will motivate the identification strategy based on that change of slope.

3.1 Review of the literature

3.1.1 Regression Discontinuity Design

Regression Discontinuity Design (hereafter RDD), first introduced by Thistlethwaite and Campbell (1960), has found increasing success in the late '90s (see for example Hahn et al., 2001 for the identification assumption or Imbens and Lemieux, 2008; Lee and Lemieux, 2010 for a review and a guide to practice). The key idea is that the treatment is (totally or partially) determined by a threshold point of an observed continuous variable (often defined *assignment*, *forcing* or *running* variable).

Following Hahn et al. (2001) we can distinguish two cases of RDD: the *sharp* and the *fuzzy* RDD. In the sharp RDD the treatment assignment is totally determined by the running variable, so all the people that have a value of the running variable greater (smaller) than the threshold point are included in the treatment group,

3. Combining Jump and Kink ratio

instead all the people that have a value smaller (greater) are in the control group. So defining as D the treatment indicator, that is equal to 1 if an observation is in the treatment group and 0 otherwise, as X the running variable and as x_0 the threshold point, we have that:

$$D = \mathbb{1}_{\{X \geq x_0\}}.$$

RDD situations are commonly known as a problem of causal inference where there is completely absence of common support in the variable that determines the exposition to the treatment. That is the reason why, without additional assumptions or informations (see for example Angrist and Rokkanen, 2015; Dong and Lewbel, 2012) it is just possible to estimate the *Average Treatment Effect* (hereafter ATE) at the threshold point. In formula:

$$ATE = \mathbb{E}[Y(1) - Y(0) | X = x_0],$$

where $Y(1)$ and $Y(0)$ represent the *Potential Outcomes* already defined in the previous section. In other words the average of the difference between the value of the outcome variable that the people at the threshold point would have if they had been treated ($Y(1)$) and the value that they would have in the control group ($Y(0)$). Following that the *observed outcome* Y could be defined as:

$$Y = Y(1)D + Y(0)(1 - D).$$

Assuming that the averages of the potential outcomes are continuous at the threshold point, therefore:

$$\begin{aligned} \lim_{x \rightarrow x_0^+} \mathbb{E}[Y(1)|X = x] &= \lim_{x \rightarrow x_0^-} \mathbb{E}[Y(1)|X = x] \\ \lim_{x \rightarrow x_0^+} \mathbb{E}[Y(0)|X = x] &= \lim_{x \rightarrow x_0^-} \mathbb{E}[Y(0)|X = x], \end{aligned} \tag{3.1}$$

ATE at the threshold point is identified by the quantity $\beta = y^+ - y^-$, where:

$$y^+ = \lim_{x \rightarrow x_0^+} \mathbb{E}[Y | X = x] \quad y^- = \lim_{x \rightarrow x_0^-} \mathbb{E}[Y | X = x]. \quad (3.2)$$

Instead, in the *fuzzy* RDD, the treatment assignment is not a deterministic function of the running variable, but it is influenced by others (observed or unobserved) variables. However the probability to be treated is discontinuous at the threshold point x_0 . In this situation we can assume that the running variable just determines the *assignment* to treatment (Z), as:

$$Z = \mathbb{1}_{\{X \geq x_0\}}.$$

As in Angrist et al. (1996); Imbens and Angrist (1994), for the general non-compliance setting, it could be assumed that there are different, unobserved, groups that have different reactions to the treatment assignment, as:

- **Always-Taker** (AT): $D = 1$, for $Z = 0, 1$
- **Never-Taker** (NT): $D = 0$, for $Z = 0, 1$
- **Compliers** (C): $D = 1$, for $Z = 1$ and $D = 0$, for $Z = 0$
- **Defiers**: $D = 0$, for $Z = 1$ and $D = 1$, for $Z = 0$.

Assuming that the Defiers group does not exist we can decompose the average of the outcome Y above (Y^+) and below (Y^-) the threshold as:

$$\begin{aligned} \mathbb{E}[Y^+] &= \mathbb{E}[Y(1)^+ | AT] \mathbb{Pr}[AT] + \mathbb{E}[Y(0)^+ | NT] \mathbb{Pr}[NT] + \mathbb{E}[Y(1)^+ | C] \mathbb{Pr}[C] \\ \mathbb{E}[Y^-] &= \mathbb{E}[Y(1)^- | AT] \mathbb{Pr}[AT] + \mathbb{E}[Y(0)^- | NT] \mathbb{Pr}[NT] + \mathbb{E}[Y(0)^- | C] \mathbb{Pr}[C], \end{aligned} \quad (3.3)$$

if we assume also the continuity at the threshold point (as in eq. (3.1)) of the potential outcomes in the three groups and the continuity of the probability of

3. Combining Jump and Kink ratio

belonging to a group:

$$\begin{aligned} \lim_{x \rightarrow x_0^+} \mathbb{E}[Y(t)|X = x \& G] &= \lim_{x \rightarrow x_0^-} \mathbb{E}[Y(t)|X = x \& G] && \text{for: } t = 0, 1 \& G = AT, NT, C \\ \lim_{x \rightarrow x_0^+} \mathbb{P}\text{r}[G|X = x_0] &= \lim_{x \rightarrow x_0^-} \mathbb{P}\text{r}[G|X = x_0] && \text{for: } G = AT, NT, C \end{aligned}$$

we have that the difference between $y^+ - y^-$ is equal to $\mathbb{E}[Y(1) - Y(0)|X = x_0 \& C]\mathbb{P}\text{r}[C|X = x_0]$.

Using similar arguments we have that the difference between $d^+ - d^-$, where:

$$d^+ = \lim_{x \rightarrow x_0^+} \mathbb{E}[D | X = x] \quad d^- = \lim_{x \rightarrow x_0^-} \mathbb{E}[D | X = x],$$

is equal to $\mathbb{P}\text{r}[C|X = x_0]$, so the ratio:

$$\beta = \frac{y^+ - y^-}{d^+ - d^-},$$

identifies the so called *Local Average Treatment Effect* (hereafter LATE) at the threshold point x_0 , that is the ATE at the threshold point for the subpopulation of compliers, in formulas:

$$LATE = \mathbb{E}[Y(1) - Y(0)|X = x_0 \& C]$$

3.1.1.1 Estimation

The *ATE* in the sharp RDD could be obtained from the coefficient β_1 of a linear regression in the form:

$$Y = \beta_0 + \beta_1 D + \beta_2 X + \beta_3 X D + \epsilon, \quad (3.4)$$

instead, in order to estimate the LATE in a fuzzy RDD, it's necessary to compute the ratio between γ_1 and α_1 in the equations:

$$Y = \gamma_0 + \gamma_1 Z + \gamma_2 X + \gamma_3 X Z + v \quad (3.5)$$

$$D = \alpha_0 + \alpha_1 Z + \alpha_2 X + \alpha_3 X Z + \zeta, \quad (3.6)$$

where the equation (3.5) is commonly called *Intention To Treat* (hereafter ITT) equation and the equation (3.6) is also known as *First Stage* (hereafter FS) equation.

As showed by Hahn et al. (2001), that ratio is *numerically* equivalent to the coefficient β_1 of the *Instrumental Variable* (hereafter IV) regression of:

$$Y = \beta_0 + \beta_1 D + \beta_2 X + \beta_3 XZ + \epsilon, \quad (3.7)$$

where D is considered *endogenous* and it is instrumented with the equation (3.6). In other words Z is the *additional instrument* for the endogenous variable D .

The main problem concerns the fact that all the equations above hold if the conditional averages of Y and D given X (above and below the threshold) are linear. That condition is too strong to be realistic in the empirical applications. This is the reason why usually the RDD could be viewed as a problem of estimation of 2 (or 4 in the fuzzy RDD) boundary points. Imbens and Kalyanaraman (2012) propose a *data-driven* algorithm to estimate the *optimal bandwidth* in a *Local Linear Regression*. Gelman and Imbens (2014) underline how using high order polynomial estimation could generate estimates with unwanted properties, as observations far from the threshold that are weighted more than observation near the threshold. Instead Card et al. (2014) show that the use of *Local Polynomial Regressions* in place of simple Local Linear Regressions in many case could significantly improve the performance of the estimates.

3.1.2 Regression Kink Design

In the recent literature some papers investigate the treatment effect estimation using the information given by the change of slope (defined as *Kink* in contrast with the *Jump* of the standard RDD) in the treatment probability or in treatment amount, the so called *Regression Kink Design* (from here RKD).

Some papers use the RKD in order to estimate the effects of a *continuous treatment*. These are the situations in which all the sample could be considered as *treated*, but the intensity of the treatment received varies across individuals. The intensity of the treatment varies following a deterministic (sharp RKD) or stochastic (fuzzy

3. Combining Jump and Kink ratio

RKD) rule on the basis of a running variable. The same rule provides that in some point the relation between the amount of the treatment and the running variable change, creating a kink. One example is the paper by Simonsen et al. (2015) studies the price sensitivity of demand for prescription drug using the kinked reimbursement schemes provided by the Danish law. Instead Card et al. (2012) investigate the effect of unemployment insurance benefits on the duration of joblessness in Austria, where the benefit schedule has kinks at the minimum and maximum benefit level. They also provide some conditions under which the RKD identifies the “treatment-on-the-treated” parameter formulated by Florens et al. (2008).

Dong (2014) investigates the information given by the Kink-ratio in a fuzzy RDD, so in a binary treatment setting (under RCM). Their key idea is to compute the derivatives with respect to X of the averages of Y (above and below the threshold) and assuming, joint with the standard RDD assumptions, the continuity in derivatives of the potential outcomes and of the probability of belonging to a specific group, so:

$$\begin{aligned} \lim_{x \rightarrow x_0^+} \frac{\partial \mathbb{E}[Y(t)|X = x]}{\partial X} &= \lim_{x \rightarrow x_0^-} \frac{\partial \mathbb{E}[Y(t)|X = x]}{\partial X} && \text{for: } t = 0, 1 \\ \lim_{x \rightarrow x_0^+} \frac{\partial \mathbb{Pr}[G|X = x]}{\partial X} &= \lim_{x \rightarrow x_0^-} \frac{\partial \mathbb{Pr}[G|X = x_0]}{\partial X} && \text{for: } G = AT, NT, C; \end{aligned} \quad (3.8)$$

we obtain that the following equations:

$$\begin{aligned} \frac{\partial(y^+ - y^-)}{\partial X} &= \mathbb{E}[Y(1) - Y(0)|C \& X = x_0] \frac{\partial \mathbb{Pr}[C|X = x_0]}{\partial X} \\ &+ \frac{\partial \mathbb{E}[Y(1) - Y(0)|C \& X = x_0]}{\partial X} \mathbb{Pr}[C|X = x_0] \\ \frac{\partial(d^+ - d^-)}{\partial X} &= \frac{\partial \mathbb{Pr}[C|X = x_0]}{\partial X}, \end{aligned} \quad (3.9)$$

so the Kink-ratio is equal to:

$$\begin{aligned} \frac{\partial(y^+ - y^-)/\partial X}{\partial(d^+ - d^-)/\partial X} &= \mathbb{E}[Y(1) - Y(0)|C \& X = x_0] \\ &+ \frac{\partial \mathbb{E}[Y(1) - Y(0)|C \& X = x_0]}{\partial X} \frac{\Pr[C|X = x_0]}{\partial \Pr[C|X = x_0]/\partial X}. \end{aligned} \quad (3.10)$$

From equation (3.10) we can obtain the following conclusions:

- In general, the Kink-ratio is a combination of the LATE and another Bias term
- If there is no Jump in the treatment probability or if the first derivative of the LATE is equal to zero or both, the Kink-ratio is an unbiased estimator for the LATE.

She also shows that the Kink-ratio, that is equal to the ratio of the coefficients γ_3 and α_3 in the equations (3.5) and (3.6), could be obtained from the coefficient β_1 of an IV regression in the form:

$$Y = \beta_0 + \beta_1 D + \beta_2 X + \beta_3 Z + \epsilon,$$

where D is considered endogenous and it is instrumented by the equation (3.6). In other words, XZ is the *additional instrument* for the endogenous variable D . Then Dong analyzes in detail the case in which the Kink-ratio is not biased because the first derivatives of LATE is equal to zero. In that case it is possible to combine the Jump and the Kink-ratio, in order to minimize the variance of the combination. She suggests to estimate an IV regression in the form:

$$Y = \beta_0 + \beta_1 D + \beta_2 X + \epsilon,$$

where the endogenous variable D is instrumented by the equation (3.6). In that case, the model is *over-identified*, because there are more additional instruments (Z and XZ) with respect to endogenous variables (D). It is easy to show that

3. Combining Jump and Kink ratio

the coefficient β_1 is a combination of the Jump and the Kink-ratio, in the form:

$$\beta_1 = \frac{w_A^J \gamma_1 + w_A^K \gamma_3}{w_A^J \alpha_1 + w_A^K \alpha_3}, \quad (3.11)$$

where the coefficients $\gamma_1, \gamma_3, \alpha_1, \alpha_3$ are the ones of the equations (3.5) and (3.6), and the weights are equal to:

$$\begin{aligned} w_A^J &= \alpha_1 \text{Var}[z] + \alpha_3 \text{Cov}[z, xz] \\ w_A^K &= \alpha_1 \text{Cov}[z, xz] + \alpha_3 \text{Var}[xz], \end{aligned} \quad (3.12)$$

where z and xz are respectively the residuals of the equations:

$$\begin{aligned} Z &= \lambda_0 + \lambda_1 X + z \\ XZ &= \kappa_0 + \kappa_1 X + xz \end{aligned} \quad (3.13)$$

3.2 A combined version of the RDD & RKD estimators

Here we propose an alternative model that combines the Jump and the Kink ratio. The aim is to find a combination that improves the estimate of the LATE (in comparison to the standard Jump ratio). The idea is to find a proper combination of the ratios when the Kink ratio is potentially biased and the Jump ratio is too inaccurate to be used.

In our model, we explicitly introduce the first derivative of the LATE (XD) in the structural model and we consider it as an endogenous variable. We use as instruments the variable Z and XZ , in this way the model is *exactly identified*. So the structural equation becomes:

$$Y = \beta_0 + \beta_1 D + \beta_2 X + \beta_3 XD + \epsilon, \quad (3.14)$$

and the first stage equations are the one in (3.6) and:

$$XD = \delta_0 + \delta_1 Z + \delta_2 X + \delta_3 XZ + \eta. \quad (3.15)$$

As usual the parameter of interest is represented by the coefficient β_1 related to D in the structural equation.

It is easy to show that the obtained coefficient (as the model proposed by Dong in equation (3.11)) is a combination of the Jump and the Kink ratio in the form:

$$\beta_1 = \frac{w_B^J \gamma_1 + w_B^K \gamma_3}{w_B^J \alpha_1 + w_B^K \alpha_3}, \quad (3.16)$$

where the weights are equal to:

$$\begin{aligned} w_B^J &= \{\delta_1 \text{Cov}[z, xz] + \delta_3 \text{Var}[xz]\} \text{Var}[z] \\ &\quad - \{\delta_1 \text{Var}[z] + \delta_3 \text{Cov}[z, xz]\} \text{Cov}[x, xz] \\ w_B^K &= \{\delta_1 \text{Cov}[z, xz] + \delta_3 \text{Var}[xz]\} \text{Cov}[z, xz] \\ &\quad - \{\delta_1 \text{Var}[z] + \delta_3 \text{Cov}[z, xz]\} \text{Var}[xz]. \end{aligned} \quad (3.17)$$

From equations (3.17) it is clear that the weights given to the two ratios depend from the variance and covariance matrix of z and xz , that depends only on some conditional and unconditional moments of the running variable (see appendix B), and the coefficients of the equation (3.15). So the value of the coefficients of the first stage of XD is crucial to determine the proportion between the Jump and the Kink ratio in the combined estimator. Assuming that the true data generating process (hereafter DGP) of the variable D is the one expressed in equation (3.6), and ζ has zero mean, we have that:

$$\mathbb{E}[D|X] = \alpha_0 + \alpha_1 Z + \alpha_2 X + \alpha_3 XZ,$$

consequently the conditional mean of XD is equal to:

$$\mathbb{E}[XD|X] = X\mathbb{E}[D|X] = \alpha_0 X + \alpha_1 XZ + \alpha_2 X^2 + \alpha_3 X^2 Z. \quad (3.18)$$

It is straightforward from equation (3.18) to see that in the conditional mean of XD there is not a term that depends just from Z , so, if the model is correctly specified (in that case it is estimated with a second order polynomial), $\delta_1 = 0$ and $\delta_3 = \alpha_1$. If we substitute zero to δ_1 in the equation (3.17) we see that w_B^K

is equal to zero and the model reduces to the standard Jump ratio. Therefore in order to obtain a combination that includes both the Jump and the Kink ratio, it is necessary that the FS for XD is “miss-specified”. It is important to underline that in the proposed model, the ITT and the two FS equations are estimated with the same polynomial order, so if the FS equation for XD is miss-specified it does not mean that the FS equation for D is miss-specified too.

3.3 Simulation Study

Here we propose a simulation study in which we can compare the performance (expressed in terms of *Root Mean Square Error*, hereafter RMSE) of the different estimators.

The data are generated from a structural equation in the form:

$$Y = \beta_0 + \beta_1 D + \beta_2 X + \beta_3 (X \cdot D) + \epsilon, \quad (3.19)$$

where D is generated from the unobserved variable D^* , that has the following DGP:

$$D^* = \alpha_0 + \alpha_1 Z + \alpha_2 X + \alpha_3 XZ + v, \quad (3.20)$$

and $D = \mathbb{1}(D^* \geq 0)$. As showed in the previous section it is important to control the polynomial order of the FS equation, so we generate v from an uniform distribution between -1 and 1 .

In order to create endogeneity in the equation (3.19), we have to impose that $\text{Cov}[\epsilon, v] \neq 0$, consequently we generate ϵ as:

$$\epsilon = F^{-1}\left(\frac{v+1}{2}\right) + N,$$

where $F^{-1}(\cdot)$ is the inverse of the distribution function of a uniform between 0 and 1 and $N \sim \mathcal{N}(0, \sigma^2)$. In that model the level of endogeneity could be controlled with the parameter σ^2 .

The choice of the uniform distribution in the error of the equation (3.20) is because

the conditional expectation of D is in the form:

$$\begin{aligned}
 \mathbb{E}[D|X] &= \mathbb{Pr}[D^* \geq 0|X =] = \mathbb{Pr}[\alpha_0 + \alpha_1 Z + \alpha_2 X + \alpha_3 XZ + v \geq 0|X] \\
 &= \mathbb{Pr}[-v \leq \alpha_0 + \alpha_1 Z + \alpha_2 X + \alpha_3 XZ|X] \\
 &= \mathbb{Pr}[v \leq \alpha_0 + \alpha_1 Z + \alpha_2 X + \alpha_3 XZ|X] \\
 &= F_{v|X}(\alpha_0 + \alpha_1 Z + \alpha_2 X + \alpha_3 XZ) \\
 &= \frac{\alpha_0 + \alpha_1 Z + \alpha_2 X + \alpha_3 XZ - 1}{2},
 \end{aligned} \tag{3.21}$$

where $F_{v|X}(\cdot)$ is the distribution function of a uniform between -1 and 1 , and in the second row is used the fact that $-v$ is also distributed as a uniform between -1 and 1 .

Substituting the equation (3.21) in the conditional average of Y , we obtain:

$$\begin{aligned}
 \mathbb{E}[Y|X] &= \beta_0 + \beta_1 \mathbb{E}[D|X] + \beta_2 X + \beta_3 X \mathbb{E}[D|X] \\
 &= \beta_0 + \frac{\alpha_0 \beta_1}{2} + \frac{\alpha_1 \beta_1}{2} Z + \left(\frac{\alpha_2 \beta_1 + \alpha_1 \beta_3}{2} + \beta_1 \right) X \\
 &\quad + \frac{\alpha_3 \beta_1 + \alpha_1 \beta_3}{2} XZ + \frac{\alpha_2 \beta_3}{2} X^2 + \frac{\alpha_3 \beta_3}{2} X^2 Z \\
 &= \gamma_0 + \gamma_1 Z + \gamma_2 X + \gamma_3 XZ + \gamma_4 X^2 + \gamma_5 X^2 Z,
 \end{aligned} \tag{3.22}$$

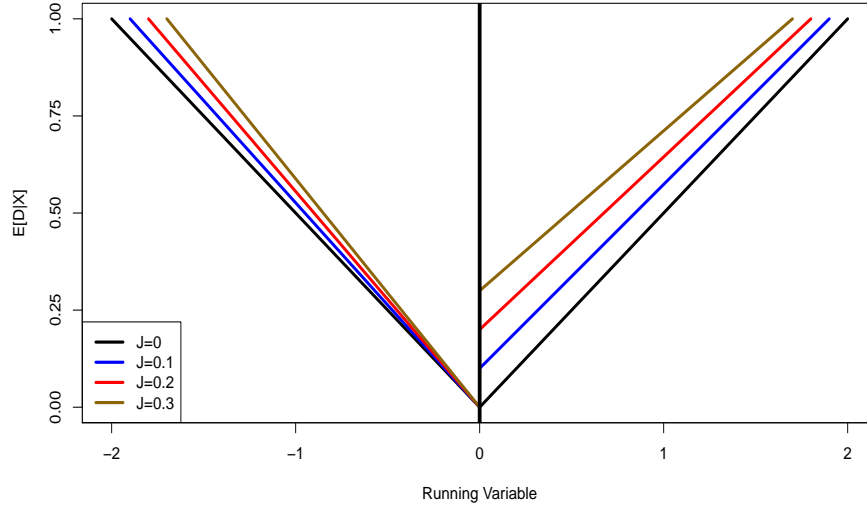
so the conditional average of Y , in order to obtain unbiased coefficients of the ITT equation have to be estimate with a second order polynomial.

We simulate the value of X as an uniform between $(-h, h)$, in order to be symmetric around the threshold point (zero). First of all we choose the value of the Jump (J) and the Kink (K) in the first stage, after that we fix four points in the conditional average of D . We want that in the point $-h$ the mean of is equal to one, in the point 0^- to zero, in the point 0^+ to J and in the point h is still equal to one. Finally we choose the value of h to fix the kink when the value of the jump is increased. It is easy to show that the corresponding values of the parameters α are:

$$\begin{aligned}
 h &= \frac{2-J}{K} & \alpha_0 &= -1 & \alpha_1 &= -\frac{2}{h} \\
 \alpha_2 &= 2J & \alpha_3 &= \frac{2}{h} (2 - J).
 \end{aligned} \tag{3.23}$$

3. Combining Jump and Kink ratio

Figure 3.1: Conditional average of the first stage equation for $K = 1$



In Figure 3.1 is showed the trend of $\mathbb{E}[D|X]$ fixing the value of K and varying the value of J . It is evident that increasing the value of J the range of the running variable X decrease in order to keep constant the change of slope at the threshold point.

In the equation of the outcome Y , described in (3.19), the bias term of the Kink ratio is controlled by the coefficient β_3 . It is easy to show that $L' = (\beta_3 \cdot J)/K$, where L' is the first derivative of the LATE at the threshold point. So, we fix the value of L' and, given the value of K and J , we choose the value of β_3 . Summarizing the parameter of the equation (3.19) are:

$$\begin{aligned} \beta_0 &= 1.5 & \beta_1 &= 2 \\ \beta_2 &= -0.4 & \beta_3 &= \frac{K \cdot L'}{J}. \end{aligned} \quad (3.24)$$

We want to understand both the finite sample properties and the asymptotic properties of the different estimators, for this reason we fix four dimension of the sample sizes (500, 1000, 3000, 5000 observations). In addition to that we want to

compare the different estimators both when the Kink ratio is or is not biased, so the simulation results are presented for $L' = 0$ (table 3.1) and for $L' = 0.5$ (table 3.2). Then for each sample size we replicate the estimates 2000 times and finally we compute the *Root Mean Square Error* of the estimators.

In Tables 3.1 and 3.2 are represented the performance of the following estimators:

- **Jump ratio:** It is the standard RDD estimator. This is the ratio between the coefficients γ_1 and α_1 respectively in the equations (3.5) and (3.6).
- **Kink ratio:** It is the Kink ratio expressed by the ratio of the coefficients γ_3 and α_3 .
- **Dong's Model:** It is the combination of the Jump and the Kink ratio obtained by the IV regression proposed by Dong (2014).
- **Two Endogenous Model:** It is the combination of the Jump and the Kink ratio described in Section 3.2. For the same estimator is presented also the percentage of rejection of the null hypothesis that the coefficient related to XD is equal to zero.
- **IK Opt. Band.:** In these cells are represented the RMSE of the estimator obtained using the algorithm proposed by Imbens and Kalyanaraman (2012) to estimate the Optimal Bandwidth in a Local Linear Regression.

As seen before in the simulation the ITT is a second order polynomial and the FS equation is linear, consequently the XD is quadratic too. For these reason if the model is estimated with a linear IV regression, the ITT equation is miss-specified, instead if we estimate a quadratic regression (controlling for X^2 and X^2Z too) the FS equation is over-specified. Moreover in the quadratic regression the coefficients δ_1 of the equation (3.15) is equal to zero and the combined estimator described in (3.16) and (3.17) reduces to the standard Jump ratio. In order to study the performance of our model (in comparison to the other model described before) have been reported the estimates with the first and the second polynomial order.

Table 3.1.: Root Mean Square Error of the different estimators with $K = 1$ and $L' = 0$.

	n=500	n=1000	n=3000	n=5000					
J=0	Jump ratio	4520.449	343.683	989.688	689.863	654.37	2570.024	1155.903	157.523
	Kink ratio	.268	1.154	.184	.768	.109	.417	.082	.329
	Dong's model	.272	1.098	.188	.757	.112	.415	.084	.329
	Two Endogenous Model	.262	64.347	.177	263.343	.106	14.679	.079	262.951
	(% reject.XD coef.)	.641	0	.9	0	1	0	1	.001
IK Opt. Band.	1011.887		212.629		674.637		579.91		
J=0.1	Jump ratio	213.824	53.592	22.999	873.755	1.386	45.267	1.005	1.698
	Kink ratio	.264	1.29	.188	.795	.107	.439	.082	.332
	Dong's model	.26	1.125	.187	.732	.106	.42	.081	.318
	Two Endogenous Model	.267	1860.874	.189	76.685	.107	15.949	.082	619.944
	(% reject.XD coef.)	.045	0	.043	0	.045	0	.052	0
IK Opt. Band.	83.915		123.578		6.008		1.825		
J=0.2	Jump ratio	74.454	59.948	1.214	2.715	.606	1.011	.479	.717
	Kink ratio	.281	1.553	.201	.89	.115	.474	.088	.358
	Dong's model	.276	1.102	.196	.734	.112	.412	.087	.318
	Two Endogenous Model	.284	1805.872	.203	31.478	.115	21.596	.088	12.699
	(% reject.XD coef.)	.041	0	.041	0	.044	0	.052	.001
IK Opt. Band.	15.681		6.737		1.129		.764		
J=0.3	Jump ratio	1.155	7.275	.724	1.181	.4	.615	.308	.471
	Kink ratio	.297	1.516	.207	.93	.121	.505	.092	.379
	Dong's model	.281	.938	.197	.662	.116	.38	.088	.297
	Two Endogenous Model	.301	79.746	.208	14.087	.121	.684	.093	.5
	(% reject.XD coef.)	.047	0	.049	0	.048	.001	.054	.002
IK Opt. Band.	7.02		1.381		.664		.515		

3. Combining Jump and Kink ratio

Table 3.2: Root Mean Square Error of the different estimators with $K = 1$ and $L' = 0.5$.

	n=500		n=1000		n=3000		n=5000	
	p=1	p=2	p=1	p=2	p=1	p=2	p=1	p=2
J=0								
Jump ratio	828.369	487.574	1850.411	1248.89	2915.816	864.14	2837.429	336.654
Kink ratio	.257	1.206	.187	.769	.105	.425	.082	.331
Dong's model	.261	1.087	.19	.753	.108	.423	.083	.33
Two Endogenous Model	.25	43.491	.181	72.022	.102	25.395	.08	21.475
(% reject.XD coef.)	.635	0	.888	0	1	0	1	.001
IK Opt. Band.	367.981		360.061		635.389		10058.88	
J=0.1								
Jump ratio	3506.039	571.759	1057.097	36.434	36.457	18.077	32.727	1.946
Kink ratio	.422	1.766	.288	1.208	.171	.794	.131	.666
Dong's model	.557	1.584	.422	1.133	.319	.765	.294	.636
Two Endogenous Model	.268	87.175	.188	67.21	.107	55.655	.085	55.187
(% reject.XD coef.)	1	.007	1	.006	1	.006	1	.007
IK Opt. Band.	265.122		74.692		83.87		3.926	
J=0.2								
Jump ratio	30.476	20.902	9.813	52.47	7.089	1.032	6.957	.748
Kink ratio	.329	1.478	.226	1.015	.128	.694	.101	.617
Dong's model	.424	1.157	.336	.867	.281	.589	.266	.511
Two Endogenous Model	.288	66.377	.202	96.775	.113	8.351	.087	4.408
(% reject.XD coef.)	.981	0	1	.001	1	.004	1	.012
IK Opt. Band.	21.594		8.306		1.501		1.162	
J=0.3								
Jump ratio	3.535	10.725	3.028	1.232	2.777	.618	2.748	.492
Kink ratio	.314	1.736	.221	1.014	.125	.692	.096	.607
Dong's model	.378	1.004	.311	.729	.259	.489	.244	.41
Two Endogenous Model	.301	12.6	.21	10.824	.119	.652	.092	.492
(% reject.XD coef.)	.738	0	.935	.001	1	.015	1	.056
IK Opt. Band.	9.318		1.566		.787		.664	

3. Combining Jump and Kink ratio

The results show as, in that choice of the parameters (very small Jump and large Kink), the standard Jump ratio has very poor performance. Instead the Kink ratio, for every sample size, value of J and polynomial order, has better performance, even if, for $J > 0$ and $L' \neq 0$, it is a biased estimator for the LATE. Moreover the estimator proposed by Dong, when the first derivative of the LATE is different from zero, is not able to improve the performance of the simple Kink ratio. Instead, when $L' = 0$ and $J > 0$ it is the estimator with the best performance.

Additional consideration have to be done for the model that include XD as an additional endogenous variable:

1. The best performances of the estimator are for $p = 1$, in other words when the XD equation is *miss-specified*, for $p = 2$ the performance of the model has very poor performance.
2. When the Kink ratio is biased, it is the best estimator, among the considered, for every value of J and n .
3. When $J > 0$ the t-test related to the XD coefficient is able to discriminate when to use the Dong's model of the model with two endogenous variables.
4. The better improvement are for small (but not zero) value of J e.g. $J = 0.1$.

So, in general, the simulation results show an improvement of the performance of the two endogenous estimator in comparison to the others estimator, when the XD equation is miss-specified and the combined estimator is different from the standard Jump ratio.

Chapter 4

Results with *Aspetti della Vita Quotidiana* Survey

In this section results will be presented and discussed on well-being using the estimators proposed in section 3. We started the first row of the tables in 2.3, extracting all the individuals with 52 y.o. in the different cohorts. Moreover we have extracted the five questions regarding their subjective personal satisfaction. These questions involve their satisfaction about:

1. economic situation,
2. health status,
3. family relation,
4. friend relation,
5. leisure time.

Each respondent has to reply according to a likert scale from 1 to 4, where 1 is *extremely* and 4 is *not at all*. We want to understand how the retirement could increase the probability to be dissatisfied in these five topics, for this reason we have created a set of dummy variables (one for each question) that are equal to 1 if the respondent answers 4 (not at all), zero otherwise. Table 4.1 summarizes the distribution of the answers of the respondents.

4. Results with *Aspetti della Vita Quotidiana* Survey

We have estimated the effects using both the model with two endogenous and

		Answer			
		Extremely	Very	Slightly	Not at all
Male	Economic Satisf.	75	1406	888	274
	Health Satisf.	443	1768	348	85
	Fam. Rel. Satisf.	983	1496	123	31
	Friend Rel. Satisf.	642	1624	313	58
	Leisure Time Satisf.	318	1223	894	201
Female	Economic Satisfaction	71	1327	983	297
	Health Satisf.	326	1724	515	119
	Fam. Rel. Satisf.	1027	1508	115	28
	Friend Rel. Satisf.	617	1629	341	86
	Leisure Time Satisf.	293	1238	917	225

Table 4.1: Frequency distribution of the outcome variables about personal satisfaction

the model proposed by Dong. Moreover we have estimated the IV-regressions using the simple *linear probability model* based on *two stage least squares* and, in order to take into account the binary nature of the dependent variable, a probit-IV. This model is also based on two stage least squares algorithm and consists of the following steps:

1. to estimate the first stage regressions of the endogenous variable on the exogenous variables and the additional instruments using ordinary least squares
 2. to compute the residuals of the FS regressions
 3. to use that in the structural equation, estimated using a standard probit regression, in addition to the exogenous and the endogenous variables,
- (see Rivers and Vuong, 1988, for details and for standard error calculation).

Table 4.2 summarizes the results for the different outcomes and estimators. For each estimator the coefficients is shown of the variable related to the retirement status, and in parenthesis its standard error. In the probit-IV regression,

	Dong				Two endogenous				
	IV		Probit-IV		IV		Probit-IV		
	Coef.(S.E.)	Coef.(S.E.)	Marg. Eff.	Coef.(S.E.)	p-value	XD	Coef.(S.E.)	Marg. Eff.	p-value
Economic Dissatisf.	.406* (.209)	2.61** (1.248)	.763 (.722)	.053 (.722)	.601	.844 (3.876)	.207	.624	
Health Dissatisf.	.105 (.107)	1.745 (1.559)	.315 (.302)	.097 (.302)	.977	1.75 (4.207)	.317	.999	
Fam. Rel. Dissatisf.	.044 (.065)	1.654 (2.251)	.173 (.306)	.178 (.306)	.638	5.741 (9.525)	.99	.65	
Friend Rel. Dissatisf.	-.037 (.087)	-.849 (1.798)	-.024 (.28)	-.114 (.28)	.771	-1.684 (5.045)	-.039	.844	
Leisure Time Dissatisf.	.239 (.172)	1.501 (1.234)	.365 (.78)	.54 (.78)	.68	4.136 (6.272)	.957	.661	

(a) Male

	Dong				Two endogenous				
	IV		Probit-IV		IV		Probit-IV		
	Coef.(S.E.)	Coef.(S.E.)	Marg. Eff.	Coef.(S.E.)	p-value	XD	Coef.(S.E.)	Marg. Eff.	p-value
Economic Dissatisf.	.867 (.559)	4.708 (3.024)	.958 (.693)	.858 (.693)	.981	4.723 (3.794)	.959	.992	
Health Dissatisf.	.21 (.286)	2.597 (3.242)	.699 (.651)	.417 (.651)	.597	5.096 (7.124)	.986	.603	
Fam. Rel. Dissatisf.	.276 (.179)	10.097 (6.561)	1 (.531)	.455 (.531)	.574	18.859 (21.966)	1	.572	
Friend Rel. Dissatisf.	.583* (.338)	8.618* (4.92)	.998 (.463)	.631 (.463)	.863	9.527 (7.002)	.998	.836	
Leisure Time Dissatisf.	.166 (.378)	1.19 (2.572)	.303 (1.21)	.601 (1.21)	.547	4.225 (7.864)	.958	.555	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

(b) Female

Table 4.2: Results Age 52

marginal effects are also reported at the threshold point, computed as:

$$\Phi(\hat{\beta}_0 + \hat{\beta}_1) - \Phi(\hat{\beta}_0),$$

where β_0 and β_1 are respectively the constant term and the coefficient related to D of the structural equation. The coefficients related to the running variable are omitted because we are interested in the marginal effect in the threshold point, so $(X - 1997) = 0$. Finally, for the model with two endogenous variables the p-values related to the significance test of the coefficients related to XD are reported, in order to understand if the model with two endogenous variables or the Dong's model have to be preferred.

In any case, the coefficient related to XD is not significant, so the second endogenous variable could be omitted. Results, obtained with the Dong's model, show a negative effect on economic satisfaction for males and on friend relations for females, both significant at the 10% level. However no difference appears in the others outcomes.

4.1 Subsample selection

In this section we show some results obtained selecting a subsample of the data. The aim is to provide some robustness check of the results or some stratification that want to try to understand how the results are heterogeneous in the population.

4.1.1 Robustness

In RDD applications is crucial to determine what is the maximum distance from the threshold that have to be included in the sample. The trade off, as usual, is between bias and variance. The inclusion of observation too far from the threshold could cause an increase of the bias of the estimates, because it is imposed a linear (or polynomial) trend for a large range of the running variable. However if the bandwidth is reduced, observations are dropped from the sample and the variance of the estimates increases.

For this reason we have decided to drop the two extreme (with respect to the threshold point) cohorts of the sample (so all the people with 52 y.o. in 1993 and 2000) in order to provide a robustness check.

Table 4.3 summarizes the results. in the first two rows are reported the coeffi-

Table 4.3: Results with bandwidth reduction

	Male		Female	
	Coeff.	S.Err.	Coeff.	S.Err.
FS: Z	.0253	(.0384)	.023	(.0317)
FS: XZ	-.0608***	(.0218)	-.0236	(.0177)
Economic disatisf.	.4966	(.3063)	.8742	(.7612)
Health disatisf.	.1102	(.1505)	.2959	(.4033)
Family relation disatisf.	-.0148	(.0954)	.353	(.278)
Friend relation disatisf.	.0058	(.1178)	.7044	(.5192)
Leisure time disatisf.	.0718	(.2271)	.1094	(.5045)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

icients and the related standard error of the Jump and the Kink in the first stage regressions. The last five rows refer to the coefficient related to D in the Dong's model.

Results are closed to the ones showed in table 4.2. They have all the same sign (excepted for the one related to the Family relation satisfaction, that is really closed to zero) and the size is comparable. However for female both the instruments lose their significance, so also the IV estimates become unreliable.

4.1.2 Stratification

As seen before the structure of the reforms involves a very specific part of the population: the workers that drop out very early from school and that have started to work really young. For this reason we have decided to exclude from the sample all the individuals with high educational attainment (high school or university graduates). These individuals are plausibly unaffected by the reform, because they do not have reached the 35 years of contribution.

4. Results with *Aspetti della Vita Quotidiana* Survey

Table 4.4: Results excluding individuals with high educational attainment

	Male		Female	
	Coeff.	S.Err.	Coeff.	S.Err.
FS: Z	.0542	(.0345)	-.0124	(.0245)
FS: XZ	-.051***	(.0152)	-.0217**	(.0106)
Economic disatisf.	.4622*	(.2425)	.8743	(.6952)
Health disatisf.	.089	(.117)	.478	(.4198)
Family relation disatisf.	.0661	(.0724)	.3641	(.2466)
Friend relation disatisf.	-.0202	(.0924)	.5964	(.4237)
Leisure time disatisf.	.1459	(.1799)	.2529	(.5072)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.4 summarizes the results. The results confirm the ones proposed in table 4.2, with a negative effect for males in the economic satisfaction (statistically significant at the 0.1 level) and no significant effects in the other outcomes.

In addition to that, another relevant stratification is between Southern and Northern Italy. As known the Northern Italy is characterized by an higher fraction of workers that are employed in the private sector. For this reason it is reasonable to suppose that this is the area more affected by the reforms.

Results are in tables 4.5. First stage regressions show how the reform has an effect in the retirement probability only in the Northern Italy, in particular for males that have both a significant Jump and Kink. The coefficients of the first stage regression about Southern Italy do not provide enough information to identify the retirement effect and this is particularly evident looking at the coefficients of the IV regressions for females. Results confirm the negative effect on the economic satisfaction for males (just in the Northern Italy).

4.1. Subsample selection

	North		South	
	Coeff.	S.Err.	Coeff.	S.Err.
FS: Z coef.	.1369**	(.0536)	-.0035	(.044)
FS: XZ coef.	-.0694***	(.0238)	-.0364*	(.0193)
Economic disatisf.	.3584**	(.1814)	.4799	(.5898)
Health disatisf.	.1055	(.0951)	-.086	(.2865)
Family relation disatisf.	.043	(.0547)	.0944	(.1854)
Friend relation disatisf.	-.0289	(.086)	.0409	(.2044)
Leisure time disatisf.	.1695	(.136)	.0442	(.445)

(a) Male

	North		South	
	Coeff.	S.Err.	Coeff.	S.Err.
FS: Z coef.	.009	(.0423)	-.0104	(.0261)
FS: XZ coef.	-.0373**	(.0186)	-.0081	(.0111)
Economic disatisf.	.1469	(.3933)	2.9837	(4.4545)
Health disatisf.	.2999	(.28)	1.9137	(2.7996)
Family relation disatisf.	.219	(.1781)	.7182	(1.1269)
Friend relation disatisf.	.2813	(.2373)	.3207	(1.33)
Leisure time disatisf.	-.0092	(.3636)	1.83	(3.0485)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

(b) Female

Table 4.5: North-South Stratification

Chapter 5

Two Sample Instrumental Variable

In this section it will be explained how to use two sample instrumental variable in a RDD setting in order to:

- a. Improve the accuracy of the estimates with administrative data
- b. Obtain delayed effects of retirement.

The *Two Sample Instrumental Variable* (hereafter TSIV) estimator was first proposed by Angrist and Krueger (1992), and more recently improved by Inoue and Solon (2010). The key idea is very simple and it consists on estimating moments on two complementary data source. In steps:

1. Estimating the FS regression using a data source
2. Using the coefficients estimated in the previous step in the other dataset to obtain the fitted values of the endogenous variables
3. Finally it possible to estimate the structural equation using the fitted value.

So defining as $Y^{(d)}$ the $(n^{(d)} \times 1)$ outcome, as $\mathcal{X}^{(d)}$ the $(n^{(d)} \times p + k)$ matrix that includes the set of the endogenous (p) and exogenous variables (k) and as $\mathcal{Z}^{(l)}$ the $(n^{(d)} \times q + k)$ (with $q \geq p$) matrix that includes the set of the additional

5. Two Sample Instrumental Variable

instruments (q) and the exogenous variables, where $d = 1, 2$ denotes if they belong to the first or to the second sample, we have the FS equations estimated with the second sample are, in matrix form:

$$\mathcal{X}^{(2)} = \alpha \mathcal{Z}^{(2)} + \mathbf{v},$$

where \mathbf{v} is a $(n^{(2)} \times p + k)$ matrix with the last k columns identically equal to zero. The previous equations could be estimated using standard OLS in order to obtain the values $\hat{\alpha}$ that have to be used to obtain the fitted values of the endogenous variables in the first sample as:

$$\hat{\mathcal{X}}^{(1)} = \hat{\alpha} \mathcal{Z}^{(1)},$$

finally the structural equation could be estimated with the regression:

$$Y^{(1)} = \beta \hat{\mathcal{X}}^{(1)} + \epsilon.$$

The previous equations show how it is necessary to observe $(Y; \mathcal{Z})$ in the first sample and $(\mathcal{X}; \mathcal{Z})$ in the second sample, so \mathcal{Z} has to be observed in both samples. In a RDD context we just need three variables: the outcome, the treatment indicator and the forcing variable, all the other variables (Z, XZ, XD) are just functions of the forcing variable and of the treatment indicator. Consequently whatever is the model, among the proposed in section 3, that we want to estimate we need to observe the outcome and the forcing variable in the first sample and the treatment indicator and the forcing variable in the second sample.

5.1 The Work Histories Italian Panel database

The *Work Histories Italian Panel* (hereafter Whip) is a statistical database constructed from administrative data of the Italian National Institute of Social Security (hereafter Inps). It includes the work histories of the private (non agricultural) sector workers. Inps has extracted all the records contained in their administrative archives related to the individuals born in 24 different days (re-

ardless of the year of birth) for a sample size of about 1:15 of the entire Inps population. For each worker we are able to observe:

1. the working career, including yearly wages, type of contract etc.
2. the unemployment periods
3. the retirement.

It allows to estimate the first stage regression with a larger sample size (table 5.1 summarizes the sample size of Whip in comparison to the AVQ survey) and to know exactly when a worker is retired.

Table 5.1: Sample size of Whip in comparison to AVQ survey

	Survey	
	AVQ	Whip
Male	2,697	117,908
Female	2,747	44,301

In order to make comparable the FS regressions with the ones obtained using survey data, we choose four cohort before and four cohorts after the introduction of the new eligibility criterion for retirement. Moreover we have used the retirement archive to understand when the individuals have retired. In this way we was able to understand how many people was retired at age 52, and we have reconstructed the treatment dummy variable.

Table 5.2 and figure 5.1 summarize the results of the FS regressions. They show a relevant reduction in the standard errors of the coefficients in comparison to the ones obtained with the AVQ survey.

Table 5.3 shows the results obtained with the TSIV estimator, in which, FS regression are estimated with the Whip data source and the structural equation is estimated with AVQ survey. Standard Errors are computed using *bootstrap* algorithms. In comparison to the results proposed in table 4.2 for females, we can see a significative negative effect (an increase in the probability to be dissatisfied) in family and in friend relations and a negative effect, significant at 10 percent

5. Two Sample Instrumental Variable

Figure 5.1: First Stage Estimation using Whip for 52

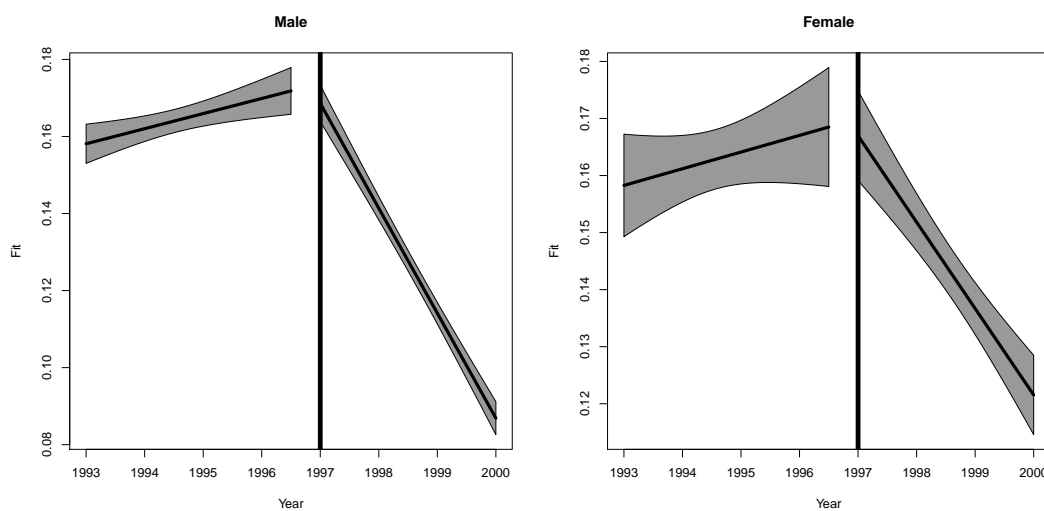


Table 5.2: First Stage Comparison

	Male		Female	
	AVQ	Whip	AVQ	Whip
X	0.0103 (0.0096)	0.0039*** (0.0014)	0.0068 (0.0079)	0.0029 (0.0024)
Z	0.0355 (0.0311)	-0.0052 (0.0044)	0.0210 (0.0257)	-0.0030 (0.0075)
XZ	-0.0520*** (0.0134)	-0.0312*** (0.0018)	-0.0216** (0.0109)	-0.0181*** (0.0031)
Constant	0.2005*** (0.0259)	0.1738*** (0.0037)	0.1155*** (0.0216)	0.1700*** (0.0064)
Observations	2697	117908	2747	44301

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5.3: Estimates for 52 y.o. population obtained combining AVQ and Whip data source

	Male		Female	
	Coef.	Stand. Err.	Coef.	Stand. Err.
Economic Dissatisfaction	.638*	(.345)	1.117*	(.652)
Health Dissatisfaction	.186	(.199)	.429	(.413)
Fam. Rel. Dissatisfaction	.109	(.107)	.499**	(.226)
Friend Rel. Dissatisfaction	-.082	(.155)	.793**	(.386)
Leisure Time Dissatisfaction	.489*	(.296)	.542	(.552)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

level, for the economic satisfaction. Moreover we can see, for males, a negative effect, significant at 10% level, for economic and leisure time satisfaction. All the coefficients confirm the direction of the estimates obtained with just the AVQ survey.

5.2 Delayed effects

Indeed in this section it will be explained how to use the TSIV estimator to identify what are the effects of retirement some years after.

Atchley (1976) first has proposed the theory that after retirement there is a positive short-term effect, often called an *honeymoon* period, on the subjective satisfaction. After that period, it starts a decline in their subjective well-being with a consequently negative effect.

In our idea, we want to observe people some years after the comparison proposed in the previous section, but, unfortunately, AVQ survey does not provide any repeated measure of the same individuals. However we are able to observe in the next waves, some individuals that are representative of the same cohort, in the following years.

For this reason we have decided to construct the dataset as in table 2.3, and to use the first row of the table (so all the people aged 52 y.o. in the different waves) to estimate the first stage regression and the other rows to estimate the structural

5. Two Sample Instrumental Variable

equations. In this way we are able to identify the effects of the retirement at 52 y.o. in their well-being from 1 to 6 year after.

It have to be underlined that the cohorts involved by the reform have just postponed their retirement, so the comparison between the two groups some years later, is a combination of two main effects:

1. The evolution of the subjective well-being for retired people. So how their status changes after some years of retirement.
2. The change of retirement status for the control group. So after some years the cohorts involved by the pension reform become eligible to retirement and change their status.

Figure 5.2 shows the results for 52 y.o. Male. On the horizontal axis is represented the year after the 52 y.o., instead the vertical axis represents the estimate of the retirement effect. Each point of the graph represent a pointwise estimate and the shaded area is the 95% bootstrap confidence interval of the estimate.

Results show a negative effect on the economic satisfaction that persists also one year after the retirement. It does not appear any effects after two years. In the third year after retirement there is a positive effect; this is probably due to a mix of an attenuation effect for the retired people at 52 and to the fact that the control group involved in the reforms has just retired. After the fourth year the distance between the two groups starts to reduce and cannot be deduced any significative difference.

Results show also a negative effect on leisure time satisfaction one year after the retirement. That difference completely disappear starting from the third year. Finally it does not appear any effect on health, family and friend relation satisfaction.

5.2. Delayed effects

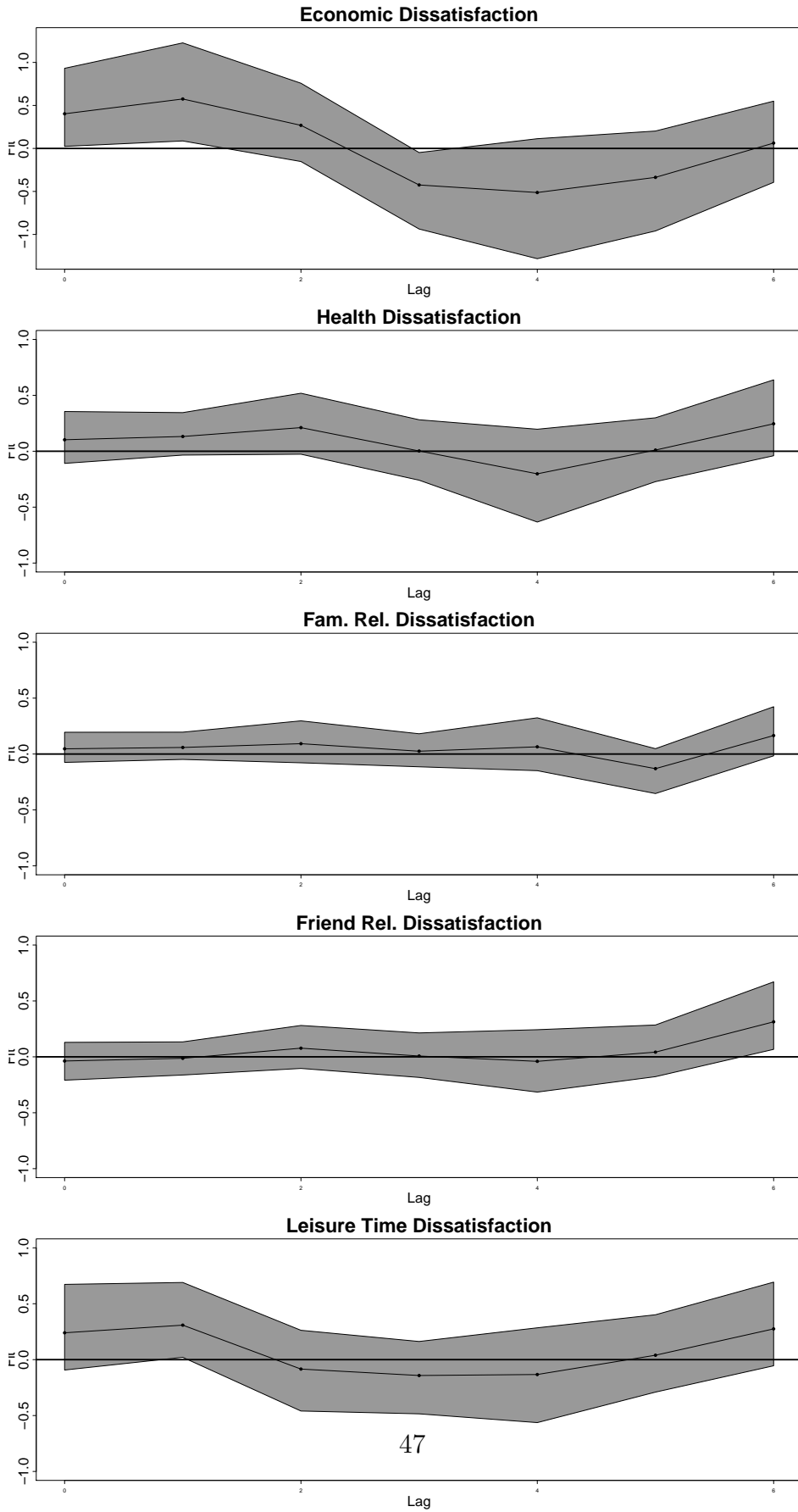


Figure 5.2: Delayed effects 52 Male

Chapter 6

Conclusions

In this thesis we have analyzed how to use in the proper way the information provided by the change of slope in a Regression Discontinuity Design.

First we have analyzed the pension reforms that took place in Italy starting from '90s. We have focused our attention on the Dini's law of the 1995 that has increased the eligibility criterion for the *seniority pension*. The law has introduced two different ways to access to retirement: the first based on an age requirement in addition to 35 years of contribution, the second just based on years of contribution. Moreover this law provided a transition period of progressive tightening of the requirements in order to increase eligibility criterion to 57 years of age and 35 of contribution or 40 years of contribution regardless of age. We have showed how this law has created a reduction of the retirement probability across cohorts comparing individuals at the same age. Further we have showed that the introduction of the law as in the standard RDD settings represents a *weak instrument* for the estimation of the retirement effects. However the progression provided by the law represents a strong instrument in order to base the identification strategy on this change of slope (Kink, in comparison to the Jump of the standard RDD) at the threshold point.

For this reason we have reviewed the recent literature on Regression Kink Design, showing that the Kink ratio, in a binary treatment setting, is a potentially biased estimator of the parameter of interest, the *Local Average Treatment Effect* (LATE) and its bias depends on the unobserved first derivative of the LATE

with respect to the running variable, computed in the threshold point. We have compared the existing models based on *Instrumental Variables* regressions based on the Jump and the Kink ratio. We have then proposed an alternative model that includes the interaction between the treatment variable and the forcing variable as an endogenous variable. The aim is to control for the first derivative of the LATE in the structural equation. We have described how the resulting estimator is a combination of the Jump and the Kink ratio.

We have proposed a simulation study to compare the different estimators when the Jump ratio is too inaccurate to be used. Results have showed how our model provides an improvement of the efficiency of the estimator when the Kink ratio is biased. Moreover the t-test related to the interaction term is able to determine when the first derivative of the LATE is different from zero.

Then, using the survey named *Aspetti della Vita Quotidiana* edited by the Italian National Statistical Office, we have estimated the effects of retirement on subjective well-being for individuals of age 52. Results show a negative effect on the economic satisfaction for males and on friend relations satisfaction for females. However retirement does not appear to affect health, family relation, friend relation and leisure time satisfaction for males and on economic, health, family relation and leisure time satisfaction for females.

Finally we have generalized the results using the *Two Sample Instrumental Variable* (TSIV) estimator, first proposed by Angrist and Krueger (1992). This is based on the use of complementary data sources to estimate an IV regression. The First Stage regression is estimated with one dataset. Then the coefficients are used to obtain the fitted values of the endogenous regressors in the other dataset. Finally the fitted values are used to estimate the structural equation. We have used that estimator in the RDD setting to estimate the first stage regression using administrative data that come from the National Institute of Social Security. These data have a larger sample size and provides information on when exactly the worker is retired. We have use administrative data with survey data in order to improve the accuracy of the IV estimator. Results confirm all the signs of the coefficients obtained with just survey data, but they show also a significant negative effect on family relation satisfaction for females.

We also used the TSIV estimator in order to construct delayed outcomes of the same cohorts using the following waves of the survey data. We wanted to understand how the retirement effects persist some years after the retirement. Unfortunately, this survey does not provide any repeated measure for the same individuals, but the individuals in the following waves are representative of the same cohorts, for this reason the use of the TSIV estimator was necessary. Results show a negative effect in economic and leisure time satisfaction that persists also one year after the retirement.

Appendix A

Alternative specifications of the IV regression

A. Alternative specifications of the IV regression

In this appendix it will be showed the correspondence between the Jump and the Kink ratio and the IV regressions. Defining the ITT as:

$$Y = \gamma_0 + \gamma_1 Z + \gamma_2 X + \gamma_3 XZ + \zeta,$$

and the FS equations as:

$$D = \alpha_0 + \alpha_1 Z + \alpha_2 X + \alpha_3 XZ + v.$$

The Jump ratio (τ_J) and the Kink ratio (τ_K) are respectively defined as:

$$\tau_J = \frac{\gamma_1}{\alpha_1} \qquad \tau_K = \frac{\gamma_3}{\alpha_3}$$

Jump ratio

The standard Jump ratio is obtained from the coefficient β_1 of the equation:

$$Y = \beta_0 + \beta_1 D + \beta_2 X + \beta_3 XZ + \epsilon,$$

where D is instrumented by:

$$D = \alpha_0 + \alpha_1 Z + \alpha_2 X + \alpha_3 XZ + v$$

According to the Frisch–Waugh–Lovell theorem, the coefficient β of the regression:

$$Y = \alpha + \beta X_1 + \gamma X_2 + \epsilon$$

could be equivalently estimated from the residuals of the regression of Y on X_2 regressed on the residuals of X_1 on X_2 . We can remove the exogenous variables taking y , d e z respectively as the residuals of the regressions of Y , D and Z on

X and XZ . So we have that the structural and the first stage regression become:

$$\begin{aligned}y &= \beta_1 d + \epsilon \\d &= \alpha_1 z + v,\end{aligned}$$

and the ITT becomes:

$$y = \gamma_1 z + \zeta,$$

therefore the IV estimator of β_1 is equal to:

$$\beta_1^{IV} = \frac{\text{Cov}[y, z]}{\text{Cov}[d, z]} = \frac{\text{Cov}[\gamma_1 z, z]}{\text{Cov}[\alpha_1 z, z]} = \frac{\gamma_1 \text{Var}[z]}{\alpha_1 \text{Var}[z]} = \frac{\gamma_1}{\alpha_1}.$$

Kink ratio

The Kink ratio is obtained with the structural equation:

$$Y = \beta_0 + \beta_1 D + \beta_2 X + \beta_3 Z + \epsilon,$$

where D is instrumented with:

$$D = \alpha_0 + \alpha_1 Z + \alpha_2 X + \alpha_3 XZ + v.$$

A. Alternative specifications of the IV regression

Defining as y , d e xz respectively the residuals of the regressions of Y , D e XZ on X e su Z , we have that the structural and the FS equations reduce to:

$$\begin{aligned}y &= \beta_1 d + \epsilon \\d &= \alpha_3 xz + v,\end{aligned}$$

and the ITT to:

$$y = \gamma_3 xz + \zeta.$$

so the IV estimator is equal to:

$$\beta_1^{IV} = \frac{\text{Cov}[y, xz]}{\text{Cov}[d, xz]} = \frac{\text{Cov}[\gamma_3 xz, xz]}{\text{Cov}[\alpha_3 xz, xz]} = \frac{\gamma_3 \text{Var}[xz]}{\alpha_3 \text{Var}[xz]} = \frac{\gamma_3}{\alpha_3}$$

Dong's estimator

The structural equation is:

$$Y = \beta_0 + \beta_1 D + \beta_2 X + \epsilon,$$

where D is instrumented with

$$D = \alpha_0 + \alpha_1 Z + \alpha_2 X + \alpha_3 XZ + v$$

As usual, defining y , d , z e xz respectively the residuals of the regressions of Y , D , Z and XZ on X , we obtain that the structural and the FS equations become:

$$\begin{aligned}y &= \beta_1 d + \epsilon \\d &= \alpha_1 z + \alpha_3 xz + v,\end{aligned}$$

and the ITT equation:

$$y = \gamma_1 z + \gamma_3 xz + \zeta.$$

The IV estimator of β_1 is equal:

$$\begin{aligned}\beta_1^{IV} &= \frac{\text{Cov}[y, \alpha_1 z + \alpha_3(xz)]}{\text{Cov}[d, \alpha_1 z + \alpha_3(xz)]} = \frac{\text{Cov}[\gamma_1 z + \gamma_3(xz), \alpha_1 z + \alpha_3(xz)]}{\text{Cov}[d, \alpha_1 z + \alpha_3(xz)]} \\&= \frac{\alpha_1 \gamma_1 \text{Var}[z] + (\alpha_1 \gamma_3 + \alpha_3 \gamma_1) \text{Cov}[z, (xz)] + \alpha_3 \gamma_3 \text{Var}[(xz)]}{\alpha_1 \text{Cov}[d, z] + \alpha_3 \text{Cov}[d, (xz)]} \\&= \frac{\gamma_1 (\alpha_1 \text{Var}[z] + \alpha_3 \text{Cov}[z, (xz)]) + \gamma_3 (\alpha_1 \text{Cov}[z, (xz)] + \alpha_3 \text{Var}[(xz)])}{\alpha_1 \text{Cov}[d, z] + \alpha_3 \text{Cov}[d, (xz)]} \\&= \frac{\gamma_1 \text{Cov}[d, z] + \gamma_3 \text{Cov}[d, (xz)]}{\alpha_1 \text{Cov}[d, z] + \alpha_3 \text{Cov}[d, (xz)]}\end{aligned}$$

So the IV estimator is a combination of the Jump and the Kink ratio in the form:

$$\tau = \frac{w_A^J \gamma_1 + w_A^K \gamma_3}{w_A^J \alpha_1 + w_A^K \alpha_3}$$

where $w_1 = \text{Cov}[d, z]$ e $w_2 = \text{Cov}[d, (xz)]$, finally substituting to d its equation we have that:

$$\begin{aligned}w_A^J &= \text{Cov}[d, z] = \alpha_1 \text{Var}[z] + \alpha_3 \text{Cov}[z, xz] \\w_A^K &= \text{Cov}[d, xz] = \alpha_1 \text{Cov}[z, xz] + \alpha_3 \text{Var}[xz]\end{aligned}$$

Model with two endogenous variables

Taking as structural equation:

$$Y = \beta_0 + \beta_1 D + \beta_2 X + \beta_3 XD + \epsilon,$$

where D and XD are instrumented with:

$$\begin{aligned} D &= \alpha_0 + \alpha_1 Z + \alpha_2 X + \alpha_3 XZ + v \\ XD &= \delta_0 + \delta_1 Z + \delta_2 X + \delta_3 XZ + \xi. \end{aligned}$$

Taking y , d , xd , z e xz that are respectively the residuals of the regression of Y , D , XD , Z and XZ on X , we have:

$$\begin{aligned} y &= \beta_1 d + \beta_3 xd + \epsilon \\ d &= \alpha_1 z + \alpha_3 xz + v \\ xd &= \delta_1 z + \delta_3 xz + \xi, \end{aligned}$$

and the ITT as:

$$y = \gamma_1 z + \gamma_3(xz) + \zeta$$

the IV estimator is equal to:

$$\begin{pmatrix} \beta_1^{IV} \\ \beta_3^{IV} \end{pmatrix} = \begin{bmatrix} \text{Cov}[z, d] & \text{Cov}[z, (zd)] \\ \text{Cov}[(xz), d] & \text{Cov}[(xz), (xd)] \end{bmatrix}^{-1} \begin{bmatrix} \text{Cov}[z, y] \\ \text{Cov}[(xz), y] \end{bmatrix},$$

taking just the first element of the vector (that is our parameter of interest) and substituting to y , d and xd their expression, we obtain that the β_1 coefficient is

still a combination of the Jump and the Kink ratio in the form:

$$\tau = \frac{w_B^J \gamma_1 + w_B^K \gamma_3}{w_B^J \alpha_1 + w_B^K \alpha_3},$$

but the weights are equal to:

$$\begin{aligned} w_B^J &= \{\delta_1 \text{Cov}[z, xz] + \delta_3 \text{Var}[xz]\} \text{Var}[z] - \{\delta_1 \text{Var}[z] + \delta_3 \text{Cov}[z, xz]\} \text{Cov}[x, xz] \\ w_B^K &= \{\delta_1 \text{Cov}[z, xz] + \delta_3 \text{Var}[xz]\} \text{Cov}[z, xz] - \{\delta_1 \text{Var}[z] + \delta_3 \text{Cov}[z, xz]\} \text{Var}[xz]. \end{aligned}$$

So the weights depend from the coefficients of the XD equation and from the covariance matrix of z and xz .

Appendix B

Conditional Moments of the Running Variable in the IV regression

B. Conditional Moments of the Running Variable in the IV regression

As seen in chapter 3 the weights of the combination of the Jump and the Kink ratios obtained with the IV regression proposed by Dong (equation (3.12)) and in this thesis (equation (3.17)) depend on the coefficients of the first stage regressions and on the covariance matrix of z and xz that are the residuals of the regressions (3.13). In this appendix it is showed why that covariance matrix depends just of some conditional and unconditional moments of the running variable. Without loss of generality we can assume that the threshold point is $x_0 = 0$, so $Z = \mathbb{1}_{\{X \geq 0\}}$ and that X is bounded in an interval $[-h, h]$. So we have that:

$$\begin{aligned} \text{Var}[z] &= \mathbb{E}[z^2] = \mathbb{E}[(Z - \lambda_0 - \lambda_1 X)^2] \\ &= \mathbb{E}[Z^2] - 2\lambda_0 \mathbb{E}[Z] - 2\lambda_1 \mathbb{E}[XZ] + 2\lambda_0 \lambda_1 \mathbb{E}[X] + \lambda_0^2 + \lambda_1^2 \mathbb{E}[X^2] \\ &= \mathbb{E}[Z] - 2\lambda_0 \mathbb{E}[Z] - 2\lambda_1 \mathbb{E}[XZ] + 2\lambda_0 \lambda_1 \mathbb{E}[X] + \lambda_0^2 + \lambda_1^2 \mathbb{E}[X^2] \end{aligned}$$

where the first equality is due to the fact that z has zero mean and the last because Z is a dummy variable, so it is idempotent. Now substituting to the coefficients λ_0 and λ_1 their expressions:

$$\lambda_1 = \frac{\text{Cov}[X, Z]}{\text{Var}[X]} \quad \lambda_0 = \mathbb{E}[Z] - \lambda_1 \mathbb{E}[X] = \mathbb{E}[Z] - \frac{\text{Cov}[X, Z]}{\text{Var}[X]} \mathbb{E}[X],$$

we obtain:

$$\begin{aligned}
\text{Var}[z] &= \mathbb{E}[Z] - 2\mathbb{E}[Z]^2 + \frac{\text{Cov}[X, Z]}{\text{Var}[X]} \mathbb{E}[X] \mathbb{E}[Z] - 2 \frac{\text{Cov}[X, Z]}{\text{Var}[X]} \mathbb{E}[XZ] \\
&\quad + 2 \frac{\text{Cov}[X, Z]}{\text{Var}[X]} \mathbb{E}[Z] \mathbb{E}[X] - 2 \frac{\text{Cov}[X, Z]}{\text{Var}[X]} \mathbb{E}[X]^2 + \mathbb{E}[Z]^2 + \frac{\text{Cov}[X, Z]^2}{\text{Var}[X]^2} \mathbb{E}[X]^2 \\
&\quad - 2 \frac{\text{Cov}[X, Z]}{\text{Var}[X]} \mathbb{E}[X] \mathbb{E}[Z] + \frac{\text{Cov}[X, Z]^2}{\text{Var}[X]^2} \mathbb{E}[X^2] \\
&= \mathbb{E}[Z] - \mathbb{E}[Z]^2 - 2 \frac{\text{Cov}[X, Z]}{\text{Var}[X]} \{\mathbb{E}[XZ] - \mathbb{E}[X] \mathbb{E}[Z]\} + \frac{\text{Cov}[X, Z]^2}{\text{Var}[X]^2} \{\mathbb{E}[X^2] - \mathbb{E}[X]^2\} \\
&= \mathbb{E}[Z] - \mathbb{E}[Z]^2 - 2 \frac{\text{Cov}[X, Z]}{\text{Var}[X]} \text{Cov}[X, Z] + \frac{\text{Cov}[X, Z]^2}{\text{Var}[X]^2} \text{Var}[X] \\
&= \mathbb{E}[Z] - \mathbb{E}[Z]^2 - 2 \frac{\text{Cov}[X, Z]^2}{\text{Var}[X]} + \frac{\text{Cov}[X, Z]^2}{\text{Var}[X]} = \mathbb{E}[Z] - \mathbb{E}[Z]^2 - \frac{\text{Cov}[X, Z]^2}{\text{Var}[X]} \\
&= \mathbb{E}[Z] - \mathbb{E}[Z]^2 - \frac{\{\mathbb{E}[XZ] - \mathbb{E}[X] \mathbb{E}[Z]\}^2}{\text{Var}[X]}.
\end{aligned}$$

Because $Z = \mathbb{1}_{\{X \geq 0\}}$, we have that:

$$\begin{aligned}
\mathbb{E}[Z] &= 0 \cdot \Pr[X < 0] + 1 \cdot \Pr[X \geq 0] = \Pr[X \geq 0] \\
\mathbb{E}[XZ] &= 0 \cdot \Pr[Z = 0] + \mathbb{E}[X|Z = 1] \Pr[Z = 1] = \mathbb{E}[X|X \geq 0] \Pr[X \geq 0],
\end{aligned}$$

so finally the expression becomes:

$$\text{Var}[z] = \Pr[X \geq 0] - \Pr[X \geq 0]^2 - \frac{\{\mathbb{E}[X|X \geq 0] \Pr[X \geq 0] - \mathbb{E}[X] \Pr[X \geq 0]\}^2}{\text{Var}[X]}.$$

B. Conditional Moments of the Running Variable in the IV regression

Using similar arguments the variance of xz is:

$$\begin{aligned}
\text{Var}[xz] &= \mathbb{E}[xz^2] = \mathbb{E}[(XZ - \kappa_0 - \kappa_1 X)^2] \\
&= \mathbb{E}[X^2 Z] - 2\kappa_0 \mathbb{E}[XZ] - 2\kappa_1 \mathbb{E}[X^2 Z] + 2\kappa_0 \kappa_1 \mathbb{E}[X] + \kappa_0^2 + \kappa_1^2 \mathbb{E}[X^2] \\
&= \mathbb{E}[X^2 Z] - 2\mathbb{E}[XZ]^2 + 2 \frac{\text{Cov}[X, XZ]}{\text{Var}[X]} \mathbb{E}[X] \mathbb{E}[XZ] - 2 \frac{\text{Cov}[X, XZ]}{\text{Var}[X]} \mathbb{E}[X^2 Z] \\
&\quad + 2 \frac{\text{Cov}[X, XZ]}{\text{Var}[X]} \mathbb{E}[X] \mathbb{E}[XZ] - 2 \frac{\text{Cov}[X, XZ]^2}{\text{Var}[X]^2} \mathbb{E}[X]^2 + \mathbb{E}[XZ]^2 \\
&\quad + \frac{\text{Cov}[X, XZ]^2}{\text{Var}[X]^2} \mathbb{E}[X]^2 - 2 \frac{\text{Cov}[X, XZ]}{\text{Var}[X]} \mathbb{E}[X] \mathbb{E}[XZ] + \frac{\text{Cov}[X, XZ]^2}{\text{Var}[X]^2} \mathbb{E}[X^2] \\
&= \mathbb{E}[X^2 Z] - \mathbb{E}[XZ]^2 - \frac{\text{Cov}[X, XZ]^2}{\text{Var}[X]} \\
&= \mathbb{E}[X^2 | X \geq 0] \Pr[X \geq 0] - \{\mathbb{E}[X | X \geq 0] \Pr[X \geq 0]\}^2 \\
&\quad - \frac{\{\mathbb{E}[X^2 | X \geq 0] \Pr[X \geq 0] - \mathbb{E}[X] \mathbb{E}[X | X \geq 0] \Pr[X \geq 0]\}^2}{\text{Var}[X]},
\end{aligned}$$

and the covariance between z and xz is:

$$\begin{aligned}
\text{Cov}[z, xz] &= \mathbb{E}[z \cdot xz] = \mathbb{E}[(Z - \lambda_0 - \lambda_1 X)(XZ - \kappa_0 - \kappa_1 X)] \\
&= \mathbb{E}[XZ] - \kappa_0 \mathbb{E}[Z] - \kappa_1 \mathbb{E}[XZ] - \lambda_0 \mathbb{E}[XZ] + \lambda_0 \kappa_0 \\
&\quad + \lambda_0 \kappa_1 \mathbb{E}[X] - \lambda_1 \mathbb{E}[X^2 Z] + \lambda_1 \kappa_0 \mathbb{E}[X] + \lambda_1 \kappa_1 \mathbb{E}[X^2] \\
&= \mathbb{E}[XZ] - \mathbb{E}[XZ] \mathbb{E}[Z] + \frac{\text{Cov}[XZ, X]}{\text{Var}[X]} \mathbb{E}[Z] \mathbb{E}[X] - \frac{\text{Cov}[XZ, X]}{\text{Var}[X]} \mathbb{E}[XZ] \\
&\quad - \mathbb{E}[Z] \mathbb{E}[XZ] + \frac{\text{Cov}[X, Z]}{\text{Var}[X]} \mathbb{E}[X] \mathbb{E}[XZ] + \mathbb{E}[XZ] \mathbb{E}[Z] - \frac{\text{Cov}[X, Z]}{\text{Var}[X]} \mathbb{E}[XZ] \mathbb{E}[X] \\
&\quad - \frac{\text{Cov}[XZ, X]}{\text{Var}[X]} \mathbb{E}[X] \mathbb{E}[Z] + \frac{\text{Cov}[X, Z] \text{Cov}[XZ, X]}{\text{Var}[X]^2} \mathbb{E}[X]^2 \\
&\quad + \frac{\text{Cov}[XZ, X]}{\text{Var}[X]} \mathbb{E}[X] \mathbb{E}[Z] - \frac{\text{Cov}[X, Z] \text{Cov}[XZ, Z]}{\text{Var}[X]^2} \mathbb{E}[X]^2 \\
&\quad - \frac{\text{Cov}[X, Z]}{\text{Var}[X]} \mathbb{E}[X^2 Z] + \frac{\text{Cov}[X, Z]}{\text{Var}[X]} \mathbb{E}[XZ] \mathbb{E}[X] - \frac{\text{Cov}[XZ, X] \text{Cov}[X, Z]}{\text{Var}[X]^2} \mathbb{E}[X]^2 \\
&\quad + \frac{\text{Cov}[X, Z] \text{Cov}[XZ, Z]}{\text{Var}[X]^2} \mathbb{E}[X]^2 \\
&= \mathbb{E}[XZ] - \mathbb{E}[XZ] \mathbb{E}[Z] - \frac{\text{Cov}[XZ, X] \text{Cov}[X, Z]}{\text{Var}[X]} \\
&= \mathbb{E}[X|X \geq 0] \Pr[X \geq 0] - \mathbb{E}[X|X \geq 0] \{\Pr[X \geq 0]\}^2 \\
&\quad - \frac{\{\mathbb{E}[X^2|X \geq 0] \Pr[X \geq 0] - \mathbb{E}[X] \mathbb{E}[X|X \geq 0] \Pr[X \geq 0]\}}{\text{Var}[X]} \\
&\quad \times \{\mathbb{E}[X|X \geq 0] \Pr[X \geq 0] - \mathbb{E}[X] \Pr[X \geq 0]\}.
\end{aligned}$$

Note that if X is symmetric around the threshold point (e.g. X is uniformly distributed in the interval $[-h; h]$), we have that:

$$\begin{aligned}
\mathbb{E}[X] &= 0, \\
\Pr[X \geq 0] &= 1/2 \\
f_{X|X \geq 0}(x) &= 2f_X(x) \quad \forall x \in [0; h],
\end{aligned}$$

and consequently $\text{Cov}[z, xz] = 0$.

References

- Aguila, E., Attanasio, O., and Meghir, C. (2011). Changes in consumption at retirement: evidence from panel data. *Review of Economics and Statistics*, 93(3):1094–1099.
- Angrist, J. D., Imbens, G. W., and Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American statistical Association*, 91(434):444–455.
- Angrist, J. D. and Krueger, A. B. (1992). The effect of age at school entry on educational attainment: An application of instrumental variables with moments from two samples. *Journal of the American Statistical Association*, 87(418):328–336.
- Angrist, J. D. and Rokkanen, M. (2015). Wanna get away? regression discontinuity estimation of exam school effects away from the cutoff. *Journal of the American Statistical Association*, (just-accepted):00–00.
- Atchley, R. C. (1976). *The sociology of retirement*. Halsted Press.
- Battistin, E., Brugiavini, A., Rettore, E., and Weber, G. (2009). The retirement consumption puzzle: evidence from a regression discontinuity approach. *The American Economic Review*, 99(5):2209–2226.
- Bertoni, M. and Brunello, G. (2014). Pappa ante portas: The retired husband syndrome in japan.

REFERENCES

- Billari, F. C. and Galasso, V. (2014). Fertility decisions and pension reforms. evidence from natural experiments in italy.
- Blau, D. M. (2008). Retirement and consumption in a life cycle model. *Journal of Labor Economics*, 26(1):35–71.
- Bonsang, E. and Klein, T. J. (2012). Retirement and subjective well-being. *Journal of Economic Behavior & Organization*, 83(3):311–329.
- Card, D., Lee, D., Pei, Z., and Weber, A. (2012). Nonlinear policy rules and the identification and estimation of causal effects in a generalized regression kink design. Technical report, National Bureau of Economic Research.
- Card, D., Pei, Z., Lee, D. S., and Weber, A. (2014). Local polynomial order in regression discontinuity designs. Technical report.
- Charles, K. K. (2004). Is retirement depressing?: Labor force inactivity and psychological well-being in later life. *Research in Labor Economics*, 23:269–299.
- Costa Dora, L. (1998). The evolution of retirement.
- Dong, Y. (2014). Jump or kink? identification of binary treatment regression discontinuity design without the discontinuity. *Unpublished manuscript*.
- Dong, Y. and Lewbel, A. (2012). Identifying the effect of changing the policy threshold in regression discontinuity models. Technical report, Boston College Department of Economics.
- Florens, J.-P., Heckman, J. J., Meghir, C., and Vytlacil, E. (2008). Identification of treatment effects using control functions in models with continuous, endogenous treatment and heterogeneous effects. *Econometrica*, 76(5):1191–1206.
- Galasso, V. (2008). Postponing retirement: the political effect of aging. *Journal of Public Economics*, 92(10):2157–2169.
- Galasso, V. and Profeta, P. (2004). Lessons for an ageing society: the political sustainability of social security systems. *Economic Policy*, 19(38):64–115.

- Gall, T. L., Evans, D. R., and Howard, J. (1997). The retirement adjustment process: Changes in the well-being of male retirees across time. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 52(3):P110–P117.
- Gelman, A. and Imbens, G. (2014). Why high-order polynomials should not be used in regression discontinuity designs. Technical report, National Bureau of Economic Research.
- Gourinchas, P.-O. and Parker, J. A. (2002). Consumption over the life cycle. *Econometrica*, 70(1):47–89.
- Hahn, J., Todd, P., and Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1):201–209.
- Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1):153–61.
- Hershey, D. A. and Henkens, K. (2013). Impact of different types of retirement transitions on perceived satisfaction with life. *The Gerontologist*, page gnt006.
- Hurd, M. D. and Rohwedder, S. (2005). The retirement-consumption puzzle: Anticipated and actual declines in spending at retirement.
- Imbens, G. and Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of Economic Studies*, 79(3):933–959.
- Imbens, G. W. and Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, pages 467–475.
- Imbens, G. W. and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2):615–635.

REFERENCES

- Imbens, G. W. and Rubin, D. B. (1997). Estimating outcome distributions for compliers in instrumental variables models. *The Review of Economic Studies*, 64(4):555–574.
- Inoue, A. and Solon, G. (2010). Two-sample instrumental variables estimators. *The Review of Economics and Statistics*, 92(3):557–561.
- Kim, J. E. and Moen, P. (2002). Retirement transitions, gender, and psychological well-being a life-course, ecological model. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 57(3):P212–P222.
- Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, 48:281–355.
- Leombruni, R., Mazzarella, G., and Mosca, M. (2012). Whip4policy - accesso alle pensioni. *IRVAPP Progress Report 2012*, 3.
- Luengo-Prado, M. J. and Sevilla, A. (2013). Time to cook: Expenditure at retirement in spain*. *The Economic Journal*, 123(569):764–789.
- Rivers, D. and Vuong, Q. H. (1988). Limited information estimators and exogeneity tests for simultaneous probit models. *Journal of econometrics*, 39(3):347–366.
- Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, 66(5):688.
- Simonsen, M., Skipper, L., and Skipper, N. (2015). Price sensitivity of demand for prescription drugs: exploiting a regression kink design. *Journal of Applied Econometrics*.
- Stancanelli, E. and Van Soest, A. (2012). Retirement and home production: A regression discontinuity approach. *American Economic Review*, 102(3):600–605.

REFERENCES

- Thistlethwaite, D. L. and Campbell, D. T. (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational psychology*, 51(6):309.
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Gianluca Mazzarella

CURRICULUM VITAE

Personal Details

Date of Birth: June, 4, 1985
Place of Birth: Napoli, Italy
Nationality: Italian

Contact Information

University of Padova
Department of Statistics
via Cesare Battisti, 241-243
35121 Padova. Italy.

Tel. +39 329 3957511
e-mail: mazzarella@stat.unipd.it

Current Position

Since January 2012; (expected completion: June 2015)

PhD Student in Statistical Sciences, University of Padova.

Thesis title: Pension Reforms, Unemployment Benefits and Labor Supply

Supervisor: Prof. Enrico Rettore

Research interests

- Causal Inference
- Evaluation of Policy
- Unemployment benefits
- Retirement effects

Education

October 2007 – March 2010

Master (laurea specialistica/magistrale) degree in Statistics, Economics and Finance.

University of Padua, Faculty of Statistical Sciences

Title of dissertation: “Identification of Unemployment Benefits using Regression Discontinuity Design” (in Italian)

Supervisor: Prof. Adriano Paggiaro

Final mark: 110/110

October 2004 – October 2007

Bachelor degree (laurea triennale) in Economics.

University of Molise, Faculty of Economics

Title of dissertation: “Diebold-Mariano test for the forecast comparison” (in Italian)

Supervisor: Prof. Claudio Lupi

Final mark: 110/110 cum laude.

Visiting periods

February 2014 – June 2014

Department of Statistics, Harvard University,
Cambridge, USA .

Supervisor: Prof. Donald B. Rubin

Work experience

June 2010 – December 2011

Junior Researcher.

Irvapp – Research Institute for the Evaluation of Public Policies.

Computer skills

- **Operating Systems:** Mac OS, Linux, Windows
- **Statistical Software:** R, Stata, SAS, SPSS
- **Programming Languages:** SQL
- **Other Software:** L^AT_EX, Microsoft Office

Language skills

- **Italian:** native
- **English:** good
- **French:** basic

Publications

Articles in journals

Mazzarella, G., Rettore, E., Trivellato, U., Zanini, N. (2014). The effect of a mixed passive and active labour market policy: Evidence from an Italian programme for dismissed workers. *Rassegna Italiana di Valutazione* **58**, 80–101.

Teaching experience

January 2013 – June 2013

Statistics

Bachelor Degree in Molecular Biology

Teaching assistant, 25 hours

University of Padua

Instructor: Prof. Alessandra Bianchi

References

Prof. Enrico Rettore

Department of Economics,

University of Padova

Via Del Santo, 33 – 35121 Padova

Phone: +39 049 827 4002

e-mail: enrico.rettore@unipd.it

Prof. Antonio Schizzerotto

Department of Sociology and Social Research

University of Trento

Via Verdi, 26 – 38122 Trento

Phone: +39 0461 281431

e-mail: antonio.schizzerotto@soc.unitn.it