UNIVERSITÀ CATTOLICA DEL SACRO CUORE
MILANO

Dottorato di Ricerca in Economia
Ciclo XXVII
S.S.D: SECS-P/02; SECS-P/06; SECS-S/06

# Essays on Social Networks and Non-Cognitive Skills

Tesi di Dottorato di: Lucia Maria Barbone
Matricola: 4010272
Anno Accademico 2014/15

Dottorato di Ricerca in Economia
Ciclo XXVII
S.S.D: SECS-P/02; SECS-P/06; SECS-S/06

# Essays on Social Networks and Non-Cognitive Skills

Coordinatore: Ch.mo Prof. Gianluca Femminis

Advisor: Prof. Maurizio Motolese

Tesi di Dottorato di: Lucia Maria Barbone
Matricola: 4010272
Anno Accademico 2014/15

# Acknowledgements

I would like to express my sincere appreciation and gratitude to the people that have supported me in various ways during these PhD years.

I wish to thank my supervisor, Prof. Maurizio Motolese, for giving me very precious advice on my research, and for always being so positive on my work. I am very grateful for the endless guidance and moral support over the years, especially during tough times. I am especially indebted to him for having encouraged me to pursue graduate studies and for his support of my decision to spend some time abroad: his vision on the importance of working and learning from different cultures and perspectives has helped to shape my idea of academia as an international community.

I would like to express my deep appreciation to Prof. Peter Dolton for the opportunity to work together over these years, and for his constant readiness to help. I am particularly thankful for the encouragement, the feedback, and the push to aim higher and higher with my work: the passion he demonstrates on a daily basis for both research and teaching has been a real motive of inspiration in my own research and teaching. I also wish to thank him for letting me use our joint work for my PhD dissertation.

I am also grateful to Prof. Luigi Campiglio, for the stimulating and exciting discussions during my years at the Catholic University of Milan, and for being the source of inspiration on the topic of my research. His ability to maintain a unitary and broad vision of all the aspects of Economics has always offered me revolutionary perspectives on various research questions.

I owe Maurizio, Peter, and Luigi all I learned over these years, and I am indebted to the three of them for the contributions and the guidance on both my PhD and my career choices. They have been invaluable and dedicated mentors, and they all have contributed to make these years an amazing experience.

I am grateful to the PhD program Director, Prof. Gianluca Femminis, for his advice, precision, and organisation, and in particular for arranging the annual PhD seminars which were a great opportunity for us to receive precious feedback on our work.

I also would like to thank the professors and researchers that I met over the years, and that all contributed to various degrees, with discussions and comments, towards the finalisation of this dissertation: Richard Tol, Lorenzo Cappellari, Andrea Galeotti, Jan Gunning, Daniel Hamermesh, Marteen Lindeboom, Steve Machin, Petra Todd, Peter Urwin, Giuseppe Migali, Matteo Sandi, Edgar Salgado, Hector Rufrancos, and Tsegay Tekleselassie. I thank the seminar participants at the University of Sussex, the Catholic University of Milan, the 2014 DEFAP Summer School, WPEG 2014 Conference, the 2014 PhD Annual Conference (University of Sussex), the 2015 RES Conference, the 2015 ESPE Conference, the 2015 Econometric Society World Congress, and the 2015 Cesifo Area Conference on the Economics of Education for helpful comments and discussions.

I gratefully acknowledge the scholarship received towards my PhD from the Catholic University of Milan, and the grants received towards the participation in various conferences and training courses from Fondazione Cariplo. I also wish to acknowledge the hospitality of the Department of Economics at the University of Sussex, and all of its Faculty, staff, and PhD student, who offered me an interesting and stimulating work environment.

Lastly, I would like to thankfully remember my uncle Luca Barbone, who could not see the end of this work. He has been, and will always be, a fundamental role model and inspirational figure: his patience, encouragement, together with his commitment for high-quality work, will never be forgotten. I can only wish to be able to continue his legacy. This thesis is dedicated to him.

# Co-Autorship Disclaimer

Chapter 2 and 3 of this thesis is a joint work with prof. Peter Dolton, Department of Economics, University of Sussex, and CEP, London School of Economics. Chapter 2 is a revised version of the working paper presented at the University of Sussex, the Catholic University of Milan, the 2014 DEFAP Summer School, WPEG 2014 Conference, the 2014 PhD Annual Conference (University of Sussex), the 2015 RES Conference, the 2015 ESPE Conference, the 2015 Econometric Society World Congress, and the 2015 Cesifo Area Conference on the Economics of Education.

Selected Media Coverage:

- BBC Radio Wales - Good Morning Wales Programme (Interview), 31 March 2015
- BBC Sussex (Interview), 31 March 2015
- The Independent, *School children with more friends earn more as adults, study claims*, by Chris Green, 30 March 2015
- Telegraph, *Why a work and no play wont pay off in the end*, by John Bingham, 31 March 2015
- Teaching Personnel, *School social skills can boost earnings in later life*, by Charlotte Michaels, 31 March 2015

# Declaration

This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (`http://www.cpc.unc.edu/addhealth`). No direct support was received from grant P01-HD31921 for this analysis.

# Contents

**Acknowledgements** i

**Co-Autorship Disclaimer** iii

**Declaration** iv

**Contents** v

**List of Figures** viii

**List of Tables** x

**1 Preface** 1

**2 Interpersonal Skills and Social Networks: A Review of the Literature** 7

2.1 Non-cognitive and Interpersonal skills . . . 9
2.1.1 Personality Traits . . . 10
2.1.2 Behavioural Issues as proxies for (a lack of) non-cognitive skills . . . 24
2.1.3 Social and Interpersonal Skills . . . 26
2.1.4 Discussion . . . 31
2.2 A Brief Introduction to Social Network Metrics . . . 33
2.2.1 Network Centrality . . . 33
2.2.2 Degree Centrality and its Derivatives: Indegree and Outdegree) . . . 34
2.2.3 Bonacich Power Centrality . . . 38
2.3 The Rationale for Using Network Metrics as a Measure of Interpersonal Skills . . . 40
2.4 A Review of the Available Network Datasets . . . 46
2.4.1 The Diffusion of Microfinance Dataset (DF) . . . 47
2.4.2 Tastes, Ties, and Time Dataset (TTT) . . . 49

2.4.3 Context of Adolescent Substance Use Study (CASU) . . . . 49
2.4.4 Wisconsin Longitudinal Study (WLS) . . . . . 50
2.4.5 AddHealth (AH) . . . . . 51
2.5 Empirical Estimation Issues and Distributions of Social Network Measures . . . . . 52
2.5.1 The Poisson Distribution . . . . . 53
2.5.2 The Negative Binomial Distribution . . . . . 56
2.6 Empirical Application . . . . . 59
2.6.1 The Diffusion of Microfinance Dataset . . . . . 59
2.6.2 Poisson . . . . . 66
2.6.3 Negative Binomial . . . . . 68
2.7 Conclusions . . . . . 75

**3 Key Players? Identifying High School Networking Effects on Earnings 77**
3.1 Introduction . . . . . 78
3.2 Interpersonal Skills and Labour Market Outcomes . . . . . 81
3.2.1 Social Networks and Earnings Outcomes . . . . . 84
3.3 Data and Summary Statistics . . . . . 86
3.3.1 Social Network Metrics . . . . . 89
3.4 Endogeneity, Estimation and Identification of Network Effects . . . 93
3.4.1 Two Stage Exclusion Restrictions Model . . . . . 95
3.4.2 Linear-in-Means Model . . . . . 102
3.4.3 Basic Network Model . . . . . 103
3.4.4 Network Centrality Model . . . . . 104
3.4.5 Network Intensity Model . . . . . 105
3.5 Results . . . . . 106
3.5.1 Two Stage Exclusion Restrictions Model . . . . . 106
3.5.2 LiM and Networks Models . . . . . 113
3.6 Observed Heterogeneity and Robustness . . . . . 116
3.6.1 Quantile Regressions and Non-Monotonicity of the effect . . 116
3.6.2 Gender and Ethnicity Effects . . . . . 120
3.6.3 Fully Sampled Schools vs Partially Sampled Schools . . . . . 122
3.7 Discussion and Conclusion . . . . . 123

**Appendices 127**
3.A Variable Descriptions, Summary Statistics, and Descriptive Graphs . . . . . 127
3.A.1 Variable Descriptions . . . . . 127
3.A.2 Summary Statistics . . . . . 128
3.A.3 Adjacency Matrix $G$ . . . . . 134

3.A.4 Adjacency Matrix $R$ . . . 134
3.B Complete Estimation Tables . . . 136
3.C Results for Alternative Identification Strategies . . . 156

**4 Smoking Behaviour over the Early Adulthood Life Course 178**
4.1 Introduction . . . 179
4.2 Previous studies . . . 182
4.2.1 Tobacco Control Policies and Smoking Behaviour . . . 182
4.2.2 Parents' Influence . . . 200
4.2.3 The Social Influence on Smoking Behaviour . . . 204
4.3 Data and Summary Statistics . . . 213
4.3.1 Social Influence Measures . . . 218
4.3.2 Tobacco Control Policies . . . 220
4.4 Empirical Methodology . . . 222
4.4.1 Cross-Sectional Models . . . 222
4.4.2 Panel Models . . . 231
4.5 Results . . . 232
4.5.1 Cross-Sectional Results . . . 232
4.5.2 Panel Data Estimation . . . 248
4.6 Conclusions . . . 252

**Appendices 254**
4.A Overdispersion Test . . . 254
4.B Variables Description . . . 255
4.C Additional Results . . . 257
4.D Additional Summary Statistics . . . 258
4.E Additional State Information . . . 263

**5 Summary and Future Research 272**

**Bibliography 275**

# List of Figures


2.1 Indegree I . . . 36
2.2 Indegree II . . . 36
2.3 Outdegree I . . . 37
2.4 Outdegree II . . . 37
2.5 BC . . . 40
2.6 Network of Village 6 . . . 62
2.7 Degree Actual Density vs. Normal . . . 63
2.8 Degree - Poisson predicted vs actual data . . . 66
2.9 Degree - Negative Binomial predicted vs actual data . . . 70
2.10 Prediction Comparison . . . 72
2.11 Compare Actual, Poisson, and NegBin probabilities . . . 72

3.1 Network of one school - Centrality . . . 92
3.2 BC Quantiles Estimates . . . 118
3.3 Indegree Quantiles Estimates . . . 119
3.A.1 Degree Observed vs. Normal Distribution . . . 128
3.A.2 Degree: Poisson, Negative Binomial and Actual . . . 129
3.A.3 In-Degree: Poisson, Negative Binomial and Actual . . . 129
3.A.4 Bonacich Power Centrality . . . 130

4.1 State and federal cigarette excise taxes, by year. December 31, 1995, to April 1, 2009. Source: CDC . . . 184
4.2 State Taxes by Region . . . 184
4.3 Consumption Levels by Region . . . 185
4.1 Average Cigarettes per Day . . . 216
4.1 Distribution of Cigs per Day . . . 224
4.2 Distribution of Cigs per Day, without zeros . . . 226
4.3 Poisson vs. Negative Binomial . . . 227
4.4 Comparison between NB and ZINB . . . 230
4.E.1 Census regions . . . 263
4.E.2 1994 . . . 264

4.E.3 1995 . . . . . 265
4.E.4 2001 . . . . . 266
4.E.5 2008 . . . . . 267

# List of Tables


1.1 HC Literature . . . . . 2

2.1 Personality Traits . . . . . 11
2.2 Description of Personality Traits . . . . . 13
2.3 Rosenberg Self-Esteem Scale . . . . . 16
2.4 Rotter's scale . . . . . 17
2.5 Socio-Emotional Skills, CASEL . . . . . 27
2.6 Dimensions of Intelligence . . . . . 28
2.7 Indegree and Outdegree Frequencies . . . . . 38
2.8 Network Datasets . . . . . 46
2.9 Poisson distribution assumptions . . . . . 55
2.10 Summary Statistics - Village 6 . . . . . 60
2.11 Summary statistics . . . . . 61
2.12 Degree Frequencies of Village 6 . . . . . 63
2.13 Degree Frequencies . . . . . 64
2.14 PRM: Predicted and actual probabilities . . . . . 67
2.15 NBRM: Predicted and actual probabilities . . . . . 71
2.16 Comparison of Poisson and Negative Binomial . . . . . 71
2.17 BIC and AIC tests . . . . . 71
2.18 Count Data Models Results . . . . . 73

3.1 Previous Literature . . . . . 87
3.2 Summary Statistics . . . . . 91
3.3 Networks Summary Statistics . . . . . 92
3.4 Two Stage Exclusion Restriction Results . . . . . 109
3.5 II Identification Strategy: Wave II data . . . . . 111
3.6 III Identification Strategy: Differences . . . . . 112
3.7 Results Comparison . . . . . 115
3.9 Quantile Estimates: BC . . . . . 117
3.10 Quantile Estimates: Indegree . . . . . 118
3.11 Monotonicity . . . . . 119

3.12 Gender and Ethnicity Estimation . . . 121
3.13 Partially Sampled vs. Fully Sampled Schools . . . 122
3.A.1 Variable Definition . . . 127
3.A.2 AddHealth Waves, Years, and Observations . . . 128
3.A.3 Std Bonacich Power Centrality - Percentiles . . . 130
3.A.4 Individual Characteristics . . . 131
3.A.5 Individual Characteristics - Wave 4 . . . 132
3.A.6 Family Characteristics . . . 132
3.A.7 Schools Characteristics . . . 133
3.A.8 Networks Variables . . . 133
3.B.1 First Stage Estimates . . . 136
3.B.2 Earnings Estimates . . . 139
3.B.3 Intensity Relations Results - Baseline . . . 142
3.B.4 Annual Earnings and Hourly Wage . . . 143
3.B.5 Estimates with OCEAN traits . . . 144
3.B.6 Earnings Estimates - Monotonicity of Effects . . . 144
3.B.7 Earnings Estimates - Quantiles Dummies . . . 147
3.B.8 Quantile Estimation Results - BC . . . 152
3.B.9 Quantile Estimation Results - Degree . . . 154
3.C.1 Wave 2 Estimates - Stand . . . 156
3.C.2 Wave 2 Estimates . . . 160
3.C.3 Wave 2 Estimates - Stand . . . 161
3.C.4 Differences . . . 164
3.C.5 Two stage Results . . . 166
3.C.6 LiM . . . 167
3.C.7 LiM Friends Earnings . . . 169
3.C.8 Undirected Network - Basic and Centrality . . . 170
3.C.9 Undirected Network Intensity - Ranking . . . 172
3.C.10 Directed Network - Basic and Centrality . . . 174
3.C.11 Directed Network Intensity - Ranking . . . 175

4.1 Region Prices and Taxes per Pack . . . 185
4.2 Elasticities of Demand - Participation . . . 188
4.3 Elasticities of Demand - Conditional Demand . . . 191
4.4 Previous Literature on Non-price Policies . . . 196
4.5 Parents' Influence . . . 202
4.6 Peer Effects Literature . . . 207
4.1 Summary statistics . . . 214
4.2 Cigarettes per Day Across Waves . . . 215
4.3 Mother Smoking Behaviour . . . 216
4.4 Father Smoking Behaviour . . . 217

4.5 ASSI . . . 218
4.6 Network Summary statistics . . . 219
4.7 AddHealth Waves, Years, and Observations . . . 219
4.8 Information in AddHealth . . . 220
4.9 Tobacco Control Policies - Summary Statistics . . . 221
4.1 Variables of Interest . . . 222
4.2 Overdispersion Test . . . 225
4.3 Model Selection Tests . . . 229
4.1 Taxes and Smoking Participation . . . 234
4.2 Taxes and Smoking Intensity . . . 235
4.3 Smoking Participation and Non Price Policies . . . 237
4.4 Smoking Intensity and Non Price Policies . . . 238
4.5 Intergenerational Transmission and Smoking Participation . . . 241
4.6 Intergenerational Transmission and Smoking Intensity . . . 242
4.7 Social Effects and Smoking Participation . . . 245
4.9 Social Effects and Intensity of Smoking . . . 246
4.10 MFX ZINB with Networks . . . 247
4.11 ZINB Panel FE with Wave III . . . 250
4.12 ZINB Panel FE with Wave IV . . . 251
4.B.1 Variables Description . . . 255
4.B.2 List of Covariates . . . 256
4.C.1 Smoking Participation Logit - networks with exclusion restrictions 257
4.C.2 Smoking Intensity - networks with exclusion restrictions . . . 257
4.D.1 Representativeness of the sample . . . 258
4.D.2 Individual summary statistics . . . 259
4.D.3 Schools by Region . . . 259
4.D.4 Schools by State . . . 260
4.D.5 Panel Variables . . . 261
4.D.6 Panel Variables with Wave IV . . . 262
4.E.1 Anti-smoking Sentiment . . . 268
4.E.2 Prices by State and Wave . . . 269
4.E.3 Nominal Taxes by State and Wave . . . 270
4.E.4 Nominal Funds by State and Wave . . . 271

# Chapter 1

# Preface

> "*Man is by nature a social animal; an individual who is unsocial naturally, and not accidentally is either beneath our notice or more than human. Society is something that precedes the individual. Anyone who either cannot lead the common life or is so self-sufficient as not to need to, and therefore does not partake of society, is either a beast or a god*". - Aristotle, Politics

The ability of the individual to socialise, to interact with others, and to follow social norms, has always been one of the most important characteristics of human nature, something already recognised by ancient philosophers such as Aristotle in the fourth century b.C. Previous literature has shown that this relational aspect is relevant for the Economics discipline, since it helps to explain the determinants of various choices and outcomes in life, including also economic attainment and health outcomes (see, between others, Gutman and Schoon (2013); Rosenbaum (2001); Borghans, ter Weel and Weinberg (2008); Lundborg (2006); Ali et al. (2012) and Carneiro et al. (2007), and the reviews provided in the following chapters). It is therefore interesting to examine to what extent this sociability contributes to the observed differences in these outcomes.

One of the aims of the scholars in Economics is to describe and explain the causes of the differences in these outcomes. In particular, the Human Capital theory (Becker, 1975; Spence, 1973; Schultz, 1961; Nelson and Phelps, 1966; Coleman, 1988; Gardner, 1993; Bowles et al., 2001) was developed to explain these observed differences, such as in wages and employment status. The leading thread of the theory is that these discrepancies among workers, or at least a good part of it, could be accounted for through the set of skills and abilities that each worker has, and the investment that he has made to develop them. For instance, earnings differences among workers could be explained as due to the differences in both

innate and acquired characteristics, such as IQ and education.

Human capital can be broadly defined as the set of skills and abilities that make an individual able to 'produce' (Acemoglu and Autor, 2009). It is considered a 'capital', because individuals can decide to invest and stock knowledge and competencies. A standard example of a human capital investment is years of education, but also training, school choice, and other factors, which can all be included in this investment decision. Many authors have examined this concept and defined human capital in different ways. Following the classification by Acemoglu and Autor (2009), it is possible to distinguish five different interpretations, summarised in Table 1.1:

Table 1.1: HC Literature

| Author | HC Definition |
|---|---|
| Becker (1975) | Input factor in the production process |
| Gardner (1993) | Multi-dimensional concept, related to the different dimensions of human intelligence |
| Schultz, Nelson-Phelps (1961; 1966) | Capacity of the individual to adapt |
| Bowles, Gintis, Osborne (2001) | Ability to follow the predominant ideology |
| Spence (1973) | Signal of abilities |

The first four interpretations of the HC model are similar since the abilities and skills of individuals are considered as providing an advantage for one's own productivity through different channels and mechanisms. In Becker's interpretation (1975), human capital is the stock of knowledge and skills, which are then 'used by the worker as an input in the production process. This stock is considered as a one-dimensional object (Acemoglu and Autor, 2009). Gardner (1993), a psychologist, emphasised that human capital should rather be considered as a multi-dimensional factor, since intelligence is composed of various dimensions. Schultz, Nelson-Phelps (1961; 1966), instead, considered human capital as the ability to adapt to changing situations and context. Bowles, Gintis and Osborne (2001) define human capital as the ability to obey to the predominant ideology. Finally, Spence's interpretation (1973) differs from previous ones, since he considers human capital a signal of abilities, rather than a factor in production.

In this literature, in particular in the studies following Becker's HC definition, the main focus has been on innate ability and cognitive skills, such as IQ measurements and years of education. A seminal example of this is the Minces earning equation (Mincer, 1974), which modelled earnings as a function of years of education and potential years of experience. Other factors, such as socioeconomic background, ethnicity, years of experience, and time preferences, were then introduced in the HC definition, and were all proven as relevant. The relation between cognitive skills and abilities, and labour outcomes, such as productivity, earnings, and employment, is particularly interesting for researchers and policy makers, since there is some consensus on how to measure and quantify these abilities. As a consequence, results can provide guidance on the educational and employment policies of governments. However, the earnings estimates obtained with these models still present a high proportion of unexplained variance. Recently, it has been argued that part of this unexplained variance is related to characteristics other than cognitive skills and background, such as behavioural and psychological factors, and that these are at least as important as cognitive abilities to individual outcomes and success (Muller, 2006).
These factors have been typically labelled in the literature as 'non-cognitive' skills, an umbrella term including all aspects not directly involved in cognition (Hill, 2012). Examples include personality traits, sociability, motivation, charm, and extraversion. Glaeser et al. (2002) consider these social characteristics of an individual as a contributing part of one's human capital, a result of both innate abilities and investments over the lifetime. As a result, the stock of human capital at each point of time is the combination of both cognitive and non-cognitive skills, i.e., both knowledge and relational abilities. However, the term 'non-cognitive' could be interpreted as too broad to be meaningful, and the definition and the measurement of these skills are still under debate. The term non-cognitive itself could be misleading, since it seems to suggest that these skills should be considered as something separated from intelligence (Muller, 2006). On the contrary, scholars in other fields, such as Psychology and Education, have shown that these skills should be considered as one of the dimensions of intelligence (Gardner, 1993), and that they can be invested in (Borghans, Meijers and Ter Weel, 2008).

As a consequence, the last decade has seen a substantial increase in studies assessing the importance of non-cognitive skills on a variety of outcomes, such as educational attainment (Heckman et al., 2006), wages (Mueller and Plug, 2006*a*), employment probability (Cobb-Clark and Tan, 2011), absenteeism (Störmer and Fahr, 2013), unemployment spells (Gallo et al., 2003), and health behaviour (Heckman et al., 2006). Results among these studies can result as very diverse, and this might be due to the number of different measures adopted by the scholars. Nev-

ertheless, these skills have been proven to be relevant and significant. Typically, the estimations of the impact of these abilities is made difficult by three main issues: a) the lack of guidance from the economic theory on the best measure of non-cognitive skills to use (Muller, 2006), b) a possible endogeneity problem related to both measurement errors in the measures, in particular those based on self-reported information, and to a potential reverse causality problem between these skills and outcomes (Borghans et al., 2011), c) the lack of datasets available for the estimations, which need to have samples large enough to obtain credible estimates, and include all the information needed.

Overall, however, this evidence suggests that non-cognitive skills, and in particular social skills, should be accounted for in human capital models, and that they can help to explain some of the unobserved heterogeneity in individual outcomes. This thesis starts from this acknowledgment and aims to assess whether interpersonal skills, measured through social network metrics, influence labour market attainment, namely earnings, and smoking behaviour. This is done using two main sources of data, the Diffusion of Microfinance dataset, and the AddHealth dataset. These two datasets are very popular for networks estimations, since respondents in both studies were asked to directly nominate people they related to, under various dimensions of relationships. This is the main reason why they were chosen for the empirical estimates in this work. Both these datasets are carefully described in Chapter 1. The Diffusion of Microfinance is used in Chapter 1 to assess which count data estimator, between the Poisson and the Negative Binomial, would fit the data better. The AddHealth dataset is then used in the other two chapters to analyse the impact of interpersonal skills on earnings and smoking behaviour.

The contribution of this thesis to the literature on non-cognitive skills is threefold. Firstly, following the intuition recently introduced in the literature by some scholars (Hill, 2012; Borghans, ter Weel and Weinberg, 2008; Conti et al., 2013), it further examines the case for using metrics borrowed from the social networks literature to directly measure interpersonal and social skills. The use of these metrics also allows us to distinguish and compare the effects of different social skills, for instance popularity compared to the ability of becoming a central player in ones network. Secondly, it provides more evidence on the impact of these skills on both individual earnings and smoking behaviour. These skills are indeed shown to influence the success that one has during young adulthood life, with an impact on earnings similar to the estimated impact of socioeconomic background. In addition, they are also shown to be fundamental for the decision regarding tobacco consumption. Interestingly, different social skills impact these outcomes in different ways. While being a key player in ones network results as giving

an advantage for labour market success, popularity is estimated to be the most influencing factor for individual smoking behaviour. Lastly, it applies different empirical methods for the estimations of this effect, comparing various alternatives proposed in previous literature. On the one hand, the effect of interpersonal skills on earnings is thoroughly examined using a number of estimators, such as a two-step estimator, the Linear-in-Means model, and social interactions models as proposed by Goldsmith-Pinkham and Imbens (2013); Bramoullé et al. (2009) and Manski (1993). On the other hand, the effects on the smoking behaviour are estimated using a range of linear and non-linear estimators, such as OLS, the Negative Binomial, and the Zero-Inflated Negative Binomial. This exercise is interesting for two reasons: firstly, it helps to clarify the true meaning of the effects obtained, and secondly, it allows a discussion of the validity of the assumptions of these estimates. The dissertation also offers a direct contribution to the health economics literature, with the discussion of the appropriateness of the standard linearity assumption of estimators used in previous studies.

The thesis includes three distinct chapters: each of them is designed to be a self-contained paper, even though they all complement each other. The leading threat of this work is that social skills can influence both labour and health outcomes, and that social network metrics could be considered as a potentially superior way to measures these skills.

Chapter 1 aims to summarise the state of the literature on the impact of non-cognitive skills, and illustrates some fundamental premises on which the rest of the thesis is built on, such as the use of social network metrics to measure social skills, and the use of count data models in the empirical estimations. Thus, the chapter is composed of two main parts. The first part reviews the existing literature on non-cognitive skills, and illustrates the most used measures of these skills: personality traits, behavioural issues, and social skills. It then introduces the social network metrics used in this dissertation, namely *Outdegree*, *Indegree*, and *Bonacich Power centrality*, and discuss their use as measure of interpersonal abilities. The second part reviews the available datasets with social network information, and it describes two count estimators, the Poisson and the Negative Binomial, as potential choices for the estimations of the network metrics. It then further compares the two estimators through an empirical exercise, using data from a micro-finance program in India.

Chapter 2 analyses empirically the impact of interpersonal skills on earnings. Following the evidence provided in Chapter 1, these skills are measured using the three social network metrics, using information collected during the high school

period. The effect on earnings is estimated comparing alternative empirical strategies, to fully examine the robustness and the meaning of the results obtained. In particular, results obtained through a two-step estimator and through various social interaction models are compared. Finally, observed heterogeneity in the estimates is further examined through both the comparison of estimates based on gender and quantile effects.

Chapter 3 examines the impact of interpersonal skills on smoking behaviour over the young adulthood life. The aim of this chapter is twofold. Firstly, it seeks to assess the importance of popularity and social skills on both the probability of smoking and the intensity of consumption of cigarettes of the individual. Secondly, it aims to compare the relative importance of these skills to other influencing factors, such as cigarette taxes, tobacco control policies, and parents smoking behaviour. This chapter also offers a methodological contribution, since it analyses the extent to which using non-linear estimators affects the significance and the magnitude of the effects of these factors.