Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN ECONOMIA

Ciclo XIX

Settore scientifico disciplinare di afferenza: P03

Essays in Applied Health Economics: Evidence on Health and Health Care in Italy and UK

Presentata da:
Silvana Robone

Coordinatore Dottorato:
Andrea Ichino

Relatori:
Daniele Fabbri
Andrew Jones
Nigel Rice

Esame finale anno 2008

Abstract

This thesis is the result of my experience as a PhD student taking part in the Joint Doctoral Programme at the University of York and the University of Bologna. In my thesis I deal with topics that are of particular interest in Italy and in Great Britain.

Chapter 2 focuses on the empirical test of the existence of the relationship between technological profiles and market structure claimed by Sutton's theory (1991, 1998) in the specific economic framework of hospital care services provided by the Italian National Health Service (NHS). In order to test the empirical predictions by Sutton, we identify the relevant markets for hospital care services in Italy in terms of both product and geographic dimensions. In particular, the Elzinga and Hogarty (1978) approach has been applied to data on patients' flows across Italian Provinces in order to derive the geographic dimension of each market. Our results provide evidence in favour of the empirical predictions of Sutton.

Chapter 3 deals with the patient mobility in the Italian NHS. To analyse the determinants of patient mobility across Local Health Authorities, we estimate gravity equations in multiplicative form using a Poisson pseudo maximum likelihood method, as proposed by Santos-Silva and Tenreyro (2006). In particular, we focus on the scale effect played by the size of the pool of enrolees. In most of the cases our results are consistent with the predictions of the gravity model.

Chapter 4 considers the effects of contractual and working conditions on self-assessed health and psychological well-being (derived from the General Health Questionnaire) using the British Household Panel Survey (BHPS). We consider two branches of the literature. One suggests that "atypical" contractual conditions have a significant impact on health while the other suggests that health is damaged by adverse working conditions. The main objective of our paper is to combine the two branches of the literature to assess the distinct effects of contractual and working conditions on health. The results suggest that both sets of conditions have some influence on health and psychological well-being of employees.

Contents

List of Tables

Chapter 2

Chapter 3

Chapter 4

List of Figures

Chapter 2

Acknowledgments

I would like to express my most profound gratitude to my supervisors Daniele Fabbri, Andrew Jones and Nigel Rice. I have been privileged to work with them for their excellent guidance and constant support throughout the time it took me to complete my PhD. I am most grateful to Peter Smith, who, as a member of my Thesis Advisory Group, offered me very helpful comments and suggestions. I also thank Alberto Zanardi, with whom I collaborated to the first part of the thesis.

I gratefully acknowledge the Italian Department of Education for providing me financial support for the first three years of my PhD. Thereafter, I completed the PhD alongside being a member of the Health, Econometrics and Data Group (HEDG) at the Centre for Health Economics, University of York. I thank Nigel Rice and Peter Smith for giving me much flexibility and scope to allow me completing the thesis alongside my work commitments.

I thank the staff of the Italian Department of Health, in particular Filippo Palumbo and his collaborators working in the "Direzione Generale della Programmazione Sanitaria", for giving me access to the data about patient flows for hospital treatments in Italy.

My work has undoubtedly benefited from the excellent working environment provided by the Department of Economics of the University of Bologna and by the Centre for Health Economics. I thank all the colleagues, senior and junior, from these two institutions that helped and supported me at various stage of my PhD. I am particularly grateful to the present and past members of the Health, Econometrics and Data Group for their useful econometric suggestions and ideas. I learnt a great deal from this group and I thank the members also for the kind support they gave me.

Special thanks to Alberto Zanardi for passing me on his enthusiasm for research in Public Economics and for being my patient and invaluable guide during my first steps in the world of academic research. His contribution has been fundamental for my professional growth. I am really grateful to him for his help and his friendship in the last six years.

I thank Martin Forster for helping me in this “adventure” of the Joint Doctoral Programme at the University of York and the University of Bologna and Matteo Lippi Bruni for all of his constant helpful and wise advice. I am grateful to Marco Borsa for his precious technical assistance with Matlab, the software I used to perform the analysis in the first chapter, and to Teresa Bago d`Uva and Silvia Balia for their helpful advices on different parts of the thesis. I thank Casey, Claire, Helen, Paolo, Peter and Veronica for proof reading the draft of my thesis, and Giuliana, John, Ling-Hsiang and Valeria for their help with the editing.

I thank all of my “old” friends in Italy, in Ivrea, Milan, Bologna and Rome, for their constant support and for their friendship, which goes beyond space and time. I also thank all of my “new” friends in York, for sharing with me the good and the bad moments in the last two years.

And last but not the least, *grazie Papi*, *Mammuz e Ric* for your continuous support and encouragement during the years of the PhD. I hope you will be always proud of me.

Declarations

The empirical analysis presented in Chapter 2 is a joint paper with Alberto Zanardi entitled "*Market Structure and Technology: Evidence from the Italian National Health Service*". It has been published in *International Journal of Health Care Finance and Economics, (2006), 6, 215-236,*. Earlier versions of the paper were presented at the *27th Conference of the Italian Society of Public Economics*, Pavia, October 2003, at the *Annual Conference of the Italian Association of Health Economics,* Rome, Italy, October 2003, at the *5th European Conference of Health Economics*, London, September 2004, and at the *69th Health Economists' Study Group*, York, July 2006. An earlier version of the paper was awarded the prize of the Italian Health Economics Association (Associazione Italiana di Economia Sanitaria) for the best paper presented by young economists at the 2003 Conference. I contributed to the original idea and the choice of methodological approach. I undertook the analysis of the data in collaboration with my co-author and I prepared many parts of the first draft of the paper and contributed to various later drafts. I prepared the version of the text that appears in this thesis.

The empirical analysis presented in Chapter 3 is a joint paper with Daniele Fabbri entitled "*The Geography of Hospital Admissions in a National Health Service with Patient Choice: Evidence from Italy*". Earlier versions of the paper were presented at the *5th World Conference of the International Health Economics Association*, Barcellona, July 2005, at the *28th Conference of the Italian Society of Public Economics*, Pavia, September 2005, and at the *6th World Conference of the International Health Economics Association*, Copenhagen, July 2007. I contributed to the original idea and to the choice of the methodological approach. I undertook the analysis of the data with the support of my co-author and I prepared most of the first draft of the paper and contributed to later drafts. I prepared the version of the text that appears in this thesis.

The empirical analysis presented in Chapter 4 is a joint paper with Andrew M. Jones and Nigel Rice entitled *"Contractual Conditions, Working Conditions, Health and Well-being in the British Household Panel Survey*". This paper was presented at the *HEDG Seminars*, University of York, June 2006. I developed the original idea and contributed to the choice of the methodological approach. I undertook the analysis of the data with the advice of my co-authors, I prepared the first draft of the paper, contributed to later drafts, and prepared the version of the text that appears in this thesis.

Chapter 1

1. Introduction

This thesis is the result of my experience as a PhD student taking part in the Joint Doctoral Programme at the University of York and the University of Bologna. Since my research activity has been conducted partly at the University of Bologna and partly at the University of York, in my thesis I deal with some topics which are of particular interest in Italy and others which are relevant in Great Britain.

Chapter 2 and Chapter 3 refer to the Italian public hospital care sector. The study of the Italian public hospital care sector is particularly interesting due to the reforms that occurred in the Italian National Health Service (NHS) in late 1990s. The Italian NHS was created in the late 1970s as a regionally based system providing universal coverage free of charge at the point of service. The regions were funded by the central government, which was also responsible for determining the amount of resources to devote to health care and for general planning. The reforms in the late 1990s promoted a certain level of competition in the public hospital sector, by the introduction of prospective payments for hospital admissions via the Diagnosis Related Groups (DRGs) classification scheme. [1] The reforms also introduced elements of regional federalism, introducing sources of autonomous financing for the regions and giving them the main responsibility for their funding.[2] Regions can choose the level of resources to spend in the health care sector and, apart from fulfilling a series of essential services, they are free to choose the quality level and amount of services to provide. Since in Italy patients are free to select the hospital in which they are treated, overall

---

[1] According to Cellini et al. (2000), the adoption of the DRG-payment system, giving the patients` right to freely choose the accredited facilities that best meet their requirements, have to some extent introduced a 'yardstick competition' scheme in the health care system Although hospitals do not compete on prices (since their services are substantially free-of-charge), they try to attract patients competing on the quality of the services provided (Mapelli, 2000; Levaggi and Zanola, 2004)

[2] To allow some fiscal equalization across regions the National Solidarity Fund has been introduced. This Fund is financed by the central government.

these changes have made the hospital care sector more similar functionally to the traditional manufacturing sectors, where consumers are free to choose their suppliers and suppliers are free to choose which product to offer and its quality. These structural changes, considered in the light of the rapid pace of technological development that generally has characterized health care markets, offer a motivation to investigate the relationship between technologies and market structure in the Italian public hospital care sector. We investigate this relationship in Chapter 2. The analysis of the hospital care market structure is relevant from the policy point of view given the possible consequences that market concentration may have on social welfare. Considering markets with a high level of product differentiation (as in the case of the hospital care markets), in the literature there is not a clear consensus about the sign (positive or negative) of the effect of competition on social welfare (Spence, 1976; Dixit and Stigliz, 1977; Salop, 1979; Anderson et al., 1992). Considering the hospital care sector, in particular, some work states that competition reduces costs and increases quality and production efficiency (Melnick et al.,1992; Dranove et al., 1992; Vistnes, 1995; Town and Vistnes, 1997), while others state that in this sector competition leads to a decrease in social welfare (Feldstein, 1971; Held and Pauly, 1983; Robinson and Luft, 1985). The presence of health insurances or regulated prices, indeed, can induce hospitals not to compete on prices to attract patients, but to start a so called "medical arms race", which implies offering un-necessary services to attract patients. The adoption of one of the two points of view leads to opposite policy implications. Adopting the former view, the policy makers should be concerned about the market concentration in the hospital care sector and they should try to reduce it. In contrast, adopting the latter view, market concentration should not cause concern but should be considered positively (Robone and Zanardi, 2006).

In recent years the choice of health care provides has been considered as a means of improving health care services and an aim to pursue by the policy makers in many European countries, as Denmark, Sweden, Norway, Holland and Britain. Considering hospital care, the literature suggests that patient freedom to choose the hospital can improve health outcomes, although the institutional setting of the health care systems is a critical element in this regard (Propper et al., 2006). Since the hospital of “choice” of patients might not be the local provider, patients might need to travel outside their area of residence to be able to choose the hospital they prefer (Exworthy and Peckham, 2006). Patient mobility across different geographical areas is a phenomenon widely spread in the hospital care sector in Italy. This phenomenon causes concern particularly if considered with respect to the reforms that took place in this sector in the 1990s. Since patient mobility is likely to be from regions offering less and lower quality hospital care services to regions providing more and better services (which are likely to be the richest ones), the former regions are required to pay the latter regions for the services received by mobile patients. This mechanism could create deep and long lasting imbalances across regions leading to the rationing of health services to patients living in the poorest regions. These considerations motivate our analysis of the determinants of patients’ mobility in Chapter 3.

Chapter 4 refers to the UK. Over the past 20 years or so, changes in the labour market have had a huge impact on the working arrangements of employees. For example, the number of “standard” full-time permanent jobs has decreased, while “atypical” contractual arrangements (temporary work, part-time contracts, unregulated work etc.) have become more common (Kivimäki et al., 2003). In the European Union, for example, workers with atypical contractual arrangements now account for 12–15% of paid employees (Virtanen et al., 2003). Some studies have underlined that workers with atypical contractual conditions are often characterized by adverse working conditions (Aronsson G et al., 2001; Artazcoz et al., 2005; Letourneux, 1998). Whether and how these changes to employment patterns affect health and well-being is a key question. This question is particularly relevant if placed in the context of Great Britain, where the labour market appears to be more flexible compared to other European countries (Bartley et al., 2004).

In the following of this introductory section I provide a short summary of the research objectives of each Chapter. In particular, I briefly discuss the data analysed, the methods used and the way these studies contribute to the literature.

Chapter 2 "*Market Structure and Technology: Evidence from the Italian National Health Service*" focuses on the empirical test of the existence of the relationship between technological profiles and market structure claimed by Sutton's theory (1991, 1998) in the specific economic framework of hospital care services provided by the Italian NHS. Sutton theorised that industries evolve into distinct market configurations in terms of concentration, depending upon product homogeneity and whether R&D and advertising are relevant in relation to set-up costs. In particular, Sutton`s empirical predictions imply that, in markets characterized by low technological intensity, the lower bound to market concentration converges monotonically to zero when the market size increases, for any level of product homogeneity. Conversely, in markets characterized by high technological intensity, the lower bound to concentration converges to a positive value different from zero when the market size increases, while the lower bound increases from zero with the level of product homogeneity. To test these predictions, we adopt Smith`s (1985, 1994) two step procedure for the estimate of lower bounds for scatters of observations. The first step of the procedure requires making some *a priori* decision about the form of the schedule that describes the lower bound. In our application we specify the lower bound for the observations about market concentration and market size as the positive section of a rectangular hyperbolic function, and as a ray through the origin when considering the observations about market concentration and market homogeneity. The second step involves the use of pseudo-maximum likelihood methods to check that the pattern of the estimated residual fits a Weibull distribution. In order to test the empirical predictions by Sutton, we identify the relevant markets for hospital care services in Italy in terms of both product and geographic dimensions. In particular, the Elzinga and Hogarty (1978) approach has been applied to data on patient flows across Italian Provinces in order to derive the geographic dimension of each market.

The data we use are provided by the Italian Ministry of Health and contain information about the services (discharges) offered by all the public hospitals operating in the Italian NHS during 1999. For any single discharge, the data-set reports the corresponding DRG, the tariffs corresponding to each DRG, the hospital where the service was provided together with the Province where that hospital was located, and the Province where the patient resided. Moreover, the Italian Ministry of Health assembles the DRGs into 25 MDCs (representing specific diagnostic groups). To get information about the technological characteristics of the relevant markets we resort to more detailed data available for a particular Italian region (Emilia-Romagna), assuming that the technology adopted in the production of each category of health care is analogous across regions.

The study provides original contributions to the literature. To our knowledge, our study is the first to test if Sutton`s empirical predictions are valid when considering the Italian public hospital care sector. Moreover, no previous study has utilized the Elzinga and Hogarty (1978) approach to derive the geographic dimension of the markets in the Italian hospital care sector.

Chapter 3 "*The Geography of Hospital Admissions in a National Health Service with Patient Choice: Evidence from Italy*" deals with the phenomenon of patient mobility in the Italian National Health Sector. In the Italian NHS patients are enrolled into health plans managed by Local Health Authorities (LHAs). Enrolment is based on patients' place of residence. A distinctive feature of the Italian NHS is that hospital care is subject to unconstrained patient choice of the provider. Each year, it is estimated that 35% out of the 10 million hospital admissions in Italy take place outside the LHAs of residence. This figure increases to almost 42% for cancer treatment and more than 58% for complex surgery. We observe geographical imbalances in patient mobility. For example, exit rates and average travelled distance to access hospital care are larger for enrolees in southern LHAs. In this paper we aim to evaluate the extent to which the observed imbalances are due to scale effects or reflect a deeper, long lasting north/south divide. In particular, we focus on the scale effect played by the size of the pool of enrolees.

The data we analyse are made by an origin/destination matrix provided by the Italian Ministry of Health and comprise all ordinary admissions to public hospitals for the year 2001. These data are similar to those used in Chapter 2, but are at a more disaggregated level, since they give information about the LHAs where hospitals providing services are located and the LHAs where the patient resides. We classify admissions into 4 product groups: cancer, basic medicine, basic surgery and complex surgery. To control for distance, contiguity, institutional barriers, presence of autonomous hospitals and supply characteristics, and to net out the effects played by the composition of patients' flows, we estimate a gravity equation for the full matrix of pair-wise flows. Following Santos-Silva and Tenreyro (2006), we estimate gravity equations in multiplicative form using a Poisson pseudo maximum likelihood method. This method is robust to different patterns of heteroskedasticy and provides a natural way to deal with the zero flows.

The analysis contributes to the literature in several ways. This is the first study that analyses the determinants of patient mobility across LHAs in Italy using a gravity approach. Whereas the study of Levaggi and Zanola (2004) has used the gravity approach to analyse patient mobility across Italian regions, our paper is original since we investigate patient mobility across Italians LHAs. Considering the pairwise matrix of flows at LHA level instead of regional level allows our paper to gain both informational and policy relevance. Moreover, to our knowledge, the scale effect played by the size of the pool of enrolees is an aspect that has not been considered before in the context of the Italian NHS. Our study is also the first that in the context of the Italian NHS estimates a gravity equation using a Poisson pseudo maximum likelihood method, as proposed by Santos-Silva and Tenreyro (2006).

Chapter 4 "C*ontractual Conditions, Working Conditions, Health and Well-being in the British Household Panel Survey*" considers the effects of contractual and working conditions on Self-Assessed Health (SAH) and psychological well-being (derived from the General Health Questionnaire - GHQ) using the British Household Panel Survey (BHPS). We consider two branches of the literature. One suggests that “atypical” contractual conditions have a significant impact on health while the other suggests that health is damaged by adverse working conditions. The main objective of our paper is to combine the two branches of the literature to assess the distinct effects of contractual and working conditions on health and psychological well-being. As part our analysis, we attempt to evaluate the role that preferences for the number of hours of work, the level of employability and family structure play in affecting the relationship between contractual/working conditions and health and psychological well-being. We also attempt to evaluate if our data provide some support to the “demand-control-support” model (Karasek et al., 1988; Karasek and Theorell, 1990) and the “effort–reward imbalance model” (Siegrist et al., 1990; Siegrist, 1996), two of the most influential models developed to investigate the possible mechanisms underling the relationship between working conditions and health and psychological well-being.

To estimate our econometric models for self-assessed health and psychological well-being we exploit the panel data available in the first 12 waves (1991–2002) of the BHPS, a longitudinal survey of private households in Great Britain, including rich information on occupational, socio-demographic and health variables. In our analysis we use an unbalanced sample, which also includes new entrants to the survey. We consider only employees, and we exclude from our sample people outside the job market or who are self- employed.

In our analysis we adopt a dynamic approach, regressing the present level of self-assessed health and psychological well-being on past values of the variables of interest. For modelling self-assessed health, since the dependent variable is categorical, we estimate non-linear dynamic panel ordered probit models, while for modelling psychological well-being we estimate a linear dynamic specification.

The study contributes to the literature in a number of ways. The evaluation of the distinct effects of contractual and working conditions on health and psychological well-being is an original point of our study. Further, many of the previous studies in the literature focus on specific occupational groups (i.e.: civil servants, nurses, etc.), which makes it difficult to generalise results to the entire workforce. In contrast, the use the BHPS allows us to generalize our results to a population of employees in Britain. Whereas other studies considering contractual and working conditions have used this dataset in the past (Bardasi and Francesconi, 2003; Bartley et al., 2004; Rodrigues, 2002), our paper is original since we analyse the effects of contractual and working conditions *jointly* and we exploit a larger number of waves in the BHPS dataset. Moreover, we estimate dynamic panel data models (linear for GHQ and non linear for SAH), which provide many advantages compared to cross sectional approaches. For example, they increase precision in estimation, account for unobserved individual heterogeneity and reduce concerns about endogeneity bias (Cameron and Trivedi, 2005).

Chapter 2

Market structure and technology: evidence from the Italian National Health Service

2.1. Introduction

In the past few decades health care markets generally have been affected by two distinct structural changes. On one hand, the rapid pace of technological development as well as the diffusion of new technologies has increased cost by improving the quality of care and introducing new and costlier products. On the other hand a clear trend has emerged nationally towards decentralization of public intervention in health care. During the 90s, reforms favouring regionalization have reshaped national health services in Canada, Italy, Spain and Sweden, determining the rise of regional health services differentiated on the basis of organization, institutions and services, and increasing the autonomy of agents operating in the health care markets on both the demand and the supply side. On the demand side, patients have gained the right to choose the facilities that best meet their needs or expectations, while on the supply side hospitals and other health structures, apart from fulfilling a series of essential services, have become more autonomous in concentrating their own resources on specific health care productions started by technological innovations.[1] As a whole, these changes have made health care markets more similar functionally to the traditional industrial sectors, where it is generally assumed that consumers are free to choose their supplier, suppliers are free to choose which product to offer and its quality, and technology plays a crucial role.

In Italy, insofar as hospital care is concerned, the adoption of the DRG-payment system and the patients` right to freely choose the accredited facilities that best meet their requirements have to some extent introduced a 'yardstick competition' scheme in the health care system (Cellini et al., 2000). Even if hospitals do not

[1] In Italy, for instance, although a standard level of health care must be offered across different regions, the new system of regional funding has greatly increased interregional differences in the quality and quantity of the services provided by the National Health Service.

compete on prices (since their services are mostly free-of-charge), they try to attract patients by competing on the quality of the services provided (Levaggi and Zanola, 2004; Mapelli, 2000). We stress that the DRG-based payment system is not applied to all Italian hospitals. The Italian National Health Service includes two categories of public hospitals: the Hospital Trusts, which have the status of quasi-independent public agencies and have financial and technical autonomy, and hospitals run directly by Local Health Units. The Local Health Unit represents the lowest level in the organization of the health care system and are organized as nonprofit firms. Local Health Units are run by a general director, responsible for ensuring sound financial management and provided by law with substantial autonomy in managing human, financial and technological resources. Hospital Trusts are financed through a DRG-based payment system whilst the hospitals included in Local Health Units use a cost reimbursement system. Even if the latter are not financed by a DRG-based payment system, (and, therefore, their financial solvency does not strictly depend on the number of services provided), they are however subject to competitive pressures. If the hospitals belonging to a Local Health Unit provide services of low quality, patients would move towards others and consequently the manager of the Local Health Unit would loose credibility as an administrator and risk not to be appointed as manager again in the future.

A number of empirical studies have stressed the relevance of the relationship between geographic mobility of patients and the technological complexity of health services provided by different regional Health Services in Italy (Degli Esposti et al., 1996; Fabbri and Fiorentini, 1996; Spampinato, 2001; Ugolini and Fabbri, 1998). Patients are more willing to travel, and therefore to bear the associated costs of transaction, if they need highly specialized care.

The industrial organization literature has largely investigated the relationship between technology and market structure. In this regard a standard reference is offered by the theoretical and empirical contribution by John Sutton. A key feature of Sutton’s (1991, 1998) work on industrial organization is that industrial markets evolve into distinct configurations of varying concentration, depending upon product homogeneity and differentiable R&D or advertising costs. Sutton develops his analysis with reference to manufacturing sectors. A number of subsequent contributions perform empirical tests for several industries in the EU and the US, all of which support Sutton's predictions about market size and market structure (Buzzacchi and Valletti, 2003; de Juan, 2003; Giorgetti, 2000, 2002, 2003; Gruber, 2002; Lyons and Matraves, 1996; Marin and Siotis, 2002; Matraves, 1999; Robinson and Chiang, 1996). We argue that the same interpretation can be given to the health care system, and similarly test the predictions due to Sutton's argument when applied to the National Health Service in Italy.

This paper is organized as follows. In Section 2.2 the theoretical framework proposed by Sutton is briefly illustrated. Section 2.3 discusses the criteria followed for the identification of the relevant markets for the analysis of hospital care services. Section 2.4 describes the data on the hospital care services provided by the Italian National Health Service (NHS) that are at the basis of this empirical application and shows how the procedures presented in Section 2.3 have been applied to these data. Section 2.5 illustrates the statistical tests developed to test Sutton's empirical predictions about the relation between market concentration and market size, and between market concentration and product homogeneity. Section 2.6 concludes.

2.2. Technology and market structure: the Sutton's approach

Sutton (1991, 1998) investigates the relationship between an industry's R&D intensity (measured by the ratio of R&D spending on sales) and its level of concentration (measured by the combined market share of some specified number of firms). Although this relationship has been debated at length, empirically and theoretically, no consensus has formed. From a purely theoretical basis the direction of causality was originally disputed.[2] During the 70's however, the endogeneity of concentration and R&D intensity became accepted, such that both should be determined simultaneously within an equilibrium system. When tested empirically, no clear consensus emerged from cross-industry analyses concerning

[2] The latter stance is mainly based on appeal to the *structure*/*conduct*/*performance* paradigm developed by Bain (1956). Within that paradigm, it is claimed that a one-way chain of causation runs in each industry from *structure* (the level of concentration) to *conduct* (the degree of collusion), and from conduct to *performance* (profitability). Structure, in this setting, is explained by the degree of scale economies in the industry and by observed levels of advertising and R&D outlays relative to industry sales.

the sign and the form of the relationship, with all of positive, negative and no correlation found.

The starting point of the theory developed by Sutton lies in the observation that R&D (and advertising) outlays can be considered as sunk costs incurred by the firm in order to enhance consumers' willingness-to-pay for the firm's product (Shaked and Sutton, 1982, 1987; Sutton, 1989, 1991, 1998). Thus R&D and advertising outlays can be chosen by firms, and their levels can be determined endogenously as part of the specification of industry equilibrium: thus they can be considered as endogenous sunk costs.

Here we are interested in two predictions put forward by Sutton (1991).

Prediction 1: In exogenous sunk-cost markets (where R&D is low) the lower bound to equilibrium concentration converges monotonically to zero as the ratio between the market size and the set up costs increases (see Figure 1).

Prediction 2: In endogenous sunk-cost markets (where R&D is intensive) concentration remains bounded away from zero as the ratio between the market size and the set-up costs increase (see Figure 2).

The intuition behind is the following. In exogenous sunk-cost markets, a continuing increase in market size gives entrants the opportunity to build a profitable scale of operations. Alternatively in endogenous sunk-cost markets a firm is willing to suffer short-term losses in order to gain long-term revenues. As market size increase, long-term revenues increase. Through this mechanism, increasing market size leads to a competitive escalation in short-term spending: eventually, when this spending becomes too high for new entrants entry is blocked. The degree to which concentration is bounded away from zero is an empirical matter.

In a later contribution, Sutton (1998) derives two distinct empirical predictions regarding the joint distribution of R&D intensity, concentration, and product homogeneity.

Prediction 3: In industries where R&D intensity is low, the lower bound to concentration converges to zero as the size of the market becomes large, independently of the degree of product homogeneity (see Figure 3).

Prediction 4: On the contrary, in industries where R&D intensity is high, the lower bound to concentration increases from zero with product homogeneity (see Figure 4).

Therefore, if R&D is ineffective in raising consumers' willingness-to-pay, it can be shown that R&D intensity is necessarily low. Following this, if we select a set of industry for which the R&D/sales ratio is high, we know that for this group R&D effectiveness is high. Whether this necessarily implies a high level of concentration depends on the strength of the linkages between sub-markets, which in turn depends on the scope economies and the degree of substitutability across products associated with different R&D trajectories. Where these are relevant (so homogeneity is high), concentration will be necessarily high since, if all firms have a low market share, an escalation of R&D spending will be profitable. A high-R&D-spending firm can capture sales from low-R&D-spending rivals on its own trajectory and on the others. On the contrary, when the scope economies and the degree of substitutability across products are limited (so homogeneity is low), concentration may be low in spite of the effectiveness of R&D spending. There are many product groups, associated with different R&D trajectories, and therefore escalation can yield only poor returns. If on the other hand R&D intensity is low, the absence of an escalation mechanism involving R&D makes the market able to support an indefinite number of firms, and therefore the theory predicts that the lower bound to concentration is zero independent of product homogeneity.

The empirical predictions by Sutton place only weak restrictions on observable industry characteristics such as R&D intensity, market size, product homogeneity and market concentration, leaving the predictions quite general. [3]

[3] We underline that Sutton`s predictions hold independently of the level of price competition present in the markets. The theoretical framework of Sutton (1991, 1998), indeed, can be adopted in the context of different industrial organization models, such as Bertrand, Cournot or the Monopoly model. The theoretical framework of Sutton, therefore, can be applied even to markets where the economical agents do not behave according to the model of perfect competition, such as the Italian hospital care market.

This generality is exploited by Sutton (1991, 1998) by following two different paths of analysis, by developing a statistical test of his empirical predictions based on data from a selection of industrial sectors, then by discussing a series of industry cases (from flow-meters to turbine generators) that, through a detailed collection of qualitative information, enable further probing of the theory and its validity.

2.3. The identification of relevant market: methodological issues

In order to test the empirical predictions claimed by Sutton in the case of the hospital care services provided by the Italian NHS the relevant markets must first be identified. This issue is crucial when considering the anticompetitive strategies firms eventually undertake after mergers and acquisitions. Indeed, much of the literature about the identification of the relevant markets has been developed with the aim of detecting such strategies. In general, a relevant market can be characterized according to two different dimensions: the "product" dimension and the "geographic" dimension.

As underlined by Sutton (1991), the product relevant to identify a market can be characterized with either reference to the demand side by considering the substitutability of products in consumption, or to the supply side by focusing on the similarities in production techniques. The demand side criterion is used more frequently, although the latter is convenient when barriers to mobility between markets are relevant in the industry under consideration (Caveas and Porter, 1978).

Two different approaches are usually seen when identifying the relevant geographic markets for a product. The first is based on the evolution of prices in different geographic areas; the second on the movement of physical quantities of the product across different areas. Since in Italy hospital care is essentially free-of-charge and, as a consequence, the market equilibrium is not reached through the price mechanism, we focus on the second approach.

Within this general perspective, the method developed by Elzinga and Hogarty (1973, 1978) is most useful. This is a shipment-based technique involving measurements of product flows across areas, which is based on the application of two different tests:

1. the Little Out From Inside (LOFI) test. The LOFI test concerns the supply side of the market and is based on the question, “What is the smallest geographic region required to account for nearly all shipments from a given producing area?”;
2. the Little In From Outside (LIFO) test. The LIFO test refers to the demand side of the market and centres around the question, “Of total purchases within the region identified by the LOFI test do nearly all emanate from within that region itself?”.

The logic behind this approach is simple. If a certain area meets both of these tests, and this occurs when outflows to other areas and inflows into that area are relatively small, then that area is taken as a relevant market according to the Elzinga and Hogarty criterion. It is usually required that those outflows and inflows do not exceed a given threshold, which is often set to 10% or 25% of the total amount of production or consumption.

The implementation of the Elzinga and Hogarty procedure requires the following steps. The first step in the definition of the geographic market requires the identification of a geographical area as the starting point. Conventionally, the area characterized by the biggest volume of production is considered as the staring point. In our work the starting point is the Province providing the largest share of hospital care treatments in a MDC. Then a “hypothetical market” is defined as the “macro” area formed by the smallest number of areas whose residents consume, in the aggregate, at least 90% of the goods produced in the area assumed as the starting point. Notice that the Elzinga and Hogarty procedure does not require the areas forming the hypothetical market to be contiguous. Those areas do not need to share some geographical border. If at least 90% of the goods consumed within that hypothetical market are produced by the areas that form that market, the LIFO test is met, and then the procedure moves to the next step, the LOFI test. If the LIFO test is not met, it is necessary to enlarge the hypothetical market area by including one (or more) additional areas until the LIFO test is satisfied. If the hypothetical market area satisfies even the LOFI test, then that area is assumed to be a relevant market for that product. Otherwise, analogously to the approach followed before, additional areas are added to the original hypothetical market till the LOFI test is passed. Annex 1 provides more details about the way the Elzinga and Hogarty procedure has been implemented with regards to the Italian hospital care markets.

The Elzinga and Hogarty method was originally developed with reference to commodity markets, like beer (Elzinga and Hogarty, 1973) or coal (Elzinga and Hogarty, 1978; Warell, 2005). It has been recently applied to the context of hospital care as well. During the 90s the US Department of Justice relied on the analysis of patients flows in order to assess a number of cases of hospital mergers (Capps et al. 2002, 2003). Moreover, Dalmau-Matarrodona and Puig-Junoy (1998) estimated geographic hospital markets in Catalunya as a first step in evaluating the potential effect of market structure on hospital technical efficiency.

The Elzinga and Hogarty method has been criticized on several grounds. Capps et al. (2001) discuss the so-called 'silent majority fallacy' of the Elzinga and Hogarty criterion, "If travellers differ significantly from non-travellers, then the presence of a minority of travellers does not imply that local firms lack market power vis-à-vis the majority of consumers who are non-travellers" (Capps et al., 2001, 2002).

Non-traveller customers can be considered a 'silent majority', which has fewer choices after a merger. Thus, in markets with heterogeneous tastes for different services the relevant market according to the definition of Elzinga and Hogarty could be too broad. Werden (1981, 1989) has put forward two other arguments. First, the Elzinga and Hogarty method does not account for potential competition because if the price cross-elasticity of demand is high the firms in the different regions are not likely to set prices independently. In this situation, the Elzinga and Hogarty criterion would identify markets too narrowly. Secondly, the Elzinga and Hogarty method follows a static, rather than a dynamic, approach to the identification of relevant markets. In fact, it fails to analyze post-merger shipment patterns and never answers the question of whether firms could profitably raise prices after a merger. This criticism implies that in some cases markets are defined too broadly.

Although we recognize that no criterion for market identification is ideal for all kinds of markets, we believe that the Elzinga and Hogarty method suits the case of the Italian hospital care market quite well, for both practical and theoretical reasons. On practical grounds, the empirical application in this paper is based on data for patients' mobility for hospital treatment across Italian Provinces that, as discussed in details in Section 2.4, are organized in the form of an origin/destination matrix. This kind of structure is particularly suitable for the implementation of the Elzinga and Hogarty approach. Moreover, since Italian hospitals do not compete directly according to price, the use of alternative methods based on price competition would not fit this case. On the theoretical ground, the 'silent majority' argument seems not to apply to our case because we consider the flows of patients across Provinces for each Main Diagnostic Category (MDC) separately, and in each MDC some degree of patient homogeneity is likely. Blackstone and Fuhr (2000) support this argument, underlining that the Elzinga and Hogarty criterion may be more appropriate if applied by Diagnosis Related Group (DRG) or by some procedural category. In order to make our analysis more robust with regards to the "silent majority" argument, we should consider also the heterogeneity of the socio-economic conditions of the patients. Unfortunately, the data about patient flows we use for our analysis are collected at aggregated level and do not contain any information about the patients at individual level. Due to these data limitations, therefore, we cannot consider this further feature. Accordingly, while a DRG may be homogenous clinically, they may mask differences across individuals in terms of their socio-economic characteristics.

We believe that the arguments put forward by Werden (1981, 1989) seem not to be relevant in this case, because our aim is not to detect anti-competitive strategies, as in most of the studies to which Werden refers. In our analysis it would not be useful to adopt a dynamic definition of the relevant markets, since we are interested only in providing a snapshot of the concentration of the market for hospital care. On a general ground, the use of longitudinal data and the adoption of a dynamic approach could provide useful insights about the relationship between market concentration and technological intensity in the Italian hospital care market. Considering that the panel dataset we have access to is very short (it comprises only the years 1999 and 2001), however, the adoption of a dynamic perspective would not increase the informative content of our analysis in a significant way. Accordingly, we specify cross-sectional models. Also, as stressed by Blackstone and Fuhr (2000), the empirical application of the dynamic definition is very demanding due to the difficulty of predicting the future of such a rapidly-changing industry.

2.4. The identification of relevant markets: empirical application

2.4.1. The data

The structure of the hospital markets in Italy is analyzed in this paper with data containing information about the services (discharges) provided by all the public hospitals operating in the Italian NHS during 1999. For any single discharge, the data-set reports the corresponding DRG according to the classification adopted by the Italian Ministry of Health, the tariffs corresponding to each DRG by which the Local Health Units were reimbursed, the hospital where the service was provided together with the Province where that hospital was located, and the Province where the patient resided. The data includes all the 492 DRGs covered by the classification by the Italian Ministry of Health as well as the 103 Provinces into which the Italian territory is partitioned.[4] Moreover, the Ministry of Health assembles the DRGs into 25 MDCs (representing specific diagnostic groups). Some observations are excluded due to incomplete information (the tests are performed considering 480 DRGs).[5] The data is in the form of an origin-destination matrix, which allows direct investigation of each patients' mobility across Italian Provinces.

2.4.2. Product and geographic dimension of the markets

Testing the empirical predictions due to Sutton (1991, 1998) requires identification of the relevant markets, in terms of both the product and the

[4] No medical care services corresponding to DRG numbers 109, 438 and 474 are reported in the 1999 data-set, even if those DRGs are included in the official classification.

[5] In order to perform statistical tests on the data, we have excluded from the data-set those observations that do not report complete information about the corresponding DRG, the hospital where the service is provided, the Province where the hospital is operative. A number of health services corresponding to DRG numbers 468, 469, 470, 476, 477, 480, 481, 482 and 483, which the official classification does not include in any MDC, have been removed as well.

geographic dimensions. For the product dimension we define the relevant markets in such a way that the criteria referring to the demand side and the supply side are, at least to some extent, both satisfied. If we consider the demand side criterion, we should only focus on the substitutability across hospital treatments, which is very low here because of the characteristics of the services in question. As a consequence, according to this criterion, each DRG should be regarded as a separate product. When considering the supply side however, the DRGs included in a specific MDC generally share many similarities in the production techniques. A hospital producing a DRG included in a specific MDC can easily provide another DRG not as yet provided in that MDC, while it is much more difficult technologically to provide a new DRG in an MDC that is not currently provided. Thus, considering these two perspectives jointly, it seems reasonable to regard the MDCs as the relevant main products that include the DRGs as sub-products.

The Elzinga and Hogarty criterion has then been applied distinctly for each MDC to the Provincial origin-destination matrix, in order to add the geographic dimension and to identify the relevant markets.[6] As a result, 1,425 distinct relevant markets have been identified. Figure 5 and 6 report respectively the geographic sub-division of Italy into the relevant markets identified in the case of Burns and of Neonatal period diseases, taken as examples of products characterized by relatively few markets (42 in the case of Burns) and by a more fragmented market structure (83 in the case of Neonatal period diseases).

Some summary statistics of the structure of relevant markets classified according to product are provided in Table. 1. The characterization of the relevant markets is much diversified across different products. First of all, the various products are sharply differentiated in terms of the number of services provided, of sub-products included and of hospitals producing those services. For example, the number of health services ranges from 1,097,901 (Cardio-vascular apparatus diseases) to 7,268 (Burns). Variability is marked also as far as the number of sub-products included in different product categories is concerned, which have a coefficient of variation 0.68, whereas dispersion is much more limited when we look at the number of hospitals operating in each product, showing a coefficient of variation of only 0.14.

[6] Annex 1 presents the empirical procedure employed in detail.

We consider the size of the relevant markets by the number of the residents in the corresponding geographic area as a measure of composition, which differs according to the products. The distribution of this index, summarized in terms of average and coefficient of variation, greatly vary across products. In general, markets that are large on average are, to some extent, associated with higher variability in that product compared to the others (compare for example Myelo-proliferative and Neonatal period diseases).

Another way to characterize the composition of the relevant markets is to look at the number and the territorial localization of the institutional entities (in this case the Provinces) composing those markets. Table 1 shows for each product the percentage of the relevant markets including more than one Province (*multiprov*), that provides a measure of fragmentation of markets in territorial terms.[7] Again a large dispersion is noticeable across products, ranging from a minimum of 18.1% (Neonatal period diseases) to a maximum of 52.1% (Factors affecting health and health services demand). Moreover, the composition of relevant markets by Provinces highlights possible territorial discontinuities in their geographic structure. The last column of Table 1 (*discontinuity*) contains the number of non-contiguous Provinces in the same market for each product. For some specific products the structure of the relevant markets shows a relatively high number (up to 5) of non-contiguous Provinces. Upon closer investigation, those Provinces are usually located in Southern Italy as opposed to the central group of the (contiguous) Provinces constituting the market in Northern or Central Italy. This result is consistent, with previous research on patient mobility that shows relevant interregional patients flowing from South to Centre-North of the country.[8] In summary, for the products where the markets are larger on average, they are also generally more differentiated in size within the same product category, they are more likely composed of more than one Province and those

[7] Given the different size of the Provinces in which the Italian territory is partitioned, it is obvious that the number of Provinces provides only a *prima facie* measure of the territorial dimension of the markets.

[8] For instance, Donatini et al. (2001), claim that Northern regions attract more interregional patients than they lose to other regions, whereas the opposite occurs for Southern regions. This is because people in Northern Italy tend to receive care in their own or nearby regions, unlike their Southern counterparts.

Provinces are sometimes not contiguous, such that patients flow generally from Provinces in South Italy to Centre-North.

2.4.3. Market concentration, product homogeneity and technology

To analyze the market structure of the Italian NHS through Sutton's theoretical framework, some correspondences between the theoretical setting and the data need to be established. Thus (i) within each market hospitals operate as firms; (ii) revenue is measured by the number of services provided, multiplied by their corresponding tariff; (iii) the share of each hospital in each market is measured by the sales revenue gained by that hospital over the total revenue gained in that market.

In order to test the consistency of our data with the empirical predictions from Sutton's theory, three indices have been derived for each market:

1. The concentration index C_1, defined as the market shares corresponding to the firm providing the largest number of health services in the market;
2. The homogeneity index h , determined as the fraction of sales revenue corresponding to the largest sub-market (in terms of number of health services produced) in the market;
3. The size index *size*, determined as the revenue realized in the market, divided by the revenue of the median hospital.

Information about the technological characteristics of the relevant markets is also required. Given that information of this kind is not included in our data, we resort to more detailed data available for a particular Italian region (Emilia Romagna). We assume that the technology adopted in the production of each category of health care is analogous across regions, unlike the composition of health care productions, which differs across regions. This data contains the number of health care services provided by every hospital distinctly by DRG and hospital ward and, on the other hand, total costs classified by main items (medical staff, paramedic staff, executive staff, non-durable goods purchases, etc.) incurred by hospital wards of each hospital. The basic cost items by hospital wards and services provided (measured in terms of hospitalization days) classified by hospital wards and DRG have then been aggregated across hospitals. Among the cost items the category of capital costs (depreciations plus capital goods hiring) have been selected. A capital intensity index has then been derived for each DRG as the product of the ‘weight index’, which measures, according to the Italian Ministry of Health, the overall amount of human and physical resources necessary to provide the services corresponding to that DRG, and the weighted average of the ratio of capital costs to total costs for each hospital ward, where the weights are given by the share of the total number of services each hospital ward provides in each DRG.[9] The capital intensity index is then considered to be a proxy for the technological intensity of each DRG.[10] Finally, the technological intensity index *tech* for each relevant market has been derived as the weighted average of the capital intensity index for the DRGs included in that market, where the weights are given by the share of the total production of that market concerning for each sub-product (that is, each DRG).[11]

The average values of *tech* , C_1, *h* and *size* for the relevant markets classified by products are given in Table 2. The values of *tech* shows limited variation across products (coefficient of variation equal to 0.094) in comparison with the other indices reported, especially in the case of *h* and C_1 (0.535 and 0.506 respectively). Since, as discussed in Section 2.2, Sutton’s theory provides different empirical predictions depending on whether the market considered is characterized by high or low technological intensity it is interesting to evaluate the relationship between the variation of the index *tech* and that of some of the other characteristics of the relevant markets presented in Table 1. In particular, the technological intensity of the market turns out to be significantly correlated with

---

[9] More precisely, for the *i*th DRG and the *j*th hospital ward, capital intensity index t_i is derived as:

$$t_i = \sum_j \frac{k_j}{c_j} \frac{s_{ij}}{\sum_j s_{ij}} * w_i$$

where k, c, s and w denote capital costs, total costs, the number of services (measured in terms of hospitalization days) and the weight index respectively.

[10] Higher research intensity might require more medical staffing (Blank and Vogelaar, 2004). To consider this we have also tried to approximate the R&D intensity of the DRGs by defining a ‘physician intensity index.’ This is computed, analogously to the ‘capital intensity index’ with reference to the ratio between the costs for physicians and the total cost. The distribution of the values of the physician intensity index across different DRGs is similar to the capital intensity index one (the correlation between the two is 0.98). Therefore, resorting to the physician intensity index instead of the capital intensity index seems not to affect the results of our analysis significantly.

[11] We are aware that the index adopted here (the share of capital costs) is not an optimal proxy for the R&D intensity since it neglects relevant components of R&D, such as that provided by hospitals as units and by university hospitals. However we stress that, as stated in Sutton (1991), the R&D intensity is considered only in order to partition the markets into two groups. Thus, possible measurement errors in this variable are not expected to affect the results of the analysis.

both market size and variation of market size within the corresponding product category, with multi-Province composition and with the probability of non-contiguous Provinces being in the same market.

2.5. Empirical results

The consistency of the empirical predictions by Sutton with the market structure of the Italian NHS is tested here on the basis of the relevant markets as identified and characterized in Section 2.4. Given that Sutton makes different predictions according to the technological intensity of the considered markets, we first partition the sample into two different groups, one with a relatively high level of technological intensity (high-*tech* markets) and a control group in which the technological intensity is relatively low (low-*tech* markets). A rather crude way to proceed is to order the markets by the *tech* index, and to choose as cut-off the average (0.062).[12] To distinguish the two groups more sharply, we exclude from each group the 30% of the total number of markets whose *tech* value is closest to the cut-off level. As a result, the high-*tech* and the low-*tech* group include 592 and 406 markets respectively.

The partitioning of the sample into the high-*tech* and low-*tech* groups makes it possible to stress some preliminary descriptive statistics. In particular, high-*tech* markets are characterized on average by lower values of h (0.284 and 0.442 respectively), higher levels of C_1 (0.508 and 0.447 respectively) and *size* (18.17 and 17.35 respectively).

2.5.1. Market concentration and market size

The relationship between C_1 and *size* in each market is given in Figure 7 and 8. Our concern is comparing these diagrams with the predicted lower bounds from Figure 1 and Figure 2, which reflect *Predictions 1* and *2* respectively. Following Sutton, we adopt Smith's (1985, 1994) use of a two-step procedure to estimate the lower bound for the scatters of observations shown in Figure 7 and Figure 8. The first step of the procedure requires making some *a priori* decision about the form of the schedule $f(z)$ that describes the lower bound. We can then obtain a

[12] Notice that our choice of the cut-off point is to some extent arbitrary and that other criteria could be adopted to define it.

consistent estimator of the actual schedule by minimising the sum of $y_i - f(z_i)$ subject to the constraint that all residuals shall be non-negative. The second step involves employing the method of pseudo-maximum likelihood to check that the pattern of the estimated residual fits a Weibull distribution.[13] [14]

The three-parameter Weibull distribution is defined on the domain $t \geq \mu$ by:

$$\text{Prob}(T \leq t) = 1 - \exp\left\{ -\left[\frac{t-\mu}{\delta} \right]^{\beta} \right\}$$

where $\beta > 0$ and $\delta > 0$. The three parameters (β, δ, μ) denote shape, scale and location respectively. The location parameter μ represents the lower bound to the support of the distribution. If $\mu = 0$ the distribution simplifies to the so called two-parameter Weibull.

The procedure rests on the assumption that the distribution of residuals is identical at all values of the independent variable. This assumption would be unrealistic if applied directly to the values assumed by C_1, considering the upper limit of C_1 is equal to 1. For this reason we take the logit transformation of C_1:

$$C_1^* = \ln\left(\frac{C_1}{1-C_1} \right)$$

The first step of the estimate procedure implies making some assumptions about the form of the schedule describing the lower bound and estimating the parameter of the schedule. A reasonable family of candidate schedules would be:[15]

---

[13] We have chosen to use a Weibull distribution as the distribution generating the level of concentration in the markets because it is a very flexible functional form. It can assume different shapes and different position in the plane according to the values given to its parameters. For this reason it has been widely used in fitting bounds to various empirical distributions (references).

[14] As underlined by Giorgetti (2003) and Robinson and Chiang (1996), the presence of outliers in the data could influence the estimate of lower bound functions. Other estimation techniques, such as Koenker et al. (1994) are more robust to the presence of outliers. However, following the approach of Robinson and Chiang (1996), we adopt Smith`s lower bound estimation method because this is the method used by Sutton to estimate lower bound functions and the adoption of this method allows us to avoid possible inconsistencies that may arise using a different estimation technique. Accordingly we can compare our results to those of Sutton.

[15] Considering the relationship between market structure and market size, Sutton (1991) states that, in markets with endogenous sunk costs, the lower bound to market concentration not only does not converge to zero as the ratio between market size and set up cost increases, but also does not necessarily follow a decreasing (or even monotonic) pattern. Insofar as product differentiation

$$C_1^* = a + \left(\frac{b}{\ln size}\right)$$

Figure 7 and Figure 8 show the estimated schedules describing the lower bound in the high-*tech* markets (the dotted line) and in the low-*tech* markets (the thick line). As shown in Table 3, the estimated parameter a is equal to –3.773 and –6.161 for the high-and low-*tech* markets respectively. The estimated values for the parameter b result in a higher slope for the lower bound for low-*tech* markets (6.418) than for high-*tech* markets (3.744). Moreover, the parameters a and b are significant at a confidence level higher than 99% for both the two groups of markets. The standard errors of a and b are obtained with a Monte Carlo simulation of 10,000 repetitions, and support the generation of error terms from a Weibull distribution characterized by the estimated parameters. The estimated schedules in the high-*tech* and low-*tech* markets give support to *Prediction 1* and *2* presented in Section 2.2.

As second step of this procedure, we check whether the associated set of residuals fits a Weibull distribution. Assuming that the set of residuals is distributed as a two-parameter Weibull, we show that residuals fit that distribution well for both high-*tech* and low-*tech* markets.[16,17] To test the hypothesis that the estimated schedules for both groups of markets converge to the same value as the market size goes to infinity, we examine the values assumed by C_1^* distinctively for high-*tech* and low-*tech* markets when $size$ goes to infinity. That value equals –3.773 (which corresponds to C_1 = 0.026) and –6.418 (C_1 = 0.002) for high-*tech* and low-*tech* markets respectively. These results are consistent again with *Predictions 1* and *2*.[18]

---

renders advertising or research costs more effective in increasing the consumer willingness to pay, the outspending in advertising or R&S could even raise the concentration levels. In the present study we have not considered this further prediction but this might be a suitable topic for further study.

[16] The assumption that the residuals follow a two-parameter Weibull is made even by previous studies in the literature (Giogetti (2003), Marin and Siotis (2002)).

[17] We define $R_i = t_i - \mu$, given a reasonable value for μ, and rank R_i in ascending order. Then we define the cumulative distribution $F(R_i)$. If the R_i is defined as having a two-parameter Weibull distribution, a plot of $y_i = \ln(1/(1 - F(R_i)))$ against $\ln R_i$ yields a straight line. Our analysis shows that the previous relationship is supported by the data (for simplicity, we omit the graphics) the set of residuals is well approximated by a Weibull.

[18] The estimated $\beta < 2$ is a known indicator of the appropriateness of Smith's procedure. The maximum likelihood is a more common method than Smith's procedure to estimate lower bounds,

2.5.2. Market concentration and market homogeneity

The relationship between C_1 and h is shown in Figure 9 and 10 for the high-*tech* and low-*tech* markets respectively. We compare these diagrams with the predicted lower bounds shown in Figure 3 and Figure 4, which represent graphically *Predictions 3* and *4*. The data from the Italian NHS does not appear to contrast those theoretical predictions at face value. The high-*tech* group does not include observations where high values of h are coupled with low values of C_1, whereas some low-*tech* markets are characterized by both low product homogeneity and low concentration. This encourages a formal statistical procedure in order to test Sutton's predictions. Sutton's theory makes no general assumptions as to the functional form of the lower bound, however in the limiting case where all sub-markets are completely independent, the theoretical framework states that the bound for the high-*tech* group takes the form of a ray through the origin.[19]

For the same matters discussed in Section 2.5.1, before running the statistical test we apply to the C_1 index the transformation:

$$C_1^* = \ln\left(\frac{1}{1-C_1}\right)$$

As far as the form of the schedule describing the lower bound, a reasonable family of candidate schedules would be:

$$C_1^* = a * h$$

When Smith's procedure is applied, the estimated parameter a is equal to 0.177 and 0.069 respectively for high-*tech* and low-*tech* markets (as shown in Table 3). Therefore, the slope of the lower bound for the low-*tech* markets, even if not exactly equal to zero, is very close to zero and is much lower than the slope of the lower bound for the high-*tech* case. The estimate of parameter a is significant

---

but when $\beta < 2$ its usual asymptotic results do not hold (Smith 1985, 1989). In our case the estimated value of the parameter β is less than 2 for both low-*tech* and high-*tech* markets (equal respectively to 1.349 and 1.795) and thus it is appropriate to use the Smith's approach.

[19] These predictions have been tested previously using a procedure due to Mann, Scheuer and Fertig (1973). This method is useful because it bypasses the estimation of the shape and scale parameters of the distribution and constructs directly a confidence interval for the lower bound. Thus it involves less computation than Smith's approach, however it requires dealing with samples of dimensions up to 100 observations, far smaller than ours. For this reason we continue with Smith's procedure.

at a confidence level higher than 99% for both groups of markets. Thus the relationship between the estimated schedules in the high-*tech* and low-*tech* markets gives support to *Predictions 3* and *4*.

As for the relationship between C_1 and $size$, considering the relationship between C_1 and h we assume that the residuals are distributed as a two-parameter Weibull and, following the procedure described in Section 2.5.1, we show that they fit that distribution both in high-*tech* and in low-*tech* markets case.[20]

2.6. Final remarks

In this paper we have tested empirically the existence of the relationship between technological profiles and market structure claimed by Sutton's theory in a specific economic framework, that of health care services provided by the Italian NHS.

In order to test the empirical predictions by Sutton, we have identified the relevant markets for hospital care services in Italy in terms of both product and geographic dimensions. In particular, the Elzinga and Hogarty approach has been applied to data on patients' flows across Italian Provinces in order to derive the geographic dimension of each market. To our knowledge, this is the first time this approach has been utilized with reference to the Italian hospital treatments sector.

Our results support the empirical predictions by Sutton, that in markets where technological intensity is low the lower bound to market concentration converges monotonically to zero when the market size increases, for any level of product homogeneity. Conversely, in markets where technological intensity is high the lower bound to concentration converges to a positive value different from zero when the market size increases, while the lower bound increases from zero with the level of product homogeneity.

These results provide some useful indications to the policy–makers about the functioning of the Italian NHS. In order to enhance the relevance of this evidence however, the institutional setting and the regulation constraints affecting the hospital treatment sector in Italy should be investigated more thoroughly.

[20] Even in this case, the estimated value of the parameter β is less then 2 for both the categories of markets (respectively equal to 0.844 and 0.952), proving the suitability of the Smith's methodology.

Moreover, the structure of hospital care markets identified here through the Elzinga and Hogarty approach could be further exploited in order to examine the potential effect on hospital technical efficiency.

Annex 1. The application of Elzinga and Hogarty procedure

The application of the Elzinga and Hogarty procedure includes a number of different steps. First of all, for each product (that is, for each MDC as discussed in the main text) the Province providing the largest share of hospital care treatments is taken as the starting point in the identification of the geographic market. We are aware that any procedure of selecting a starting point is somewhat arbitrary. Indeed, Elzinga and Hogarty (1978) recognize that the role of subjective judgment is relevant in postulating the "potential" market area, especially when the anti-competitive strategies carried out after a merger is not a central concern.

Then a "hypothetical market" is defined as the area formed by the smallest number of Provinces whose residents consume, in the aggregate, at least 75% of the hospital care services produced in the Province assumed as the starting point. We adopt here the 75% threshold (and not a more binding condition, such as 90%) to reduce the risk, as stressed by Elzinga and Hogarty (1978), that markets overlap and are not independent. This would not be acceptable for our analysis since Sutton's framework requires dealing with independent markets.

If at least 75% of the services consumed within that hypothetical market are produced by the Provinces that make up that market, the LIFO test is met, and then the procedure moves to the next step, the LOFI test. If it is not the case, it is necessary to enlarge the hypothetical market area by including one (or more) additional Provinces up to the LIFO test is satisfied. The LOFI test requires that at least 75% of the services produced in the market is consumed within that area. If the hypothetical market area satisfies even the LOFI test, then that area is assumed to be a relevant market for that product. Otherwise, analogously to the approach followed before, additional Provinces join the original hypothetical market till the LOFI test is passed. The LIFO test is then applied again on the new area, and so on. Once a relevant market is identified, another round of the Elzinga and Hogarty procedure just described is applied to the remaining Provinces not included in that market, until all the Provinces are allocated in the corresponding relevant markets.

Figure 1. Lower bound for concentration in low R&D intensity industries

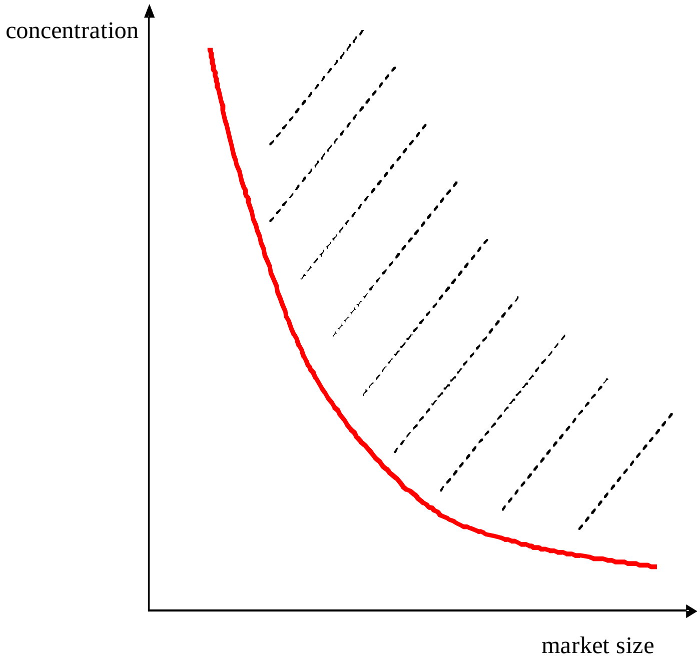


Figure 2. Lower bound for concentration in high R&D intensity industries

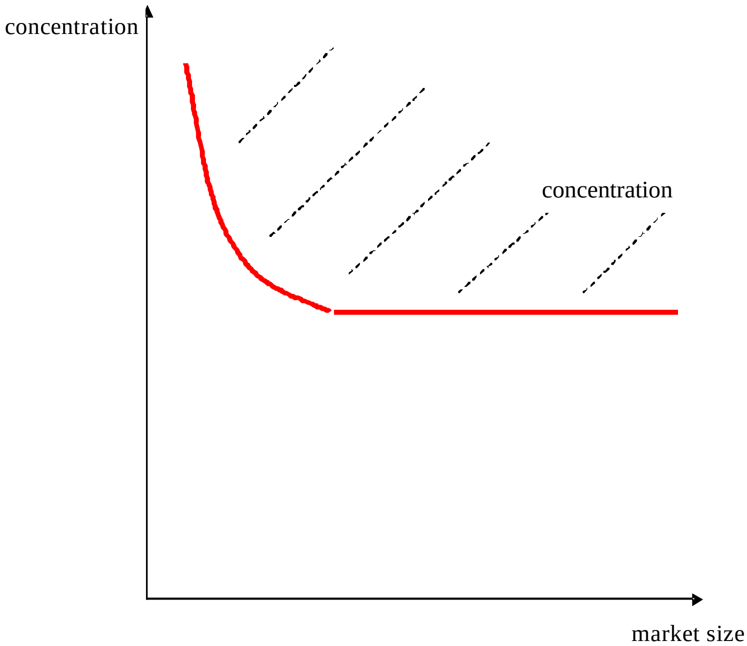


Figure 3. Lower bound for concentration in low R&D intensity industries

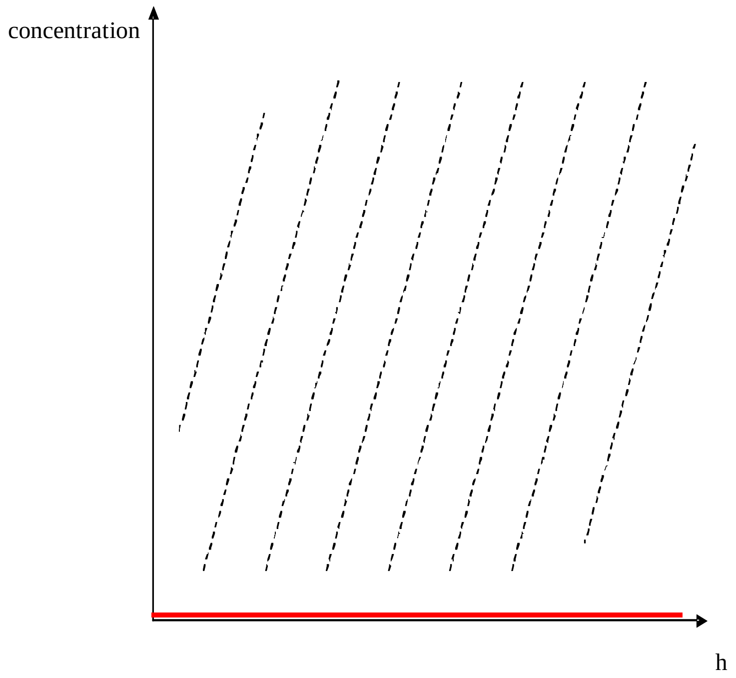


Figure 4. Lower bound for concentration in high R&D intensity industries

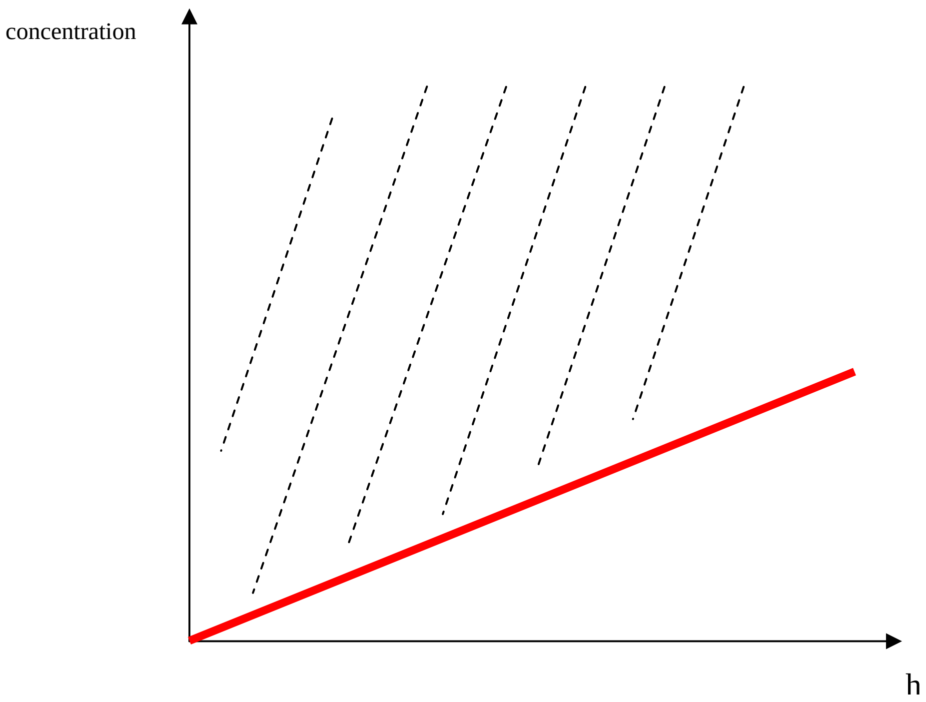
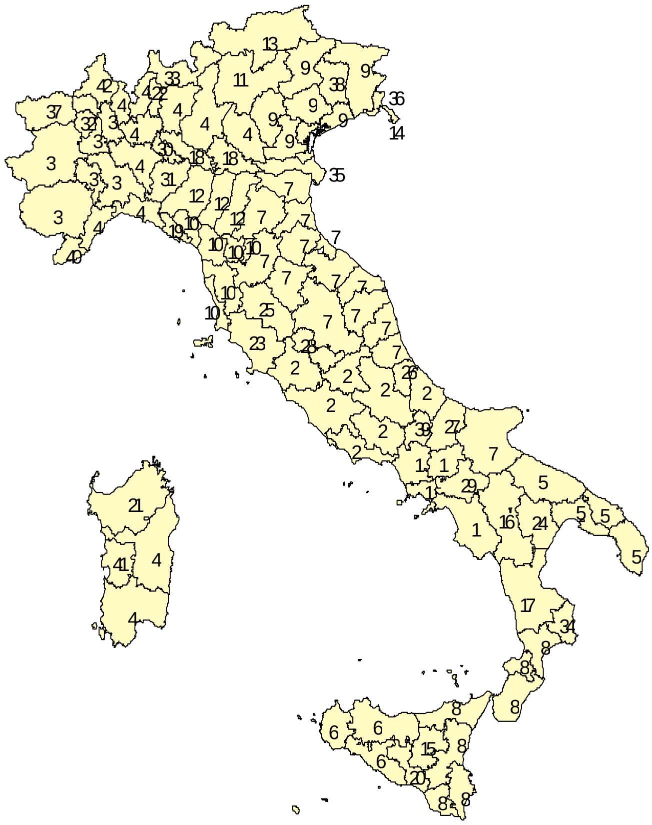
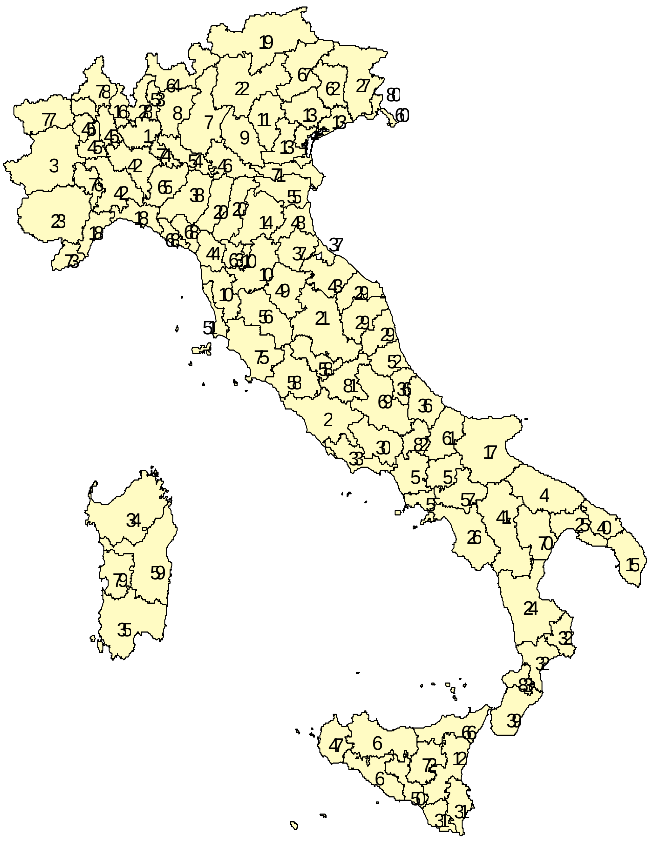


Figure 5. The relevant markets for "Burns"



Note: the reported numbers denote the inclusion of Provinces in the relevant markets.

Figure 6. The relevant markets for "Neonatal Period Diseases"



Note: the reported numbers denote the inclusion of Provinces in the relevant markets.

Figure7: Market concentration and size in high-tech markets

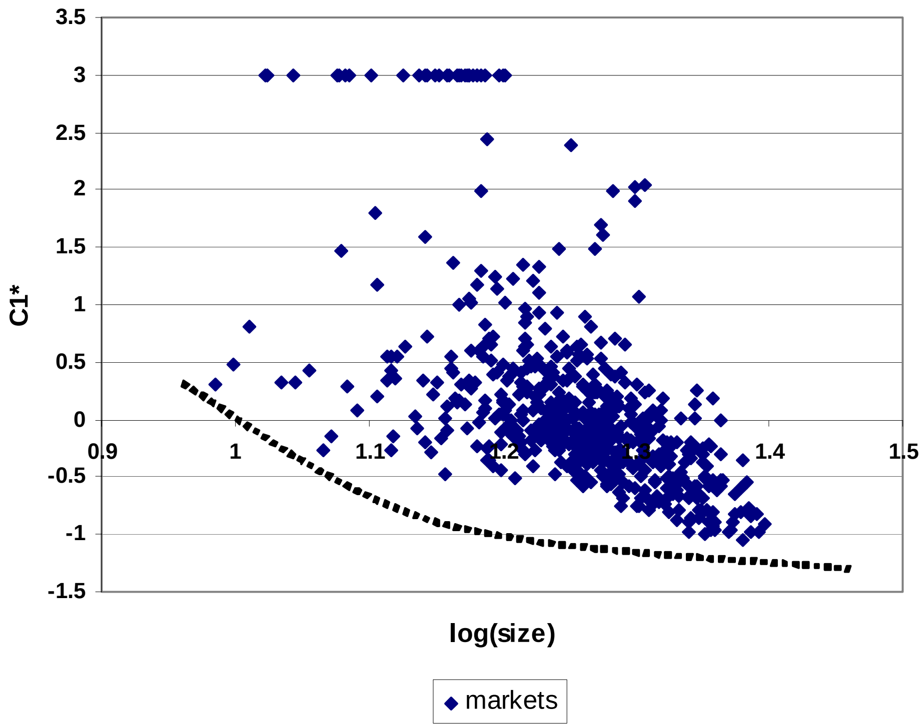


Figure 8: Market concentration and size in low-tech markets

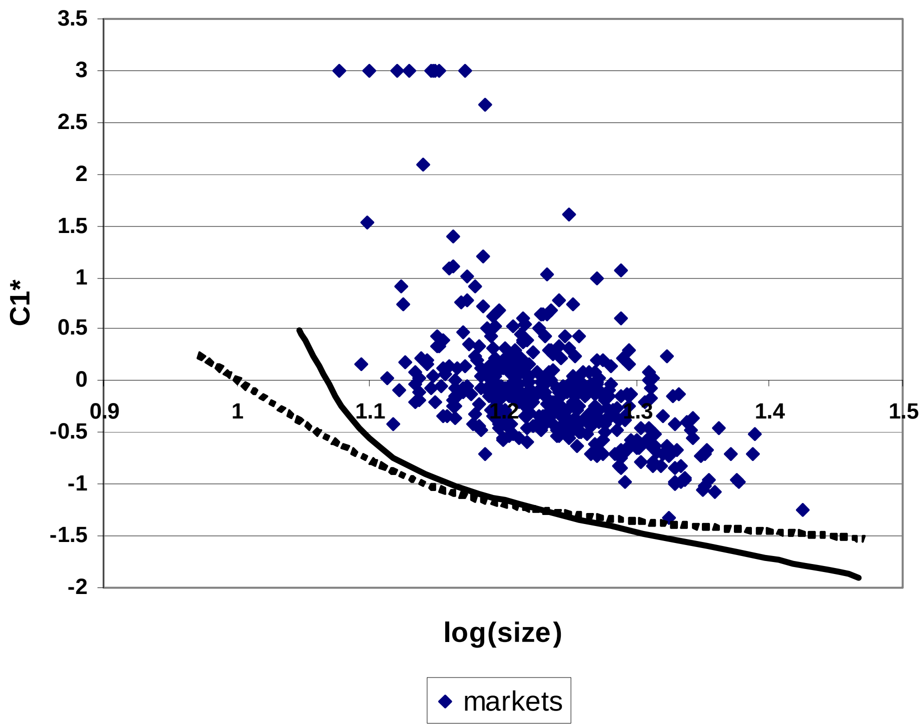


Figure 9: Market concentration and product homogeneity in high-tech markets

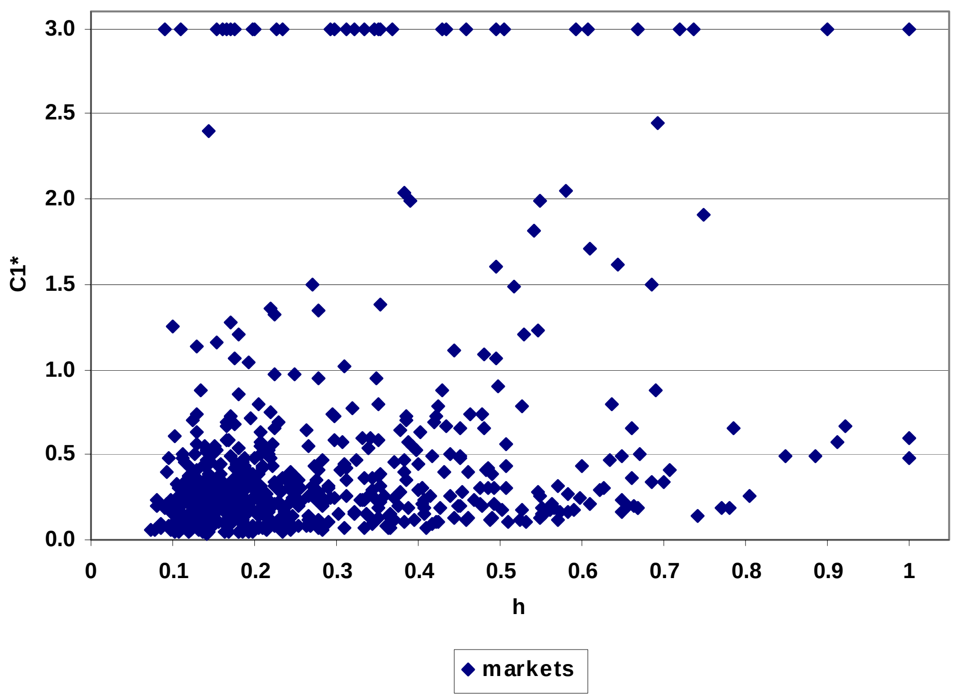


Figure 10: Market concentration and product homogeneity in low-tech markets

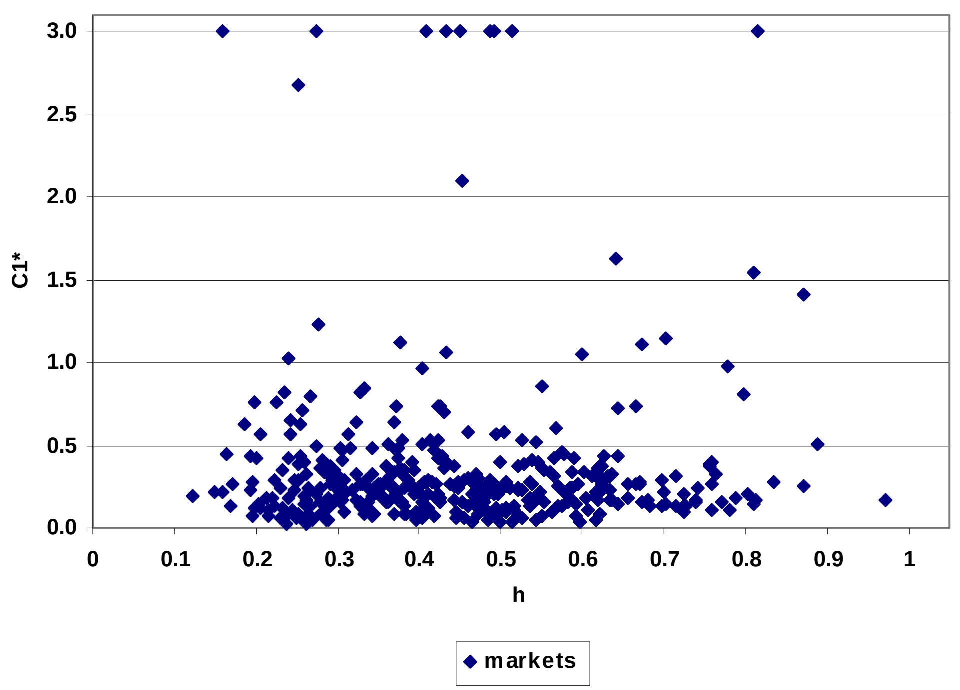


Table 1: relevant market description

Product number	Product	number of services #	number of services %	number of DRGs #	number of DRGs %	number of hospitals #	number of hospitals %	size average	size coeff variation	multiprov	discontinuity
1	Nervous system diseases	643,878	7.63	35	7.29	817	99.63	977,625	2.02	27.1%	0
2	Eye diseases	312,153	3.70	13	2.71	760	92.68	1,253,911	1.55	37.0%	2
3	Ear, nose, mouth and throat diseases	436,194	5.17	31	6.46	806	98.29	961,332	1.17	38.3%	1
4	Respiratory apparatus diseases	583,960	6.92	29	6.04	810	98.78	769,065	1.29	24.0%	0
5	Cardio-vascular apparatus diseases	1,097,901	13.01	44	9.17	813	99.15	812,393	1.39	25.4%	0
6	Digestive system apparatus	888,333	10.53	40	8.33	809	98.66	720,999	1.29	21.3%	0
7	Hepatobiliary and pancreatic diseases	331,566	3.93	18	3.75	797	97.20	1,011,928	1.69	29.8%	5
8	Muscle-skeletal system diseases and disesas of connective tissue	707,931	8.39	40	8.33	811	98.90	1,281,776	1.53	37.8%	1
9	Skin, subcutaneous tissue and breast diseases	583,268	6.91	38	7.92	809	98.66	961,332	1.32	30.0%	1
10	Endocrine, nutritional and metabolic diseases	172,431	2.04	17	3.54	808	98.54	1,441,997	1.67	42.5%	5
11	Kidney and urinary system diseases	385,963	4.57	32	6.67	798	97.32	1,068,146	1.61	31.5%	4
12	Male reproductive system diseases	146,877	1.74	19	3.96	778	94.88	994,481	1.28	29.3%	1
13	Female reproductive system diseases	277,077	3.28	17	3.54	779	95.00	1,088,300	1.35	41.5%	1
14	Gestation and birth	680,852	8.07	15	3.13	652	79.51	769,065	1.13	25.3%	0
15	Neonatal period diseases	412,222	4.89	7	1.46	648	79.02	694,939	1.11	18.1%	0
16	Diseases of the blood, of the haemopoietic organs and of the immune system	80,442	0.95	8	1.67	801	97.68	915,554	1.55	27.0%	5
17	Myelo-proliferative diseases	198,833	2.36	17	3.54	802	97.80	1,558,916	2.29	40.5%	5
18	Infectious diseases	70,018	0.83	9	1.88	792	96.59	1,130,978	1.10	33.3%	0
19	Mental diseases	179,980	2.13	9	1.88	803	97.93	873,938	1.38	24.2%	3
20	Alcohol/drugs abuse and induced organic mental diseases	32,289	0.38	5	1.04	756	92.20	860,894	1.52	23.9%	1
21	Traumatisms, poisonings and tossic effects of medicines	96,211	1.14	17	3.54	801	97.68	994,481	1.43	31.0%	2
22	Burns	7,268	0.09	6	1.25	616	75.12	1,373,331	1.66	28.6%	0
23	Factors affecting health and health services demand	83,820	0.99	7	1.46	798	97.32	1,696,468	1.04	52.9%	5
24	Serious multiple traumatisms	9,401	0.11	4	0.83	601	73.29	1,088,300	1.49	41.5%	1
25	HIV infections	18,035	0.21	3	0.63	360	43.90	1,478,798	0.98	52.6%	0
	Total	8,436,903	100.00	480	100.00	820	100.00	1,071,158	1.43	32.6%	1.72

Table 2: Technological intensity, product homogeneity, concentration and size of the markets

Product number	Product	*tech*	*h*	C_1	*size*
1	Nervous system diseases	0.0474	0.4666	0.4597	17.9059
2	Eye diseases	0.0519	0.6654	0.4247	15.5411
3	Ear, nose, mouth and throat diseases	0.0549	0.3639	0.4800	17.3861
4	Respiratory apparatus diseases	0.0576	0.2716	0.5387	18.6868
5	Cardio-vascular apparatus diseases	0.0599	0.5115	0.4259	17.1099
6	Digestive system apparatus	0.0602	0.3383	0.4072	17.9716
7	Hepatobiliary and pancreatic diseases	0.0605	0.3323	0.3878	17.8155
8	Muscle-skeletal system diseases and disesas of connective tissue	0.0605	0.6164	0.6059	17.2347
9	Skin, subcutaneous tissue and breast diseases	0.0606	0.2067	0.4516	18.9330
10	Endocrine, nutritional and metabolic diseases	0.0618	0.2193	0.4279	16.8962
11	Kidney and urinary system diseases	0.0623	0.2051	0.4338	18.4210
12	Male reproductive system diseases	0.0625	0.1522	0.4132	18.8856
13	Female reproductive system diseases	0.0625	0.3954	0.3881	18.1526
14	Gestation and birth	0.0628	0.3733	0.3828	17.9747
15	Neonatal period diseases	0.0634	0.1810	0.4550	18.9333
16	Diseases of the blood, of the haemopoietic organs and of the immune system	0.0636	0.1262	0.4419	18.4780
17	Myelo-proliferative diseases	0.0639	0.2623	0.5014	19.0146
18	Infectious diseases	0.0641	0.1551	0.4782	19.7165
19	Mental diseases	0.0642	0.2530	0.4122	19.3412
20	Alcool/drus abuse and induced organic mental diseases	0.0648	0.4865	0.3724	18.0296
21	Traumatisms, poisonings and tossic effects of medicines	0.0656	0.6366	0.6326	14.6717
22	Burns	0.0664	0.4236	0.5137	16.0294
23	Factors affecting health and health services demand	0.0690	0.1840	0.4927	18.9380
24	Serious multiple traumatisms	0.0713	0.2021	0.4751	17.8962
25	HIV infections	0.0744	0.5396	0.5706	20.2641
	average value	0.0617	0.3345	0.4584	17.9151

Table 3: Estimated parameters (standard errors in parenthesis)

	size		*homogeneity*
parameter	*a*	*b*	*a*
High-*tech*	-3.773	3.744	0.178
	(0.200)	*(0.251)*	*(0.007)*
Low-*tech*	-6.161	6.418	0.069
	(0.558)	*(0.688)*	*(0.015)*

Chapter 3

The Geography of Hospital Admission in a National Health Service with Patient Choice: Evidence from Italy

3.1. Introduction

In the Italian NHS patients are enrolled into health plans managed by Local Health Authorities (LHAs). Enrolment is based on a patient's place of residence, while financing accrues from general taxation according to a capitation payment per enrolee. With the resources available the LHAs are responsible for the health care consumption of the enrolees. A distinctive feature of the Italian NHS is that within this institutional framework (similar to many other "decentralized" tax-funded NHS systems, for example, Spain, Norway, Denmark, and the UK) hospital care is subject to the patient's unconstrained choice of provider. Patients are entitled to completely free hospital treatments and interregional flows are compensated for according to centrally set prices. Each year, it is estimated that 35% of the 10 million hospital admissions in Italy take place outside the patients' LHA of residence. This figure goes up to almost 42% for cancer treatment and more than 58% for high surgery.

This situation raises policy concerns. Equity of access and financial sustainability are the main issues at stake. Exit rates and average distance travelled to access hospital care are, for instance, much larger for enrolees in southern LHAs. Observed imbalances in patient mobility make the distribution of private mobility costs uneven and promote the accumulation of financial resources towards the already better endowed LHAs. In this paper we aim to evaluate the extent to which the observed imbalances are due to scale effects, depend on a core/periphery equilibrium, or reflect a deeper, long lasting north/south divide. In particular, we focus on the scale effect played by the size of the pool of enrolees.

25% of Italian LHAs have less than 150,000 enrolees, while 20% have more than 400,000 enrolees. Since funds accrue on a capitation basis, and smaller LHAs suffer larger outflows while receiving smaller inflows, this policy variable is crucial in determining LHA financial stability.

We work on an origin/destination matrix provided by the Italian Ministry of Health, comprising of all ordinary admissions to public hospitals for the year 2001. We classify admissions into 4 product groups. To control for distance, contiguity, and supply characteristics, and to net out the effects played by the composition of patients' flows, we estimate a gravity equation for the full matrix of pair-wise flows. We estimate gravity equations in multiplicative form using a Poisson pseudo maximum likelihood method, as proposed by Santos-Silva and Tenreyro (2006). This method is robust to different patterns of heteroskedasticity and provides a natural way to deal with the zero flows.

Our results suggest that the gravity model is a good framework for explaining the patient mobility phenomenon for complex surgery and cancer. Indeed, with reference to the main features of this model, the number of enrolees of the LHAs of origin and destination affects patient flows in a positive way, while the distance between LHAs affects patient flows in a negative way.

3.2. Institutional Background and Motivation

Patient mobility causes concern, particularly in relation to the reforms that occurred in the Italian health care sector in the late '90s. The National Health Service (otherwise known as SSN - Servizio Sanitario Nazionale) was created in 1978 as a regionally based system providing universal coverage free of charge at the point of service. The central government was responsible for determining the amount of resources to devote to health care, for general planning, and for funding Regions through general taxation and compulsory health contributions.[1] The 20 Regional authorities were responsible for local planning according to the objectives specified by the central government, and for allocating resources to the third level of the system, the Local Health Units. These were operational agencies

[1] Depending on a citizen's income, age and health condition, co-payments are also charged for drugs, out-patient treatment, some diagnostic and laboratory tests, and medical appliances.

in charge of providing services to patients through their own facilities or through contracts with private providers. They provide a wide range of services, at hospital and community level, in geographical areas with populations of about 300,000. The LHAs do not have revenues collection responsibilities, but they are funded by the Regions through a capitation system (France et al., 2005)

The Italian National Health System is quite fragmented in terms of organization of the regional services. At one end of the spectrum there is the “LHA-centred model”, where the LHAs have substantial freedom in negotiating service agreements with public and accredited private providers. At the other end of the spectrum there is the “Region-centred model”, where the regions exercise a purchaser role and fund the providers directly, while the LHAs have little organizational freedom and act mainly as providers (France et al., 2005)

The SSN guaranteed the provision of hospital treatments at a given level of quality and free of charge. On the supply side, the SSN largely relies on public production supplemented by privately licensed hospitals. Public hospitals are run by LHAs or by autonomous public trusts (Aziende Ospedaliere, Policlinici and Istituti di Ricovero e cura a carattere scientifico (IRCCS)). Privately licensed hospitals can treat patients within the SSN, i.e. free of charge, and are refunded afterwards by the LHA to which the patient is enrolled. Patients are completely free to choose their hospital; it may be publicly or privately licensed, and within or outside the LHA where they are enrolled. Since patients are totally unaware of treatment costs and can choose freely between publicly financed hospitals, choice is essentially determined by distance from home, hospital specialization, waiting lists, and perceived quality.

The functioning of such a decentralized public health care service has recently been reformed through the approval of Legislative Decrees 502/1992 and 517/1993 with the purported aim of introducing some elements of internal market competition. Significant managerial autonomy has been devolved to larger hospitals and LHAs, and, at the same time, a partial split between health care production and purchasing has been introduced. Competition has been promoted by the introduction of prospective payment for hospital admission through the Diagnosis Related Groups (DRGs) classification scheme. Since the year 2000, public hospital treatments have been priced through the DRG scheme according to fixed prices set at national level. Regions sets their tariffs referring to national tariffs rates, which represents a ceiling and allow flexibility downward (so far they have been reduced by up to – 30% of national tariffs) (France et al. 2005). The Legislative Decrees have also introduced some elements of regional federalism, which were strengthened by Legislative Decree 446/1997 (introducing sources of autonomous financing for the regions) and Decree 56/2000 (stating that the funding of the NHS is mainly the responsibility of the Regions).[2] [3] Regions can choose how many resources to spend in the health care sector, subject to some constraints. They have to guarantee a minimum level of health care (*livelli essenziali di assistnza* – LEAs) but they are free to choose the quality level and the amount of services to provide.[4] This implies that the quality and quantity of health care might vary across Regions. Regions are free to provide non-LEA services, but they must be financed with resources raised autonomously. Since the patients' right to refer themselves to any hospital has not been limited so far, in the emerging environment publicly financed hospitals, including those run by the LHAs, are strongly motivated to invest in quality and to establish a good reputation in order to attract patients and to generate a stable cash flow.

We would expect patient mobility to cause expenditure uncertainty for the LHAs' planners, at least in the short term, since it is hard to predict. Regions providing treatments to patients who are resident in other Regions receive financial compensation for the treatment offered. National tariffs, which are DRG-based, are used to fund interregional patient flows. It is likely that Regions providing fewer and lower quality health care services have to pay the regions providing more services (which are probably the richer ones) because of the treatment received by mobile patients. This mechanism could create deep and long lasting imbalances across regions and bring rationing of health services to

[2] A portion on national income taxes (the IRPEF) was transferred to regions (the regional IRPEF), and health insurance contributions were replaced by regionally collected taxes on the value added by companies and on the salaries of public sector employees (the IRAP)

[3] The National Solidarity Fund has been introduced in order to allow fiscal equalization across regions; it is financed by funds coming from central government.

[4] The bunch of LEAs was defined on the basis of effectiveness, appropriateness and efficiency criteria.

patients living in the poorest regions. For these reasons, we believe that it is important to analyse the determinants of patient mobility.

Patient mobility is partly unavoidable. It would be inefficient to provide more specialized or very rare treatments at every hospital. On the contrary, it is efficient, to some extent, to concentrate their production in a few centres. Apart from these treatments, it makes sense to consider the decision to move as the manifestation of dissatisfaction for the local supply of health care, as suggested by Tiebout's "voting with the feet" mechanism. In this paper, we are not going to formulate considerations about the welfare consequences of patient mobility, but we aim to provide some insights into the determinants of this behaviour in order to let health care planners make more informed policy decisions.

3.3. The Geography of Hospital Admission in the Italian NHS

3.3.1. Mobility flows data

We analyse patient mobility across Italian LHAs using data on hospital admissions that occurred in all public hospitals during the year 2001. Patient flows are reported in an origin/destination matrix provided by the Italian Ministry of Health. For each single DRG the matrix reports the corresponding flow of patients occurring between each pair of origin-destination LHAs, where the origin refers to the LHA where the patients are enrolled, and the destination refers to the LHA where the patients receive the hospital treatment.

In the reference year, the Italian territory is partitioned into 197 LHAs. However, only 190 LHAs are present in our dataset due to the fact that we were forced to aggregate those LHAs operating in a single municipality.[5] Moreover, we disregarded all the flows generated and directed to the LHAs of Sardinia and Sicily because these two regions are islands and the patient flow to and from them may follow peculiar patterns. We end up with a set of 173 LHAs.

To provide a comprehensive and significant analysis of mobility patterns in hospital admission, while maintaining manageability, we reduced the dimension

[5] The municipalities of Turin and Rome are articulated into 4 and 5 LHAs, respectively.

of the matrix by aggregating the DRGs into 7 broad groups (PRODUCT). In order to account for different patient severity we considered the following groups of clinical procedures and/or conditions: complex surgery (CS), emergencies (EM), cancer (CA), HIV, delivery (DE), basic surgery (BS) and basic medicine (BM). Table A1 of the appendix details the aggregation, reporting the sub-product constituting each product. It also reports the share of hospital treatments occurring in each main category at national level. The share of these categories varies significantly, from the largest categories, BM (47.6%) and BS (23.6%), to the smallest ones, HIV (0.5%) and complex surgery (1.7%).

In the following analysis, we disregard the HIV, DE and EM products. There are too few observations present in HIV to make the analysis reliable. For DE and EM, we expect patient flow to be affected by “occasional mobility”. This kind of mobility is not due to a deliberate decision by the patient to search for hospital treatment in a LHA different from the assisting one, but to the patient’s need for hospital admission when she/he happens to be far from home for reasons of holiday, working, or family links. In order to exclude flows potentially flawed by “occasional mobility” we also excluded the observation corresponding to “small flows” for the product categories included in the analysis: that is flows made by 1 patient for CS and CA and by 2 patients in the case of BM and BS. The resulting dataset is made up of 119,716 observations, referring to 173 LHAs.

3.3.2. Summary Indicators of Patient Mobility

Our preliminary analysis is based upon a set of summary indicators of patient flows across the LHAs. We consider the following indicators:

- Exit rate: the share of hospital admissions (in a certain product) received by LHA enrolees outside that LHA.
- Inflow rate: the share of hospital admissions (in a certain product) in a given LHA that are provided to non-enrolees.
- Accessibility: the weighted average of the distance covered by all the enrolees travelling outside a specific LHA of residence to receive hospital treatments in another LHA (in a certain product), where the weights are the percentage of enrolees from the LHA of origin travelling to each LHA

of destination. This represents travelling *outside* a LHA for receiving treatments. Note that this is a negative measure of accessibility, i.e. the higher its value the poorer the accessibility.

- Attractiveness: the weighted average of the distance travelled by enrolees outside the LHA of residence to receive hospital treatment in a specific LHA of destination (in a certain product), where the weights are the percentage of enrolees from each LHA of origin entering the LHA of destination. This represents travelling *into* a LHA for receiving treatments.

Table 1 reports some descriptive statistics about the overall summary indicators of the LHAs, and of the LHAs classified according to two dimensions: the presence of an autonomous hospital (Azienda Ospedaliera, Policlinico or IRCCS) and the dimension of the pool of enrolees. To this purpose, we grouped the LHAs into 4 classes, approximately close to the quartiles.[6] As expected, the exit rate from the LHAs with an autonomous hospital (13.8%) is lower than the overall rate (18.3%), while the exit rate from the LHAs without an autonomous hospital is higher (20.4%). Inversely, the inflow rate for the former LHAs (18.1 %) is higher than the overall rate (11.4%) and lower for the latter LHAs (8.2%). As expected, attractiveness is higher for the former LHAs (56.00) than for the latter (42.05), indicating that patients cover longer distances to access LHAs with an autonomous hospital. Accessibility is also higher for the former (90.15), than for the latter (55.94) but the interpretation is less straightforward. This may be explained by considering that there are fewer patients living in an LHA with an autonomous hospital that decide to exit and therefore those that do are more severe cases than those exiting from an LHA without an autonomous hospital. More severe patients usually seek appropriate care further. In relation to the dimension of the pool of enrolees, we observe that the LHAs with more enrolees have larger inflow rates (13.7% for more than 400,000) than the LHAs with fewer enrolees (9.8 % for less than 150,000), as expected. What is puzzling and

[6] For the computation of these summary indicators we have not considered the LHAs of Rome and Turin. However, since patient flows involving these LHAs are not negligible, for the estimate of the gravity model we consider them by introducing a dummy variable for Rome and another one for Turin.

unexpected is that the exit rates are also higher for the former (15% for more than 400,000) than for the latter (24%).

In Table 2 the summary indicators of patient flows are considered by grouping the flows according to the PRODUCT defined above. Results are reported at national level, and then broken down into the four Italian macro geographical areas. Considering Italy as a whole, the exit and inflow rates are very high for complex surgery (57.6% and 21.1% respectively) and cancer (36.0% and 16.7%) while they are lower for basic surgery (20.7% and 13.9%) and basic medicine (12.7% and 8,2%). This basic fact applies to all geographical areas. As expected, the exit rates from the South are always much higher than the average values. However, it is surprising that in the North-West the exit rates are also higher than the average for basic surgery and basic medicine. In contrast, the exit rates are always lower than average in the North-East and in the Centre (except for the complex surgery in the Centre). Another unexpected observation is that not only the exits but also the inflows are higher than the average in the South. In the Centre the inflows are always lower than the average, while within the North the results are mixed: inflows are higher than the average for complex surgery in the North-West and High and basic surgery in the North East, but lower otherwise. The other two summary measures, accessibility and attractiveness, are useful for interpreting these results.

At the national level, accessibility for complex surgery (109.7) and cancer (99.5) is almost double that for basic surgery (51.0) and basic medicine (54.0). Accessibility is below the average value for every macro product in the North West, the North East and the Centre. In particular, it is about half the average value in the North West and the East for complex surgery (51.7 and 53.1, respectively) and cancer (54.2 and 54.6, respectively). This measure is always above the average in the South, where it assumes very high values for all the products: 226.9 for complex surgery, 182.0 for cancer, 74.8 for basic surgery and 83.3 for basic medicine. This striking difference between the North-Centre and the South seems to suggest that patients in the North-Centre, if they exit their LHA, do not travel very long distances (particularly in the North), while patients in the South might do.

At the national level, attractiveness is similar for three of the products (41.4, 42.7 and 41.3 for complex surgery, basic surgery and medicine, respectively) while it is higher for cancer (60.8). For complex surgery and cancer, the North (particularly, the West) has much higher values of attractiveness than the South. Indeed, for complex surgery the values are 50.7 and 45.9 in the North West and the East respectively, and 94.6 and 55.8 for cancer respectively. In contrast, they are just 36.4 and 42.5 in the South. The Centre shows values higher that the average for basic surgery and medicine (48.3 and 45.0 respectively). Thus, the North, for complex surgery and cancer, and the Centre, for basic surgery and medicine, seem to be more “attractive” than the South.

Based on these results, the following conclusions could be drawn. Patients in the North and Centre do not exit much from their LHAs. Exceptions to this are patients in the North West for basic surgery and basic medicine, since their exit flows seem to be above the average, but they just cover short distances. In contrast, LHAs in the South are subject to both high levels of exit and inflows. Since the patients in the North and Centre are not travelling long distances, patient flows in the LHAs of the South are most likely to be coming from other LHAs in the South. Thus, there are two different types of patient mobility present in the South: short distance mobility, towards other LHAs in the South, and long distance mobility, towards LHAs in the North and Centre. The flows seem to be directed to the North for complex surgery and cancer and to the South for basic surgery and medicine.

In order to gain some additional insights into the factors behind the observed mobility patterns we rely upon the estimation of gravity models for patient flows. After a review of the most relevant literature related to gravity models and hospital choice, the remaining part of the paper details our empirical strategy.

3.4. Literature Review

3.4.1. Gravity Models

For our study we will adopt the framework of the gravity models. The social gravity model of spatial interaction has been developed as an analogy to the

Newtonian gravity model of Physics. In this model it is hypothesized that a greater level of spatial interaction should occur between two points in space the greater are the two population masses at those points and the lesser is the spatial distance between them (Stewart, 1948).

These models have been widely used in the field of Migration, where the baseline gravity model specification implies that the flow of residents between two areas is directly proportional to the population present in those areas and inversely related to the distance between the areas (Congdon, 1991). They have also been extensively used in the field of Trade, in order to explain the commercial flows between countries on the basis of their dimension (measured as GDP or population) and their distance (examples are given by Dalgin et al., 2004, and Rauch, 1999).

We can distinguish between the approaches based on micro or individual decision-making and those that deal with macro effects on aggregate flows. This has been clearly pointed out by Stillwell (2005) referring to the migration models literature. Micro approaches relate to the individual migrating unit (person, group or household) and involve the identification of those factors that influence this decision-making process. The factors having a bearing on these decisions include both the characteristics of individual persons (such as age, marital status, household status), or wider family units (such as family size and structure), and the wider characteristics of the potential destinations (such as regional relativities of unemployment, wages or house prices). In contrast, macro approaches relate to aggregate migration flows and investigate the relationships between migration and macro variables such as population size, unemployment rates or environmental conditions.

Another important distinction we have to consider is between the so called “push” and “pull” factors influencing the aggregate flows. The former are the elements of “propulsiveness” of an area (the characteristics of an area that make the residents want to exit that area) and the latter are the elements of “attractiveness” (the characteristics that make the area appealing). In this regard, three separate types of gravity models can be identified (Fotheringham and O`Kelly, 1989):

1. the “unconstrained” of “total flow constrained” models yield insights into interaction patterns by providing information on the attributes of both the origins and the destinations of the interactions;
2. the “production constrained” models only provide information on the destination characteristics, and thus do not consider the pushing factors of the origins;
3. the “attraction constrained” models only give insights into the origin characteristics, and thus do not pay attention to the pulling factors of the destinations.

Researchers could think that the unconstrained models are the most useful since they provide greater amounts of information. Unfortunately, as shown by Fotheringham and O`Kelly (1989), in the spatial interaction models there is often a trade-off between the quality and the quantity of information provided. Thus, according to the finalities of the study, attraction or production constrained models could be more appropriate.

3.4.2. Hospital Choice

We believe it is important to consider the past literature on hospital choice models in order to be aware of the main variables that can influence patient choice and how they influence it. This analysis allows us to form some expectations about the signs with which these variables will enter into our gravity equation.

One issue that has to be considered in modelling hospital choice is the role that is played by physicians. In many of the countries considered, the patient needs the request of the physician to be admitted to hospital, or may strictly follow the suggestions of the physician regarding which hospital to choose. Some authors have assumed that the physician acts as an agent for the patient and thus she is the key decision maker for hospital admission (Burns and Wholey, 1992; Luft et al., 1990; McGuirk and Porell, 1984). In most of the studies, however, the individual is assumed to play an active role in making their choice of hospital. Indeed, patients could choose physicians on the basis of the doctor’s privileges of admission (Dranove et al., 1992) and they may also directly express their preferences over hospitals to the physicians (Tay, 2003).

The variables that are considered most may be grouped under some categories: distance, hospital characteristics, area characteristics, and individual characteristics. In many of the earliest models of hospital choice the gravity model framework was adopted and the variables considered were mainly distance and hospital characteristics. Since McFadden`s (1974) pioneering work, conditional choice models have become very popular among researchers studying hospital choice. These models have the advantage that they allow for an estimation of effects of patients` and hospitals` attributes on hospital choice. More recently, most of the studies take into account not only the hospitals` and areas` variables but also the patients` individual characteristics.

Distance

Earlier studies on patients’ choice of hospital have already focused on distance to the hospital facilities, highlighting the “distance decay” pattern phenomenon: people tend to use the service of closer, over more distant, health providers so that the number of persons using particular providers declines at greater distance (Bashshur et al., 1971; Morill and Earickson, 1968; Morrill et al., 1970; Roghmann and Zastowny, 1979; Studnicki, 1975). In general, distance has been widely recognised as a powerful predictor of hospital choice (Basu and Friedman, 2001; Burns and Wholey, 1992; Dranove and Shanley, 1989; Goodman et al., 1997; Seniger, 1999; Tai et al., 2004; Tay, 2003)[7]. More recently, other variables, such as travel time and travel costs, have been used instead of distance in order to represent the difficulty for the patient in reaching the hospital. They have also been shown to negatively affect the probability of choosing the hospital (Bessho, 2003; McNamara, 2003; Varkevisser and van der Geest, 2006).

Hospital Characteristics

Between the various hospital characteristics the quality of the hospital is usually taken into consideration. Variables capturing hospital quality include both input and output measures. Between the more used input measures, there are number of nurses/doctors per bed (McNamara, 2003; Tay, 2003), the range of

---

[7] In studies where the physician is considered the decision maker, the distance of the physician to the hospital is often considered and it is found to negatively influence admission to the hospital (McGuirk and Porell, 1984; Burns and Wholey, 1992).

specialized services offered (Tai et al., 2004); and teaching status (Basu and Friedman, 2001; Burns and Wholey, 1992; Goodman et al.,1997; Tay, 2003), while between the more popular outcome measures, there are mortality rates (Burns and Wholey, 1992; Tay, 2003) and complications of the patients admitted to the hospital (Tay, 2003). There is strong evidence showing that patients tend to choose higher quality hospitals.

These variables are often hard to measure and if just considered individually could be misleading. As outlined by Tay (2003), the use of patient outcomes as a proxy for quality is complicated for two reasons. Firstly, outcome measures can be very noisy, especially if considered for hospitals with low patient volume. Secondly, a selection bias problem could arise: good quality hospitals could attract sicker patients, with the higher probability of having complications or dying, and thus they may report lower outcome performance. A way to address this problem is to adjust the outcome for the differences in the case-mix of hospitals. The use of the input measures as indicators of hospital quality can be problematic, as well. To some extent, indeed, these variables represent the amount of resources that are utilized and the “effort” a hospital is putting in, but this doesn’t automatically imply a good quality result. Thus, it seems very important to consider not just a single indicator, but many indicators simultaneously.

Hospital size (measured as the number of beds) is another common hospital level variable used in this literature (Goodman et al., 1997; McGuirk and Porell, 1984; McNamara, 2003; Tai et al., 2004; Tay, 2003; Varkevisser and van der Geest, 2006). It is often considered as a proxy of hospital quality and it positively affects the probability of choosing the hospital. As stressed by Varkevisser and van der Geest (2006), the use of this variable may raise a problem of endogeneity: is it the larger hospital size that increases the likelihood of selecting that hospital, or do the high selection rates lead to the larger hospital size?

Some recent studies have considered waiting times as a relevant hospital characteristic that may influence patient choice (Bessho, 2003; Varkevisser and van der Geest, 2006). Low levels of waiting time seem to strongly attract patients.

Area Characteristics

Some studies have tried to take into consideration how the characteristics of the area where people live could affect hospital choice. Some studies have considered metropolitan vs. non-metropolitan areas, and rural vs. urban (Basu and Friedman, 2001; Goodman et al., 1997; Varkevisser and van der Geest, 2006). The evidence from these studies is mixed. For example, in the study by Goodman et al. (1997) belonging to a rural area does not significantly affect the probability of referring to a hospital further away, while in the study by Basu and Friedman (2001) a patient’s residence in a rural county adjacent to a metropolitan area increased the likelihood that they would cross the county boundary. Other studies have considered median household income (Basu and Friedman, 2001; Goodman et al., 1997), showing that a high level of this variable increases the probability of referring to a hospital further away.

A critical issue in the hospital choice model regards the definition of the areas to consider as geographic entities. Arbitrary jurisdictional geographic boundaries (e.g., counties or zip code clusters) are often used in order to define the market in which the hospitals are operating, without checking the consistency of market area definitions adopted with the economic principles and/or the more accepted methods for deriving market definitions. Our choice of using the LHA as the geographical area of reference seems the most appropriate choice in relation to the Italian context. As mentioned above, indeed, the LHAs are the administrative entities with significant managerial autonomy at the lowest level of the organization of the Italian NHS.

Individual Characteristics

Among individual level characteristics, the most common ones included in the studies are gender, age, education, race, income, insurance status (Basu and Friedman, 2001; Bessho, 2003; Goodman et al., 1997; Tai et al., 2004). In general, women and old people are less likely to travel long distances to receive hospital services, but the results are not conclusive. High levels of personal income and belonging to the white ethnic group seem to increase the probability of referring to hospitals further away.

A very important variable considered by many studies is the severity of the conditions of the patients referring to the hospitals. Most of the authors dealing with this issue develop their analysis by considering DRGs of different severity separately, for example orthopaedics vs. neurosurgery (Basu and Friedman, 2001; Goodman et al., 1997; Tai et al., 2004; Varkevisser and van der Geest, 2006), while others consider whether the health conditions of the patients are life threatening or not (McNamara, 2003). This variable has been proved very significant for the choice of hospital. The literature, indeed, suggests that there is a wide consensus that willingness to pay for referring to higher quality hospitals further away is significantly higher for patients with more severe conditions.

4.3. Patient Mobility and the Gravity Approach

Some papers have recently dealt with the patient mobility phenomenon by developing a gravity model. We are going to briefly describe the study of Congdon (2001) and Levaggi and Zanola (2004).

The central focus of Congdon (2001) is on modelling patient flows to emergency units and describing how such models may be adapted to allow for unit closures and expansion, or the opening up of other units. The paper deals with five boroughs in North East London (part of North Thames Health Region) and an adjacent LHA in Essex, where the total resident population (1991 Census) was 1.1 million. This area has been subdivided into 127 electoral wards, and the analysis relates to 84,500 patient flows (resulting in inpatient admissions) from these small areas of residence to eight hospitals with emergency patient facilities. The estimation of the gravity model is based on simulation based Bayesian methods. The main regressors are the distance from the hospital, the population and the hospital mass (in terms of number of beds).[8] The first regressor negatively affects the hospital inflows, while the other two positively affect inflows.

Levaggi and Zanola's (2004) article is a benchmark for our study because it is the first one analysing patient mobility across Italian Regions by adopting a gravity approach. In this paper, the number of people moving from one region to

---

[8] In an extension of the model the author also considers the health needs (measured by the York Acute Needs score) and the percentage of the population aged over 65.

another is related to the quality of the service offered and to the distance between them. The structure of the empirical specification of the cross-migration measure is the following:

$$\mu_{ir} = \alpha_0 + \alpha_1 \left(\frac{y_i}{y_r} \right) + \alpha_2 z(.) + \varepsilon_i$$

where μ_{ir} = net patient flow, i.e. inflows – outflows, from each region (i is the region of origin or the flow while r is the rest of Italy), y_i = per capita income in region i, y_r = national per capita income, z(.) = function of relative regional hospital quality attributes of the region and the rest of Italy and ε_i is the error term.

The regressors considered in the models are per capita income, the percentage of people aged over 65 years (considered as a proxy of the regional need of health care) and hospital quality, measured in terms of structure indicators (number of beds for 1000 people, number of hospitals for 1000 people, public expenditure at regional levels in nominal values), outcomes indicators (ratio between the number of inpatients and the number of beds) and process measurement (index of turnover).

The dataset used is made up of a sample of panel observations covering regional mobility and other indicators over the period 1994-1997. The assumption of fixed coefficients over time and over cross-section units has been checked through an F-test. Since the null hypothesis of equal coefficients for each year could not be rejected the data are considered as a pool. All the regressors (apart from the number of beds for 1000 people and regional expenditures) are statistically significant. The authors find that per capita income has a positive impact on inflows (thus it has to be interpreted as ability to pay for quality at the regional level), older people seem less prone to travel and the quality variables have a positive impact on the net inflow of patients.

3.5. Econometric Model and Estimation

As far as the empirical estimation of the gravity models is concerned, there is a long tradition in the literature of taking the stochastic version of these models, making a log-linearization of them and using the OLS model to estimate the parameters of interest. This procedure is appealing because it is very simple but it can have serious drawbacks.

First, this way of proceeding can create problems when the flow between two areas is zero and thus the dependent variable in the regression is zero (Porell and Adams, 1995; Stilwell, 2005; Santos-Silva and Tenreyro, 2006). Several methods have been developed to deal with this problem. In a number of studies, the pairs with zero flows have just been dropped from the dataset and the log linear form of the gravity model has been estimated by OLS. Rather than throwing away the observations with zero flows, some authors have attributed the value of 1 to the dependent variables with zero flows. Others have used a Tobit estimator (for a more complete description of the various procedures see Frankel, 1997). These procedures will generally lead to biased estimators for the parameters of the model, and the bias will be particularly strong when there are a large number of zero flows.[9]

Secondly, as pointed out by Santos-Silva and Tenreyro (2006), when the error term in the log gravity equation is heteroskedastic, the OLS method leads to inconsistent estimates. The expected value of the logarithm of a random variable depends both on its mean and on the higher-order moments of the distribution. Hence, if the variance of the error term in the gravity equation depends on the regressors, the expected value of the logarithm of the error term will also depend on the regressors, violating the condition of consistency of OLS.

To address these two problems Santos-Silva and Tenreyro (2006) proposed to use a pseudo-maximum likelihood estimator, which they show to be robust to different patterns of heteroskedasticity and to provide a natural way to deal with the zero flows.[10]

Using a different approach, some authors make reference to the Non-inear Least Squares method (NLS) to estimate the gravity models in their multiplicative form instead of the log-linear form (see, for instance, Frankel and Wei, 1993). The

[9] An example of how biased these sorts of estimates could be is provided by Santos-Silva and Tenreyro (2006).

[10] For the sake of completeness, we underline that the Poisson regression has already proposed in the literature as a way to address the problem of zero flows (See for instance, Goodman et al. (1997)). However, to our knowledge, Santos-Silva and Tenreyro (2006) are the first ones to use this method to address the issue of heterogeneity.

NLS estimator is an asymptotically valid estimator for the gravity equation, but it can be inefficient because it does not address the issue of the heteroskedasticity that might characterize the data. The use of the pseudo likelihood maximum method, therefore, seems preferable because it allows an estimator that is more efficient than the NLS estimator to be obtained (Santos-Silva and Tenreyro, 2006).

Under the assumption that the conditional variance is proportional to the conditional mean, the pseudo maximum likelihood estimator is numerically equal to the Poisson Pseudo Maximum Likelihood (PPML) estimator which is often utilized to estimate count data models (Santos-Silva and Tenreyro, 2006). The only requirement to ensure that the estimator based on the Poisson likelihood function is consistent is the correct specification of the conditional mean. The data do not need to follow a Poisson distribution and the dependent variable does not even need to be an integer (Gourieroux et al., 1984)

In case the assumption that the conditional variance is proportional to the conditional mean does not hold (which is often likely), the estimator does not fully account for the heteroskedasticity in the model. For this reason, the inference has to be based on an Eicker-White robust covariance matrix estimator (Eicker, 1963; White, 1980).

In our empirical analysis we are forced to adopt a macro approach, since we work on a matrix of aggregated flows. We develop an “unconstrained” model for patient mobility across LHAs, in order to allow for a flexible analysis of the interaction between pushing and pulling factors. This does, however, result in a low predicting power. Moreover, we do not explicitly account for the role played by the referring physician in the choice of the treating hospital.

We assume that:

$$M_{ijp} = Push_i^{a_k} \times Pull_j^{b_k} \times f(D_{ij}, X_{ij}) \times F_i^{c_k} \times F_j^{d_k} \tag{1}$$

where:

M_{ijp} = the number of patients living in LHA$_i$ moving to LHA$_j$ for a hospital admission in the PRODUCT group p, with $p = CS, CA, BS, BM$;

$Push_i$ = set of push factors in origin LHA$_i$;

$Push_i$ = set of pull factors in destination LHA$_j$;

f = deterrence function;

D_{ij} = distance between LHA$_i$ and LHA$_i$;

X_{ij} = pair-specific impeding factors other than distance;

F_i and F_j = other controls for the origin and destination characteristics;

In our analysis we estimate the stochastic version of equation (1):

$$M_{ijp} = Push_i^{a_k} \times Pull_j^{b_k} \times f(D_{ij}, X_{ij}) \times F_i^{c_k} \times F_j^{d_k} * \eta_{ij} \qquad (2)$$

where:

η_{ij} is an error term assumed to be statistically independent of the regressors. We estimate this model with the PPML estimator utilized by Santos-Silva and Tenreyro (2006), and adopting the Eicker-White robust covariance matrix estimator.

3.6. Data and Model Specification

We consider, as dependent variables, the flows occurring from each possible LHA of origin to all the LHAs of destination. The original matrix includes all the 489 DRGs present in the classification made by the Italian Ministry of Health. After applying our aggregation along the product categories defined above, and also our exclusion criteria, the final dataset comprises 6808 non-null pair-wise flows. Since our aim was to analyse all the possible origin-destination LHA pair-wise flows we enlarged the dataset, attributing a zero value flow to the LHA combinations for which no patient flows occur. We therefore end up with four matrices, one for each product, each made up of 29,929 flows.

Our specification includes a set of control variables for spatial deterrence, and a set of push and pull factors plus some additional controls. We provide some details in the following. Table 3 contains the definition of each variable used in the empirical analysis, while Table 4 provides summary statistics.

3.6.1. Spatial Deterrence

The DISTANCE between each pair of LHAs has been calculated as the Euclidean norm between LHAs' centroids.[11] Geo-referenced coordinates of the centroids were constructed from ESRI Dataset reporting geographical coordinates (in metres) for each municipality of Italy. The variable was finally expressed in 10 kms. In our model we let DISTANCE enter the deterrence function according to a power instead of a, more common, exponential function. To allow for more flexibility we specify the power function as a cubic polynomial.

CONTIGUITY is a dummy variable, specific to each combination of LHAs of origin and destination, which assumes a value of 1 when the LHA or origin and destination share a border (or if the LHA or origin and destination coincide), and 0 otherwise. This variable is often included in the gravity models (see, for instance, Santos-Silva and Tenreyro, 2006) because the contiguity of the two "trading areas" may be an important factor to facilitate trade. Finally, to control for the presence of "institutional flow barriers" that can affect patient flows, we inserted the dummy for BARRIERS, assuming a value of 1 if the LHA of origin is different from the LHA of destination, and 0 otherwise.

3.6.2. Push and Pull Factors

We are particularly interested in analysing the effects on patient flows of some LHA specific variables. Note that all these variables enter the model both as pushing factors (i.e. referred to the LHAs of origin) and pulling factors (i.e. referred to the LHAs of destination).

POP indicates the number of enrolees of the LHA. The Ministry of Health provides this information. We will consider this measure in 100,000 units. Since enrolment is basically defined on the place of residence we might consider Italian LHAs as quite similar to US health maintenance organizations (HMOs) but for the absence of any adverse selection. From a theoretical perspective, the greater the

[11] Other measure of distance could have been adopted in our analysis. We could have considered, for example, the "road distance" between the LHAs` centroids or the "driving time" required to travel from one LHA to another. The computation of these measures of distance, however, is more complex and requires the formulation of more assumptions (for example, travelling routes, average driving speed) than the Euclidean distance among LHAs` centroids. This also require adding data not available to our study. Therefore, we have chosen to rely on the Euclidean measure as a more objective measure of distance.

pool of enrolees the greater should be the possibility of risk sharing among them, leading to economies of scale in insurance cost. This implication has found empirical support in the analysis of Wholey et al. (1996). In our exercise we try to test the presence of scale effect due to the size of the pool of enrolees. We expect the population of the LHAs of origin and destination to have a positive impact on patient flows, as predicted by the standard gravity models.

HOSPITAL indicates the presence of Aziende Ospedaliere (AOs), Policlinici or IRCCSs in the LHA. The AOs mainly provide tertiary care and they comprise many teaching hospitals. The Policlinici are multi-ward acute trusts where wards corresponding to the different medical or surgical specialties are present. The IRCCSs are hospitals of excellence aimed at performing scientific research in the areas of bio-medicine and health services management. All of them are public enterprises with a legal status broadly similar to that of the British trust hospitals. Information about the number of AOs, Policlinici or IRCCSs present in each LHA is provided by the Italian Ministry of Health. Since these hospitals are providing highly specialized health services, we expect them to have a high potential attraction to patients. We therefore expect their presence in an LHA to positively affect the inflows and negatively affect the exits.

3.6.3. Additional Controls

First, we consider some variables representing hospital quality. The information about the values assumed by these variables in 2001 is provided by the Italian Ministry of Health. We consider the BED_POP and the DOC_BED. They are defined as the number of hospital beds for 1,000 inhabitants of the LHAs, and the number of doctors for 100 hospital beds, respectively.[12] As shown in the literature on hospital choice, these characteristics of the LHA of origin should negatively influence the exit flows and positively influence the inflows. We also consider THEIL, a variable representing the Theil concentration index of the sub-PRODUCTs in PRODUCTs. This index is a proxy for the ability of the

[12] We have also considered the per capita expenditures of the LHAs (in 1,000 euros) as a possible indicator of hospital quality. However, it has been impossible to include it in our analysis due to the high level of collinearity with the population of the LHAs, which is one of the main variables of interest for our study.

LHA to provide all the typologies of treatment present in PRODUCT. The lower the value of Theil, the greater the ability of the LHA, and the lower the exits from the LHA, and the higher the inflows into it will be.

Secondly, we consider some socio-economic characteristics of the population, such as age, education and employment, for which information is provided by the Italian National Institute of Statistics (ISTAT). To approximate these variables, we consider the ratio of the LHA population over 65 and under 14 (expressed in %) (ELDERLY), the index of not completing compulsory studies in the LHA (ILLITERACY), and the unemployment rate in the LHA (UNEMPLOYMENT).

Thirdly, we consider some of the LHAs’ geographical characteristics, which can indicate the difficulty of exiting for the residents in the LHAs. We control for the average altitude (in metres) of the centroid (ALTITUDE), derived from the data about the altitude of each municipality provided by ISTAT. This variable picks up additional deterrence effects due to the poor accessibility of LHAs located in mountain or hill areas. Moreover, we consider the dimension (SURFACE) of the LHAs, measured in 10 square kms.

The variable REMOTENESS is another geographical control we include in our analysis. It is defined as the mean distance (in kms) of the LHA from all other LHAs, weighed by the number of inhabitants of each LHA. The REMOTENESS variable is expected to negatively affect patient mobility, because for people living in an LHA relatively far from another LHA the costs of moving are higher. Results in this direction are also found in the literature on hospital choice (on this point, see Congdon, 2001).

Lastly, some of the main variables include interaction terms with dummies indicating the Italian macro geographical areas AREA_NW, AREA_NE, AREA_C, AREA_S). AREA_NE is the base category. The classification of the Regions of continental Italy as four macro areas was done by ISTAT. Moreover, we control for the belonging of the LHAs to the 18 Italian Regions considered including regional dummy variables.[13][14]

3.7. Results

After some initial cleaning of the data, as detailed in Section 3.4.1, we performed the PPML estimate on the sub-samples of complex surgery, cancer, basic surgery and basic medicine separately.[15] The models have been estimated using the STATA 9.2 software.

3.7.1. Specification Analysis

As a first step, we checked the correct specification of our models by performing the RESET test (Ramsey, 1969). This test is performed by computing the predicted values from the regression function, calculating the square of those values, and estimating again the original model adding this new variable to the regressors. If this variable is not significant then the model should be considered correctly specified. As shown by Table 5, the models regarding complex surgery and cancer successfully pass the RESET test, the model for basic surgery is most likely considered "on the edge of rejection" (the p-value of the square of the predicted is 0.03), while basic medicine does not pass the test.

Another test that we performed to verify the correct specification of our model is the Link test (Pregibon, 1980). This is similar to the RESET one, but it implies adding not only the square values of the predicted ("predicted^2") as regressors but also the predicted values themselves ("predicted"). If the latter are insignificant, then the model has serious problems of specification, while if only the former are insignificant the problem of specification is milder. Table 6 shows the results of this test. Cancer is the model with best performance: "predicted" is strongly significant while "predicted^2" is not significant at a level of 95%. Complex surgery does not pass the test in strict terms, since its "predicted^2" is

[13] In our analysis we also include a dummy variable for the LHAs of Rome and of Turin in order take into consideration the peculiarity of these two "artificial" LHAs, which, as explained above, in reality are made up of more than one LHA.

[14] It would be interesting

[15] We tried to estimate a pooled model, taking into consideration all four categories at the same time. This model was rejected by the specification tests we are adopting in our analysis. Therefore, the results are not presented in the paper.

significant at a level higher than 99%. However, the misspecification problems are not very severe since the z value of "predicted^2" is moderate. The basic surgery and the basic medicine model definitely do not pass the test.

Therefore, by considering both the results of the RESET test and the Link test, we believe that our gravity model provides an adequate frame to explain the patient mobility phenomenon in the case of complex surgery and cancer, but not in the case of basic surgery and basic medicine. [16]

Although the PPML estimator is consistent even if the variance function is not well specified, it is interesting to test the specific pattern of heteroskedasticity assumed by the estimator. For example, if the relationship between the conditional variance and the conditional mean is:

$$V\left[y_i|x\right]=\lambda_0 E\left[y_i|x\right]^{\lambda_1} \tag{3}$$

it is possible to estimate λ_1, following the approach suggested by Manning and Mullahy (2001). Assuming y_i* denotes the estimated value of $E[y_i| x]$, λ_1 can be estimated from

$$\left(y_i-y_i^*\right)^2=\lambda_0 (y_i^*)^{\lambda_1}+\xi_i \tag{4}$$

using the appropriate PPML estimator. This approach is asymptotically valid and the inference about λ_1 can be made using the Eicker-White robust covariance matrix estimator.

If we want to check the assumption that the conditional variance is proportional to the conditional mean, a simpler method based on the Gauss-Newton regression is available (Davidson and MacKinnon, 1993). Indeed, to check if $\lambda_1=1$ (that would imply proportionality between the conditional variance and mean) equation (4) can be expanded in a Taylor series around $\lambda_1=1$

$$\left(y_i-y_i^*\right)^2=\lambda_0 y_i^*+\lambda_0(\lambda_1-1)+y_i^*\ln(y_i^*)+\xi_i \tag{5}$$

[16] To try to improve the model specification we have considered also alternative specifications. For instance, we have run models without including the square or the cubic term of DISTANCE. The specification we present in this chapter is the one that presents the better performances with regards to the Reset and Link tests.

Therefore, checking the significance of the parameter $\lambda=\lambda_0(\lambda_1-1)$ allows for testing the hypothesis of proportionality between the conditional variance and mean. Use of the weighted least square is recommended because the error term ξ_i is unlikely to be homoskedastic.

Table 7 reports the results for the Gauss-Newton regression, estimated separately for each of the four product categories. For each product we cannot reject the null hypothesis of proportionality between the conditional variance and the conditional mean. In fact, in all four cases, the estimates for the λ coefficient are not significantly different from zero at a level superior to 95%.

3.7.2. Estimated Results

Table 8 reports the results of the gravity models for High Surgery, cancer, basic surgery and basic medicine. The estimated coefficients, robust standard errors and level of significance for the main variables, hospital characteristics, and remoteness of the LHAs are presented. The results for the other controls are available on request. Since we are dealing with a non-linear model, the coefficients presented in Table 8 have only qualitative informational content. We focus our attention on the models for complex surgery and cancer since the models for basic surgery and basic medicine do not pass the specification tests. They are presented, however, for the sake of completeness. For complex surgery and cancer, the Pseudo-R^2 is 0.92 and 0.94, respectively.

We also include the interaction terms of the logarithm of POP the dummy variables indicating the four macro geographical areas. North-East is the base category. For both complex surgery and cancer in all areas this variable has a positive effect on patient flows (the only exception is POP for the LHAs of origin in the Centre for cancer). This result is consistent with the predictions of the traditional gravity model according to which the dimension of both the LHAs of origin and destination have a positive effect on the flows.

In the model of High Surgery, the population of the LHAs of origin has a greater effect in the North West and South than in the North East (the coefficients are 1.41, 1.32 and 1.15, respectively) while in the South it has a lower effect (its coefficient is 0.95). The same ranking of the coefficients also hold when considering the population of the LHAs of destination. Notice that only the coefficients for the North West and East are significant for both the LHAs of origin and destination.

In the model for cancer, the effect of the population of the LHAs of origin is greater in the North West than in the North East (the coefficients are 1.10 and 0.21, respectively) but lower in the South (its coefficient is 0.09). This effect is not significant in the North East and South. Considering the population of the LHAs of destination, the greatest effect is for the South (1.46), followed by the Centre (0.93), the North East (0.61) and the North West (0.46). The effect is significant in all areas, apart from the North West.

In the estimated gravity models, we also use the interaction term of DISTANCE for the macro geographical areas of origin and of destination of the flows. If we consider the availability of patients to travel a certain distance as a proxy of their willingness to pay for the transaction costs linked to the exit, the interactions of DISTANCE with the macro-geographical areas represent the different willingness to pay patients have in different macro-areas. The category of reference is the distance patients travel from an LHA in the North East to reach another LHA in the North East. The distance always has a negative effect for both complex surgery and cancer. This effect is always significant apart from when the area of origin is the North West or the area of destination is the Centre for High Surgery, and when the area of destination is the North West for cancer. Considering the interactions with the areas of origin, the effect of distance is greater in the North West and the North East, while it is lower in the Centre and South. However, the opposite holds when the interactions with the areas of destination are considered. These results are valid for both complex surgery and cancer. The results imply that when patients are considering moving from their LHA of origin, distance is a much higher determining factor if the LHA of origin is in the North West or East than in the Centre or South. If the availability for patients to travel a certain distance is equivalent to the willingness to pay for the transaction costs linked to the exit, then patients living in the Centre and South have a higher willingness to pay. In contrast, when patients are considering where to be treated, the deterrence power of distance is lower if the LHA of destination is in the North West or East, rather than in the Centre or South. This can also be interpreted as a patient's higher willingness to pay to be treated in the former areas than in the latter ones.

In our regression models we also estimate the quadratic and cubic terms of distance ($DISTANCE^2$ and $DISTANCE^3$). The sign of these terms (positive for the quadratic and negative for the cubic ones) confirm that distance has a negative effect overall on patient flows, but the relationship between the two variables is not linear. In particular, when the distance is increasing, patient flows tend to decrease at a decreasing rate until a saddle point, beyond which they decrease at an increasing rate. The value assumed by the distance in correspondence to the saddle point depends on which PRODUCT and which geographical area is considered.

As expected, in both the models of complex surgery and cancer, CONTIGUITY has a positive effect on patient mobility across LHAs (the coefficients are 0.24 and 0.03, respectively) while BARRIER has a negative effect (the coefficients are -1.52 and -0.92, respectively). However, the effect of the former variable is significant only in the case of High Surgery, while the effect of the latter is significant for both the macro categories of products. The interaction between the two variables is positive but insignificant for both complex surgery and cancer.

In the model for High Surgery, HOSPITAL has a clear negative and significant effect on the exits and a positive and significant effect on the inflows. Indeed, the coefficient relating to the variable of LHA of origin assumes a value of -0.25 while the one relating to the LHA of destination assumes a value of 0.63. Therefore, as expected, the presence of autonomous hospitals in an LHA is a strong attraction factor. For cancer, unexpectedly, this variable is not significant and, when referred to the LHAs of destination, its sign is different from that expected.

When we consider the controls shown in Table 4 relative to the hospital characteristics, they have the expected signs and tend to behave as good proxies of the hospital quality of the LHAs. Indeed their sign is positive if they are referred to the LHAs of destination ("BED_POP" assumes a value of 0.47 for complex surgery and 0.42 for cancer, "DOC_BED" 1.94 for complex surgery and 1.17 for cancer) and their sign is negative if they are referred to the LHAs of origin (the former variable is -0.13 for complex surgery and -0.17 for cancer, the latter is -0.42 for complex surgery and -0.65 for cancer). These effects are all significant, with the exception of “DOC_BED” for the LHAs of origin). For THEIL, complex surgery reflects our expectations. The corresponding coefficient is positive and highly significant for the LHAs of origin (0.30) while it is negative and highly significant for those of destination (-0.40). For cancer, this variable is not significant when referred to the LHA of origin, while for the LHA of destination it has a negative and significant effect on patient mobility; a result which is quite puzzling.

For both complex surgery and cancer, REMOTENESS has the expected negative sign when referred to the LHAs of destination even if the coefficient is significant just for cancer. Unexpectedly, it has a positive and significant sign when referred to the LHAs of origin. This means that patients in the more remote LHAs tend to exit more. However, this unexpected result is coherent with the results of the descriptive statistics shown in Section 3.4.2. These statistics highlighted the presence of patient flows from the South to the North and Centre of Italy, showing a relevant exit phenomenon from the “remote” LHAs of the South.

3.8. Final Remarks

In most of the cases our results are consistent with the predictions of the gravity model. In particular, the number of enrolees of the LHAs of origin and destination affects patient flows in a positive way, while the distance between LHAs affects patient flows in a negative way. The contiguity of the LHAs plays a positive role, while the presence of institutional barriers to mobility plays a negative one. For High Surgery, the presence of autonomous hospitals has a negative role on exits and a positive role on inflows. Therefore, in general, the gravity model can be regarded as a good framework for explaining patient mobility for hospital care treatments across LHAs in Italy.

The analysis conducted so far has some limits.

The models proposed just seem to be reliable for treatments regarding complex surgery and cancer. Indeed, in the case of basic surgery and basic medicine the specification test adopted suggests that the gravity model is not correctly specified. Maybe the decision process ruling the patient mobility phenomenon for these treatments has peculiarities requiring the development of a different model. Moreover, focussing on complex surgery and cancer, some of the variables considered are not significant and (in very few cases) they do not have the expected sign.

In our analysis we have not considered the role played by waiting times in hospital choice, which has been shown in the literature to be of great importance for the choice of hospitals and patient mobility. Data about waiting times are only available at the regional level. Therefore it has been impossible to include this variable in the present exercise.[17] Perhaps the omission of this variable is particularly relevant for the case of basic surgery and basic medicine. This could be a path for future research.

Some of the choices we made for the estimation procedures may be questionable. In our econometric specification we did not include LHA fixed effects. In a recent paper Anderson and van Wincoop (2003) argue that the traditional gravity equation suffers from a problem of omitted variable since it does not take into account the multilateral resistance term, which is very relevant in trade. Following these authors, many recent studies estimating a gravity equation include importers and exporters fixed effects as regressors in order to take into account the presence of these multilateral resistance terms (for instance Cheng and Wall, 2004; Egger and Pfaffermayr, 2004; Westerlund and Wilhelmsson, 2006; de Frahan and Vancauteren, 2006). The problem with the inclusion of "importer" and "exporter" fixed effects is that if we work with cross-sectional data the model cannot be fully identified and we cannot estimate the effect on the dependent variable of the regressors that are importer or exporter specific.

Some authors have addressed this issue by exploiting the panel nature of their data. For example, Egger and Pfaffermayr (2004) estimate a linear gravity model with fixed effect using the Hausman and Taylor (1981) estimator. Others

[17] If we had included them in our regression model, their effects would have been captured by the regional dummies.

have modelled the importer and exporter individual effects as random (uncorrelated) effects instead of fixed (correlated) effects (de Frahan and Vancauteren, 2006). Others, like Wei and Frankel (1997), have not included individual dummies in their gravity model. Including individual dummies would have undermined the efforts taken to estimate the effects of the variables that do not have variability over time (or other dimensions).

In our paper we take a position similar to Wei and Frankel (1997) and adopt the traditional gravity equation, which does not include "individual effects". Since we are mainly interested in studying the effects of the LHAs' populations and the presence of AOs and IRCCSs on patient mobility, we do not include any LHA dummy variables, because their presence would not allow the study of the main effect of interest.[18] We recognize, however, the importance played in our model by the heterogeneity of the LHAs. Thus, "as a compromise", even if we do not include LHA specific effects, we include Regional specific effects in our model in order to account for geographical heterogeneity.[19] Moreover, we include many controls at LHA level to try to represent the more relevant characteristics of the LHAs.

In future work we could try to perform some formal tests to detect the presence of spatial autocorrelation in the data regarding patient flows (using, for instance, the indexes developed by Moran (1950) and Geary (1954)). Testing for the presence of spatial autocorrelation in the data could be relevant because this sort of autocorrelation leads to spatial correlation in the residual of our model, and it leads to the violation of the assumption of independence of the residuals (Fisher and Griffith, 2006). In case we detect the presence of spatial autocorrelation in the data, to account for it we could include additional terms in our model, either in the deterministic part of the model or in the stochastic one (some recent applications of spatial models are, e.g., Sacerdote (2001), Kelejian and Prucha (2004) and Baltagi et al. (2007))

[18] We underline that the main aim of Anderson and van Wincoop's (2003) paper is studying the trade border effect and for their research hypothesis it does not matter if they cannot estimate individual specific variables.

[19] We include dummies at regional level because the Region is the administrative entity at the level higher than the LHA in the organization of the Italian NHS.

One aspect of our model that we could develop in the future relates to the role played by the hospital dimension in the patient mobility phenomenon. As shown in Section 3.5.2, in the literature about hospital choice this variable is often assumed to be an indicator of the quality of hospital, and thus an attraction factor. This assumption should be viewed with caution. By taking into account some "empirical" considerations, this variable can be thought of as either a pulling or a pushing factor. Generally, bigger hospitals should be able to sustain greater fix costs (because these costs are distributed amongst a higher quantity of consumers) and thus they should be able to guarantee better quality (for example, adopting more up-to-date technologies). Moreover, doctors working in these hospitals can treat a greater number of patients of the same type and thus they should be able to specialize and obtain better clinical results. For these reasons quality could be higher in larger hospitals and so patients should be attracted to them (Taroni, 2001). However, due to congestion problems in larger hospitals, doctors and nurses may provide less care to patients and, in general, more managerial problems might arise. This could decrease the quality in larger hospitals and so potential patients may be incentivised to refer to other hospitals (Taroni, 2001). Moreover, large hospitals could also sustain higher average costs. A study of the University of York on the NHS, for example, has shown that scale economies on costs can be reached only until 200 beds for hospitals, above that threshold diseconomies appear (the average costs increase) (NHS Centre for Review and Dissemination, 1996). As a future extension of our paper, we could verify if our data corroborate the hypothesis of a U shape (inverse U-shape) relationship between the size of the hospitals and the exit flows (inflows) to address the potential issue of provider congestion.

Table 1. Performance measures

	Exit rate	Inflow rate	Accessibility	Attractiveness	N° obs.
OVERALL	18.3%	11.4%	66.9	46.5	171
REGIONAL					
North-West	19.8%	10.8%	45.3	49.2	39
North-East	13.5%	11.1%	42.0	47.2	45
Centre	17.1%	9.0%	54.9	47.9	36
South	22.1%	13.8%	114.0	43.0	51
ENROLLEES					
less than 150	15.0%	9.8%	61.4	45.1	44
150 - 250	18.1%	9.8%	62.3	42.0	57
250 - 400	17.8%	13.6%	81.3	50.0	41
more than 400	24.3%	13.7%	64.1	52.8	29
HOSPITAL					
NO	20.4%	8.2%	55.9	42.1	116
YES	13.8%	18.1%	90.1	56.0	55
PRODUCT					
complex surgery	57.6%	21.1%	109.7	41.4	171
cancer	36.0%	16.7%	99.5	60.8	171
basic surgery	20.7%	13.9%	51.0	42.7	171
basic medical	12.7%	8.2%	54.0	41.3	171

Table 2. Performance measures by PRODUCT and AREA

		ITALY	North-West	North-East	Centre	South
Exit rate	**PRODUCT**					
	complex surgery	57.6%	55.4%	47.3%	62.8%	64.8%
	cancer	36.0%	33.8%	23.8%	32.8%	50.5%
	basic surgery	20.7%	22.9%	16.2%	20.5%	23.0%
	basic medical	12.7%	13.7%	8.4%	10.6%	17.1%
Inflow rate	**PRODUCT**					
	complex surgery	21.1%	20.8%	23.7%	14.3%	23.8%
	cancer	16.7%	15.5%	15.7%	12.1%	21.6%
	basic surgery	13.9%	12.1%	14.4%	13.1%	15.4%
	basic medical	8.2%	8.3%	6.9%	5.7%	11.1%
Accessibility	**PRODUCT**					
	complex surgery	109.7	51.7	53.1	77.5	226.9
	cancer	99.5	54.2	54.6	87.6	182.0
	basic surgery	51.0	41.9	37.1	44.6	74.8
	basic medical	54.0	43.0	36.6	46.0	83.3
Attractivness	**PRODUCT**					
	complex surgery	41.4	50.7	45.9	32.9	36.4
	cancer	60.8	94.6	56.8	54.9	42.5
	basic surgery	42.7	41.1	40.5	48.3	41.8
	basic medical	41.3	42.7	38.8	45.0	39.9

Table 3. Definition of variables

VARIABLE	DEFINITION	SOURCE	UNIT OF MEASURE
<u>**DEPENDENT**</u>			
$FLOW_{ij}$	Number of patient enrolled to LHA i admitted in a public hospital of LHA j	Ministry of Health	
<u>**REGRESSORS**</u>			
PAIRWISE SPECIFIC			
DISTANCE	Straight line distance between origin and destination LHA centroids	ESRI (elaboration)	10 km
CONTIGUITY	=1 if the origin LHA and the destination LHA share a border	Our elaboration	
BARRIERS	=1 if the origin and destination LHAs belong to different Regions	Our elaboration	
LHA SPECIFIC			
POP	Number of enrolees in the LHA	Ministry of Health	10,000 inhabitants
HOSPITAL	=1 if a big hospital (Azienda Ospedaliera, Policlinico, Irccs) is present in the LHA territory	Ministry of Health	
BED_POP	Number of hospital beds for 1,000 enrollees in the LHA	Ministry of Health	N°/1000 inhabitants
DOC_BED	Number of doctors for 100 hospital beds in the LHA	Ministry of Health	N°/100 hospital beds
THEIL	Theil concentration index of the product lines in macro product lines	Ministry of Health	
REMOTENESS	Mean distance of the LHA from all other LHAs (weighed by the share of inhabitants)	ESRI (elaboration)	KM
UNEMPLOYMENT	Unemployment rate in the LHA	ISTAT 2001	%
ILLITERACY	Share of the population not having completed the compulsory studies in the LHA	ISTAT 2001	%
ELDERLY	Ratio of the elderly (+65) above the young (-14) in the LHA	ISTAT 2001	%
ALTITUDE	Height above sea level of LHA centroid	ISTAT 2001	10 meters
SURFACE	Geographical surface of the LHA	ISTAT 2001	10 square km
AREA_NW	=1 if the LHA is locate in the North-West		
AREA_NE	=1 if the LHA is locate in the North-East		
AREA_C	=1 if the LHA is locate in the Centre		
AREA_S	=1 if the LHA is locate in the South		

Table 4. Descriptive statistics

VARIABLE	MEAN	STD. DEV.	MIN	MAX
FLOW	42.53	925.38	0.00	142883.00
DISTANCE	38.85	244.77	0.10	110.64
CONTIGUITY	0.03	0.17	0.00	1.00
BARRIER	0.93	0.26	0.00	1.00
POP	29.23	27.07	1.40	254.08
HOSPITAL	0.35	0.48	0.00	1.00
BED_POP	3.72	1.31	0.21	8.36
DOC_BED	0.45	0.12	0.24	1.05
THEIL	6.56	0.64	5.15	7.63
REMOTENESS	404.50	77.95	308.07	641.85
UNEMPLOYMENT	10.14	7.84	1.93	37.65
ILLITERACY	9.91	3.32	4.47	24.30
ELDERLY	153.80	50.84	42.49	284.22
ALTITUDE	23.13	19.95	0.12	104.20
SURFACE	144.94	94.50	8.82	620.71
AREA_NW	0.23	0.42	0.00	1.00
AREA_NE	0.26	0.44	0.00	1.00
AREA_C	0.21	0.41	0.00	1.00
AREA_S	0.29	0.46	0.00	1.00

Table 5. Reset Test

	chi2(1)	Prob > chi2
complex surgery	0.06	0.804
cancer	0.76	0.384
basic surgery	4.70	0.030
basic medical	7.17	0.007

Table 6. Link test

	z	P>|z|
complex surgery		
predicted	292.32	0.000
predicted^2	-8.08	0.000
cancer		
predicted	377.26	0.000
predicted^2	1.66	0.097
basic surgery		
predicted	2095.62	0.000
predicted^2	767.96	0.000
basic medicine		
predicted	2650.03	0.000
predicted^2	1534.46	0.000

Table 7. Gauss Newton Regression

	chi2(1)	Prob > chi2
complex surgery	1.00	0.317
cancer	1.09	0.278
basic surgery	1.04	0.299
basic medical	1.02	0.307

Table 8. **Gravity model POISSON estimates**

	COMPLEX SURGERY		CANCER		BASIC SURGERY		BASIC MEDICINE	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
DISTANCE	-9.72E-01***	5.67E-02	-1.23E+00***	4.45E-02	-1.48E+00***	8.21E-02	-1.90E+00***	9.42E-02
$DISTANCE^2$	5.07E-02***	6.62E-03	6.16E-02***	5.60E-03	1.09E-01***	1.94E-02	1.36E-01***	2.00E-02
$DISTANCE^3$	-9.34E-04***	1.91E-04	-9.82E-04***	1.56E-04	-2.85E-03***	9.12E-04	-3.31E-03***	8.41E-04
DISTANCExAREA_NW_o	-7.08E-02	1.26E-01	2.09E-01**	9.64E-02	-9.28E-03	1.27E-01	2.01E-01	1.39E-01
DISTANCExAREA_C_o	4.51E-01***	1.02E-01	5.92E-01***	8.02E-02	2.38E-01*	1.26E-01	4.61E-01***	1.59E-01
DISTANCExAREA_S_o	7.05E-01***	9.49E-02	1.08E+00***	6.87E-02	1.05E+00***	1.08E-01	1.44E+00***	1.26E-01
DISTANCExAREA_NW_d	1.25E-01**	5.49E-02	3.79E-02	4.89E-02	7.14E-02	6.82E-02	2.07E-01***	5.35E-02
DISTANCExAREA_C_d	-4.39E-02	7.85E-02	-2.12E-01***	5.63E-02	8.84E-02	5.59E-02	1.01E-01	6.65E-02
DISTANCExAREA_S_d	-2.75E-01***	8.90E-02	-4.94E-01***	6.18E-02	-6.42E-01***	9.29E-02	-7.18E-01***	9.16E-02
$DISTANCE^2$xAREA_NW_o	2.17E-03	1.66E-02	-2.73E-02***	7.85E-03	1.35E-02	3.61E-02	-2.75E-02	2.95E-02
$DISTANCE^2$xAREA_C_o	-3.67E-02***	7.18E-03	-4.12E-02***	6.21E-03	-3.61E-02	2.42E-02	-6.45E-02***	2.50E-02
$DISTANCE^2$xAREA_S_o	-4.72E-02***	7.02E-03	-6.04E-02***	5.90E-03	-1.02E-01***	1.95E-02	-1.29E-01***	2.04E-02
$DISTANCE^2$xAREA_NW_d	-1.69E-03	1.38E-03	-2.64E-04	1.21E-03	-2.66E-03	1.63E-03	-4.34E-03***	1.50E-03
$DISTANCE^2$xAREA_C_d	3.08E-03	2.22E-03	7.18E-03***	1.94E-03	-1.37E-03	1.81E-03	6.66E-04	2.51E-03
$DISTANCE^2$xAREA_S_d	1.53E-02***	5.21E-03	2.05E-02***	4.03E-03	5.91E-02***	1.05E-02	5.72E-02***	1.21E-02
$DISTANCE^3$xAREA_NW_o	-1.10E-04	5.94E-04	6.28E-04***	1.82E-04	-7.23E-04	1.98E-03	1.06E-03	1.26E-03
$DISTANCE^3$xAREA_C_o	8.01E-04***	1.93E-04	7.74E-04***	1.59E-04	1.42E-03	1.04E-03	2.19E-03**	9.35E-04
$DISTANCE^3$xAREA_S_o	9.17E-04***	1.93E-04	9.79E-04***	1.58E-04	2.81E-03***	9.12E-04	3.28E-03***	8.43E-04
$DISTANCE^3$xAREA_NW_d	8.33E-06	1.01E-05	7.82E-07	8.48E-06	2.21E-05**	1.13E-05	2.64E-05**	1.23E-05
$DISTANCE^3$xAREA_C_d	-3.87E-05**	1.97E-05	-6.55E-05***	1.94E-05	4.94E-06	1.68E-05	-2.34E-05	2.41E-05
$DISTANCE^3$xAREA_S_d	-2.40E-04**	1.12E-04	-2.43E-04***	7.79E-05	-1.40E-03***	3.96E-04	-1.13E-03***	4.21E-04
CONTIGUITY	0.239**	0.116	0.031	0.116	0.531***	0.121	0.391**	0.163
BARRIER	-1.516***	0.219	-0.919***	0.216	-2.106***	0.258	-2.276***	0.502
CONTIGUITYxBARRIER	0.331	0.270	0.014	0.232	1.504***	0.253	1.492***	0.486
ln(POP)_o	1.146***	0.147	0.662***	0.138	1.109***	0.175	0.890***	0.235
ln(POP)_d	0.389**	0.160	0.215	0.133	-0.132	0.170	0.084	0.235

ln(POP)xAREA_NW_o	0.263*	0.177	-0.561***	0.172	-0.841***	0.203	-0.399	0.246
ln(POP)xAREA_C_o	-0.191	0.186	-0.127	0.202	-0.330	0.213	-0.253	0.273
ln(POP)xAREA_S_o	0.170	0.211	-0.147	0.192	-0.758***	0.234	-0.572**	0.276
ln(POP)xAREA_NW_d	0.396*	0.198	0.610***	0.170	0.872***	0.197	0.396	0.244
ln(POP)xAREA_C_d	-0.133	0.194	0.320*	0.179	0.449**	0.204	0.303	0.268
ln(POP)xAREA_S_d	0.076	0.245	0.854***	0.187	0.928***	0.230	0.688**	0.271
HOSPITAL_o	-0.250**	0.104	-0.088	0.095	-0.044	0.122	-0.235*	0.131
HOSPITAL_d	0.627***	0.116	-0.036	0.099	0.016	0.124	0.211	0.130
BED_POP_o	-0.134***	0.043	-0.175***	0.041	-0.222***	0.051	-0.236***	0.055
DOC_BED_o	-0.422	0.353	-0.650	0.415	-0.891***	0.338	-0.631	0.426
BED_POP_d	0.472***	0.042	0.418***	0.037	0.382***	0.049	0.361***	0.053
DOC_BED_d	1.941***	0.478	1.168**	0.466	0.387	0.357	0.297	0.445
THEIL_o	0.299**	0.146	-0.391	0.280	-0.495*	0.282	-0.039	0.762
THEIL_d	-0.399**	0.193	1.710***	0.293	1.205***	0.289	1.407*	0.765
REMOTENESS_o	2.556**	1.092	5.408***	0.905	2.529*	1.327	1.319	1.821
REMOTENESS_d	-0.595	1.425	-5.992***	0.986	-1.517	1.384	-0.777	1.865
CONSTANT	-14.020***	5.188	-5.946*	3.436	-9.414***	1.946	-11.008***	2.615
Log likelihood	-69521.27		-103357.76		-202132.24		-342091.22	
N° obs.	29929		29929		29929		29929	
Wald chi^2(90)	43344.700		41796.340		43750.240		42980.080	
Pseudo R^2	0.924		0.941		0.972		0.977	

***, ** and * denotes significance at 1%, 5% and 10% levels respectively. Reported standard errors are robust. Each equation contains socio-economic and geographic controls, plus 19 dummies for region of origin fixed effects and 19 dummies for region of destination fixed effects. Considering the LHA specific regressors, we use the suffix "_o" to refer to the LHA of origin and the suffix "_d " to refer to the LHA of destination.

Annex 1

Table A1. Composition of PRODUCT groups

PRODUCT	Overall share of hospital treatments	DESCRIPTION
CS = complex surgery	1.7	Surgical Neurology Pulmonary Surgery Cardiovascular Surgery Transplants
CA = cancer	7.8	Surgical Oncology Medical Oncology Chemotherapy and Radiotherapy
BS = basic surgery	23.6	Surgical Ophthalmology Surgical Otorhinolaryngology Surgical Gastroenterology Orthopedic Surgery Surgical Endocrinology Urologic Surgery Vascular Surgery General Surgery
BM = basic medicine	47.6	Medical Neurology Medical Ophthalmology Medical Otorhinolaryngology Pulmonary Medicine Cardiology Medical Gastroenterology Orthopedic Medicine Medical Endocrinology Urologic Medicine Psychiatry Vascular Medicine General Medicine Rehabilitation
EM = emergency	4.0	Surgical traumatology Major traumatology Minor traumatology
HIV	0.5	HIV
DE = delivery	14.8	Gynecology Surgical obstetrics Medical obstetrics Neonatology

Chapter 4

Contractual Conditions, Working conditions, Health and Well-Being in the British Household Panel Survey

4.1. Introduction

Over the past 20 years or so, changes in the labour market have had a substantial impact on the working arrangements of employees. For example, the number of "standard" full time permanent jobs has decreased, while non-standard work arrangements (temporary work, part-time contract, unregulated work etc.) have become more common (Kivimäki et al., 2003). In the European Union, for example, non-standard employment now accounts for 12–15% of paid employment (Virtanen et al., 2003). A key question is whether and how these changes to employment patterns affect health and well-being. Working conditions have also undergone significant changes over the past decades. The decline of manufacturing jobs, the growth of service oriented work and computerization appear to have made the "traditional" sources of adverse physical and environmental working conditions less relevant and have increased the scope for psychosocial job stressors (Cappelli et al., 1997). Given these changes, it is relevant to evaluate how working conditions affect health and psychological well-being in society today. Moreover, some studies have underlined that workers with atypical contractual conditions are often characterized by adverse working conditions (Aronsson G., 2001; Artazcoz et. al., 2005; Letourneux, 1998). Given this evidence, it is important to take into consideration contractual conditions as well as working conditions.

The analysis of the influence of contractual and working conditions on health and psychological well-being can also provide an empirical contribution to the debate about the neoclassical labour economics framework, which suggests that workers are induced to accept jobs with undesirable non-wage characteristics by compensating differences in their wage rates. The theory of compensating wage differentials has proved to be an important tool in economists' attempts to understand labour markets. But empirical research about the determinants of individual earnings has provided limited support for the theory (Brown, 1980). Workers experiencing adverse conditions for health and well-being, indeed, should be compensated through wage premiums, that they might invest in order to maintain their health and well-being stable. Therefore, the empirical investigation of the relationship between contractual and working conditions and health and psychological well-being can provide some evidence in favour or against the neoclassical theory of compensating wage differentials.

The influence of contractual conditions on health and well-being has been analysed in previous studies with respect to part-time and temporary employment (see, for instance, Benavides et al., 2000, Gash et al., 2006, and Kivimaki et al., 2003). A common hypothesis in the literature when investigating contractual conditions is that people with a temporary or a part-time job constitute a "disadvantaged" group, who experience poor conditions of work, including, for example, low wages, few benefits, job insecurity, little training and no possibility of promotion, and are forced to take these jobs because they cannot find more "traditional ones" (Kunda et al., 2002; Segal and Sullivan, 1997). These characteristics have often been assumed to have a negative influence on health and well-being. However, some empirical evidence suggests that the reality is more complex. Not all individuals with these types of job contracts can be considered to be "disadvantaged" (Silla et al., 2005), and a worsening of health occurs only if these jobs are associated with low levels of employability, are involuntary or offer no contractual certainty (Artazcoz et al., 2005; Price and Burgard, 200; Silla et al., 2005). We attempt to investigate empirically the impact that temporary and/or part-time contractual employment has on health and well-being and how these effects vary across individuals. We also evaluate how the effects of contractual and working conditions on health and well-being are influenced by family structure.

When considering the effects of working conditions on health and psychological well-being, we refer to the "demand-control-support" model (Karasek et al., 1988; Karasek and Theorell, 1990) and the "effort–reward imbalance model" (Siegrist et al., 1990; Siegrist, 1996). These are two of the most influential models developed to investigate the possible mechanisms underlying these effects.

We consider the effects of contractual and working conditions on self-assessed health (SAH) and psychological well-being (derived from the General Health Questionnaire (GHQ)) using twelve waves (1991/92 – 2002/2003) of the British Household Panel Survey (BHPS). While one branch of the literature suggests that "atypical" contractual conditions have a significant impact on health and well-being, another suggests that health is damaged by adverse working conditions. As far as we are aware, previous studies have not explicitly considered the two factors jointly. Our aim is to combine the two branches of the literature to assess the distinct effects of contractual and working conditions on health and psychological well-being.

Since the effect of contractual and working conditions on health and well-being is unlikely to be immediate and takes time to manifest itself, we adopt a dynamic approach, regressing the present level of self-assessed health and psychological well-being on past values of the variables of interest. We also condition the present level of health and of psychological well-being on their past values. This allows us to reduce concerns about reverse causality. For modelling self-assessed health, since the dependent variable is categorical, we estimate non-linear dynamic panel ordered probit models, while for modelling psychological well-being we estimate a dynamic linear specification.

The structure of the paper is as follows. Section 4.2 gives an overview of the existing literature on the associations between contractual and working conditions and health and psychological well-being. Section 4.3 introduces the BHPS data and describes the sample and variables we use for estimation. In Section 4.4 we describe the econometric models and the estimation strategy. Section 4.5 reports and discusses the results of our estimates and some conclusions and discussion are provided in Section 4.6.

4.2. Literature Review

4.2.1. Contractual conditions

The empirical evidence regarding the influence of contractual conditions on health is mixed. If we consider fixed versus permanent jobs, studies have reported that fixed-term workers have worse physical health than permanent workers (see for example, Klein Hesselink and Van Vuure, 1999, Benavides et al., 2000, and Gash et al., 2006, who make reference to Dutch data, to a set of 15 European countries and to Spanish and German data, respectively). Kivimaki et al. (2003) show that temporary employment is associated with higher mortality among respondents in the Town Study in Finland. In other studies fixed-term contracts have been shown to have either no influence (Virtanen et al., 2003) or positive influences on health (Sverke et al., 2000).

As far as we are aware, few studies make reference to the association of general self-assessed health with having a part-time rather than a full time job. One study that does is Benach et al. (2004) who analyse data from 15 EU countries and report that full time workers have worse indicators of health compared to part-time workers.

Psychological well-being is traditionally considered to be negatively affected by fixed-term employment. These contracts are considered stressful since they imply job insecurity (Bohle et al., 2001; Burchell 1994, 1999). The stressful influence of fixed term contracts could be explained, for instance, by considering the inability of fixed-term contract workers to plan and control their lives given the short-term nature of their jobs (Burchell, 1994).[1] This traditional assumption is confirmed by several studies (Klein Hesselink and van Vuuren, 1999; Lasfargues et al., 1999; Martens et al., 1999). It is unlikely, however, that fixed-term contracts have the same impact on all workers. Individual characteristics, such as tolerance for ambiguity and self-control, play a relevant role in influencing responses to stress and the selection process into permanent employment (Bauer and Truxillo, 2000). Therefore, we are not surprised to find other studies in the

[1] In this regard, for instance, there is evidence that having a non-permanent employment contract is associated with a lower probability of getting married and creating a family (Artazcoz et al., 2005).

literature that report no differences in psychological well-being between fixed term contract and permanent contract workers (Sverke et al., 2000).

Evidence from recent papers suggests that people with atypical contractual conditions cannot be considered as a homogeneous group when comparing their health and well-being with that of permanent workers. Referring to temporary versus permanent work arrangements, for instance, Artazcoz et al. (2005) consider the Catalonian Health Survey and differentiate between workers with fixed-term contracts, with non-fixed term temporary contracts and working with no contract at all. They find fixed term contracts are not associated with poorer mental health status, while working with non-fixed term temporary contracts or with no contract at all is positively associated with poor mental health for many categories of workers. Silla et al. (2005) also consider Spanish workers, but differentiate temporary workers according to their preferences for contractual arrangements and level of employability. "Traditional" temporary workers (with low employability and low preference for temporary contracts) have lower well-being than permanent workers, but other temporary workers report higher well-being than permanent workers. Referring to part-time and full-time contracts, Price and Burgard (2006) consider the Americans` Changing Lives study in the US and show that women with a part-time contract have lower depressive symptoms and lower body mass indices than full-time workers, but only if the reduced working time is voluntary.

Some caution should be exercised when considering the influence of atypical contractual employment arrangements on health across countries. Differences in national employment rates and employment regulations, for example, will determine what can be considered typical and atypical employment contracts and may serve to moderate their impact on health (Benach et al., 2004). Accordingly, generalising relationships formed at national level is often difficult (Virtanen et al., 2003).

If we focus on British data, Bardasi and Francesconi (2004) use the first 10 waves of the BHPS to investigate the influence on general health status and mental health of atypical employment (defined as temporary and part-time jobs). The authors find that atypical employment does not appear to be strongly associated with adverse general health. However, there is some evidence that those in seasonal or casual jobs have poorer mental health. Rodrigues (2002) considers data from the BHPS for the years 1991-1993 and finds that the health status of part-time workers with permanent contracts is not significantly different from those who are employed full-time. However, part-time casual work without a contract is reported to be associated with poorer health.

4.2.2. Working conditions

Several studies present evidence that adverse working conditions have negative effects on health and psychological well-being. Many of these studies make reference to the broad categories of working conditions present in two of the most influential models developed to investigate these effects, the “demand-control-support” model developed by Karasek et al. (1988) and Karasek and Theorell (1990) and the “effort–reward imbalance model” of Siegrist et al. (1990) and Siegrist (1996). The demand-control-support model considers the categories of job demand, decision latitude (job control, ie. high levels of decision authority and skill utilization) and social support at work. Job demand can be physical (regarding manual work), psychological (regarding pace of work, quantity of work and conflicts at work) and contractual (considering the number of working hours and irregular work schedule) (Marchand, 2005). The model postulates that negative health effects derive not from a single aspect of the work environment, but from the joint effect of the demands of a work situation and the range of discretion in decision-making available to the workers facing those demands. In particular, high job demand and low job control is the worst combination for health. Job demands place the individual in a state of “stress”. If no action can be taken by the individuals, then the unreleased stress may lead them to suffer from adverse health consequences (Henry and Cassell, 1969).[2] Social relationships are a second analytical level added to this model. “Individuals who are ‘socially integrated’ link together their capacities for accommodating stress. […]social support buffering should reduce the strength of association between task characteristics and strain symptoms” (Karasek et al., 1982, pag. 182).

[2] “The implied logical structure is roughly analogous to an 'energy reservoir' model; either the energy summoned by the individual to cope with his environment is used for active behavior, or it is focused internally with deleterious consequences” (Karasek et al., 1982, pag. 182).

The effort–reward imbalance model considers the categories of effort, such as the demands of the job and the motivation of workers in challenging situations, and reward at work in terms of salary, esteem, job stability and available career opportunities. It predicts that a negative impact on health occurs when there is an imbalance between the two dimensions. The motivation behind this prediction is the following. Work has positive effects for individuals on their emotional and motivational level. Indeed, "occupational status is associated with recurrent options of contributing and performing, of being rewarded or esteemed, and of belonging to some significant group (e.g. work colleagues)" (Siegrist, 1996, pag. 29). However, these potentially beneficial effects of work are contingent on the presence of reciprocity in the social exchanges. "The effort-reward imbalance model claims that lack of reciprocity between costs and gains (i.e., high-cost/low-gain conditions) define a state of emotional distress with special propensity to autonomic arousal and associated strain reaction" (Siegrist, 1996, pag. 30)

Several empirical studies considering physical and mental health have provided evidence in favour of the two models (see, for instance, Pikhart et al., 2004, and Theorell and Karasek, 1996). Other studies, however, have failed to support the theories (for example, Vermeulen and Mustard, 2000) and overall there does not appear to be a clear consensus on the empirical validity of these models.

Considering studies that focus on general health and that perform prospective analysis [3] Cheng et al. (2000), using a cohort of women working as nurses in the United States in 1992 and in 1996, show that those in jobs with high demands, low control, and low social support show the greatest declines in health status between 1992 and 1996. Niedhammer et al. (2003) make reference to the GAZEL cohort (a survey of workers at Electricité De France–Gaz De France) by considering individuals who were working and responded to both the 1997 and 1998 questionnaire. The analysis shows that significant predictors of poor self-reported health are psychological demands for both men and women, decision authority for men only, and social support and physical demands for women only.

[3] An extensive literature exists about the effects of workplace factors on specific diseases like cardiovascular disease, muscle-skeletal problems etc (see e.g. Juul-Kristensen and Jensen, 2005; Andersen et al., 2004). We choose not to report these studies in this short review and instead concentrate on studies that investigate effects on more general measures of health.

Warren et al. (2002) use the Wisconsin Longitudinal Study (WLS) and consider the influence of composite indices of physical and psychosocial job characteristics in 1975 and 1992 on respondents health outcomes in 1992.[4] Respondents with adverse physical conditions, for example unclean working environment, and psychosocial conditions, for example limited authority and autonomy, were less likely to describe their health as excellent.

The work by Datta Gupta and Kristensen (2007) is particularly significant for our study due to the methodology adopted. The authors make reference to the European Community Household Panel (ECHP) and consider data from Denmark, France and Spain for 1994-2001. By specifying a dynamic panel data model they evaluate the impact on self-assessed health of the variable "satisfaction with the work environment/working condition". This variable is considered as a proxy of both the physical aspect (including chemical, ergonomic and climatic harmful exposures) and the psychological aspects of the work environment (for example, relations with co-workers, conflict resolution and discretion over work). The results of the analysis show that, in the three countries, having a satisfactory work environment significantly promotes employee health.

Ferrie et al. (2002) use the Whitehall II survey, which covers London-based office staff, aged 35–55. They show that individuals with job insecurity report poorer self-rated health and well-being than those who have secure employment. Bartley et al. (2004) use the BHPS for 1991–2001 to analyze the relationship between employment status and occupational social class on limiting long-term illness. Social class, however, is defined on the basis of employment conditions such as autonomy, job security, and career prospects. Their analysis shows that men and women in routine occupations have a greater probability than those in higher professional or managerial classes to experience a spell of limiting illness together with a lower probability of recovery from such illness. More recently

[4] The physical characteristics include whether the job always or frequently involved physical effort; whether the respondent ever got dirty on the job; how many hours per week they spent doing the same things over and over; how many hours per week they spent working with their hands; and whether they were ever exposed to dangerous conditions. Different psychological characteristics included whether the respondent had the authority to hire or fire other employees; whether employees controlled other employees' rates of pay; whether they supervised other employees' work; whether they were themselves supervised by someone else; whether their job always or frequently involved working under pressure of time; and how many hours per week they spent dealing with people about work (Warren et al., 2002).

Ahmed-Little (2007) shows that the increase in shift working of junior doctors in the UK caused great dissatisfaction, even if accompanied by a reduction in working hours. Fatigue and poor performance on night shift have been reported by junior doctors and there is evidence from outside medicine that suggests that shift work may have long term consequences on health**.**[5]

Considering psychological well-being, Tennant (2001) and Michie and Williams (2003) provide reviews of studies that have investigated the impact of working conditions. The majority of the studies showing evidence that adverse working conditions affect negatively well-being are cross sectional (see, for instance, Kawakami et al., 1990, 1992, and Martens, 1999). Among the more recent studies, Godin and Kittel (2004) consider a sample of Belgian workers employed in four different firms and show that low control and low social support negatively affect mental health. Pikhart et al. (2004) analyse data from the Health, Alcohol and Psychosocial factors in Eastern Europe Study (HAPIEE) based on population samples from three cities in Russia, Poland and the Czech Republic. They find that depressive symptoms are strongly associated to effort–reward imbalance at work, material deprivation and marital status, but that they are not associated with job control.

Few studies have performed a prospective analysis. Niedhammer et al. (1998) use the French GAZEL cohort and make reference to the "demand-control-support model". They show that high psychological demands, low decision authority, and low social support at work predict the onset of depression over a 12-month follow-up. Rugulies et al. (2006) analyse data from the Danish Work Environment Cohort Study, considering the baseline variables in 1995 and the incidence of severe depressive symptoms in 2000. Low influence at work and low social support from supervisors, among women, and job insecurity, among men, seem to increase the risks of depression. Mixed evidence is found by Marchand (2005), who considers four cycles of Statistics Canada's National Population

[5] Shift work has been considered in a systematic review performed by Bambra et al. (2007). 66 studies, set in Western Europe, the USA, Canada, Japan or Australia are considered in the review. The review focus on two types of organizational changes: compressed work week (CWW) and schedule redesign. The review suggests that while CWW interventions might not always improve the health of shift workers, they are seldom detrimental. Moreover the CWW may have positive effects on well-being, especially in terms of work/life balance. Considering schedule redesign, the review shows that interventions which changed the rotation of shift work have positive or null effects on health and well-being.

Health Survey (NPHS). Job insecurity appears to increase the probability of psychological distress, but, surprisingly, higher decision authority has the same effect.

Our own study contributes to the literature in a number of ways. We assess the distinct effects of contractual and working conditions on health and psychological well-being, combining the two distinct branches of the literature. As far as we are aware, the two sets of conditions have not been considered by previous studies *explicitly* in a joint way.[6] Notice that the working conditions we consider are numerous. Many of the previous studies in the literature focus on specific occupational groups (i.e.: civil servants, nurses, etc.), which makes it difficult to generalise results to the entire workforce. We use the BHPS, which is a longitudinal dataset providing rich information on occupational, sociodemographic and health variables of a sample representative of the British population. The use of this dataset allows us to generalize our results to a population of employees in Britain. Other studies have used this dataset (Bardasi and Francesconi, 2004; Bartley et al., 2004; Rodrigues, 2002), however, we analyse the effects of contractual and working conditions jointly and we exploit a larger number of waves in the BHPS dataset. Moreover, we estimate *dynamic* panel data models (linear for GHQ and non-linear for SAH). Dynamic panel models have many advantages compared to cross sectional ones. For example, they: increase precision in estimation; account for unobserved individual heterogeneity; and reduce concerns about endogeneity bias (Cameron and Trivedi, 2005, p. 697, 743)

4.3. The BHPS Dataset

We use panel data from the first 12 waves (1991/92–2002/2003) of the BHPS, a longitudinal survey of private households in Great Britain. This survey includes rich information on occupational, socio-demographic and health variables. The dataset was designed as an annual survey and the initial sample was

[6] The only study we are aware of that considers both contractual and working conditions is Martens et al. (1999). In their study, however, only temporary and occasional contracts are considered as contractual conditions, and only rotating shifts and irregular working hours are considered as working conditions. In our study we also consider the effects of temporary employment and we refer to a much broader set of working conditions.

collected in 1991.[7] It contains observations about each adult member (16+) of a nationally representative sample of more than 5000 households. Approximately 10,000 individuals were interviewed in the first wave, and the same individuals were re-interviewed in successive waves. In case they split off from their original households, they were re-interviewed along with all adult members of their new households.[8]

In our analysis we use an unbalanced sample, which contains all the available observations at each wave that provide complete information on the variables used in the model. The sample also includes new entrants to the survey. Given the objective of our analysis, we consider only employees, and we exclude from our sample people outside the job market or self-employed. The final sample consists of 45,658 observations (23,309 for women and 22,349 for men). We only use observations for which two consecutive waves are available, since we are conditioning health and psychological well-being on one-year lagged values. Table 1 summarizes the variables used in our empirical models.

4.3.1. Dependent variables

We use two dependent variables, SAH, a categorical variable for self-assessed health, and the GHQ, which is a measure of psychological well-being.

Self-assessed Health

SAH has been used widely as a measure of health status (e.g. Adams et al., 2003; Benzeval et al., 2000; Ettner, 1996; Smith, 1999) and has been shown to be a powerful predictor of subsequent use of medical care (van Doorslaer et al., 2000) and mortality (Idler and Benyamini, 1997; Idler and Kasl, 1995). As a self-reported subjective measure of health, SAH may be prone to the measurement error often referred to as ‘state-dependent reporting bias’ (Kerkhofs and Lindeboom, 1995), ‘scale of reference bias’ (Groot, 2000) and ‘response category cut-point shift’ (Sadana et al., 2000; Murray et al., 2001). This sort of measurement error occurs if the mapping of ‘true health’ into SAH categories

---

[7] A two-stage stratified systematic sampling procedure was used to do the initial selection of households for inclusion in the survey. The procedure was designed to give each address an approximately equal probability of selection.
[8] For further details see Taylor et al. (1998).

vary with respondent characteristics, that is, if sample subgroups adopt different cut point levels in a systematic way when reporting their SAH, although they have the same level of ‘true’ health. To attempt to surmount this problem researchers could model the reporting bias making reference to more ‘objective’ indicators of true health (Kerkhofs and Lindeboom, 1995; Lindeboom and van Doorslaer, 2004) or to ‘anchoring vignettes’ (Murray et al., 2001).[9] We do not pursue the potential issue of reporting error further in the study. The BHPS has limited information on objective health variables and does not contain vignettes.

The use of all the first 12 waves of the BHPS could be problematic. For waves 1-8 and 10-12 the SAH variable represents “health status over the last 12 months”. Respondents are asked: “Compared to people of your own age, would you say your health over the last 12 months on the whole has been: excellent, good, fair, poor, very poor?”. The SF-36 questionnaire was included in wave 9. In this questionnaire, the SAH question was re-worded and included a modification to the response categories. The SAH variable for wave 9 represents the “general state of health”, using the question: “In general, would you say your health is: excellent, very good, good, fair, poor?”. To make wave 9 comparable to the other waves we collapse the original SAH variable to create a categorisation that has common support over the two versions of the question. The final SAH is a categorical variable that represents the following four health categories: “poor or very poor”, “fair”, “good or very good”, “excellent”.[10]

In our analysis we always divide our sample by gender. Dividing the sample by gender is quite common in empirical studies about contractual conditions, working conditions and health and reflects the differential trends in health over time between men and women together with any differences in working arrangements between the sexes (Artazcoz et al., 2005; Bardasi and Francesconi, 2004; Benach et al., 2004; Kivimäki et al., 2003; Rugulies et al., 2006).

Psychological Well-being (GHQ)

---

[9] Anchoring vignettes can be defined as descriptions of fixed levels of a latent construct. They are very useful sources of information because any systematic variation across individuals in the rating of the vignettes can be attributed to reporting heterogeneity.

[10] For further details about this procedure see Hernandez-Quevedo et al. (2005).

We use the reduced version of the General Health Questionnaire (GHQ) (Goldberg and Williams, 1988) present in the BHPS. The GHQ is often used as an indicator of well-being but it was at first developed to detect psychiatric illness. For the 12 individual items present in the GHQ, respondents indicate on a four-point scale (where 0 is the best scenario and 3 the worst) how they recently felt in relation to each item.[11] In our analysis our dependent variable is the Likert scale (Likert, 1952), which reports an overall score summing the individual components of the GHQ. The Likert scale, therefore, ranges from 0 to 36. We rescale the original variable so that it is increasing in good psychological health.

4.3.2. Independent variables

Contractual conditions

In our analysis the variables regarding contractual conditions are represented by having a part-time contract (*part-time job*), that is working less than 30 hours a week, and having a non-permanent contract (*temp job*).[12] Both variables are dummy variables and the reference categories are having a full-time and a permanent contract.

Working Conditions

The BHPS is a rich source of information about the working conditions of the individuals interviewed. In order to facilitate the comparison between our results on the influence of these variables on health and psychological well-being and the evidence found in the previous literature, we group the working conditions according to some of the broad categories present in the models of Karasek et al. (1988) and Siegrist et al. (1990), mentioned in Section 4.2.2.[13] We underline that the variables that we use to represent working conditions are only

[11] The 12 items are: concentration, sleep loss due to worry, perception of role, capability in decision making, whether constantly under strain, perception of problems in overcoming difficulties, enjoyment of day to day activities, ability to face problems, loss of confidence, self-worth, general happiness, whether suffering depression or unhappiness.

[12] To define if workers have a permanent or temporary job, they are asked "Leaving aside your own personal intentions and circumstances, is your current job permanent or non-permanent?"

[13] From waves 2 to 4 individuals were not asked information about some variables (*not daytime*, *unions*, *payrise* and *promotion opportunities*) if they were still in the same job as the previous year. For these cases, we assume that the value of these variables did not change from the last year it was recorded.

proxies for the conceptual categories used in the literature (described in Section 4.2.2). Most of the previous studies are based on datasets that contain more detailed information on the working conditions of the employee. However, these studies usually consider only a very small sample of employees, (i.e. employees working in a particular firm, or a particular city) and their conclusions cannot be generalized at a national level. In contrast, the use of the BHPS, which comprises observations on workers from all over Great Britain, makes our conclusions more general and valid for all Great Britain. Accordingly there is a trade-off between accuracy of measurement and generalization of results. Here, we compromise by using proxy measures to obtain results that can be generalized to a wider population.

Demanding job conditions

We consider not working during regular office hours and working unpaid overtime hours as conditions of high job demands. Therefore, we use the variable *not daytime*, which is equal to 1 if respondents do not work during day time or if they have rotation shifts, and 0 otherwise.[14] The variable *overtime hours* represents the number of overtime hours that are *not paid* that respondents work in a normal week. It is the difference between the total number of overtime hours and the number of *paid* overtime hours.[15] We expect these two variables to have a negative relationship with health and psychological well-being, as shown in the literature on high job demand (Cheng et al., 2000; Lindberg et al., 2006; Martens et al., 1999).

Control and social support

To proxy the presence of social support at the workplace, we use the variable *unions*, which is equal to 1 if a union or staff association is present at the workplace of the respondent, and 0 otherwise. The extent of control over work is

[14] Precisely, considering the categories present in the BHPS questionnaire, “not daytime” is equal to 0 if respondents report to be in the category “day”, while it is equal to 1 if they report to be in the categories "night", "rot shift", "day and night", "morning, afternoon or evening only" or "other".

[15] Considering just the total number of overtime hours worked a week could be misleading. Often, if the overtime hours are paid, they are paid more than the standard hours of work. Therefore, in this case, financial benefits could compensate the stress effect of overtime work.

approximated by the variable *managerial supervision*, which is equal to 1 if the respondent has a managerial or supervision role, and 0 otherwise. We expect *unions* and *managerial supervision* to have a positive relationship with health and psychological well-being, as shown for other variables denoting high job control and social support (Cheng et al., 2000; Godin and Kittel, 2004; Lindberg et al., 2006)

Reward

To consider the possible rewards that respondents can have in their work, we include in our analysis the variable *payrise*, which is equal to 1 if the pay of respondents includes an annual increment, 0 otherwise, and *promotion opportunities*, which assumes value 1 if there are opportunities for promotion in the current job, and 0 otherwise.[16] In the literature, the variables denoting positive reward at work have been shown to have a positive influence on health and well-being (Borg et al., 2000; Marchand, 2005; Rugulies et al., 2006). Therefore, we expect *payrise* and *promotion opportunities* to have a positive influence on well-being.

Working environment

We include the location/venue and the size of the company/institution where respondents work. The original variable representing the first condition is categorical, indicating if people are working at the employer's location, at home, travelling or other. Given this variable, four dummy variables are derived as follows: *workplace_employer*, *workplace_home*, *workplace_travel* and *workplace_other*.

The size of the company/institution is approximated by the number of employees. The original variable present in the BHPS is a categorical one, representing, for example, if respondents work in a place where there are 1-2 employees, 3-9 employees, 10-24 employees, etc. We have created a continuous variable to reduce the number of dummy variables to include in our analysis,

[16] To measure the variable *payrise* respondents to the BHPS are asked "Some people can normally expect their pay to rise every year by moving to the next point on the scale, as well as receiving negotiated pay rises. Are you paid on this type of incremental scale?". To evaluate *promotion opportunities* respondents are asked "In your current job do you have opportunities of promotion?"

following the procedure described in Contoyannis and Rice (2001).[17] Working in big companies/institutions might have a positive influence on the health and well-being of employees, since it might allow better career opportunities or the use of better facilities.

Work satisfaction

We use *satisfaction_total pay*, *satisfaction_security* and *satisfaction_work*, which assume the value 1 if respondents are satisfied with the total pay for the job, the security of the job and the work itself, respectively, and 0 otherwise.

We consider also the preferences for the number of hours worked each week. The BHPS contains a categorical variable reporting if individuals would prefer to work fewer hours (*preference less hrs*) or more hours (*preference more hrs*). Preferring the amount of hours actually worked is our base category.

Other covariates

We control for age, which is included in a cubic form (*age*, *age^2* and *age^3*) to allow for a non-linear relationship with health and psychological well-being, and for marital status, including the categories divorced or separated (*divsep)*, never married (*nevermar)*, or widowed (*widowed)* (the base category is being married or living as a couple).

We take into consideration the ethnic origin of respondents (*race*), the number of individuals living in the household (*household size*), and the presence of children in the household (*children*). The base category is being white and not having children.

We also control for individual income (*income*), which is measured as gross annual (labour and non-labour) income. For education and social class we include the variables *high education* and *lower social class* in the model.[18] [19] [20] Year

[17] "We created a continuous variable by taking the midpoint of each category for each individual. For those who could not report the category into which their establishment fell, but were able to report whether it was above or below a particular value, we estimated their observation as a weighted average of the midpoints of the relevant categories. The weights used are the proportions of the relevant sub-sample which are in the relevant categories" (Contoyannis and Rice, 2001, p.610).

[18] Income is transformed to natural logarithms to allow for the concavity of the health–income relationship (E.g. Ettner, 1996).

dummies are included to account for aggregate health shocks, time-varying reporting changes and possible effects of age which are not captured by the age variables we include in the model.

Table 2 presents the mean, standard deviation, minimum and maximum for the regressors used in our empirical models for the sub-samples of women and men. Overall, the differences in means between the two sub-samples appear very small. A much higher percentage of women, however, have a part-time job than men. Moreover, a higher percentage of men than women work in a place different from the employer's or at home, have a managerial or supervision role and have promotion opportunities.

4.4. The Econometric Models and Estimation Strategy

4.4.1. The econometric models

To model self-assessed health we use a dynamic panel ordered probit specification with random effects. The ordered probit can be used to model discrete dependent variables taking ordered multinomial outcomes. Therefore, it applies well to our measure of self-assessed health, which has categorical outcomes "poor or very poor", "fair", "good or very good", "excellent".

The latent variable specification of the model that we estimate can be written as:

$$h^*_{it} = \beta' x_{it-1} + \gamma' h_{it-1} + \alpha_i + \varepsilon_{it} \qquad (1)$$

$i = 1,\ldots.N$ (number of individuals in the sample)

$t = 2,\ldots.T$ (number of waves of the survey)

[19] *High education* is equal to 1 if people have a qualification equal or superior to A level, and 0 otherwise. *Lower social class* is equal to 1 if people belong to the BHPS categories "skilled manual" "armed forces", "partly skilled", "unskilled", and 0 otherwise.

[20] There are many variables that are worth trying to include in our regression model as controls, such as the worker employment sector. However, the use of the STATA program reoprob.ado, written by Frechette (2001) imposes restrictions on the number of variables that it is possible to include in the regression model. Therefore, we choose to adopt as controls only the basic socio-economic conditions. In future work we could try to include in our analysis other controls, such as the worker employment sector.

where x_{it-1} is a set of observed variables which may be associated with the health indicator. These variables are lagged one period to account for delays between contractual conditions and working conditions impacts on health.[21] Moreover, the SAH variable makes reference to health status over the last 12 months, while many of the job related variables makes reference to the present time. Therefore, to try to ensure that measures of contractual and working conditions are obtained before measures of health and well-being it is necessary to use the one-year lag of the former variables. α_i is an individual-specific and time invariant random component while ε_{it} is a time and individual-specific error term. This error term is assumed to be normally distributed, uncorrelated across individuals and waves, and uncorrelated with α_i. The x_{it-1} are assumed to be uncorrelated with ε_{is} for all *t* and *s*. We restrict the variance of the idiosyncratic error term to be equal to one, since we do not have a natural scale for the latent variable.

The estimation of the effects of contractual and working conditions on health and psychological well-being may raise concerns about the presence of endogeneity bias, unless one can establish that the causation is unidirectional (as has been proposed, for example, by Adams et al., 2003). To reduce concerns about reverse causality (i.e. health and psychological well-being affecting contractual and working conditions) and following previous studies (Chapman and Hariharan, 1994; Contoyannis et al., 2004), we include previous health status lagged one period, h_{it-1}, in our empirical models. This further allows us to identify the impact of working conditions and contractual conditions on changes in health status.

Since we consider self reported data, we do not observe the latent level of health h^*_{it}, but only an indicator of the category in which the latent indicator falls, h_{it}:

$$h_{it} = j \quad \text{if} \quad m_{j-1} < h^*_{it} < m_j \quad j = 1,\ldots.4 \tag{2}$$

where $m_0 = -\infty$, $m_{j-1} <= m_j$, $m_4 = +\infty$.

[21] This has been stressed, for example, by Bartley et al. (2004).

Assuming that the error term is normally distributed, the probability of observing the particular category of SAH reported by individual *i* at time t (h_{it}), conditional on the set of regressors and the individual effect, can be expressed as:

$$P_{itj} = P(h_{it} = j) = \Phi(m_j - \beta' x_{it-1} - \gamma' h_{it-1} - \alpha_i) - \Phi(m_{j-1} - \beta' x_{it-1} - \gamma' h_{it-1} - \alpha_i) \qquad (3)$$

Where $\Phi(.)$ is the standard normal distribution function. This formulation shows that it is not possible to identify separately an intercept in the linear index (β_0) and the cut points (m), since the model only allows identification of ($m_j - \beta_0$). We adopt a conventional normalization, setting $\beta_0 = 0$, to deal with this issue. The random effect ordered probit specification is estimated with STATA (release 9.0, Stata Corporation) using the program *reoprob.ado*, written by Frechette (2001).

In our analysis psychological well-being is measured using the GHQ. We model psychological well-being with a linear model. This model can be expressed by equation (1), keeping all the related assumptions. In this case, however, we need to underline that h^*_{it} does not represent a latent variable but the observed one**.**

To allow for the possibility of a correlation between the observed regressors and the individual effect we parameterize the individual effect (Chamberlain, 1984; Mundlak, 1978; Wooldridge, 2005), allowing it to be correlated with the within-individual means of the time-varying regressors. Moreover, since we estimate dynamic models, the problem of initial conditions needs to be considered. Two assumptions are typically made concerning a discrete time stochastic process with binary outcomes (Heckman, 1981). The same issues arise when we deal with ordered categorical outcomes (Contoyannis et al., 2004). The first assumption is that the initial observations are exogenous variables, while the second assumption is that the process is in equilibrium, meaning that the marginal probabilities have approached their limiting values and can be considered time-invariant. If the error process is not serially independent and the first observation is not the true initial outcome of the process the first assumption is not valid, and the estimators we obtain are inconsistent. In our case, we know that the latter condition does not hold, since the first year for which we have observations does not coincide with the start of individuals` health trajectory. The second assumption is not valid if non-stationary variables such as age and time trends are included in the model, as they are in our study.

For both the linear and non-linear model, we deal with the initial condition problem by adopting the conditional maximum likelihood approach of Wooldridge (2005), modelling the distribution of the unobserved effect conditional on the initial value and any exogenous explanatory variables.[22] The likelihood function resulting from Wooldridge's approach is based on the joint distribution of the observations conditional on the initial observations. A limitation of this approach is that it requires specifying a complete model for the unobserved effects. This approach, therefore, may be sensitive to misspecification.

Following the approach of Wooldridge, we parameterize the distribution of the individual effects as:

$$\alpha_i = \alpha_0 + \alpha_1 h_{i1} + \alpha_2 \bar{x}_i + u_i \tag{4}$$

where $\bar{x}_i$ is the average over the sample period of the observations on the time-varying socio-economic controls variables and u_i is assumed to be distributed N(0, σ^2_u), independent of the regressors, the idiosyncratic error term (ε_{it}), and the initial conditions.

4.4.2. Estimation Strategy

Our approach to estimation is to first assess the distinct effects of contractual conditions and working conditions separately on health and psychological well-being. We then estimate a model containing both working and contractual conditions and further assess the impact of employees' job satisfaction and the effect of variables relating to the "demand-control-support" model of Karasek and the "effort–reward imbalance" model of Siegrist.

The first models regress SAH and GHQ, respectively, on the contractual conditions variables only and the set of controls (*age*, *age^2, age^3, divsep*,

[22] Wooldridge's (2005) approach for dealing with linear panel models is also used by Hauck and Rice (2004).

neverm, *widowed*, *race*, *household size*, *children*, *income*, *high education*, *lower social class*, *preference less hrs*, *preference more hrs*). To assess the effects of voluntary or involuntary part-time employment, we introduce variables representing the joint effect of having a part-time job and being unsatisfied with the number of hours worked. Accordingly we introduce the interaction terms *part-time job*preference less hrs* and *part-time job*preference more hrs*, with the baseline representing a part-time job where the employee is satisfied with the hours worked. We expect these interaction terms to be negatively related to health and well-being.

For workers on temporary contracts we consider the potential impact of the individuals` chances of finding alternative employment (should they wish to do so) by considering variables representing potential employability (Silla et al., 2005). In this study we consider higher levels of education as a proxy for higher employability. We, therefore, introduce the interaction terms *temp job*high edu*. We expect this term to have a positive effect on health and well-being, since temporary workers with a high level of employability should be less concerned about a lack of job security.

We also try to consider the role that family structure plays on the health and psychological well-being of workers in part-time and temporary jobs, by allowing contractual conditions to have different effects according to whether an employee has children. Therefore, we introduce the interaction terms *part-time job* children* and *temp job*children*. The rationale for introducing the latter term is that the stress due to insecurity of job could be worse for workers who have children to support and accordingly we expect to observe a negative sign on the effect of this variable. It is harder to predict the influence of *part-time job* children*. Indeed, employees with children could benefit from having a part-time job, due to having more time to spend with their family. However, for workers who have to maintain children, the lower income associated with a part-time job could have stressful effects. Therefore, the final effect on health and psychological well-being could depend on which of these effects prevails.

In the second model we regress SAH and the GHQ, respectively, on the working condition variables as set out in Section 4.3.2, conditional on the set of controls.

In the third model we use both the contractual and the working conditions variables as regressors, to assess the robustness of the results obtained in the previous models when we consider the two sets of regressors together. We further attempt to evaluate if our data provide support to the theoretical frameworks of Karasek et al. (1988) and Siegrist et al. (1990) (from now on, referred to as Karasek and Siegrist). Considering Karasek's framework, we introduce the interaction term *no daytime*managerial/supervision*, where working outside normal daytime hours is an indicator of high job demand, and having a managerial supervision role is an index of high control. A positive coefficient would provide support for Karasek`s framework. Considering Siegrist's framework, we introduce the interaction term *no daytime*promotion opportunities*, where not working during the day is an indicator of high effort while having promotion opportunities is an indicator of reward.[23] A positive coefficient would lend support to Siegrist`s framework. We also introduce the variables about satisfaction with working conditions (*satisfaction_total pay*, *satisfaction_security* and *satisfaction_work*), because they reflect, at least partly, a balance between the effort and reward at work. We are aware that work satisfaction may be influenced not only by objective working conditions but can also reflect personality characteristics of respondents (i.e. if individuals have a tendency to be pessimistic or optimistic). Although not offering conclusive results, positive coefficients on these variables would provide some support for Siegrist`s model.

We underline that in the model for SAH, since we are dealing with a non-linear ordered categorical model, the estimated coefficients have only qualitative content. To provide information about the magnitude of the effects we present partial effects (Wooldridge, 2002). In particular, we report the change in the probability of reporting excellent health due to a marginal change for continuous

[23] Actually, the interaction terms *no daytime*managerial supervision* and *no daytime*promotion opportunities* are not introduced contemporaneously in the third model, because the two terms are highly correlated. We estimate a model where we introduce *no daytime*managerial supervision* only and another model where we introduce *no daytime*promotion opportunities* only. Since the results of these two models are extremely similar, we have chosen to present the results of the latter model, and to report only the results related to *no daytime*managerial supervision* for the former model.

variables and to a discrete change for binary variables.[24] We compute the effects for a hypothetical representative agent with "average characteristics".[25] Inference on the significance of the estimated coefficients is undertaken using the Wald test.

When dealing with non-linear models, attention should be given to interaction terms, as highlighted by Ai and Norton (2003) and Norton et al. (2004). First, the partial effect for an interaction term could be non-zero even if the directly estimated coefficient of the interaction term is zero. Secondly, we cannot rely on standard tests on the coefficients of the interaction term to test the statistical significance of the interaction effect. Thirdly, the interaction effect is conditional on the independent variables and may have different signs for different values of the covariates. Therefore, to compute the magnitude of the interaction effects it is necessary to compute the cross derivative (for continuous variables) or differences (for categorical ones). Moreover, the statistical significance of the interaction effect must be tested for the cross partial derivative (difference) of the dependent variables and not for the directly estimated coefficient of the interaction term. In our analysis we adopt the strategy proposed by Ai and Norton (2003) and Norton et al. (2004) to compute the partial effect and the standard errors for the interactions.[26]

4.5. Results

For all of the models described in the previous Section, the coefficients for the lagged and the initial value of the dependent variable are statistically significant at the 1% level and substantial in magnitude. This result supports our

[24] Notice that the direction of the effect of the covariates on the probabilities of reporting the extreme outcomes ("poor or very poor" and "excellent" health, in our study) is unambiguously determined by the sign of the coefficients (Wooldridge, 2002)

[25] We attribute the mean value to the covariates that are continuous and the modal value to the covariates that are categorical. To make the partial effects meaningful, when we compute the partial effects of *part-time*pref less hrs*, *part-time*pref more hrs* and *part-time*children* the representative agent is assumed to have a part-time job, while computing the partial effects for *temp job*high education* and *temp job*children* this individual is assumed to have a temporary job. Considering the partial effects for *not daytime*manag-sup* and *no day time*prom opp*, the representative agent is assumed not to work during the daytime.

[26] The standard errors of the interactions are computed by applying the delta method (Norton et al., 2004).

use of a dynamic model and indicates that current health is a function of previous period health.[27]

To verify if the parametrisation of the individual effect described in Section 4.4.1 holds, we assess if the coefficient α_2 in equation (4) is significantly different from zero. Since income is the only time-varying variable used as a control we include its within-individual mean as $\bar{x}_i$. In all of the models presented below, the coefficient of this variable is statistically significant at the 1% level. The only exception is the model for females for psychological well-being. In all the other cases the correlated effects model has to be preferred to the pure random effects model. Accordingly, below we report the estimates derived from the correlated effects model.

4.5.1. Self-assessed Health

Table 3 reports the estimated coefficients and the standard errors for females and Table 4 for males, while Table 5 and Table 6 report the partial effects (the change in the probability of reporting excellent health) and standard error for women and men, respectively.[28] The first column presents the model considering contractual conditions only; the second column considers working conditions only; and the third column, the full model, where we use both contractual and working conditions variables. To conserve space, we report the results only for the variables related to contractual and working conditions. Estimated coefficients for the control variables present in the model are available on request. Inference on the statistical significance of the relationship between the main terms and SAH is done by referring to the standard errors of the estimated coefficients, while considering the interaction terms we refer to the standard errors of the partial effects (as suggested by Ai and Norton, 2003 and Norton et al., 2004).

The first column of results suggests that there is a positive and statistically significant relationship between health and having a part-time job (compared to

[27] In our paper we do not consider the problem of health-related attrition. Contoyannis et al. (2004) and Jones et at. (2006), however, show that although health-related attrition exists in the BHPS data, it does not appear to distort the magnitude of the effects of socioeconomic variables when modelling the determinants of health. This result allows us to think that attrition should also not be relevant in the estimate of the effect of contractual and working conditions.

[28] To compute the partial effects and the related standard errors we use the STATA command "nlcom", which implies the use of the delta method for calculations.

having a full time job), for employees satisfied with the number of hours worked or who do not have children. This result is consistent with previous studies (Price and Burgard, 2006). However, and as expected, not being satisfied with the number of hours worked or having children has a negative influence on the health of part time workers. This holds for both women and men. However, these effects are not significant at conventional levels. The magnitude of the computed partial effects, overall, is larger for men than for women.[29] Notice that the partial effects for *part-time*pref less hrs* and *part-time*pref more hrs* are larger than that for *part-time*children.* This suggests that the health of part-time workers is influenced more by preferences for hours worked than by the demands of a family.

Consistent with previous literature (Silla et al., 2005), our analysis reveals a negative relationship between health and having a temporary job (compared to having a permanent job) for women with a low level of education. Unexpectedly, however, this relationship appears positive for men. Having a high level of education positively influences health for women and men with a temporary job. A further asymmetry between women and men relates to the presence of children. The relationship between having a temporary job and health is positive for women with children, while the relationship is negative for men. Notice that the magnitude of the partial effect (in absolute terms) for *temp job*children* is smaller than that for *temp job*high education* for women, while the opposite result holds for men.[30] Moreover, *temp job*high education* is statistically significant for women, while *temp job*children* is statistically significant for men. These results suggest that, for men with a temporary job, the family structure has a larger influence on reporting excellent health than their level of employability, while for women the opposite holds. The asymmetries we observe between women and men about the influence of temporary jobs on health could perhaps derive from the

[29] For people satisfied with the number of hours worked or without children, a shift from full time to part-time increases the probability of reporting excellent health by 2.3% for women and 7.5% for men. If the person with the part-time job wishes to work less hours, the probability of reporting excellent health reduces by about 6% for both women and men, while if she/he wishes to work more hours the probability reduces by 2% and 4.8 % for women and men, respectively. If the part-time workers have children, the probability reduces by 0.7% and 3.2% for females and males, respectively.

[30] In fact, the partial effect for the probability of reporting excellent health for temporary workers with a high level of education is 0.052 for females and 0.032 for males, while that partial effect for temporary workers with children is 0.021 and -0.091.

different roles women and men play within the family structure. Our results, indeed, could be explained in the light of a "traditional" view of the family, where taking care of children is mainly a responsibility of women while men are the ones with the main responsibility for child material sustenance.

In the second model, where we consider working conditions, the partial effects of some conditions have the sign that we expect according to the previous literature (*payrise* and *overtime hours* for women and men, *not daytime* and *managerial supervision* for women, and *promotion opportunities* for men). However, the partial effect of other variables (*promotion opportunities* and *unions* for females, *not daytime*, *managerial supervision* and *unions* for males) does not exhibit the expected sign.[31] The magnitude of all of these effects, however, is small.[32]

Considering the variables related to the working environment, we observe an asymmetry between men and women. The relationship between working at home (compared to working at the employer's workplace) and health is positive for women while it is negative for men, while the opposite holds for workers who travel or workers in places different from the employer's workplace. For women, the magnitude of these partial effects seems to be large compared to that of most of the other working conditions variables. Notice that the partial effect of working at home for females is particularly high and is statistically significant. This could be due to the fact that working at home may allow a high level of flexibility (people can organize their time in a better way, they can reduce the travel time to work, they can take care of family or the house at the same time, etc). The relationship between the number of employees at the workplace and health is negative for women and positive and statistically significant for men. The magnitude of the related partial effects, however, appears very small.[33]

In the third model, where we consider both contractual and working conditions, the magnitude of the partial effects of these conditions and their level

[31] Notice that the partial effect of *promotion opportunities* and *unions* for females is also statistically significant.
[32] The variation in the probability of reporting excellent health induced by any of these working conditions is smaller than 2.5%.
[33] The value reported in the table is about 0.00005 for females and 0.00006 for males, meaning that, for example, the presence of 100 more people at the workplace increases the probability of reporting excellent health by 0.5% for females and 0.6% for males.

of statistical significance change little compared to the models where contractual and working conditions are considered separately.[34] This result suggests that our model is robust and that there is little collinearity between the regressors. An interesting finding, however, regards the magnitude of the partial effect of the variable *temp job*. For females, in the third model the partial effect is still negative but is smaller than in the model where only contractual conditions are considered. For males, the partial effect is still positive but becomes larger than in the first model. Although these results are not conclusive, they appear to suggest that when contractual conditions are analysed without considering working conditions, the estimate of the effect of having a temporary job (and a low level of education) might be biased by the omission of some working condition variable in the model.[35] The fact that workers with adverse contractual conditions are often characterized by adverse working conditions has been underlined in the literature, for example, by Artazcoz et al. (2005), Aronsson (2001) and Letourneux (1998).

Our findings appear to provide some support in favour of the Karasek`s framework for women, but not for men, since the partial effect of *no daytime*managerial supervision* is positive for females and negative for males (none of them, however, is statistically significant). Regarding the Siegrist`s framework, the interpretation of our findings is more ambiguous, since some of the factors considered are in favour of the framework while others are not. For both women and men, *satisfaction_security* and *satisfaction_work itself* appear to show a positive and highly statistically significant relationship with health, while *no daytime*promotion opportunities* exhibits a negative (and statistically non-significant) relationship. Due to data limitation and the use of proxies to represent the working conditions present in the Karasek and Siegriest models, we are unable to test directly the theoretical foundations of those models. Instead we use a general framework t that allows us to examine their predictions to the extent that our data permit. Should more detailed information become available in the BHPS

[34] In the model for males, the partial effect of *promotion opportunities* changes sign, but its magnitude continues to be very small and it is not statistically significant compared to first and second model.

[35] Notice that having a temporary job (in general, for people with both low and high education) is slightly correlated with some working conditions. For example, the correlation coefficient between having a temporary job and having promotion opportunities, or a managerial/supervision role, or pay rise is about 0.1.

about the working conditions of employees, a more precise evaluation of the two models could performed.

4.5.2. Psychological Well-being

Table 7 reports the coefficients and standard errors for the models for females, while Table 8 reports results for males.

In the first model, the results concerning the relationship between contractual conditions and psychological well-being are similar to those found when analysing self-assessed health. Some differences, however, are apparent. Considering part-time employment, the coefficient of the interaction term for preferring less hours has an unexpected positive sign for men, and for preferring more hours has an unexpected positive sign for both women and men (the sign of the partial effects of these interactions is negative in the model for health). For men, the coefficient of the interaction term *part-time*children* is statistically significant and has a high magnitude compared to *part-time*, *part-time*pref less hrs* and *part-time*pref more hrs*. This suggests that having children and having a part-time job has a particular negative influence on the psychological well-being of men. Considering having a temporary job, for females the coefficient of the variable *temporary job*children* is negative (while the partial effect for this variable in the model for self-assessed health is positive) and its magnitude, in absolute terms, is bigger than that of *temp-job* and *temp-job*high education.* These differences suggest that family structure plays a larger role for psychological well-being than for general self-assessed health.

In the second model, where we consider contractual conditions only, the results are generally similar to those found analysing self-assessed health. Even for this model, however, some differences can be seen.[36] The relationship between psychological well-being and working unpaid overtime hours appears to be more relevant than in the case of self-assessed health. For both women and men, the coefficient of *overtime hours* is non-negligible and is statistically significant. The magnitude of the coefficients of the variables relating to the work place continues to be large and that of those relating to the number of employees at work place is

[36] An interesting difference is that in the model for psychological well-being the coefficient of *not daytime* has the expected negative sign for both women and men.

still negligible. However, all of these coefficients are not statistically significant. For women, it is interesting to note that working at home has a positive relationship with health, while it seems to be negatively related to psychological well-being. The opposite holds for the variable *workplace other*, which is negatively related to health, and positively related to psychological well-being.

When we consider both contractual and working conditions simultaneously, the magnitude of the coefficients and their level of statistical significance change little compared to the relevant coefficients in the first and second model. Similar results found for self-assessed health in support or otherwise of Siegrist`s and Karasek`s models hold for psychological well-being.[37] [38]

In Annex 1 we report sensitivity analyses performed to assess the robustness of our results. First, the negative relationship between *unions* and health and psychological well-being suggested by our estimates is unexpected. To assess the robustness of this result, we define the variable *unions* in a different way (making reference to individual union membership instead of the presence of a union at the workplace) and re-estimate our models using this new variable. Secondly, we are aware that our results regarding Karasek`s and Siegrist`s frameworks could be affected by the choice of the specific variables we use to approximate the broad conceptual categories of "job demand" and "job reward" used in these models. Therefore, we evaluate alternative specifications of the interaction terms introduced to investigate these theoretical models, using *overtime hours* instead of *not daytime* to represent the conceptual category of "job demand", and *payrise* instead of *promotion opportunities* to represent the category of "job reward". Overall, the sensitivity analysis suggests that our results are not sensitive with regard to the definition of the *union* variable and the way we approximate the conceptual categories of "job demand" and "job reward".

[37] Notice that in the model for psychological well-being, the work satisfaction variables that are positive and highly statistically significant are *satisfaction_total pay* and *satisfaction security*.

[38] In a further specification (not reported here), we have run the model for GHQ not including the variables denoting satisfaction with work. The sign and the level of statistical significance of *not daytime*promotion opportunites* do not change**.** Therefore, the results in the main specification do not appear to be influenced by the presence of multicollinearity between the two sets of regressors related to Siegrist`s framework.

4.6. Final Remarks

Our study analyses the influence that contractual and working conditions have on self-assessed health and psychological well-being of employees using twelve waves (1991/92 – 2002/2003) of the British Household Panel Survey. The results suggest that both contractual and working conditions have some influence on health and psychological well-being of employees. These empirical results, therefore, provide some evidence against the neoclassical theory of equalizing differences. It is interesting to note that asymmetries among women and men exist with regard to these effects.

In our analysis we attempt to evaluate the role that preferences for the number of hours of work, the level of employability and family structure play in affecting the relationship between contractual/working conditions and health and psychological well-being. Our estimates show that being unsatisfied with the number of hours worked has a negative influence on the health of individuals who have a part-time job. Having a high level of employability (in our study approximated by having higher levels of education) appears to influence positively the health and psychological well-being of individuals with temporary job arrangements. Family structure appears to influence the health and well-being of workers with atypical contractual conditions. If workers have part-time or temporary work arrangements, the presence of children in the family is negatively related to health and psychological well-being (with the exception of women with temporary jobs).

The results concerning the relationship between contractual/working conditions and psychological well-being are similar to those found when analysing self-assessed health. Some differences, however, are observed. In particular family structure and working unpaid overtime hours appear to play a larger role for psychological well-being than for general self-assessed health.

Our results appear to provide limited support in favour of Karasek's model for women (but not for men). The interpretation of our findings with regards to Siegrist's model is more ambiguous, since this model does not receive direct support by the inclusion of the interaction term *no daytime*promotion opportunities* but it receives some indirect support considering the effects of the working satisfaction variables.

We assess the distinct effects of contractual and working conditions on health and psychological well-being, combining two distinct branches of the literature. As far as we are aware, previous studies have not considered *explicitly* the two factors jointly. Secondly, the analysis of the effects of contractual and working conditions on both self-assessed health and psychological well-being, for both women and men, allows us to highlight interesting asymmetries in these effects. Thirdly, most of the previous studies in the literature have focused on specific occupations (i.e. civil servants, nurses, etc.), and this makes the generalisation of their results to the entire workforce problematic. In our study we use the BHPS, a dataset containing a representative sample of the British population. Fourthly, the methodology we adopt for our analysis has several advantages compared to that one used by other studies in the literature. Indeed, we estimate dynamic panel data models, which allow us to account for the presence of individual specific effects and reduces concerns about reverse causality.

Our study suggests that, under certain circumstances, adverse contractual and working conditions can have a negative influence on the health and psychological well-being of workers in Great Britain. Improving the health and psychological well-being of workers could not only improve population health and reduce health inequalities, but could also have positive implications for the wider economy (Bartley et al., 2004). Workers with better health and psychological well-being, indeed, are likely to suffer less from illnesses limiting their working capacity and to have better work performance and less sickness leave. The implications at a macro-economic level of an improvement in the health conditions of workers can be particularly relevant in Great Britain, given that this country reports a low level of labour productivity compared to the other G7 countries (Office for National Statistics, 2008). Policy makers, therefore, should make some efforts to consider the cost, both at a social and economic level, of the health limitations that might derive from adverse contractual and working conditions.

Annex 1. Sensitivity analysis [39]

Unions

The estimates presented in Section 4.5.1 and 4.5.2 suggest a negative relationship between *unions* and health and psychological well-being. This result is unexpected since studies in the literature suggest that social support at work positively influences health and psychological well-being (Cheng et al., 2000; Godin and Kittel, 2004) and we consider the presence of unions at workplace as an element of social support for employees. To check the robustness of this result we consider individual union membership instead of the presence of unions at workplace. The former variable is also present in the BHPS. This variable is equal to 1 if a worker is a member of a union or association, and 0 otherwise. We re-estimate our models using this new variable for both self-assessed health and psychological well-being. The results for this alternative specification are extremely similar to those of the original specification. In particular, in the models for self-assessed health the coefficient has the same negative sign and the same level of significance as in the original models, while in the models for psychological well-being the coefficient remains negative but is not statistically significant. Therefore, our results appear not to be sensitive to the definition of the *union* variable.

Karasek`s and Siegrist`s frameworks

The results presented in Sections 4.5.1 and 4.5.2 regarding Karasek`s and Siegrist`s frameworks could be affected by the choice of the specific variables we use to approximate the broad conceptual categories of "job demand" and "job reward" used in these models. Therefore, we evaluate alternative specifications of the interaction terms introduced to investigate these theoretical models, using *overtime hours* instead of *not daytime* to represent the conceptual category of job demand, and *payrise* instead of *promotion opportunities* to represent the category of job reward. First, we re-estimate our models substituting the interaction term *not daytime*managerial supervision* with *overtime hours*managerial supervision*

[39] The results for these sensitivity checks are not reported here but are available on request

and *not daytime*promotion opportunities* with *overtime hours*promotion opportunities* in the model for both SAH and the GHQ, for both women and men. Generally, the results found with this new specification are very similar to those found with the original one, for both females and males, and for both self-assessed health and psychological well-being. In all cases the interaction terms are not statistically significant (as in the original specification), and some of them (*overtime hours*promotion opportunities* in the model for SAH for females, and *overtime hours*managerial supervision* and *overtime hours*promotion opportunities* in the model for the GHQ for males) change sign. Secondly, we re-estimate our models substituting the interaction term *not daytime*promotion opportunities* with *not daytime*payrise* and *overtime hours*promotion opportunities* with *overtime hours*payrise*. The results are very similar to those in the original models, for both females and males, and for both self-assessed health and psychological well-being. In all the cases the interaction terms are not statistically significant (as in the original specification), and just one interaction term (*overtime hours*payrise* in the model for the GHQ for female) changes sign. Given that none of these effects are statistically significant, a change in sign is not of great concern. These sensitivity checks suggest that our specification is not sensitive to the way we approximate the broad conceptual categories of "job demand" and "job reward", particularly for the models for SAH.

Table 1. Variable definitions

Self-assessed health	1 if “poor or very poor”, 2 if “fair”, 3 if “good or very good”, 4 if “excellent” health
GHQ	Psychological well-being (0-36, where 0 is the worst level, 36 the best)
CONTRACTUAL CONDITIONS	
part-time job	1 if current job is part-time, 0 otherwise
temp job	1 if current job is temporary, 0 otherwise
WORKING CONDITIONS	
Demanding job conditions	
not daytime	1 if not working during the day or having rotation shift, 0 otherwise
overtime hours	number of overtime hours in normal week
control	
unions	1 if there is a union or staff association at workplace, 0 otherwise
managerial supervision	1 if managerial or supervision duties, 0 otherwise
reward	
payrise	1 if pay includes annual increment, 0 otherwise
promotion opportunities	1 if opportunities of promotion in current job, 0 otherwise
working environment	
workplace home	1 if working at home, 0 otherwise
workplace travel	1 if working travelling, 0 otherwise
workplace other	1 if NOT working at the employeer, home or travelling, 0 otherwise
employed at workplace	number of people employed at the workplace
WORK SATISFACTION	
satisfaction_total pay	1 if satisfied with total pay of the job, 0 otherwise
satisfaction_security	1 if satisfied with security of the job, 0 otherwise
satisfaction_work itself	1 if satisfied with the work itself, 0 otherwise
preference less hrs	1 if preferred working fewer hours, 0 otherwise
preference more hrs	1 if preferred working more hours, 0 otherwise
CONTROLS	
age	Age in years at 1st December of current wave
divsep	1 if divorced or separated, 0 otherwise
nevermar	1 if never maried, 0 otherwise
widowed	1 if widowed, 0 otherwise
race	0 if white, 1 otherwise
household size	n. of people in the household including the respondent
children	1 if in the household there is at least one child (less than 16), 0 otherwise
income	log of Annual labour income (in pounds)
high education	1 if people have a qualification equal or superior to A level, 0 otherwise
lower social class	1 if "skilled manual" “armed forces”, "partly skilled", "unskilled”, 0 otherwise.

Table 2. Regressors` mean, standard deviation, minimum and maximum

	Females N= 23,309				Males N= 22,349			
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
CONTRACTUAL CONDITIONS								
part-time job	0.352	0.478	0	1	0.046	0.210	0	1
temp job	0.070	0.256	0	1	0.050	0.218	0	1
WORKING CONDITIONS								
Demanding job conditions								
not daytime	0.350	0.477	0	1	0.297	0.457	0	1
overtime hours	1.540	4.349	0	71	2.296	5.485	0	80
control								
unions	0.525	0.499	0	1	0.501	0.500	0	1
managerial supervision	0.325	0.468	0	1	0.426	0.495	0	1
reward								
payrise	0.499	0.500	0	1	0.435	0.496	0	1
promotion opportunities	0.454	0.498	0	1	0.555	0.497	0	1
working environment								
workplace home	0.011	0.104	0	1	0.008	0.088	0	1
workplace travel	0.032	0.177	0	1	0.132	0.338	0	1
workplace other	0.047	0.212	0	1	0.096	0.295	0	1
employed at workplace	212.189	318.103	1	1000	257.645	329.004	1	1000
WORK SATISFACTION								
satisfaction_total pay	4.865	1.627	1	7	4.655	1.616	1	7
satisfaction_security	5.442	1.540	1	7	5.156	1.628	1	7
satisfaction_work itself	5.572	1.337	1	7	5.404	1.376	1	7
preference less hrs	0.309	0.462	0	1	0.356	0.479	0	1
preference more hrs	0.082	0.274	0	1	0.075	0.263	0	1
CONTROLS								
age	37.604	11.598	15	76	37.470	11.686	16	81
divsep	0.085	0.279	0	1	0.043	0.203	0	1
nevermar	0.176	0.381	0	1	0.216	0.412	0	1
widowed	0.019	0.138	0	1	0.005	0.070	0	1
race	0.027	0.163	0	1	0.028	0.164	0	1
household size	2.989	1.196	1	10	3.085	1.264	1	11
children	0.370	0.483	0	1	0.375	0.484	0	1
(log) income	9.041	0.832	0.693	12.472	9.591	0.759	0	13.082
high education	0.514	0.500	0	1	0.589	0.492	0	1
lower social class	0.279	0.448	0	1	0.470	0.499	0	1

Table 3. Correlated random effects model for self-assessed health. Estimated coefficients. (FEMALES)

	contractual conditions only		working conditions only		full model	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
contractual conditions						
part-time job	0.066 *	0.037			0.074 **	0.037
temp job	-0.036	0.066			-0.020	0.067
part-time*pref less hrs	-0.115 **	0.051			-0.113 **	0.052
part-time*pref more hrs	-0.089	0.068			-0.091	0.068
temp job*high educat	0.114	0.072			0.117	0.072
part-time*children	-0.013	0.044			-0.015	0.044
temp job*children	0.056	0.072			0.065	0.072
working conditions						
working environment						
workplace home			0.228 **	0.095	0.220 **	0.095
workplace travel			-0.049	0.053	-0.052	0.053
workplace other			-0.056	0.043	-0.066	0.043
employed at workplace			-1E-05	3E-05	-2E-06	3E-05
demanding job conditions						
not daytime			-0.022	0.022	-0.040	0.028
overtime hours			-0.001	0.002	-3E-04	0.002
control and social support						
unions			-0.058 **	0.024	-0.054 **	0.024
managerial supervision			0.007	0.023	0.011	0.023
reward						
payrise			0.019	0.020	0.019	0.020
promotion opportunities			-0.041 **	0.020	-0.047 *	0.024
satisfaction_total pay					-0.008	0.006
satisfaction_security					0.015 **	0.006
satisfaction_work itself					0.022 ***	0.007
pref less hrs	-0.041 *	0.023			-0.036	0.023
pref more hrs	0.038	0.055			0.045	0.055
not daytime*manag-sup					0.021	0.042
no day time*prom opp					0.019	0.038
Log Likelihood	-22417.2		-22409.26		-22396.224	
N	23309		23309		23309	

***, ** and * denote significance at 1%, 5% and 10% levels respectively.

The other regressors included in our model are: the lagged and the initial value of SAH, age, age^2, age^3, divsep, nevermar, widowed ,race, household size, children, income, high education , lower social class.

The interaction terms "no daytime*managerial/supervision" and "no daytime*promotion opportunities" are not introduced contemporaneously in the third model. We estimate a model where we introduce "no daytime*managerial/supervision" only and another model where we introduce "no daytime*promotion opportunities" only. Since the results of these two models are extremely similar, we present the results of the latter model, and report only the results related to the variable "no daytime*managerial/supervision" for the former model.

Table 4. Correlated random effects model for self-assessed health. Estimated coefficients. (MALES)

	contractual conditions only			working conditions only			full model		
	Coef.		Std. Err.	Coef.		Std. Err.	Coef.		Std. Err.
contractual conditions									
part-time job	0.189	***	0.068				0.176	***	0.069
temp job	0.003		0.080				0.051		0.080
part-time*pref less hrs	-0.096		0.151				-0.107		0.151
part-time*pref more hrs	-0.053		0.106				-0.053		0.106
temp job*high educat	0.030		0.090				0.026		0.090
part-time*children	-0.078		0.140				-0.073		0.140
temp job*children	-0.221	**	0.103				-0.215	**	0.103
working conditions									
working environment									
workplace home				-0.006		0.111	-0.023		0.111
workplace travel				0.055	*	0.032	0.051		0.032
workplace other				0.012		0.034	0.008		0.034
employed at workplace				9E-05	**	4E-05	1E-04	***	4E-05
demanding job conditions									
not daytime				0.031		0.024	0.031		0.030
overtime hours				-5E-04		0.002	-4E-04		0.002
control and social support									
unions				-0.027		0.024	-0.015		0.024
managerial supervision				-0.011		0.024	-0.017		0.024
reward									
payrise				0.005		0.020	0.004		0.024
promotion opportunities				0.014		0.021	-0.003		0.021
satisfaction_total pay							-0.009		0.007
satisfaction_security							0.035	***	0.006
satisfaction_work itself							0.033	***	0.007
pref less hrs	-0.060	***	0.021				-0.045	***	0.210
pref more hrs	-0.073	*	0.038				-0.066	*	0.038
not daytime*manag-sup							0.010		0.044
no day time*prom opp							-0.023		0.041
Log Likelihood	-20545.725			-22409.26			-20510.123		
N	22349			22349			22349		

***, ** and * denote significance at 1%, 5% and 10% levels respectively.

The other regressors included in our model are: the lagged and the initial value of SAH, age, age^2, age^3, divsep, nevermar, widowed ,race, household size, children, income, high education , lower social class.

The interaction terms "no daytime*managerial/supervision" and "no daytime*promotion opportunities" are not introduced contemporaneously in the third model. We estimate a model where we introduce "no daytime*managerial/supervision" only and another model where we introduce "no daytime*promotion opportunities" only. Since the results of these two models are extremely similar, we present the results of the latter model, and report only the results related to the variable "no daytime*managerial/supervision" for the former model.

Table 5. Correlated random effects model for self-assessed health. Partial effect on probability of reporting excellent health. (FEMALES)

	contractual conditions only		working conditions only		full model	
	Part Eff	Std. Err.	Part Eff	Std. Err.	Part Eff	Std. Err.
contractual conditions						
part-time job	0.023	0.021			0.026	0.023
temp job	-0.013	0.020			-0.007	0.022
part-time*pref less hrs	-0.062	0.019			-0.059	0.019
part-time*pref more hrs	-0.020	0.016			-0.019	0.016
temp job*high educat	0.052	0.028			0.055	0.028
part-time*children	-0.007	0.017			-0.008	0.017
temp job*children	0.021	0.031			0.024	0.031
working conditions						
working environment						
workplace home			0.089	0.036	0.074	0.062
workplace travel			-0.019	0.021	-0.018	0.022
workplace other			-0.022	0.017	-0.024	0.020
employed at workplace			-5E-05	2E-05	-5E-05	2E-05
demanding job conditions						
not daytime			-0.009	0.002	-0.020	0.015
overtime hours			-3E-04	1E-03	-2E-04	2E-03
control and social support						
unions			-0.015	0.007	-0.019	0.014
managerial supervision			0.003	0.009	0.004	0.008
reward						
payrise			0.008	0.008	4E-04	0.008
promotion opportunities			-0.023	0.009	-0.014	0.011
satisfaction_total pay					-0.018	0.016
satisfaction_security					0.032	0.023
satisfaction_work itself					0.047	0.030
pref less hrs	-0.016	0.009			-0.013	0.011
pref more hrs	0.015	0.022			0.016	0.026
not daytime*manag-sup					0.010	0.015
no day time*prom opp					-0.011	0.013
N	23309		23309		23309	

***, ** and * denote significance at 1%, 5% and 10% levels respectively.

The other regressors included in our model are: the lagged and the initial value of SAH, age, age^2, age^3, divsep, nevermar, widowed ,race, household size, children, income, high education , lower social class.

The partial effects indicates the change in the probability of reporting excellent health due to a marginal change for continuous variables and to a discrete change for binary variables

We compute the partial effects for a hypothetical representative agent with "average characteristics". We attribute the mean value to the covariates that are continuous and the modal value to the covariates that are categorical.
To compute the partial effect of "part-time*pref less hrs", "part-time*pref more hrs" and "part-time*children" we make reference to the representative individual with a part-time job. To compute the partial effect of "temp job*high educat" and "temp job*children" we make reference to the representative individual with a temporary job.

The interaction terms "no daytime*managerial/supervision" and "no daytime*promotion opportunities" are not introduced contemporaneously in the third model. We estimate a model where we introduce "no daytime*managerial/supervision" only and another model where we introduce "no daytime*promotion opportunities" only. Since the results of these two models are extremely similar, we present the results of the latter model, and report only the results related to the variable "no daytime*managerial/supervision" for the former model.

Table 6. Correlated random effects model for self-assessed health. Partial effect on probability of reporting excellent health. (MALES)

	contractual conditions only		working conditions only		full model	
	Part Eff	Std. Err.	Part Eff	Std. Err.	Part Eff.	Std. Err.
contractual conditions						
part-time job	0.075	0.029			0.070	0.026
temp job	0.001	0.031			0.020	0.030
part-time*pref less hrs	-0.060	0.058			-0.061	0.059
part-time*pref more hrs	-0.048	0.038			-0.046	0.038
temp job*high educat	0.032	0.035			0.034	0.035
part-time*children	-0.032	0.054			-0.031	0.055
temp job*children	-0.091	0.042			-0.089	0.042
working conditions						
working environment						
workplace home			-0.002	0.044	-0.009	0.043
workplace travel			0.021	0.013	0.020	0.013
workplace other			0.005	0.013	0.003	0.013
employed at workplace			6E-05	2E-05	5E-05	7E-05
demanding job conditions						
not daytime			0.012	0.009	0.012	0.012
overtime hours			-5E-04	1E-03	-2E-04	9E-04
control and social support						
unions			-0.011	0.009	-0.005	0.010
managerial supervision			-0.007	0.009	-0.005	0.002
reward						
payrise			0.002	0.008	0.001	0.009
promotion opportunities			0.006	0.008	-0.001	0.008
satisfaction_total pay					-0.021	0.016
satisfaction_security					0.081	0.030
satisfaction_work itself					0.077	0.031
pref less hrs	-0.024	0.009			-0.019	0.010
pref more hrs	-0.029	0.017			-0.026	0.019
not daytime*manag-sup					-0.004	0.016
no day time*prom opp					-0.011	0.019
N	22349		22349		22349	

***, ** and * denote significance at 1%, 5% and 10% levels respectively.

The other regressors included in our model are: the lagged and the initial value of SAH, age, age^2, age^3, divsep, nevermar, widowed ,race, household size, children, income, high education , lower social class.

The partial effects indicates the change in the probability of reporting excellent health due to a marginal change for continuous variables and to a discrete change for binary variables

We compute the partial effects for a hypothetical representative agent with "average characteristics". We attribute the mean value to the covariates that are continuous and the modal value to the covariates that are categorical.
To compute the partial effect of "part-time*pref less hrs", "part-time*pref more hrs" and "part-time*children" we make reference to the representative individual with a part-time job. To compute the partial effect of "temp job*high educat" and "temp job*children" we make reference to the representative individual with a temporary job.

The interaction terms "no daytime*managerial/supervision" and "no daytime*promotion opportunities" are not introduced contemporaneously in the third model. We estimate a model where we introduce "no daytime*managerial/supervision" only and another model where we introduce "no daytime*promotion opportunities" only. Since the results of these two models are extremely similar, we present the results of the latter model, and report only the results related to the variable "no daytime*managerial/supervision" for the former model.

Table 7. Correlated random effects model for psychological well-being. Estimated coefficients. (FEMALES)

	contractual conditions only			working conditions only			full model		
	Coef.		Std. Err.	Coef.		Std. Err.	Coef.		Std. Err.
contractual conditions									
part-time job	0.127		0.126				0.078		0.128
temp job	-0.035		0.244				0.012		0.245
part-time*pref less hrs	-0.356	*	0.191				-0.349	*	0.191
part-time*pref more hrs	0.327		0.252				0.300		0.251
temp job*high educat	0.013		0.262				0.020		0.262
part-time*children	-0.080		0.150				-0.077		0.150
temp job*children	-0.141		0.262				-0.115		0.262
working conditions									
working environment									
workplace home				-0.429		0.315	-0.499		0.315
workplace travel				-0.205		0.183	-0.197		0.183
workplace other				0.093		0.151	0.088		0.151
employed at workplace				2E-05		1E-04	5E-05		1E-04
demanding job conditions									
not daytime				-0.001		0.073	0.068		0.097
overtime hours				-0.022	***	0.008	-0.020	**	0.008
control and social support									
unions				-0.235	***	0.077	-0.219	***	0.077
managerial supervision				0.042		0.077	0.044		0.077
reward									
payrise				0.088		0.071	0.065		0.071
promotion opportunities				-0.188	***	0.070	-0.137	***	0.084
satisfaction_total pay							0.056	***	0.021
satisfaction_security							0.062	***	0.023
satisfaction_work itself							-0.008		0.026
pref less hrs	-0.230	***	0.083				-0.195	**	0.084
pref more hrs	-0.372	*	0.206				-0.336		0.206
not daytime*manag-sup							0.052		0.147
no day time*prom opp							-0.212		0.135
N	23309			23309			23309		

***, ** and * denote significance at 1%, 5% and 10% levels respectively.

The other regressors included in our model are: the lagged and the initial value of GHQ, age, age^2, age^3, divsep, nevermar, widowed ,race, household size, children, income, high education , lower social class.

The interaction terms “no daytime*managerial/supervision” and “no daytime*promotion opportunities” are not introduced contemporaneously in the third model. We estimate a model where we introduce “no daytime*managerial/supervision” only and another model where we introduce “no daytime*promotion opportunities” only. Since the results of these two models are extremely similar, we present the results of the latter model, and report only the results related to the variable “no daytime*managerial/supervision” for the former model.

Table 8. Correlated random effects model for psychological well-being. Estimated coefficients. (MALES)

	contractual conditions only		working conditions only		full model	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
contractual conditions						
part-time job	0.455 **	0.196			0.444 **	0.199
temp job	0.137	0.245			0.299	0.247
part-time*pref less hrs	0.205	0.471			0.189	0.470
part-time*pref more hrs	0.251	0.326			0.238	0.326
temp job*high educat	0.248	0.275			0.245	0.275
part-time*children	-0.811 *	0.419			-0.853 **	0.419
temp job*children	-0.587 *	0.321			-0.566 *	0.321
working conditions						
working environment						
workplace home			-0.092	0.317	-0.134	0.317
workplace travel			0.116	0.086	0.109	0.086
workplace other			-0.025	0.096	-0.030	0.096
employed at workplace			1E-04	9E-05	1E-04	9E-05
demanding job conditions						
not daytime			-0.100	0.064	-0.045 *	0.098
overtime hours			-0.010 *	0.006	-0.011 *	0.006
control and social support						
unions			-0.217 ***	0.063	-0.174 ***	0.063
managerial supervision			-0.036	0.066	-0.056	0.066
reward						
payrise			0.084	0.059	0.064	0.059
promotion opportunities			0.006	0.061	0.025	0.071
satisfaction_total pay					0.015	0.019
satisfaction_security					0.113 ***	0.019
satisfaction_work itself					0.064 ***	0.023
pref less hrs	-0.057	0.061			-0.007	0.062
pref more hrs	0.052	0.117			0.085	
not daytime*manag-sup					-0.020	0.127
no day time*prom opp					-0.190	0.124
N	22349		22349		22349	

***, ** and * denote significance at 1%, 5% and 10% levels respectively.
The other regressors included in our model are: the lagged and the initial value of GHQ, age, age^2, age^3, divsep, nevermar, widowed ,race, household size, children, income, high education , lower social class.
The interaction terms “no daytime*managerial/supervision” and “no daytime*promotion opportunities” are not introduced contemporaneously in the third model. We estimate a model where we introduce “no daytime*managerial/supervision” only and another model where we introduce “no daytime*promotion opportunities” only. Since the results of these two models are extremely similar, we present the results of the latter model, and report only the results related to the variable “no daytime*managerial/supervision” for the former model.

Chapter 5

Concluding Remarks

This thesis has been developed in the context of the Joint Doctoral Programme at the University of York and the University of Bologna. Since my research activity has been conducted partly in Great Britain and party in Italy, my thesis deals with topics which are relevant in these two countries.

In the following section I give a short summary of the main research findings of Chapter 2, Chapter 3 and Chapter 4. Moreover, I discuss some research topics that have the potential for further development.

In Chapter 2 "*Market Structure and Technology: Evidence from the Italian National Health Service*" we test empirically the existence of the relationship between market structure and technological profiles claimed by Sutton's theory in a specific economic framework, that of hospital care services provided by the Italian National Health Service (NHS). Our results provide evidence in favour of the empirical predictions by Sutton. In markets where technological intensity is low the lower bound to market concentration converges monotonically to zero when the market size increases, for any level of product homogeneity. Conversely, in markets where technological intensity is high the lower bound to concentration converges to a positive value different from zero when the market size increases, while the lower bound increases from zero with the level of product homogeneity.

Our results offer some useful indications for policy–makers about the functioning of the Italian NHS. To enhance the relevance of this evidence, however, we could investigate more thoroughly the institutional setting affecting the hospital care sector in Italy. The greater autonomy provided to regions by the reforms of the 1990s has increased the differences across regions in terms of organization of the regional services and funding of providers (France et al., 2005). These regional differences could have an influence on the relationship between market structure and technological profiles that we investigate. However, since each region has virtually its own organizational model (Mapelli, 2000), the analysis taking the institutional setting into consideration appears too complex to be included in this thesis. Moreover, we could further exploit the structure of the hospital care markets identified here through the Elzinga and Hogarty (1978) approach in order to examine the potential effect of market structure on hospital technical efficiency. This would provide an empirical contribution regarding the Italian NHS to the literature on competition and hospital performance (i.e. Dalmau-Matarrodona and Puig-Junoy, 1998; Robinson and Luft, 1985). This sort of analysis would require the use of hospital performance measures such as, for example, technical efficiency scores computed through the application of the Data Envelopment Analysis (DEA) technique. The application of the DEA technique would require the use of data about hospital inputs and outputs, not available in the dataset utilized in this thesis. Therefore, we leave this analysis for future research.

Chapter 3 "*The Geography of Hospital Admissions in a National Health Service with Patient Choice: Evidence from Italy*" deals with patient mobility across LHAs in Italy. To analyse the determinants of this phenomenon, we estimate gravity equations in multiplicative form using a Poisson pseudo maximum likelihood method, as proposed by Santos-Silva and Tenreyro (2006). In most of the cases our results are consistent with the predictions of the gravity model. In particular, patient flows are affected positively by the number of enrolees of the LHA of origin and of destination, and negatively by the distance among LHAs. The contiguity of the LHAs plays a positive role for patient flows while the presence of institutional barriers to mobility a negative one. For complex surgery, the presence of autonomous hospitals has a negative (positive) role on exits (inflows). In general, the gravity model can be regarded as a good framework for explaining patient mobility for hospital care treatment across LHAs in Italy. The specification tests adopted in our analysis, however, suggest that only the models for complex surgery and cancer are correctly specified. The decision process governing patient mobility for basic surgery and basic medical treatments could have peculiarities requiring the development of different models, given the specific characteristics of these treatments. In our analysis we have not considered the role played by the waiting times in the hospital choice, which has been shown in the literature to be of great importance for the choice of hospitals. Italian data about waiting times are available just at the regional level. Therefore, it has not been possible to include this variable in the present analysis (if we had included them in our regression model, their effects would have been captured by the regional dummies). Perhaps the omission of this variable is particularly relevant for the case of basic surgery and basic medicine. This could be a path for future research when more disaggregated data on waiting times become available.

Another possible future development of our model relates to the role played by the hospital size in the patients` mobility phenomenon. In the literature about hospital choice this variable is often assumed to be an indicator of the quality of hospitals, and thus an attraction factor. This assumption, however, should be viewed with caution, since, according to some empirical considerations, this variable can be also thought of as a push factor. In hospitals whose size is above a certain threshold there could be congestion problems, doctors and nurses may provide less care to patients and, in general, more managerial problems can arise. These factors could decrease the quality of the services offered by larger hospitals and give potential patients the incentive to refer to other hospitals (Taroni, 2001). Therefore, as a future extension of our work, we could verify if our data do corroborate the hypothesis of a U shape (inverse U-shape) relationship between the dimension of the hospitals and the exit flows (inflows).

Under the econometrical point of view, in the future we could try to perform some formal tests to detect the presence of spatial autocorrelation in the data regarding patient flows (using, for instance, the indexes developed by Moran (1950) and Geary (1954)). In case we detect the presence of spatial autocorrelation in the data, we could include additional terms in our model, either in the deterministic part of the model or in the stochastic one, to account for it

Chapter 4 "*Contractual conditions, Working conditions, Health and Well-being in the British Household Panel Survey*" analyses the influence that contractual and working conditions have on self-assessed health and psychological well-being of employees using the BHPS. The results suggest that both contractual and working conditions have some influence on health and psychological well-being of employees. Our estimates show that being unsatisfied with the number of hours worked has a negative influence on the health of individuals who have a part-time job, while having a high level of employability has a positive influence on the health and psychological well-being of individuals with temporary job arrangements. Family structure also appears to influence the health and well-being of workers with atypical contractual conditions and it appears to play a larger role for psychological well-being than for the general self-assessed health. As far as the two theoretical models investigating why working conditions affect health and psychological well-being, our results appear to provide some support in favour of Karasek's model for women (but not for men), while the interpretation of our findings with regards to Siegrist`s model is more ambiguous, since some of the factors considered are in favour of the framework while others not.

The analysis could be extended in future research in several directions. First, previous studies have shown that factors having a negative effect on health and psychological well-being can be different from those having a positive effect (Bartley et al., 2004; Borg et al., 2000; Lindberg et al, 2006). Accordingly, we could verify if our data support the hypothesis that contractual and working conditions have asymmetric effects on increases and decreases in the levels of health and psychological well-being. Secondly, it has been shown that the nature of the impact of socio-economic conditions on health may vary considerably over the life cycle (Smith, 2004). Considering this aspect, in the future we could attempt to develop a model for the consequences of contractual and working conditions on health and psychological well-being over the life cycle of workers. Considering the life cycle of workers we could take into consideration if the age of the children that the workers take care for plays any role in the way contractual and working conditions influence health and psychological well being. Thirdly, we could evaluate if characteristics at aggregate (and not individual) level, such as the percentage of workers in a certain area that have atypical contractual arrangements, affect the relationship between contractual/working conditions and health and psychological well-being. Shields and Wheatley Price (2005), for example, show that the negative effect of unemployment on psychological well-being is greater in areas characterized by low employment deprivation. This result could be due, for instance, to stigma effects linked to unemployment status (Clark, 2003). Similar stigma effects could be linked to atypical contractual arrangements in areas where the majority of workers have traditional contractual arrangements. To evaluate this third hypothesis we would need additional information about contractual conditions at area level. Currently, these are not available in the BHPS.

In recent years the availability of new data sources and the development of advanced econometric methods have offered huge scope for testing theoretical and empirical predictions regarding health and health care systems in different national contexts. The work contained in this thesis has attempted to contribute to this agenda in the context of two specific countries: Italy and Britain.

Abbreviations

BHPS: British Household Panel Survey

BM: basic medicine

BS: basic surgery

CA: cancer

CWW: compressed work week

DE: delivery

DEA: Data Envelopment Analysis

DRGs: Diagnosis Related Groups

ECHP: European Community Household Panel

EM: emergencies

GAZEL: Electricité De France–Gaz De France

GHQ: General Health Questionnaire

HAPIEE: Health, Alcohol and Psychosocial factors in Eastern Europe Study

HMOs: health maintenance organizations

IRCCS: Istituti di Ricovero e cura a carattere scientifico

ISTAT: Italian National Institute of Statistics

LHAs: Local Health Authorities

NHS: National Health Service

NPHS: Statistics Canada’s National Population Health Survey

PPML: Poisson Pseudo Maximum Likelihood

SAH: Self-Assessed Health

SSN: Servizio Sanitario Nazionale

WLS: Wisconsin Longitudinal Study

References

Adams P., Hurd M. D., McFadden D., Merrill A., Ribeiro T., (2003), "Healthy, wealthy, and wise? Tests for direct causal paths between health and socioeconomic status", *Journal of Econometrics*, 112, 3-56.

Ahmed-Little Y., (2007), "Implications of shift work for junior doctors", *British Medical Journal*, 334, 777-778.

Ai C., Norton E.C., (2003), "Interaction terms in logit and probit models", *Economic Letters*, 80, 123-129.

Andersen I., Burr H., Kristensen T.S., Gamborg M., Osler M., Prescott E., Diderichsen F., (2004), "Do factors in the psychosocial work environment mediate the effect of socioeconomic position on the risk of myocardial infarction? Study from the Copenhagen Centre for Prospective Population Studies", *Occupational and Environmental Medicine*, 61, 886-892.

Anderson J.A., van Wincoop E., (2003), "Gravity with gravitas: a solution to the border puzzle", *The American Economic Review*, 93, 170-192.

Anderson S.P., DePalma A., Thisse J.F., (1992), *Discrete choice theory of product differentiation*, Cambridge, MIT Press.

Aronsson G., (2001), "A new employment contract", *Scandinavian Journal of Work, Environment and Health*, 27, 361-4.

Artazcoz L., Benach J., Borrell C., Cortes I., (2005), "Social inequalities in the impact of flexible employment on different domains of psychological health", *Journal of Epidemiology and Community Health*, 59, 761-767.

Baltagi B.H., Egger P. and Pfaffermayr M., (2007), "Estimating Models of Complex FDI: Are There Third-Country Effects? Department of Economics", Journal of Econometrics, 140 (1), 260-281.

Bain J.S., (1956), *Barriers to New Competition*, Cambridge, Harvard University Press.

Bambra C., Petticrew M., Whitehead M., Sowden A., Akers J., (2007), "The health and wellbeing effects of changing the organisation of shift work: a systematic review”, Public Health Research Consortium, Project Final Report, University of York.

Bardasi E., Francesconi M., (2004), “The impact of atypical employment on individual well-being: evidence from a panel of British workers”, *Social Science and Medicine*, 58, 1671-1688.

Bartley M., Sacker A., Clarke P., (2004), “Employment status, employment conditions, and limiting illness: prospective evidence from the British household panel survey 1991–2001”, *Journal of Epidemiology and Community Health*, 58, 501-506.

Bashshur R.L., Shannon G.W., Metzner C.A., (1971), “Some ecological differentials in the use of medical services”, *Health Services Research*, 6, 61-75.

Bauer T.N., Truxillo, D.M., (2000), “Temp-to-permanent employees: a longitudinal study of stress and selection success”, *Journal of Occupational Health Psychology*, 5, 337-346.

Benach J., Gimeno D., Benavides F.G., Martinez J.M, Tornè M., (2004), “Types of employment and health in the European Union: changes from 1995 to 2000”, *European Journal of Public Health*, 14, 314-321.

Benavides F.G., Benach J., Diez-Roug AV., Roman C., (2000), “How do types of employment relate to health indicators? Findings from the second European Survey on Working Conditions”, *Journal of Epidemiology and Community Health*, 54, 494-501.

Benzeval M., Taylor J., Judge K., (2000), “Evidence on the relationship between low income and poor health: is the Government doing enough?”, *Fiscal Studies*, 21, 375–399.

Bessho S., (2003), “What are the determinants of hospital choice? Time, costs, option demands and communication”, The University of Tokyo, *mimeo*.

Blackstone E.A., Fuhr J.P., (2000), “Economic issues in hospital mergers and health systems”, *Journal of Science and Health Policy*, 1, 82-109.

Blank J.L.T., Vogelaar I., (2004), “Specifying technical change: a research on the nature of technical change in Dutch hospital industry”, *The Quarterly Review of Economics and Finance*, 44, 448-463.

Bohle, P., Quinlan M., Mayhew C., (2001), "The health and safety effects of job insecurity: an evaluation of the evidence", *Economic and Labour Relations Review*, 12, 32-60.

Borg V., Kristensen T.S., Burr H., (2000), "Work environment and changes in self-rated health: a five year follow-up study", *Stress Medicine*, 16, 37-47.

Brown C., 1980, "Equalizing Differences in the Labor Market", *The Quarterly Journal of Economics*, 94(1), 113-134.

Burchell B.J., (1994), "The effects of labour market position, job insecurity and unemployment on psychological health", in: Duncan, G., Marsh C., Vogler C. (eds.): *Social Change and the Experience of Unemployment*, Oxford, Oxford University Press, 188-212.

Burchell B.J., (1999), "The unequal distribution of job insecurity, 1966-86", *International Review of Applied Economics*, 13, 437-458.

Burns L.B., Wholey D.R., (1992), "The impact of physician characteristics in conditional choice models for hospital care", *Journal of Health Economics*, 11, 43-62.

Buzzacchi L., Valletti T.M. (2003), "Firm size distribution: testing the "Independent Submarkets Model" in the Italian motor insurance industry", London School of Economics, *mimeo*.

Cameron A. C., Trivedi P. K., (2005), *Microeconometrics: Methods and Applications*, Cambridge University Press, New York.

Cappelli P., Bassi L., Katz H., Knoke D., Osterman P., Useem M., (1997), *Change at work*, New York, Oxford University Press.

Capps C., Dranove D., Greenstein S., Satterthwaite M., (2001), "The silent majority fallacy of the Elzinga-Hogarty criteria: a critique and new approach to analyzing hospital mergers", *NBER Working Paper*, 8216, Cambridge, MA.

Capps C., Dranove D., Greenstein S., Satterthwaite M., (2002), "Antitrust policy and hospital mergers: recommendations for a new approach", *Antitrust Bulletin*, 17, 677-714.

Capps C., Dranove D., Greenstein S., Satterthwaite M., (2003), "Geographic market definition in hospital merger cases", (Joint Statement) *Federal Trade*

Commission and Department of Justice Hearings on Health Care and Competition Law and Policy.

Caveas R.E., Porter M.E., (1978), “Market structure, oligopoly, and stability of market shares,” *Journal of Industrial Economics*, 26, 289–313.

Cellini R., Pignataro G., Rizzo I., (2000), “Competition and efficiency in health care: an analysis of the Italian case”, *International Tax and Public Finance*, 7, 503–519.

Chamberlain G., (1984), “Panel data”, in Griliches Z., Intriligator M.D. (eds) *Handbook of Econometrics*, 1, North Holland, Amsterdam, 1247–1318.

Chapman K.S., Hariharan G., (1994), “Controlling for causality in the link from income to mortality”, *Journal of Risk and Uncertainty*, 8, 85-93.

Cheng I., Wall H., (2004), “Controlling for heterogenerity in gravity models of trade and integration”, The Federal Reserve Bank, *Working Paper Series*.

Cheng Y., Kawachi I., Coakley E.H., Schwartz J., Colditz G., (2000), “American women: prospective study characteristics and health functioning in association between psychosocial work”, *British Medical Journal*, 320, 1432-6.

Clark A., (2003), “Unemployment as a social norm: psychological evidence from panel data”, *Journal of Labour Economics*, 21, 323-351.

Congdon, P., (1991) “An application of general linear modelling to migration in London and the South East”, in Stillwell J.C.H, Congdon, P. (eds.) *Migration Models: Macro and Micro Perspectives*, Belhaven Press, London, 113-136.

Congdon P., (2001), “The development of gravity models for hospital patient flows under system change: a Bayesian modelling approach”, *Health Care Management Science*, 4, 289-304.

Contoyannis P., Rice N., (2001), “The impact of health on wages: evidence from the British Household Panel Survey”, *Empirical Economics*, 26, 599-622.

Contoyannis P., Jones A.M., Rice N., (2004), “The dynamics of health in the British Household Panel Survey”, *Journal of Applied Econometrics*, 19, 473-503.

Dalgin M., Mitra D., Trindade V., (2004), “Inequality, nonhomotetic preferences and trade: a gravity approach”, *NBER Working Paper*, 10800, Cambridge, MA.

Dalmau-Matarrodona E., Puig-Junoy J., (1998), “Market structure and hospital efficiency: evaluating potential effects of deregulation in a National Health Service”, *Review of Industrial Organization*, 13, 447-66.

Datta Gupta N., Kristensen N., (2007), “Work environment satisfaction and employee health: panel evidence from Denmark, France and Spain, 1994-2001”, *European Journal of Health Economics.*

Davidson R., MacKinnon J.G., (1993), *Estimation and Inference in Econometrics*, Oxford, Oxford University Press.

Degli Esposti G., Rimondi M., Virgilio G., Ugolini C., (1996), “Matrici di mobilità per DRGs: analisi descrittiva ed applicazioni per la programmazione e le politiche sanitarie regionali”, *Mecosan*, 19, 53-62.

Dixit A., Stiglitz J.E., (1977), “Monopolistic competition and optimum product diversity”, *American Economic Review*, 67, 297-308.

Donatini A., Rico A., D’Ambrosio M.G., Lo Scalzo A., Orzella L., Cicchetti A., Profili S., (2001), *Health Care System in Transition*, *Italy*, European Observatory on Health Care Systems.

Dranove D., Shanley M., (1989), “A note on the relationship aspects of hospital markets definitions”, *Journal of Health Economics*, 8, 473-478.

Dranove D., Shanley M., Simon C., (1992), “Is hospital competition wasteful?”, *RAND Journal of Economics*, 23, 247-262.

Egger P., Pfaffermayr M., (2004), “Distance, trade and FDI: a Hausman-Taylor SUR approach”, *Journal of Applied Econometrics*, 19, 227-246.

Eicker F., (1963), “Asymptotic normality and consistency of the Least Squares Estimators for families of linear regressions”, *The Annals of Mathematical Statistics*, 34, 447-456.

Elzinga K., Hogarty F., (1973), “The problem of geographic market delineation in antimerger suits”, *Antitrust Bulletin*, 18, 45-81.

Elzinga K., Hogarty F., (1978), “The problem of geographic market delineation revisited”, *Antitrust Bulletin*, 23, 1-18.

Ettner S., (1996), “New evidence on the relationship between income and health”, *Journal of Health Economics*, 15, 67–85.

Exworthy M., Peckham S., (2006), “Access, Choice and Travel: Implications for Health Policy”, *Social Policy & Administration*, 40, 267–287.

Fabbri D., Fiorentini G., (1996), “Mobilità e consumo sanitario: metodi per la valutazione di benessere”, *Mecosan*, 19, 37-52.

Feldstein M.S., (1971), “Hospital price inflation: a study of nonprofit price dynamics”, *American Economic Review*, 61, 853-872.

Ferrie J.E., Shipley M.J., Stansfeld S., Marmot M.G., (2002), “Whitehall II study related behaviours in British civil servants: the morbidity, physiological measures, and health security on self reported health, minor psychiatric effects of chronic job insecurity and change in job”, *Journal of Epidemiology and Community Health*, 56, 450-454.

Fisher M.M., Griffith D.A., (2006), “Modelling spatial autocorrelation in spatial interaction data”, *mimeo*

Fotheringham A.S., O’Kelly M.E., (1989), *Spatial interaction models: formulations and applications*, Kluwer Academic Publishers.

de Frahan B.H., Vancauteren M., (2006), “Harmonization of food regulations and trade in single market: evidence from disaggregated data”, *European Review of Agricultural Economics*, 33, 337-360.

France G., Taroni F., Donatini A., (2005), “The Italian health-care system”, *Health Economics*, 14, S187–S202.

Frankel J., Wei S., (1993), “Trade Blocs and Currency Blocs,” *NBER Working Paper*, 4335, Cambridge, MA.

Frankel J., (1997), *Regional Trading Blocs in the World Economic System*, Washington DC, Institute for International Economics.

Frechette G., (2001), “Random effects ordered probit”, *Stata Technical Bulletin*, 59, 23-27.

Gash V., Mertens A., Romeu Gordo L., (2006), “Are fixed-term jobs bad for your health? A comparison of West-Germany and Spain”, *IAB Discussion Paper*, 8/2006, German Federal Employment Agency.

Geary R.C., (1954), “The Contiguity Ratio and Statistical Mapping”, *The Incorporated Statistician*, 5, 115-145

Giorgetti M.L., (2000), “La verifica empirica del modello di Sutton per i settori manufatturieri italiani”, *L'industria*, 21, 543-564.

Giorgetti M.L., (2002), “Concentration and scope economies in the pharmaceutical sector”, *Discussion Paper*, 205, University of Brescia.

Giorgetti M.L., (2003), “Low concentration and scope economies in the pharmaceutical sector bound estimation - quantile regression and simplex method: an application to Italian manufacturing sectors”, *Journal of Industrial Economics*, 51, 113-120.

Godin I., Kittel F., (2004), “Differential economic stability and psychosocial stress at work: associations with psychosomatic complaints and absenteeism”, *Social Science & Medicine*, 58, 1543–1553

Goldberg D., Williams P., (1988), *A user’s guide to the General Health Questionnaire*, Nfer-Nelson, Windsor.

Goodman D.C., Fisher E., Stukel T., Chang C., (1997), “The distance to community Medical Care and the likelihood of hospitalization: is closer always better?”, *American Journal of Public Health*, 87, 1144-1150.

Gourieroux, C., Monfort A., Trognon A., (1984), “Pseudo maximum likelihood methods: applications to Poisson models”, *Econometrica*, 52, 701-720.

Groot W., (2000), “Adaptation and scale of reference bias in self-assessments of quality of life”, *Journal of Health Economics*, 19(3), 403–420.

Gruber H., (2002), “Endogenous sunk costs in the market for mobile telecommunications: the role of license fees”, *Economics and Social Review*, 33, 55-64.

Hauck K., Rice N., (2004), “A longitudinal analysis of mental health mobility in Britain”, *Health Economics*, 13, 981-1001.

Hausman G.H., Taylor W.E., (1981), “Panel data and unobservable individual effects”, *Econometrica*, 49, 1377- 1398.

Heckman J., (1981), “The incidental parameters problem and the problem of initial conditions in estimating a discrete time–discrete data stochastic process”, in Manski C.F., McFadden D. (eds), *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press, Cambridge, MA.

Held P.J., Pauly M.V., (1983), “Competition and efficiency in the end stage renal disease program”, *Journal of Health Economics*, 2, 95-118.

Henry J., Cassel J., (1969), 'Psychological factors in essential hypertension, recent epidemiological and animal experimental evidence', *American Journal of Epidemiology*, 90, 171-200.

Hernández-Quevedo C., Jones A.M., Rice N., (2005), “Reporting bias and heterogeneity in self assessed health. Evidence from the British Household Panel Survey”, *HEDG Working Paper 05/04*, University of York.

de Juan R., (2003), “The independent submarkets model: an application to the Spanish retail banking market”, *International Journal of Industrial Organization*, 21, 1461-1487.

Idler E.L., Benyamini Y., (1997), “Self-rated health and mortality: a review of twenty-seven community studies”, *Journal of Health and Social Behaviour*, 38, 21–37.

Idler E.L., Kasl S.V., (1995), “Self-ratings of health: do they also predict change in functional ability?”, *Journal of Gerontology*, 50B, 344–353.

Jones A.M., Koolman X., Rice N., (2006), “Health-related non-response in the British Household Panel Survey and European Community Household Panel: using inverse-probability-weighted estimators in non-linear models”, *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 169, 543-569.

Juul-Kristensen B., Jensen C., (2005), “Self-reported workplace related ergonomic conditions as prognostic factors for musculoskeletal symptoms: the "BIT" follow up study on office workers”, *Occupational and Environmental Medicine*, 62, 188-194.

Karasek R.A., Konstantinos P.T., Chaudhry S.S., (1982), “Coworker and Supervisor Support as Moderators of Associations between Task Characteristics and Mental Strain”, Journal of Occupational Behaviour, 3 (2), 181-200.

Karasek R., Theorell T., Schwartz J.E., Schnall P.L., Pieper C.F., Michela J.L, (1988), “Job characteristics in relation to the prevalence of myocardial infarction in the US Health Examination Survey (HES) and the Health and Nutrition Examination Survey (HANES)”, *American Journal of Public Health*, 78, 910–918.

Karasek R, Theorell T., (1990), *Healthy work: stress, productivity, and the reconstruction of working life*, New York, NY, Basic Books

Kawakami N., Haratani T., Araki S., (1992), “Effects of perceived job stress on depressive symptoms in blue-collar workers of an electrical factory in Japan”, *Scandinavian Journal of Work, Environment and Health*, 18, 195–200.

Kawakami N., Araki S., Kawashima M., (1990), “Effects of job stress on occurrence of major depression in Japanese industry: a case-control study nested in a cohort study”, *Journal of Occupational Medicine*, 32, 722– 5.

Kelejian H.H., Prucha I.R., (2004), “Estimation of Simultaneous Systems of Spatially Interrelated Cross Sectional Equations”, *Journal of Econometrics*, 118, 27-50.

Kerkhofs M., Lindeboom M., (1995), “Subjective health measures and state dependent reporting errors”, *Health Economics*, 4, 221–235.

Kivimäki M., Vahtera J., Virtanen M., Elovainio M., Pentti J., Ferrie J.E., (2003), “Temporary employment and risk of overall and cause-specific mortality”, *American Journal of Epidemiology*, 158, 663-668.

Klein Hesselink D.J., van Vuuren T., (1999) “Job flexibility and job insecurity: the Dutch case”, *European Journal of Work and Organizational Psychology*, 8, 273-93.

Koenker R., Ng P., Portnoy S., (1994), 'Quantile Smoothing Splines', *Biometrika*, 81, 673-680.

Kunda G., Barley S.R., Evans J., (2002), “Why do contractors contract? The experience of highly skilled professionals in a contingent labour market”, *Industrial and Labor Relations Review*, 55, 234–61.

Laffont J.J., Tirole J., (1993), *A theory of incentives in regulation and procurement*, Cambridge, MA, MIT Press.

Lasfargues G., Doniol-Shaw G., Deriennic F., Bardo F., Huez D., Rondeau du Noyer C., (1999), “Précaritié de l’emploi et santé: differences hommes-femmes”, *Archieves des Maladies Professionnelles et de Médecine du Travail*, 8, 273-293.

Levaggi R., Zanola R., (2004), “Patient’s migration across regions: the case of Italy”, *Applied Economics*, 36, 1751-1757.

Letourneux V., (1998), *Precarious employment and working conditions in the European Union*. Dublin, European Foundation for the Improvement of Living and Working Conditions. Luxembourg, Office for Official Publication of the European Communities.

Likert R., (1952), “A technique for the development of attitude scales”, *Education and Psychological Measurement*, 12, 313-315.

Lindberg P., Josephson M., Alfredsson L., Vingard E., (2006) “Promoting excellent work ability and preventing poor work ability: the same determinants? Results from the Swedish HAKuL study”, *Occupational and Environment Medicine*, 63, 113-120.

Lindeboom M., Doorslaer E., (2004), “Cut-point shift and index shift in self-reported health”, *Journal of Health Economics*, 23(6), 1083-1099.

Luft H.S., Granick D.W, Mark D.H., Pelzman D.J., Phibbs C.S., Lichtenberg E., McPhee S.J., (1990), “Does quality influence choice of hospital”, *Journal of the American Medical Association,* 263, 2899-906.

Lyons B.R., Matraves C., (1996), “Industrial concentration”, in S. Davies and B. Lyons, *Industrial organization in the European Union: structure, strategy and the competitive mechanism,* Oxford University Press, 86-104.

Mann N., Scheuer E., Fertig K., (1973), “A new goodness-of-fit test for the two-parameter weibull or extreme-value distribution with unknown parameters”, *Communications in Statistics*, 2, 383-400.

Manning W.G., Mullahy J., (2001), “Estimating log models: to transform or not to transform?”, *Journal of Health Economics,* 20, 461-494.

Mapelli V., (2000), “Tre riforme sanitarie e ventuno modelli regionali”, in Bernardi L. (eds.), *La finanza pubblica italiana, Rapporto 1999*, il Mulino.

Marchand A., Demers A., Durand P., (2005), “Do occupation and work conditions really matter? A longitudinal analysis of psychological distress experiences among Canadian workers”, *Sociology of Health and Illness*, 27, 602-627.

Marin P.L., Siotis G., (2002), “Innovation and market structure: an empirical evaluation of the 'Bound Approach' in the chemical industry”, *CEPR Discussion Paper*, 3162, London.

Martens M.F.J., Nijhuis F.J.N., Van Boxtel M.P.J, Knotterus J.A., (1999) “Flexible work schedules and mental and physical health. A study of a working population with non-traditional working hours”, *Journal of Organizational Behaviour*, 20, 35-46

Matraves C., (1999), “Market structure, R&D and advertising in the pharmaceutical industry”, *Journal of Industrial Economics*, 47, 169-194.

McFadden D., (1974), “Conditional logit analysis of qualitative choice behavior”, in Zarembka P. (eds), *Frontiers in econometrics*, Academic Press, New York, USA, 105-142.

McGuirk M.A, Porell F.W., (1984), “Spatial patterns of hospitalization utilization: the impact of distance and time”, *Inquiry*, 21, 84-95.

McNamara P.E., (1999), “Welfare effects of rural hospital closures: a nested logit analysis of the demand for rural hospital services”, *American Journal of Agricultural Economics*, 81, 686-691.

Melnick G., Zwanziger J., Bamezai A., Pattison R., (1992), “The effects of market structure and bargaining position on hospital prices”, *Journal of Health Economics*, 11, 217-234.

Michie S., Williams S., (2003), “Reducing work related psychological ill health and sickness absence: a systematic literature review”, *Occupational and Environment Medicine*, 60, 3-9.

Moran P., (1950), “Notes on continuous stochastic phenomena”, *Biometrika*, 37,17-23.

Morill R.L, Earickson R., (1968), “Hospital variation and patient travel distances”, *Inquiry*, 5, 26-35.

Morrill R.L, Earickson R., Rees P., (1970), “Factors influencing distances travelled to hospitals”, *Economic Geography*, 46, 161-71.

Mundlak Y., (1978), “On the pooling of time series and cross-section data”, *Econometrica*, 46, 69–85.

Murray C.J.L., Tandon A., Salomon J., Mathers C.D., (2001), “Enhancing cross-population comparability of survey results”, *GPE Discussion Paper*, 35, WHO/EIP, Geneva.

NHS Centre for Review and Dissemination – Nuffield Institute for Health, (1996), “Hospital volume and Health Care Outcomes, Costs and Patient Access”, Effective Health Care, 2 (8).

Niedhammer I., Chea M., (2003) “Prospective analyses of the French GAZEL cohort health: comparative results of cross sectional and psychosocial factors at work and self reported”, *Occupational and Environment Medicine*, 60, 509-515

Niedhammer I., Bugel I., Goldberg M., Leclerc A., Gueguen A., (1998), “Psychosocial factors at work and subsequent depressive symptoms in the Gazel cohort”, *Scandinavian Journal of Work, Environment and Health*, 24, 197–205.

Norton E.C., Wang H., Ai C., (2004), “Computing interaction effects and standard errors in logit and probit models”, *The Stata Journal*, 4, 154-167.

Office for National Statistics, published on 19 February 2008, (http://www.statistics.gov.uk/cci/nugget.asp?id=160)

Pikhart H., Bobak M., Pajakb A., Malyutin S., Kubinov R., Topor R., Sebakova H., Nikitin Y., Marmot M., (2004), “Psychosocial factors at work and depression in three countries of Central and Eastern Europe”, *Social Science & Medicine*, 58, 1475–1482.

Porell F.W., Adams E.K., (1995), “Hospital choice models: a review and assessment of their utility for policy impact analysis”, *Medical Care Research and Review*, 52, 158-195.

Pregibon D., (1980), “Goodness-of-link tests for generalized linear models”, *Applied Statistics*, 29,15–24.

Price R. H., Burgard S., (2006), “Non-standard work and health: who is at risk and who benefits?”, National Poverty Centre, *mimeo.*

Propper C., Wilson D., Burgess S., (2006), “Extending choice in English health care: the implications of the economic evidence”, *Journal of Social Policy*, 35, 537-557.

Ramsey, J.B., (1969), “Tests for specification errors in classical linear least squares regression analysis”, *Journal of the Royal Statistical Society B*, 31, 350-371.

Rauch J.E., (1999), “Network versus markets in international trade”, *Journal of International Economics*, 48, 7-35.

Robinson J.C., Luft H.S., (1985), “The impact of hospital market structure on patient volume, average length of stay, and the cost of care”, *Journal of Health Economics*, 4, 333-356.

Robinson W.T., Chiang J., (1996), “Are Sutton's predictions robust? Empirical insights into advertising, R&D, and concentration”, *Journal of Industrial Economics*, 44, 389-408.

Robone S., Zanardi A., (2006), “Struttura di mercato e tecnologia: un’analisi empirica del Servizio sanitario nazionale italiano”, *Politiche sanitarie*, 7, 1-14.

Rodrigues E., (2002) “Marginal employment and health in Britain and Germany: does unstable employment predict health?”, *Social Science and Medicine*, 55, 963–79.

Roghman K.J., Zastowny T.R., (1979), “Proximity as a factor in the selection of health care providers: emergency room visits compared to obstetric admissions and abortions”, *Social Science and Medicine*, 13, 61-9.

Rugulies R., Bultmann U., Aust B., Burr H., (2006), “Psychosocial work environment and incidence of severe depressive symptoms: prospective findings from a 5-Year follow-up of the Danish Work Environment Cohort Study”, *American Journal of Epidemiology*, 163, 877-887.

Sacerdote B., (2001), “Peer Effects with random Assignment: Results of Dartmouth Roommates”, *Quarterly Journal of Economics*, 116, 681-704.

Sadana R., Mathers C.D., Lopez A.D., Murray C.J.L., Iburg K., (2000), “Comparative analysis of more than 50 household surveys on health status”, *GPE Discussion Paper*, 15, EIP/GPE/EBD, World Health Organisation, Geneva.

Salop S.C., (1979), “Monopolistic competition with outside goods”, *Bell Journal of Economics*, 10, 141-156.

Santos-Silva J., Tenreyro S., (2006), “The log of gravity”, *The Review of Economics and Statistics*, 88, 641-658.

Segal L.M., Sullivan D.G., (1997), “The growth of temporary services work”, *The Journal of Economic Perspectives*, 11, 117–36.

Seninger S.F., (1999), “Information availability and market power in export competition: the case of regional health care markets”, *Journal of Regional Science*, 39, 479-495.

Shaked A., Sutton J., (1982), “Relaxing price competition through product differentiation”, *Review of Economic Studies*, 49, 3-13.

Shaked A., Sutton J., (1987), “Product differentiation and industrial structure”, *Journal of Industrial Economics*, 36, 131-146.

Shields M.A., Wheatley Price S., (2005), “Exploring the economic and social determinants of psychological well-being and perceived social support in England”, *Journal of the Royal Statistical Society A.*, 168 (3), 513-537.

Siegrist J., (1996), “Adverse health effects of high-effort/low reward conditions”, *Journal of Occupational Health Psychology*, 1, 27–41.

Siegrist J., Peter R., Junge A., Cremer P., Seidel, D., (1990), “Low status control, high effort at work and ischemic heart disease: prospective evidence from blue-collar men”, *Social Science & Medicine*, 31, 1127–1134.

Silla I., Gracia F. J., Peiro´ J. M., (2005), “Job insecurity and health-related outcomes among different types of temporary workers”, *Economic and Industrial Democracy*, 26, 89–117.

Smith J.P., (1999), “Healthy bodies and thick wallets: the dual relationship between health and economic status”, *Journal of Economic Perspectives*, 13, 145–166.

Smith J.P., (2004), “Unraveling the SES-health connection”, *Population and Development Review*, 30, 108-132

Smith R.L., (1985), “Maximum likelihood estimation in a class of non-regular cases”, *Biometrika*, 72, 67-92.

Smith R.L., (1989), “Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone”, *Statistical Science*, 4, 367-393

Smith R.L., (1994), “Nonregular regression”, *Biometrika*, 81, 173-183.

Spampinato G.C., (2001), “Mobilità sanitaria regionale e federalismo”, *Politiche sanitarie*, 2, 44-56.

Spence A.M., (1976), “Product selection, fixed costs, and monopolistic competition”, *Review of Economic Studies*, 43, 217-235.

Stewart J.Q., (1948), “Demographic gravitation: evidence and application”, *Sociometry*, January, 31-58.

Stillwell J., (2005), “Inter-regional migration modelling: a review and assessment”, paper prepared for the 45th congress of the European Regional Science Association, The Netherlands, 23-27 August 2005.

Studnicki J., (1975), “The minimization of travel efforts as a delineating influence for urban hospital service area”, *International Journal of Health Service*, 5, 679-93.

Sutton J., (1989), “Endogenous sunk costs and industrial structure”, in Bonanno G., Brandolini D. (eds.), *Market structure in the new industrial economics*, Oxford, Oxford University Press.

Sutton J., (1991), *Sunk costs and market structure*, Cambridge, MIT Press.

Sutton J., (1998), *Technology and market structure*, Cambridge, MIT Press.

Sverke M., Gallagher D.G., Hellgreen J., (2000), “Job stress, well-being, and work attitudes among employees with different employment contracts”, in Isaksson K., Hogstredt C., Eriksson, C., Theorell T. (eds.) *Health effects of the new labour market*, New York, Kluwer Academic/ Plenum Publishers, 145-67.

Tai W.C., Porell F. W., Adams K., (2004), “Hospital choice and rural Medicare beneficiaries: patients, hospital attributes, and the patient-physician relationship”, *Health Services Research*, 39, 1903-1922.

Taroni F., (2001), “Programmazione e regolazione dell`assistenza ospedaliera: un`agenda infinita”, in Fiorentini G. (eds), *I servizi sanitari in Italia*, Il Mulino, Bologna.

Tay A., (2003), “Assessing competition in hospital care markets: the importance of accounting for quality differentiation”, *The RAND Journal of Economics*, 34, 786-814.

Taylor M., Brice J., Buck N., Prentice E., (1998), *British Household Panel Survey user manual volume A: introduction, technical report and appendices*, University of Essex, Colchester.

Tennant C., (2001), “Work-related stress and depressive disorders”, *Journal of Psychosomatic Research*, 51, 697– 704

Theorell T., Karasek R.A., (1996), “Current issues relating to psychosocial job strain and cardiovascular disease research”, *Journal of Occupational Health Psychology*, 1, 9–26.

Town R., Vistnes G., (1997), “Hospital competition in HMO networks: an empirical analysis of pricing behaviour”, mimeo.

Ugolini, C., Fabbri D., (1998), “Mobilità saniatria ed indici di entropia”, *Mecosan*, 26, 9-24.

van Doorslaer E., Wagstaff A., van der Burg H., Christiansen T., De Graeve D., Duchesne I., Gerdtham U.G., Gerfin M., Geurts J., Gross L., Häkkinen U., John J., Klavus J., Leu R.E., Nolan B., O'Donnell O., Propper C., Puffer F., Schellhorn M., Sundberg G., Winkelhake O., (2000), “Equity in the delivery of health care in Europe and the US”, *Journal of Health Economics*, 19, 553-583.

Varkevisser M., van der Geest S., (2006), “Why do patients bypass the nearest hospital? An empirical analysis for orthopaedics care and neurosurgery in the Netherlands”, *HEDG Working Paper 06/01*, University of York.

Vermeulen M., Mustard C., (2000), “Gender differences in job strain, social support at work, and psychological distress”, *Journal of Occupational Health Psychology*, 5, 428–40.

Virtanen M., Kivimaki M., Elovainio M., Vahtera J., Ferrie J. E., (2003), “From insecure to secure employment: changes in work, health, health related behaviours, and sickness absence”, *Occupation and Environmental Medicine*, 60, 948–953.

Vistnes G., (1995), “Hospital mergers and antitrust enforcement”, *Journal of Health Politics, Policy and Law*, 20, 175-190.

Warell L., (2005), “Defining geographic coal markets using price data and shipments data”, *Energy Policy*, 33, 2216-2230.

Warren L.R., Hoonakker P., Carayon P., Brand J., (2004), “Job characteristics as mediators in SES-health relationship”, *Social Science & Medicine*, 59, 1367-1378.

Wei S.J., Frankel J.A., (1997), “A open versus closed trading blocks”, in Ito T., Krueger A.O. (eds.) *Regionalism versus multilateral trade agreements*, Chicago University Press

Werden J., (1981), “The use and misuse of shipments data in defining geographic markets”, *Antitrust Bulletin*, 26, 719-737.

Werden J., (1989), “The limited relevance of patient migration data in market delineation of hospital merger cases”, *Journal of Health Economics*, 8, 363-376.

Westerlund J., Wilhelmsson F., (2006), “Estimating the gravity model without gravity using panel data”, Lund University, *mimeo*

White H., (1980), “A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity”, *Econometrica*, 48, 817-838.

Wholey D., Feldman R., Christianson J.B., Engberg J., (1996), “Scale and scope economies among health maintenance organizations”, *Journal of Health Economics*, 15, 657-684.

Wooldridge J., (2002), *Econometric analysis of cross section and panel data*, MIT Press, Cambridge, MA.

Wooldridge J., (2005), “Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity”, *Journal of Applied Econometrics*, 20, 39–54.