

Physician dual practice and the public health care provision

Review of literature

Karolina Socha

The Research Unit of Health Economics
University of Southern Denmark
www.healtheconomics.dk

Health Economics Papers
2010:4



Abstract

Objectives: A combination of public and private clinical practice by physicians, referred to as physician dual practice, is present in numerous health care systems. The phenomenon has been receiving attention in connection with arguments about its negative impact for the public provision of health care. A ban or restrictions on dual practice is widely advocated. The aim of this paper is to review and critically discuss the empirical and theoretical findings on the subject of physician dual practice and its effects for the provision of public health care.

Methods: A systematic literature review, using electronic and manual searches, identified 23 positions on the subject of physician dual practice consisting of peer-reviewed journal articles, including two review articles on selected aspects of dual practice, academic working papers, book chapters, and WHO publications.

Results: The subject is short on hard evidence. Theoretical analyses of dual practice impact on public health care provision shows that dual practice might bring about both positive and negative effects. Some of the conclusions on the negative effects of dual practice, however, depend on assumptions, which are questioned in broader economic literature. Moreover, while it seems that dual practitioners take up private practice predominantly to increase their income, it does not automatically imply that dual practice as a whole is a profit-maximising combination. Physicians seem to face promising opportunities in the private market. Still, they spend relatively little time in the private practice and supply labour to the lower paid public sector job. Eventually, the potential costs and effectiveness of dual practice regulation are rarely considered.

Conclusions: Further research is needed into why physicians engage in dual practice and whether potential costs of dual practice outweigh the costs of enforcing restrictions on dual practice. A promising direction appears to be an analysis comparing the behaviour of the public sector physicians with and without dual practice, respectively, in a given institutional settings and under uniform assumptions about the objectives of the physicians in these two groups. Moreover, it seems useful to employ the labour economics framework for multiple job holding in the analysis of the dual practice effects. This framework helps to recognise that physicians holding two jobs might be different from single job holding physicians with regard to the most preferred combination of income and leisure, which implies differences in labour market participation constraints and marginal costs of labour/effort between the two groups.

Keywords: Dual practice, moonlighting physician, public-private practice.

Contents

1.	Introduction	4
2.	Motivation behind dual practice	6
3.	Dual practice and physicians' labour supply	9
	3.1. <i>"Public versus private health care in a National Health Service."</i> (Brekke & Sjørgard, 2007)	10
4.	Dual practice and the costs of the health care provision in the public sector	13
	4.1. <i>"Should physicians' dual practice be limited? An incentive approach."</i> (González, 2004)	13
	4.2. <i>"Regulating providers' reimbursement in a mixed market for health care."</i> (Rickman & McGuire, 1999)	16
	4.3. <i>"On the policy of transferring public patients to private practice."</i> (González, 2005)	19
	4.4. Dual practice and misappropriation of the public sector supplies/resources	22
5.	Dual practice and quality of health care	22
	5.1. <i>"Moonlighting: public service and private practice."</i> (Biglaiser & Ma, 2007)	24
	5.1.1. More on dual practice, consumer welfare and quality of public health care	26
	5.2. Dual practice and the public sector waiting lists	29
	5.3. <i>"Physician dual practice"</i> (Eggleston & Bir, 2006)	31
	5.3.1. <i>"Private medical services in the Italian public hospitals: the case for improving HRM."</i> (De Pietro, 2006)	32
6.	Regulatory response to dual practice	33
7.	Concluding remarks	36
	Acknowledgments	37
	List of references	38

1. Introduction

Physician dual practice has a diversity of meanings in the literature: The expression is used to indicate a physician with multiple medical specialties; It is also used to describe health professionals who combine clinical practice with other health-related activities such as research, teaching, or management. A combination of clinical practice with a job not related to health care is also sometimes referred to as dual practice (Ferrinho et al., 2004). In most of the economic literature, however, dual practice is understood as a situation where a physician combines clinical practice in the public sector with a clinical practice in the private sector. This paper adopts this last definition of physician dual practice, and the physicians involved in dual practice are referred to as dual practitioners.

Physician dual practice is a widespread phenomenon, and it is commonly believed to have direct implications for physicians' labour supply, volume of health care production, and quality of health care services (García-Prado & González, 2007). Particular attention is paid to dual practice effects for the public provision of health care. The fact that a single physician provides health care in both the public and the private sector is the main source of controversy in the policy debate, where the existence of physician dual practice is perceived as being at odds with the optimal labour market arrangements regarding employer-employee relation. Consequently, a need for regulatory measures towards physician dual practice is commonly stressed (WHO, 2000). The economic literature on the subject of physician dual practice and its effects for the public health care sector is relatively young and not voluminous. The general conclusion from this literature appears to be that dual practice has both positive and negative effects for the public health care provision, and the net effect is difficult to determine (Eggleston & Bir, 2006). The aim of this paper is to review and discuss the economic literature on the subject of physician dual practice effects for the public health care sector.

There is evidence that dual practice is prevalent in numerous health care systems of low, middle and high-income countries (García-Prado & González, 2007; Jan et al., 2005; Ferrinho et al., 2004; Bir & Eggleston, 2003; WHO, 2000). Physicians engaged in dual practice (dual practitioners) are generally expected to have better income raising opportunities in the private practice than in the public one. For this reason they are suspected to concentrate their attention and work effort on the private practice at the expense of the public one. A common assumption is that a dual practitioners' commitment to the patients is compromised by the pursuit of profit-maximisation such that they under-represent the interests of the public patients and the public payers (Biglaiser & Ma, 2007; Ferrinho et al., 2004; WHO, 2000). Dual practitioners are also suspected to favour long waiting times in the public sector in order to boost the demand for the private services (Humphrey & Russell, 2004; Morga & Xavier, 2001; Iversen, 1997). Physicians' engagement in dual practice is also argued to result in cream-skimming of profitable patients from the public sector waiting lists to the private practices, which increases private practice profits at the expense of the public sector (González, 2005). Further, an increase in costs in the public sector might occur due to an over-provision of health care by dual practitioners who want to earn a good reputation to support their private practice (González, 2004), or because the public resources might be used for private practice purposes without adequate payment (Ferrinho et al., 2004; Gruen et al., 2002). Eventually,

self-referrals of patients from the public to the private practice of the dual practitioners might be a form of demand inducement (Bir & Eggleston, 2003).

There are also positions which stress advantages of allowing dual practice such as an opportunity for the government to retain the most skilled physician in the public service when the physicians are allowed to support their public salaries with extra earnings from the private practice (Eggleston & Bir, 2006; Jan et al., 2005; Bir & Eggleston, 2003). Prevalence of dual practice is said to have the potential to constrain the over-provision of public health care services when the providers are financed under the full-cost reimbursement mechanism (Rickman & McGuire, 1999). Dual practice is also presented to be an efficient tool for selecting those public sector patients who can be served better in the private market (Biglaiser & Ma, 2007; Eggleston & Bir, 2006). Eventually, the physicians themselves report that dual practice offers a wide range of opportunities other than just extra money, including clinical autonomy, skills development, and the possibility to offer the type of health care that is unavailable in the public sector (Jumpa, 2007; Humphrey & Russell, 2004).

There are several interesting observations one can make about the state-of-the-art analysis of dual practice effects for the public health care provision. According to the agency literature the physicians might anyhow prioritise their financial interests over the interest of the patients and the payers, regardless of the engagement in dual practice (McGuire, 2000; Pratt & Zeckhauser, 1991). The single job physicians are also expected to provide minimum observable effort (McGuire, 2000), favour long waiting times (Cullis & Jones, 2000; Cullis & Jones, 1985), cream-skin profitable patients, and skimp the costly, severely ill ones (Martinussen & Hagen, 2009; McGuire & Hughes, 2003; Ellis, 1998; Newhouse, 1996), and over-provide health care (Gal-Or, 1999; Ellis & McGuire, 1993; Blomqvist, 1991; Ellis & McGuire, 1990; Ellis & McGuire, 1986). Hence, an interesting question is whether, when involved in dual practice, physicians intensify these actions as compared to physicians who are not involved in dual practice. In the existing literature on dual practice, however, the physicians who are not involved in dual practice are usually decoupled from the analysis or assumed to be mission-oriented and automatically representing the interests of the public sector better than the dual practitioners. Such approaches might produce misleading conclusions about the dual practice effects.

Furthermore, in the broader economic literature, the prevalence of dual practice (dual job holding) is not seen as being at odds with optimal employer-employee (principal-agent) relations. Common agency theory (multiprincipals incentive theory) indicates that restrictions on the agents' involvement with more than one principal are costly means of behaviour monitoring and when implemented might be inefficient due to asymmetric information (Mezzetti, 1997; Martimort, 1996). Hence, one can argue that unregulated dual practice can be seen as the second-best solution.

Another feature of the literature on physician dual practice is that the motives behind dual practice are not much investigated. It is usually taken for granted that the motivation behind dual practice is exclusively of a financial nature. One should notice, however, that empirical studies focus on developing countries, where public salaries are often below costs of living and dual job holding is a coping strategy (Jumpa, 2007; Jan et al., 2005; WHO, 2000). Moreover, it should be noticed that even if the private practice is only a source of extra income, it does not directly imply that dual practice is solely a profit-maximising activity. According to the theory of labour economics, dual job holding is a profit-maximising activity only when the dual jobholder works as much as possible in the better paid job, i.e. reaches an income constraint in that job (Perlman, 1966; Mabry, 1973;

Shisko & Rostker, 1976). If, however, an individual holds two jobs although he/she could achieve the most preferred combination of income and leisure concentrating exclusively on the better paid job, rational-profit maximisation fails to explain dual job holding, dual practice included.

The positions reviewed in this article have been selected in a search strategy involving searching of electronic data bases: ISI Web of Knowledge, PubMed. The search terms used were: dual practice, moonlighting, public private practice. The search of the electronic data bases, supported by a cross-check with other electronic search engines such as Google-Scholar, identified 10 positions on the subject of dual practice effects for the public health care provision. Subsequently, a manual reference lists search revealed 13 additional contributions on the subject, including two review articles focusing on selected aspects of dual practice. The 23 contributions consist of articles published in peer-reviewed economic and health sciences journals, academic working papers, book chapters, and publications of the World Health Organisation (WHO).

The article reviews and discusses the findings presented in the literature on the physician dual practice effect for the public health care provision. In comparison to the existing paper (Eggleston & Bir, 2006), which overviews 4 theoretical contributions on the subject, this paper offers a more comprehensive compilation of findings and an in-depth discussion of the conclusions. The implications of the dual practice literature are also discussed in the light of labour economics studies on dual/multiple job holding. Following the observations listed above, it is scrutinised in how far the differences in work performance between the dual practitioners and the physicians working exclusively in the public sector are demonstrated, whether the potential costs of dual practice are set against the potential costs of restrictions on dual practice, and how much the motivation behind dual practice has been investigated.

Nearly all of the contributions regard the supply side of the health care system. The discussion of the literature is organised in thematic sections. Some of the reviewed articles provide insights that could be listed under more than one section. Nevertheless, the thematic division is employed for the ease of organising the content of the paper. The following sections discuss findings with regard to: Motivation behind dual practice (2); Dual practice and the physician labour supply (3); Dual practice and the costs of the public health care provision (4); Dual practice and the health care quality (5); Regulatory response to dual practice (6); Concluding remarks (7).

2. Motivation behind dual practice

A common assumption is that public sector physicians who engage in dual practice (take up a fee-for-service private practice) are motivated by self-interest, with the commitment to the patient compromised by the pursuit of profit-maximisation. The reasons for dual practice are not well studied though, and most evidence regards developing countries.

For low-income and middle-income countries there exists evidence that physicians engage in dual practice as a result of low public sector salaries, which do not allow for a comfortable standard of living or even do not exceed the minimum costs of living. Thus, dual practice can be viewed as a possible system solution to problems such as very limited financial resources in the public health care sector (Jan et al., 2005; Ferrinho et al., 2004; Gruen, 2002; Macq et al., 2001; WHO, 2000).

Gruen et al. (2002) examine dual practice in Bangladesh through interviews with a purposively selected sample of 100 public sector physicians representing all levels of service (primary, secondary, tertiary), seniority, and different geographical areas of the country. Dual practice is

concluded to be a widespread phenomenon and a response to the low public sector salaries. The financial motivation is further supported among the physicians in primary health care, who answered unanimously that they would give up the private practice for a salary increase in the public sector. In the secondary and tertiary sector, however, the proportion of physicians who would give up the private practice for a corresponding increase in the public sector salary was only 54 per cent. Such statements do not entirely fit into the assumption of an exclusively financial motivation behind dual practice. It is also interesting to notice that 65 percent of the interviewees did not consider giving up their public practice. The authors do not, however, provide sufficient information with regard to the conditions in the private health care market, which would allow assessing whether the lack of intention to give up the public practice is associated with the lack of sufficient opportunities in the private market. Still, some indication of non-financial motivation appears in other interviewees' statements that they value the public sector job for the social status it gives and the opportunities such as teaching or serving the poor.

Jumpa et al. (2007) report results of qualitative interviews with a sample of 20 physicians in Lima representing the public sector, the private providers, the Ministry of Health, and the Peruvian College of Physicians. The study indicates that dual practice is a well-accepted, widespread activity. The dominating motivation behind dual practice seems to be financial and is related to the need to supplement very low public sector salaries. Indeed, as a result of very low public sector salaries physicians are often involved in multiple job holding, i.e. three or four jobs. It has also been underlined that an apparent over-supply of physicians in Peru results in high competition in the private market, which means that the private practice is not necessarily a guarantor of sufficient income. The financial motivation is concluded to be dominating, but the private practice was also perceived as offering a possibility to provide health care of higher quality than in the public sector and a greater sense of being valued, which, as a consequence, increases the physicians' professional satisfaction. Moreover, the involvement in the private practice was also seen as possibility to improve one's skills.

From the above-mentioned evidence, dual practitioners appear to take up private practice exclusively or mainly for financial reasons. This, however, does not directly indicate that the engagement in dual practice is a purely profit-maximising activity. Ferrinho et al. (2004) notice that with extremely low public sector salaries in many countries, it is surprising that many physicians remain in the public service when they could earn much more in a private practice. Further, among those who are engaged in dual practice, many spend comparatively little time in their private practice. The authors state, it is unlikely that the reason is lack of opportunities in the private market. It appears that physicians, including dual practitioners, have other sources of motivation than profit-maximisation. Dual practice appears to be at odds with the prediction that individuals would prefer to concentrate exclusively on the higher paid job instead of holding two jobs with different payoffs. It might mean that physicians engage in dual practice due to complementarities that exist between the public and the private job (Eggleston & Bir, 2006). In other words, private practice might be of help with regard to financial needs, while public service might be better for fulfilling professional aspirations such as, e.g. teaching and/or providing service to the poor.

While many public sector physicians in low-income and even middle-income countries have to resort to dual practice or multiple job holding as a means of survival, this does not seem to apply to public sector physicians in developed countries, whose income opportunities in the public and the

private sector exceed the subsistence level (Gruen et al., 2002). The evidence on motivation behind dual practice in developed countries is provided only in a single study by Humphrey and Russell (2004). The authors report results of interviews with 60 dual practitioners in south-east England on the subject of physicians' motivation behind dual practice and their perception of dual practice benefits. Private practice was seen as offering a range of rewards, including financial benefits, increase in strategic influence, clinical autonomy, a greater sense of being valued, and more opportunities to realise one's individual aspirations as a clinician. The opportunity to do private practice was seen as a fair reward, and the dual practitioners judged their financial expectations against earnings they could get in jobs alternative to the National Health Service (NHS). Only for a few respondents' money appeared to be the dominating motivation, in that they would gladly give up the private practice in exchange for a pay rise in the NHS. For the remaining majority, however, money was only a part of the motivation. The respondents underlined that the private practice is offering considerable control over whom they work with, the timing of the treatments, and the number of patients, which they cannot enjoy in the public practice. For most, the private practice was also gratifying thanks to the evidence of peer approval, which was said to be harder to obtain in the NHS, and better contact with patients. The latter was seen as a result of the fact that patients who use the private health care sector choose their physician, while in the NHS patients are randomly matched with a physician and happen to express their dissatisfaction with the match. Moreover, a number of surgeons expressed their dissatisfaction with the limited availability of theatre time in the NHS and valued highly the hands-on operating opportunities in the private practice. Others perceived private practice as a source of additional patients with particular conditions, enabling them to gain extra experience in shorter time than in the NHS. Some of the respondents reported to take up the private practice in a response to the inability to achieve services improvement or introduce innovations in the NHS. Most importantly, the respondents indeed saw the private practice more as a complement to their NHS work, and not an alternative. Very few respondents contemplated to leave the NHS entirely.

The qualitative studies on the subject of the motivation behind dual practice might not be seen as representative for the population of physicians in the respective countries. The truthfulness of the respondents' accounts can also be questioned, but the fact that the studies rely on self-reported data does not mean that the findings do not reflect the reality. It is interesting to notice that the findings indicating a motivation other than financial behind dual practice are supported by management and labour economics studies of dual/multiple job holding.

Empirical studies in the management literature on dual job holding suggest that the phenomenon might contribute to producing less frustrated employees because it opens possibilities for the highly skilled professionals to put their knowledge and abilities into full use and experience a full range of benefits, financial and non-financial, associated with it (Jamal et al., 1998; Bennet et al., 1994; Jamal, 1992; Jamal & Crawford, 1981).

The labour economics literature underlines that the assumption that a dual job holder is a rational profit-maximiser is consistent when there is a time constraint on the number of hours a dual job holder can work in the job that pays better per hour (Perlman, 1966; Mabry, 1973; Shisko & Rostker, 1976). In other words, if the most preferred number of hours (income) could be supplied (earned) in the higher paid job there would be no reason to engage in dual practice. When the individual is not facing a time constraint in the higher paid job, i.e. is supplying hours to the lower

paid job, while he or she could spend this time in the higher paid job, the financial motives fail to explain dual practice or any other kind of multiple job holding. The empirical studies of hours supplied to different jobs by dual/multiple job holders suggest the existence of motives other than financial behind dual job holding, especially among highly specialised/skilled and high income professionals (Conway & Kimmel, 1998; Lundborg, 1995; Kimmel & Conway, 1995). Put differently, some of the dual jobholders do not divide time between the two jobs according to a profit-maximisation principle. There is also qualitative evidence revealing variations in the motivations of multiple job holders with education, income, and gender. While low income individuals engage in multiple jobs occasionally to make ends meet, some high income, highly specialised individuals have regular multiple job holding spells and combine different jobs to realise their professional aspirations (Averett, 2001; Amirault, 1997).

In summary, while it appears that private practice provides dual practitioners with financial resources, it cannot be directly concluded that physicians involved in dual practice are rational profit-maximisers, who compromise their patients' needs whenever an opportunity for earning an extra unit of income appears. The question needs to be asked not only about the major benefits of the private practice, but also about why dual practitioners do not choose to focus exclusively on the lucrative job in the private health care sector. There appears to be a need for more empirical inquiries into the matter of the motivation behind physician dual practice. It also seems reasonable to argue that different motives might be valid for physicians in low and in high-income countries.

3. Dual practice and physician labour supply

Dual practice means that a physician is involved in dual job holding. This raises the interesting issue of how such a situation affects the total labour supply of the physician. As already mentioned above, the individual is willing to engage in dual job holding only if his/her most preferred combination of leisure and income cannot be realised when holding only a single job (Perlman, 1966; Mabry, 1973; Shisko & Rostker, 1976). In other words, the individual is engaged in dual job holding when he/she faces an income (working time) constraint in his/her job denoted as primary job. This implies that an individual is willing to take up a second job, even if the hourly wage in the second job is below the hourly wage in the primary job. In such a situation, the individual's total labour supply increases. It might also be that the individual in search for an additional job finds one that offers an hourly wage higher than the hourly wage in the primary job. Then, the individual might quit the primary job for the new one, but only if the total income that can be earned in the new job can satisfy the individual's most preferred combination of income and leisure. If this is not the case, the individual engages in dual job holding and tries to work as much as possible in the job offering the higher hourly wage and then supply labour in the lower paid primary job. This means, however, that the individual might substitute some of the work hours in the primary job with hours in the better paid second job. Whether the increase in labour supply to the second job outweighs the decrease in labour supply to the primary job depends on the relative magnitude of the substitution and the income effect of the increased hourly wage. In other words, depending on the individual's taste for money in relation to leisure, the engagement in dual job holding can result in either an increase or a decrease in the total labour supplied by an individual. Since the various individuals' tastes might differ widely, the question whether dual job holding increases or decreases the total

labour supply in the labour economics literature is regarded as a matter for empirical investigation (Bosworth et al., 1996).

In the literature on dual practice, Brekke & Sørsgard (2007) contribute with a theoretical analysis of the effect dual practice has for the total labour supply of a physician. Their model is discussed in the following subsection.

3.1. “Public versus private health care in a National Health Service.”(Brekke & Sørsgard, 2007)

Brekke & Sørsgard (2007) analyse dual practice effects for the physicians’ labour supply in a model for a health care system, where the public health care is provided through a National Health Service (NHS) free of charge at the point of consumption, but subject to rationing. The patients can also consume private health care, for which they pay out-of-pocket. All physicians are assumed to be dual practitioners who decide on the division of their labour supply between the NHS and the private practice depending on the relation between the public sector wage and the private sector profits per unit of labour. The public sector wage per unit of labour is set by a health authority, while the private sector profits per unit of labour are determined by the demand for and the supply of private health care. The authors point out that the production of the health care is very labour intensive and thus, that the dual practitioners’ labour supply to the public and the private sector can be assumed to directly determine the total supply of health care. The case of interest is when the public and the private health care services are substitutes from the consumer’s point of view. A conclusion drawn from the model is that dual practice induces a crowding-out effect on the public provision of the health care and also on the total available volume of the health care. Put differently, according to the model, when all physicians are involved in dual practice, less health care is available in the system than when they work exclusively in the public sector. The authors postulate that the health authority can discourage dual practice by increasing the public sector wage per unit of labour (in relation to the profits per unit of labour in the private practice), but since this increases the public expenditures, a ban on dual practice might be a better solution when the profits from the private practice are very high. The details of the model are presented below.

There is a fixed number of physicians who are all involved in dual practice. The population of patients is represented by a representative consumer whose utility is given by the function

$$U(X, Y, Z) = X + Y - \frac{1}{2}(X^2 + Y^2 + 2bXY) + Z, \quad (1)$$

where X is the amount of public health care, Y is the amount of private health care, where $X + Y < 1$, Z is the amount of a numeraire good, and $b \in [0,1]$ is interpreted as a parameter that measures the degree of substitutability between the public and the private health care services. According to the model, $b=1$ means that the public and the private health care are perfect substitutes from the consumer’s point of view, i.e. the consumer cares only about the total amount of health care received. The consumption choice of the representative consumer is said to be restricted by a budget constraint,

$$m - pY - Z \geq 0, \quad (2)$$

where m is the consumer’s disposable income and $p > 0$ is the unit price of the private health care. The perceived price of the public health care is zero. The authors assume the budget constraint (2) to be binding and insert the budget constraint into the consumer’s utility function (1) obtaining

$$U(X, Y) = X + Y - \frac{1}{2} (X^2 + Y^2 + 2bXY) - pY + m. \quad (3)$$

The authors assume that the demand for the public health care always exceeds the public sector capacity and that there is efficient rationing according to severity in the public sector. The effect of the efficient rationing is explained in the following way: “in the case of the representative consumer, efficient rationing would mean that since a single patient values public treatment more (because it is free), he demands as much treatment as he can in the public sector, and then he demands private care”. Extending the reasoning to heterogeneous consumers, efficient rationing is interpreted to mean that “those who value more to be treated (because they are more severe) are treated in a public hospital, while those who are rationed demand private care”. In other words, the amount of the public health care X is not a choice variable for the representative consumer. In the following, the authors state that the inverse demand function for the private health care can be obtained by maximizing the consumer’s utility function (3) with respect to the amount of private health care Y . Consequently, the inverse demand function for private health care is stated to be

$$p = I - Y - bX. \quad (4)$$

The consumer’s maximisation problem is said to be solved at this stage. From (4) the authors conclude that the market price p of the private health care is decreasing in the overall supply of both the private health care Y and the public health care X . Since the labour supplied by the physicians is assumed to reflect directly the supply of the health care, the authors conclude that the dual practitioners face an incentive to cut down on the labour supply to the public sector and also to the private sector in order to increase their private practice profits per unit of labour. The private health care price p is stated to increase by one unit with one unit decrease in the supply of the private health care and further increase by b units with one unit decrease in the supply of the public health care. Hence, there is a crowding-out effect of dual practice on the physician labour supply. The crowding-out effect is strongest when the public and the private health care services are close substitutes from the consumer’s point of view, i.e. $b \approx I$. The authors write that the health authority can partially mitigate the crowding-out effect of dual practice by increasing the public sector wage per unit of labour (in relation to the private practice profits per unit of labour) which induces a shift of physician labour from the private to the public sector. Yet, when the private practice profits per unit of labour are very high, i.e. when there are only few physicians in the market, it is optimal for the health authority to ban dual practice. The ban means, that the physicians have to choose to practice either exclusively in the public or in the private sector, and in the former case they are not compensated for the foregone private practice profits. The authors conclude that if the health authority decides to introduce a ban on dual practice the physicians’ income from the public sector is sufficiently high to keep all the physicians in the public sector.

The above-recalled assumptions and the formulation of the representative consumer’s utility problem can give rise to objections. It can, for example, be noticed, that the authors suggest that the representative consumer, when facing rationing in the public sector, automatically demands private health care. Rationing can mean that the admissions to the public sector waiting lists are rationed, e.g. some patients are denied surgery and are offered medication instead. In such a case, the representative patient must prefer costly private health care over the free alternative treatment in the public sector. It is also interesting to notice that in such a case only the physicians’ labour supply to

the private sector matters, in terms of the model. The physicians' labour supply to the public sector is irrelevant because the patient is rationed according to some clinical criterion irrespective of the physician labour supply in the public sector. Rationing can also mean that the consumer is facing a waiting time in the public sector. Then the representative consumer must prefer consumption of the costly private health care to the delayed provision of the public health care. More importantly, as the representative consumer has to pay for the private health care out-of-pocket the expenses cannot exceed the consumer's disposable income. In the model, however, the representative consumer's budget constraint is mentioned, but not imposed in the analysis of the supply side of the health care market. Consequently, the analysis of how high the dual practitioners are setting the unit price for the private health care does not account for the demand constraint. The model in fact requires that the representative consumer is cash-unconstrained and that the demand for the private health care is price inelastic. Thus, the question arises how much the representative consumer in the model represents the general population. Without taking into account the consumer's budget constraint it cannot be taken for granted that for the high unit price of the private health care the dual practitioners would be able to sell enough units of health care in order to preserve (increase) their total income. Eventually, the negative dual practice effect for the physicians' labour supply is depending on the public and the private health care being close substitutes. The authors support their assumption that the public and the private health care are (imperfect) substitutes for the consumer by a reference to the empirical paper by McAviney & Yannopoulos (1994). The latter using the UK time-series data (1980-1985) for a group of acute specialties analyse the consumers' decisions to consume the public or the private health care in a framework in which the consumption of all the other goods gives a measure of overall opportunity costs. McAviney & Yannopoulos (1994) construct a shadow price for the public health care and a shadow price for the private health care. Following an argument that the consumers' income at the margin can be affected negatively by the fact that they have to wait for the treatment¹, the time spent on the public sector waiting list is combined with a measure of earnings and a time rate of discount. The shadow price of the private health care is the price of the private health care insurance premium. The general consumption price is represented by the retail price index. McAviney & Yannopoulos (1994) find that in years 1980-1985 the cross-price elasticities of the demand for the private health care were much closer to zero than unity - between 0.0596 and 0.0623. Thus, their findings illustrate that if the waiting time in the public sector increases by one unit the corresponding increase in the demand for the private health care increases by much less than a unit. In other words, the rate of substitutability between the public and the private health care is low. The meaning of the latter, in terms of the model by Brekke & Sørsgard (2007), is that dual practitioners would have to reduce their labour supply to the public sector by much more than one unit in order to achieve a one unit increase in the demand for private health care².

Tacking stock, the effects presented in the model by Brekke & Sørsgard (2007) require a market with a very particular type of consumers: cash unconstrained, incurring high disutility from waiting on the public sector waiting lists, and with price inelastic demand for the private health care.

¹ Waiting for treatment is assumed to impose costs due to lowered ability to earn. A reduction in working time is assumed, particularly by a reduction in over-time and an increased likelihood of part-time work replacing full-time work. Moreover, on the labour demand side, employers are assumed to reduce the work intensity of individuals facing a hospitalisation.

² For such an effect to take place one requires also an assumption that the decrease in labour supply in the public sector proportionally increases the public sector waiting times.

Hence, the policy implications of the model cannot be taken as general. Besides, it seems worthwhile to recall that contractual relations in the public sector are characterised by a particular rigidity such that, the public sector employees enjoy little freedom with regard to the number of hours they would like to work (Sæther, 2005). Consequently, even if dual practitioners face an incentive to cut down on labour supply to the public sector they might face strong institutional constraints and might be unable to realise their preferences.

4. Dual practice and the costs of the health care provision in the public sector

Part of the literature on physician dual practice focuses on the question how the costs of the public health care provision are affected when the public sector physicians are involved in dual practice (González, 2005; González, 2004; Ferrinho et al., 2004; Rickman and McGuire, 1999). Each of the positions on this topic analyse a different issue. A theoretical paper by González (2004) provides arguments that dual practitioners might be using their public practice to earn a good reputation as a medical provider and increase prestige of their private practice, and analyses how such a behaviour might affect the costs of the public health care provision. Rickman and McGuire (1999) model the effects of dual practice under different reimbursement mechanisms for public hospitals. The model by González (2005) illustrates that dual practitioners might cream-skin low severity (low cost) patients from the public practice to the private one and as a result, cause an increase in the average severity of patients in the public sector and average cost of treatment. Ferrinho et al. (2004) discuss evidence that is claimed to suggest that dual practitioners use the public sector resources (supplies, equipment, sites) in their private practice without an adequate payment to the public sector. Since each of these contributions focuses on a different aspect of dual practice they can be seen as complementary. Below, the contributions are discussed in more detail. First, the theoretical contributions are presented in the following order: González, 2005 – subsection 4.1.; Rickman and McGuire, 1999 – subsection 4.2.; González, 2004 – subsection 4.3. The last subsection 4.4 presents the subject of dual practice and misappropriation of the public sector supplies/resources.

4.1 “*Should physicians’ dual practice be limited? An incentive approach.*” (González, 2004)

González (2004) argues that good reputation in the public practice has a positive influence on the number of private patients and thus, the income from the private practice. Hence, dual practitioners use their success rate in the public sector to earn good reputation as a medical provider. The idea is formulated in a model of a representative dual practitioner’s behaviour in a public practice in which the positive influence of the good reputation for the private practice is introduced as extra earnings of the amount μ , which the dual practitioner can obtain whenever a patient is cured successfully. The conclusions from the model are that dual practitioners face an incentive to systematically over-provide health care services to the public patients. The incentive arises because the over-provision of services increases the probability that the public patient treated by a dual practitioner will be successfully cured in the first attempt and thus, allows the dual practitioner to build and retain a good reputation. Based on the above reasoning, the article discusses whether physician dual practice should be limited. The author points out circumstances under which dual practice might be welfare improving and others under which it might be welfare reducing. Dual practice is concluded to be welfare reducing when the social costs of the over-provision of health care to the public patients are higher than the social costs of under-treatment. In such circumstances, the author

advocates that an introduction of exclusive contracts in the public sector with compensation for foregone private practice profits can successfully help to contain costs.

The analysis is carried out in a model with one patient, one physician (dual practitioner) and a health authority. The patient suffers from a given illness and seeks medical attention in the public sector. For simplicity, it is assumed that the patient's illness can be only of two levels of severity, either low or high, and both severities occur with the same probability. The physician, in attending to the patient, performs two tasks: (a) the diagnosis of the patient's severity and (b) the provision of a treatment. The accuracy of the diagnosis depends positively on the physician's effort e , which is taken to be $e \in [0,1]$. For $e=0$ after the diagnosis the physician does not have any information about the patient's severity except the common knowledge that the low and the high severity are equally likely. The physician incurs disutility from exerting effort, and the effort is not contractible. Thus, when the physician's payment does not depend on the correctness of the diagnosis, the physician exerts minimum effort, which in the model is $e=0$. In the model, the patient can always be given two types of treatment: a mild treatment, which can only cure the low severity, and a more expensive strong treatment, which cures both levels of severity in the first attempt. The physician is paid a fee for providing the treatment to the patient. Yet, if the patient is not cured in the first attempt the physician does not receive a fee for providing the second round of treatment. The patient is assumed to be able to observe the success of the treatment received: He/she realises when the treatment was unsuccessful, as the treatment has to be repeated. To the contrary he/she has no means to discover that the treatment has been excessive. This is why the dual practitioner is assumed to realise the extra earnings μ from good reputation (on top of the fee) whenever the patient is cured in the first attempt.

González (2004) concludes that dual practice reduces welfare when the social costs caused by the over-provision of health care services are higher than the social costs of under-provision of health care. In the model, the social costs include direct costs of the provision of the unnecessarily strong or too weak treatment³ and its indirect costs in form of health losses suffered by the patients due to the too strong or too weak treatment⁴. When the health authorities observe that the social costs of the over-provision of health care are higher than when the public physicians systematically under-treat the patients, the authorities' preferred strategy is to provide an economic incentive for the public physicians such that they systematically prescribe the mild treatment regardless of the patient's severity. Such an incentive is created by introducing two different fees: one for treating the patient with the strong treatment and another for treating the patient with the mild treatment, with the latter being higher than the former. In conditions without dual practice, such an incentive scheme is stated to induce the public sector physicians to systematically prescribe the mild treatment. The physician's involvement in dual practice, however, makes it less probable that the incentive scheme works. This is because, the positive difference between the fee for providing the mild treatment and the fee for providing the strong treatment must compensate the dual practitioner for the probability that the patient will not be cured in the first attempt and that the extra earnings μ

³ The direct costs of the unnecessarily strong treatment are expressed as the difference between the costs of the strong treatment and the costs of the cheaper mild treatment. The direct costs of the too weak treatment are the cost of the unsuccessful mild treatment, since the high severity patient can only be cured with the strong treatment.

⁴ The patient incurs health loss due to over-treatment when his/her severity is low and the treatment is strong. The patient may also suffer health loss due to delayed recovery when his/her severity is high and the first treatment is mild and thus has to be repeated. For simplicity it is assumed that the patient is always cured after the second round of treatment.

from gaining a good reputation will not be realised. The probability that μ will not be realised is stated to be $Pr=0.5$ since the low and the high severity are equally likely and the physician exerts the minimum effort $e=0$. In other words, the author concludes that it is more expensive to induce the public sector physicians to provide the mild treatment systematically when they are involved in dual practice. Thus, the author postulates that it might be optimal for the health authority to offer an exclusive employment contract for the dual practitioner. The exclusive contract includes extra financial compensation for the foregone profits from the private practice, which consist of $(\Pi+\mu)$, where Π is a profit the dual practitioner earns in the private practice regardless of the reputation. The physician can accept the exclusive contract or reject it and continue dual practice. The author analyses also an alternative to the exclusive contract, i.e. a ceiling on the profits from the dual practitioner's private practice. Such a ceiling is concluded to make it cheaper for the health authority to induce the dual practitioner to systematically prescribe the mild treatment. This is because, the dual practitioner might be less interested in the extra earnings μ from the good reputation.

A second scenario analysed by González (2004) is a situation in which the health authority observes that the social costs of the over-provision of health care services and the social costs of the under-provision are both very high. In such circumstances, the author concludes that the health authority decides to induce the physician to exert positive effort in the diagnosis such that the physician knows better whether the patient is of the low severity or of the high severity and provides a treatment most appropriate for the patient's severity, i.e. the mild treatment for the low severity and the strong treatment for the high severity. The author postulates that in order to increase the physician effort in the diagnosis the authority can introduce a positive bonus β , which is paid to the physician on top of the fee in case the physician prescribes the mild treatment and the patient is successfully cured in the first attempt. The bonus β is interpreted by the author as a premium for cost containment. Such reasoning can be, however, questioned since the positive level of effort does not influence the probability that the patient will be cured with the mild treatment, which is always $Pr=0.5$. In other words, if the physician exerts positive effort he/she knows more about the patient's true severity but does not increase the probability of receiving the bonus β beyond the initial probability $Pr=0.5$ for the level of effort $e=0$. The introduction of the bonus β is effectively an equivalent of the situation in which the health authority sets the physician's fee for providing the mild treatment to the patient higher than the fee for providing the strong treatment. For that reason, the analysis of the second scenario does not produce any new findings⁵.

The model by González (2004) points at reputation as a link between the public and the private practice of the dual practitioner, which is an interesting idea indeed. The conclusions derived from the model, however, might raise some objections. First, if the physician is assumed not to receive a fee for the second round of treatment he/she faces an incentive to systematically prescribe a strong treatment regardless of the involvement in dual practice. If one changes this assumption of the model such that the physician receives a fee for providing the second round of treatment, the matters become much more complicated. This is because whether the dual practitioner would prefer

⁵ In fact, González (2004) concludes from the analysis of the second scenario, in contrast to her conclusions from the first scenario, that it is never optimal for the health authority to introduce exclusive contracts in the public sector. Yet, as mentioned above, the conditions in the second scenario are in fact not different from the conditions in the first scenario and hence; the findings can be questioned.

to provide the strong or the mild treatment would depend on the relative sizes of the fee for the second round of treatment on one hand and the extra reputation earnings μ on the other. Further complications arising from such a change in the assumptions is that a physician who is not involved in dual practice might face an incentive to systematically under-treat the patients in order to maximise the total income by increasing the number of treatments. Besides, for the health authority, the introduction of exclusive contract is always inferior to the strategy of inducing the physician to systematically under-treat. This is because the extra financial compensation under the exclusive contract has to compensate the physician for all he/she could earn in the private practice, i.e. both the profit Π earned regardless of the reputation, and the reputation gain μ , while under the strategy of inducing the physician to systematically under-treat, the health authority needs to compensate the physician only for the probability of losing the reputation gain μ . Regarding the ceiling on the dual practitioner's profits from the private practice, which is concluded to make it cheaper for the health authorities to induce the physician to systematically under-treat the patients, it seems worth noticing that the analysis does not account for the costs of the actual enforcement of such a ceiling, i.e. the possible administrative costs.⁶

The model by González (2004) introduces an interesting idea that the reputation in the public practice matters for the profits in the dual practitioner's private practice. Yet, the modeling of the reputation effect leaves some room for improvement and the question whether dual practice is welfare improving or reducing remains open. This is because the policy implications of the model are difficult to apply in practice, since the social costs of the over-treatment and the under-treatment are not observable; in particular, the social costs of the health losses suffered by patients due to a too strong or a too weak treatment.

The idea that dual practitioners use their success rate in the public practice to increase demand for the health care they offer in the private practice is very much worth attention and could be developed further by, e.g. analysing how it influences the dual practitioner's effort in recognising the type of illness not only the severity of a given illness. Eventually, it seems worthwhile to recall the positions in the general health economics literature, which illustrate that all physicians might over-provide health care when its costs do not affect their earnings (Ellis & McGuire, 1986; Ellis & McGuire, 1990; Blomqvist, 1991). The physicians might also face an incentive to over-treat their patients in order to avoid malpractice costs (Gal-Or, 1999). Thus, a relevant question is whether dual practice magnifies this general tendency to over-treat. Rickman & McGuire (1999) investigate the answers to the question in a model of public providers' reimbursement when physicians are involved in dual practice. In comparison to González (2004), their model focuses on a situation in which a dual practitioner faces the possibility to provide both the public and the private services to the same patient. The model by Rickman and McGuire (1999) is discussed in greater detail in the next sub-section.

4.2. "Regulating providers' reimbursement in a mixed market for health care." (Rickman & McGuire, 1999)

The point of departure if Rickman & McGuire (1999) is the literature on the subject of optimal reimbursement of health care providers, such as hospitals and physicians, by third party payers,

⁶ This fact is also recognised by the author of the model.

such as governments and insurers. The health care providers can be reimbursed in several ways and the literature shows that, e.g. a retrospective full-cost reimbursement rule, might lead to over-supply of health care, while prospective fixed budgets might cause under-provision of health care. A general conclusion in this branch of literature is that some degree of cost-sharing between a health care provider and a third party payer – through a combination of a fixed fee and an incentive payment - can lead to the socially optimal level of health care provision (Newhouse, 1996; Ellis & McGuire, 1986). The model by Rickman & McGuire (1999) investigates whether cost-sharing remains the socially optimal reimbursement rule for the public hospitals when their physicians are involved in dual practice. The authors analyse also what are the effects of the prospective fixed-price and the retrospective full-cost reimbursement mechanisms in the public health care sector when physicians are involved in dual practice. The conclusions are that the cost-sharing between the health care provider and the third party payer still leads to the socially optimal level of health care provision when physician dual practice is accounted for. Yet, the degree of cost-sharing might be different – higher or lower – depending on whether the private health care offered by the dual practitioners is complementary or a substitute to the public health care. Moreover, the authors conclude that when dual practice is taken into account, the full-cost reimbursement rule might not necessarily lead to an over-supply of public health care services if the public and the private health care are substitutes. Under the prospective reimbursement mechanism, dual practice might alter for the under-provision of the public health care when the public and the private health care services are complementary. The details of the model are the following:

The model is based on the model of hospital reimbursement by Ellis & McGuire (1986). The main difference is that the physician providing the health care is a dual practitioner; an employee in a public hospital and an owner of a private practice. For a given medical case (for a given patient), the dual practitioner is facing the possibility to provide both the public and the private health care services. In support of the credibility of such setting, the authors recall the fact that the dual practitioners' private practices are often located within the public facilities, e.g. private beds in the public hospitals, which makes it possible for the patients to explore the range and availability of services across both the public and the private sector. The physician derives positive marginal utility from the patient's satisfaction, the income from providing the private health care services, and also from the financial performance of the public hospital in which he/she works. The authors recognise that in the public sector there is no direct link between the physician's utility and the hospital financial performance (the hospital reimbursement mechanism) as physicians receive an exogenously fixed salary. Still, the authors point out that the public hospitals are often encouraged to generate a surplus, and that the physician can become a residual claimant of such a surplus, i.e. derive positive marginal utility from the fact that the surplus can be used for the acquisition of better facilities, equipment or staff. In such a way, the dual practitioner's utility is linked to the public hospital performance and thus, to its reimbursement mechanism.

For the patient, the public health care is free at the point of consumption, but there is a fee for the private health care. The patient derives benefit from receiving both the public and the private health care services. The patient's benefit has a unique maximum in the amount of the public and the private health care consumed. The fee the patient pays for the private health care equals the marginal benefit he/she derives from the private health care.

The dual practitioner in his/her decision on the quantity of health care provided to the patient takes into account the marginal utility he/she can derive from the patient's benefits, the public hospital's surplus, and the income from the private practice. The public hospital surplus and the income from the private practice depend on the costs of the health care provided to the patient. The health care services are assumed to be produced at a positive and constant marginal cost in both the public and the private sector. The dual practitioner trades the patient's benefits, the public hospital surplus and the private practice income at his/her own subjective rate.

If the public hospital is reimbursed under the full-cost reimbursement mechanism, the hospital cannot generate either loss or surplus and thus the dual practitioner cannot derive utility from the public hospital financial performance. In such a situation, the dual practitioner, in the decision on the quantity of the health care, takes into account the marginal utility he/she derives from the patient satisfaction and from the private practice income. The opportunity to earn income from providing the private services to the patient influences the dual practitioner provision of the public health care services such that, when the public and the private services are substitutes, the dual practitioner provides less of the public health care than if he/she was not involved in dual practice. In this way, dual practice restrains the incentive to over-provide the public health care under the full-cost reimbursement mechanism. It may even produce a socially optimal level of public health care provision, i.e. the level of public health care provision when the patient's marginal benefit from the consumption of the public health care equals the marginal cost of the public health care provision (with both the marginal benefit and the marginal cost greater than zero). In comparison, under the full-cost reimbursement mechanism, a physician not involved in dual practice would derive utility only from the patient's satisfaction and thus, in the decision on the quantity of the health care would maximise the patient's marginal benefit from the consumption of the public health care until it reaches zero. If, however, the public and the private health care services are complements the optimisation effect will not take place because, facing the incentive to supply the private health care, the dual practitioner might provide the patient with more public health care than the patient requires and hence, more public health care than is socially optimal.

When the public hospital is reimbursed with a fixed prospective price for each patient treated, the dual practitioner can derive utility from the public hospital surplus, the private practice income, and the patient's satisfaction. The disincentive for increasing the amount of public health care, inherent to the prospective reimbursement mechanism, restrains the amount of public health care provided to the patient by the dual practitioner. When, however, the public and the private health care are complementary for the patient, a socially optimal provision of the public health care might be reached because the increased provision of public health care increases private practice profits. In the earlier model by Ellis & McGuire (1986), which does not account for the fact that the public sector physicians might be dual practitioners, the prospective reimbursement mechanism might be socially optimal only when the physician in his/her decision places equal weight on the patient satisfaction and the hospital surplus, i.e. under the assumption of perfect agency. When dual practice is taken into account, the prospective reimbursement system might be socially optimal without the physician being a perfect agent. If the public and the private health care are substitutes the prospective reimbursement mechanism results in a lower amount of public and private health care being provided than under the full-cost reimbursement mechanism, and the socially optimal provision is not reached in the public market.

Under a cost-sharing reimbursement mechanism, the public hospital is forced to bear a fraction of its own costs. When the public hospital physicians are dual practitioners, the fraction of the costs that has to be internalised by the hospital can be set lower than without dual practice, if the public and the private health care services are substitutes. This is because the dual practitioner's incentive to over-provide services in the public sector is already restrained by the opportunity to earn profits from providing some private health care instead of the public health care. When, however, the public and the private health care are complements, the dual practice brings about an opposite effect, i.e. the fraction of the costs that has to be internalised by the public hospital has to be set higher than without dual practice. This is due to the fact that the private practice income depends on the amount of the public health care provided to the patient and the dual practitioner faces an extra incentive to increase the public sector provision.

The contribution of Rickman & McGuire (1999) illustrates that dual practice might bring about a decrease, an increase, or a social optimum in the costs of the public health care provision. Which case it is depends on the type of reimbursement mechanism for the public hospitals and whether the public and the private health care are perceived as substitutes or complements by the patients. The issue whether the public and the private health care are substitutes or complements is, however, rather difficult to resolve. It depends on such characteristics of a health care system like the basket of the publicly funded services and how the functions of the health care system are divided between the public and the private providers. As already mentioned above, in the section discussing the model by Brekke & Sørsgard (2007), the substitutability depends on the relation between the price of the private health care and the (shadow) price of the public health care. Consequently, whether the public and the private health care services are substitutes depends on the individual consumer's preferences and budget constraints. Thus, it is difficult to make a clear cut conclusion with regard to the dual practice effects for the social costs of the public health care provision. Moreover, when considering the public and the private health care as substitutes the question might be posed whether in a particular medical case the patient consumes both, the public and the private services, or rather receives a treatment in one of the two sectors only. The first scenario requires the treatment to be dividable into a number of sub-services, which might be impossible in certain cases, e.g. surgery. If the treatment can be provided only in the public or only in the private sector, then the dual practice effects discussed by Rickman & McGuire (1999), for the case of the public and the private health care being substitutes, will not take place.

4.3. “On the policy of transferring public patients to private practice.” (González, 2005)

The model by González (2005) analyses dual practice implications for the policy of transferring public patients to the private sector. Under such a policy, public health care sector concludes contracts with private providers for buying treatment for some of the publicly insured patients. The usual motivation for such a policy is to shorten the public sector waiting times. The public authorities and the private providers negotiate the fee per treatment provided to public patients as well as the contingent (number) of public patients that can be transferred from the public sector waiting lists and accepted by the private providers. The fee could be thought of as an average cost of providing a particular type of treatment, e.g. Diagnosis Related Group (DRG) charge for a group of patients with a particular condition. The transfer does not entail any (additional) costs for the public patients. González (2005) underlines that in countries with domination of the public health

care provision, and where such policies function in reality, most or a substantial share of the private health care services is provided by physicians who work primarily in the public sector, i.e. by dual practitioners. The author points out that the prevalence of dual practice brings about drawbacks of the policy of transferring public patients to the private sector. This is because, under such a policy, dual practitioners manipulate the public sector waiting lists to their private benefit. In the public sector it is common for the physician to receive an exogenously fixed salary, while the income from the private practice is usually linked to the private practice profits, even if a physician is an employee (as opposed to a shareholder) in the private sector. The private practice profits depend positively on the fee paid by the public sector for each treatment minus the actual costs of the treatment. Since patients differ with regard to severity, with high severity entailing cost of the treatment higher than the average cost, dual practitioners are expected to face an incentive to cream-skim patients from the public sector waiting lists. In other words, when some of the public patients can be transferred for treatment to the private facilities, dual practitioners, whose financial interests are linked to the performance of these private facilities, manipulate the transfer such that only the low severity (low cost) patients are transferred to the private sector. Consequently, the mean severity of patients remaining in the public sector increases along with the average costs of treatment. The author concludes that dual practice diminishes incentives to introduce a policy of transferring public patients to the private sector or may even cause such a policy to be completely disadvantageous. The details of the model are presented below.

In the model, the public sector is assumed to be capacity constrained and hence, there are waiting times for elective treatments. The private sector, to the contrary, is assumed not to have any waiting times. There is an exogenously given set of patients, who all seek health care in the public sector, and a representative dual practitioner. The patients are homogeneous except for the severity of their conditions, which is measured by a random variable s . The number of patients treated in the public sector is x , while the number of patients transferred to the private sector is $(I - x)$. A patient is assumed to derive a benefit Q from the treatment. The benefit derived from the treatment in the public sector Q^{pb} is assumed to be lower than the patient's benefit from a treatment in the private sector Q^{pv} . The author explains this assumption by stressing that there are no waiting times in the private sector. Thus, $Q^{pv} > (Q^{pb} - \delta)$, where δ is the patient's marginal loss associated with waiting for a treatment. For each treatment provided to a public patient by a private provider the health authority pays an amount w , which is based on an average cost of a particular treatment. The dual practitioner incurs costs while treating patients. These costs depend on the number of patients and on the patients' average severity. In the public sector, the dual practitioner has to treat x patients and in exchange receives a payment that covers his costs and on top of that, gives a fixed amount of money. In the private sector, the dual practitioner income is a share π of the private practice (hospital) profits. The private profits depend on the fee w , the number of transferred patients $(I-x)$, and the unit cost of each treatment. The lower is the unit cost of providing a treatment in the private practice the higher are the practice profits and the corresponding contribution to the dual practitioners' income. The author assumes that the dual practitioner can decide which of his/her patients to treat in the public practice and which to transfer to the private one, subject to the condition that x patients have to be treated in the public sector. The dual practitioner faces a strong incentive for cost containment in the private practice and consequently,

decides to transfer only low severity (low cost) patients to the private practice. In other words, the dual practitioner faces an incentive to cream-skim the low severity patients from the public practice to the private one. As a result, the average severity of patients remaining in the public sector increases together with the average cost of treatment in the public sector. This, however, does not affect the dual practitioner as his/her costs incurred while treating the patients in the public sector is assumed to be always fully reimbursed. The increase in costs affects, however, the public sector gains from introducing the policy of transferring patients to the private practice. The author concludes that the presence of cream-skimming by dual practitioners can lead to a situation, in which the health authority is no longer willing to implement the policy.

The model by González (2005) indicates the potential gains from cream-skimming for dual practitioners. The described technology of cream-skimming seems, however, to over-estimate the physicians' discretionary powers. The policy of transferring public patients to the private facilities usually is aimed at these patients who are already waiting longest. In Ireland the policy was first introduced for the patients who had waited one year or more, and currently, the patients are eligible for the transfer to the private facilities only if the waiting time exceeds certain thresholds (Wiley, 2005). Moreover, the policy as implemented in Denmark, Norway, and Sweden, makes the transfer to the private facilities subject to patients' preferences, i.e. each patient, who fulfils certain criteria with regard to the length of waiting time, has an option to choose a treatment in the private sector (Vrangbæk et al., 2007). Whether the option is used is the patient's decision. Cream-skimming might still be possible, one, however, has to recognise that it is not a straightforward activity and thus, it involves effort such as, e.g. manipulating the order of patients on waiting lists or making persuasive attempt to convince the patient to chose the private sector treatment or dissuade the patient from making such a choice. The manipulative efforts cause a disutility to a dual practitioner. Hence, one can pose the question whether cream-skimming of low severity patients is as feasible and as profitable as argued by González (2005).

Furthermore, the assumption that physicians are reimbursed for all costs incurred while treating patients in the public sector does not reflect well the reimbursement mechanism in the public sector, where physicians usually receive an exogenously fixed salary. The salary does not depend on the severity of patients and thus, does not vary with the physician effort exerted while treating the patients. The effort, to the contrary, varies with the patients' severity such that the more severely ill are the patients the higher are the costs of effort (Barros & Olivella, 2005). In this context, it seems interesting to recall that González (2005) states, that if one allows for the health authority to strategically select patients, the outcome would be that only the most severe patients were transferred to the private sector. The author concludes that the health authority cannot observe the severity of each patient and thus, cannot select them strategically. Yet, the public sector physicians who are not involved in dual practice can observe severity of each of their patients. Consequently, the physicians who are not involved in dual practice could also be seen as able to select which patients they want to treat and which not to treat and transfer to the private sector instead. These physicians might be interested in such a strategic selection in order to decrease the costs of effort exerted while treating patients. The patient selection by the physicians who are not involved in dual practice would take the form of moving the most severe (most costly) patients from the public to the private sector. The effect of the action would be opposite to the effect of cream-skimming by dual practitioners and hence, the final effect for the average severity of the

public patients (and the average cost of treatment in the public sector) is difficult to settle without further investigation.

In summary, it seems that answering the question of dual practice effects for the patients' average severity in the public sector requires further investigation into the details of cream-skimming mechanism. The question could be whether the transfer of the public patients to the private facilities can be indeed manipulated by dual practitioners, and if yes, what the costs of the necessary manipulative effort for dual practitioners are. Another issue for investigation could be whether the strategic selection of patients by the physicians who are not involved in dual practice could neutralise the effect of cream-skimming by dual practitioners.

4.4. Dual practice and misappropriation of the public sector supplies/resources

Dual practice is often associated with the unauthorised use of public sector resources such as medical supplies, equipment, means of transportation, office infrastructure, and support personnel of the public hospitals (Eggleston & Bir, 2006; Jan et al., 2005; Ferrinho et al., 2004; Bir & Eggleston, 2003). Pilfering of pharmaceuticals in the public health care facilities is also speculated to be connected to the prevalence of dual practice. Dual practitioners are suspected of redirecting the pharmaceutical supplies to their private practices (Ferrinho et al., 2004). In the study by Gruen et al. (2002) – see section 2 - in depth interviews with 28 physicians revealed that transfer of subsidised public resources to the private practices takes place in Bangladesh. Still, there is no hard evidence to support that such practices are widespread. It has also been observed that members of the medical personnel in the public sector in low income countries sell pharmaceuticals or other supplies to which they have access in the public sector on the black market. Such a behaviour has been identified as a response to low public sector salaries (Ferrinho et al., 2004). In such a context, dual practice can be seen as an alternative to the trade in stolen pharmaceuticals. Most importantly, it is difficult to say how much such practices can be blamed on dual practice and how much to the poor management and unrealistically low salaries in the public sector. As most promising solution to the problem the literature proposes to improve contracting arrangements and monitoring capacity by, e.g. transparent rental contracts for using public facilities or equipment (Eggleston & Bir, 2006; Bir & Eggleston, 2003).

5. Dual practice and quality of health care

There are widespread concerns about negative effects of dual practice for the quality of the publicly provided health care (Morris et al., 2008; Jan et al., 2005; Ferrinho, 2004; Bir & Eggleston, 2003; Macq et al., 2001, WHO, 2000). The concerns are closely linked to the general observation that the dual practitioners can earn more per unit of labour in the private practice than in the public sector, which finds support in studies carried out especially in developing countries (Jan et al., 2005; Ferrinho, 2004; Gruen, 2002; Macq et al., 2001). Moreover, being paid an exogenously fixed salary in the public sector, physicians enjoy little possibility to increase their public sector income, while in the private practice they have more influence on their profits. Thus, dual practitioners are expected to prefer to concentrate their effort and attention away from the public to the private practice, which offers more flexible income raising opportunities. The results are usually named to be lower effort and absenteeism in the public practice (Ferrinho et al., 2004; Ensor & Duran-Moreno, 2002).

In the above context, self-referrals are seen as particularly problematic. The dual practitioners might purposefully provide a low quality of services in the public sector to encourage their patients to opt out and consume the private health care (Jan et al., 2005; Gruen, 2002). Fears have also been expressed that self-referrals by dual practitioners could be a form of demand inducement (Jan et al., 2005; Bir & Eggleston, 2003; Gruen et al., 2002). The issue of self-referrals and the dual practice effect for the quality of health care is analysed closer in the model by Biglaiser and Ma (2007). Biglaiser and Ma (2007) indicate that in the discussion over the dual practice effects for the quality of the public health care services one must address the question why a dual practitioner would provide higher quality services in the public sector when not engaged in dual practice anymore or before the engagement in dual practice. Dual practitioners are most often assumed to be rational profit-maximisers. If a high quality of the public service is understood to mean that a physician provides a lot of costly personal effort in the treatment, then rational profit-maximisers provide only some minimum acceptable quality in order to minimise the costs of effort, regardless of the involvement in dual practice.

As a measure for the public health care quality one can also employ information on the public sector waiting times (Iversen, 1997). It is often argued that dual practitioners might favour long waiting times in the public sector in order to boost demand for the private health care services (Morris et al., 2008; Jan et al., 2005). An English study by Morris et al., (2008) finds a positive association between the mean number of patients per NHS consultants waiting three or more months and private practice income of dual practitioners across specialties. However, the causality of such an association could not be inferred. It might be that the private income of dual practitioners across specialties is positively related to the length of waiting times because, profitable private practice induced consultants to shirk in the private sector. It might, however, also be that the existence of long waiting times in the NHS encourages patients to use private health care services and thus, encourages the private health care supply. The issue of dual practice prevalence and the public sector waiting times is analysed in greater detail in the models by Morga & Xavier (2001) and Iversen (1997).

There are also arguments that dual practice might in fact contribute to an increase of quality in the public sector (Eggleston & Bir, 2006; Jan et al., 2005; Bir and Eggleston, 2003). Eggleston & Bir (2006) propose a total compensation approach to the analysis of dual practice effects for the public health care quality. The authors argue that a right to involve in dual practice can be perceived by the public sector physicians as a non-monetary compensation for the public sector job. In such a case, allowing dual practice might be advantageous to the public health care systems, especially when the public sector is financially constrained.

In the following subsections, the contributions to the above-delineated subjects are presented in greater detail. First, the model by Biglaiser and Ma (2007) is discussed, followed by a subsection discussing related contributions regarding dual practice and the quality of the public health care. Next come contributions regarding dual practice and the public sector waiting times. The final subsections present the total compensation approach to the dual practice problem by Eggleston & Bir (2006), followed by a related contribution by De Pietro (2006).

5.1. “Moonlighting: public service and private practice.” (Biglaiser & Ma, 2007)

Biglaiser & Ma (2007) examine the effects of dual practice on the public health care service quality and consumer welfare. The focus is on the dual practitioners’ self-referrals of patients from the public to the private practice. The model illustrates that consumers gain when they are referred from the public to the private practice by dual practitioners. The reason behind the increase is that there are consumers, whose individual valuation of health care is so high that they are willing to buy health care services in the private market because there are faster and of higher quality. When these consumers initially seek health care in the public sector, dual practitioners can offer them a referral to the private practice where they receive health care services corresponding better to their individual valuation of health care. Consequently, consumer welfare is increased. The authors emphasise, however, that dual practice might also induce negative behavioural reactions among the public sector physicians, e.g. shirking, which might cause a decrease in the quality of the public health care. The adverse behavioural reactions of the dual practitioners are assumed to be positively linked with the profits from the private practice. Thus, the authors advocate a price regulation in the private health care market to limit the possible decrease in the public health care quality. The supporting argument is that dual practitioners facing a limit on the profits from the private practice might be less prone to the negative behavioural reactions. Moreover, the model illustrates that a ban on dual practice is not advantageous. The details of the model are the following:

There is a set of consumers, half of whom, for the sake of simplicity, are assumed to be cash constrained (poor consumers) and half cash unconstrained (rich consumers). All consumers initially seek health care in the public sector, i.e. there are no consumers who completely opt out from the public health care system. Each consumer would like to receive one unit of health care service and each consumer has his/her individual positive valuation v of the quality of the health care services received. A consumer’s benefit from receiving health care depends on the quality of the health care q and the consumer’s valuation of this quality vq . The consumers’ demand for the health care always exceeds its provision. The consumers do not choose their physicians in the public sector, but are randomly matched with one of them. The public health care is free at the point of consumption, but the consumers have to pay out-of-pocket for the private health care

There are two types of physicians all of whom work in the public sector: a fixed number of dedicated physicians and a fixed number of profit-maximising physicians. The dedicated physicians work exclusively in the public sector. The physicians’ preferences are represented by a combination of their own monetary payoffs and patients’ benefits. The dedicated physicians value the patient benefit more than their own monetary payoff and thus, they want to provide health care of highest quality. The profit-maximising physicians, on the contrary, are assumed to be predominantly concerned about their own monetary payoffs, i.e. the patient’s benefit matters little to them and thus, they provide health care of minimum quality in the public sector. The physicians choose the quality level to be provided to the patients; a higher quality of health care costs the physician more. The physicians in the public sector are paid by the authorities according to the costs of the health care provision they report. The costs depend on the quality of the health care provided. The health authorities allows the quality costs $c(q)$ for $q \in (q^m, q^h)$. The physicians might report their quality costs $c(q)$ untruthfully, i.e. higher than the costs of the real quality provided, and earn a rent which is a difference between the reported costs and the real costs. The health authority may randomly

choose to audit the reported quality at a cost γ .⁷ In case the physician is audited and discovered to report his/her quality costs untruthfully, the physician's payment is adjusted downwards to the payment for the true level of quality.

The physicians can observe whether the consumer is poor or rich. In the public sector, a consumer matched with a dedicated physician receives health care of high quality q^h . This is because a dedicated physician values consumers' benefits more than his/her own monetary payoff and thus, provides the highest quality as long as he/she is reimbursed for the costs of that quality. Consequently, a dominant strategy of a dedicated physician is to provide the highest quality q^h allowed in the public sector and report the true costs $c(q^h)$. Since the profit-maximising physicians have quite opposite preferences to the dedicated physicians, a consumer who is matched with a profit-maximiser in the public sector receives health care of the minimum quality q^m . The profit-maximising physicians report quality costs higher than the true quality costs and earn a rent, when not audited, or break even, if an audit is conducted. Since the audits are costly and thus, are not always conducted, the profit-maximising physician cannot be deterred from untruthful reporting (shirking). If a profit-maximising physician becomes a dual practitioner, a consumer he/she treats in the public sector receives the minimum quality q^m , if he/she is poor, or is offered a referral to the private practice, if he/she is rich. After the referral has been offered, the consumer and the dual practitioner bargain over the price and the quality of the private health care. The agreement on the referral is given by a Nash bargaining solution. Hence, both, the consumer who agreed to the referral and the dual practitioner, earn a surplus over the disagreement point, i.e. when the referral is not accepted by the consumer. The authors argue that the consumer must be better off when accepting the referral even under asymmetric information because the consumer always has an option of remaining in the public sector. Moreover, since the consumer has to pay out-of-pocket for the private health care he/she has stronger incentives to verify the quality of services received in the private practice. Overall, dual practice increases aggregate consumer welfare while having no influence on the quality of the public health care services.

If, however, one assumes that the involvement in dual practice induces reactions, such as an increase in shirking, dual practice might, on one hand, increase the welfare of these consumers who are referred to the private practice, but on the other hand decrease welfare of consumers remaining in the public sector. The authors link the magnitude of the adverse behavioural reactions to the profitability of the private practice. Dual practitioners might face an incentive to reduce quality in the public sector below the minimum quality q^m in order to save time and energy for the very profitable work in the private sector. This might be a result of increasing opportunity costs of providing services in the public practice. In such circumstances, the authors advocate price regulation (or alternatively an income ceiling) in the private health care market, which should control the strength of the incentive to shirk. The private market price regulation is controlling for the quality deterioration in the public sector, but also limits the private market efficiency. Still, the authors' final conclusion is that the positive controlling effect of the private market price regulation is of a higher magnitude (because it benefits all consumers in the public sector) than the negative effect it has for the private market efficiency (because it harms only some rich consumers

⁷ The consumers in the public sector are assumed to have weak incentives to verify the quality of the health care services because they do not pay for them.

demanding very high quality). Hence, when the private practice profits are subject to a limitations the aggregate consumer welfare always increases when dual practice is allowed, even if one assumes negative behavioural reactions.

The contribution by Biglaiser and Ma (2007) produces very important findings with regard to self-referrals and the dual practice influence on the quality of the public health care. It also leaves some issues open. One can, for example, notice that a price regulation imposes some costs on the society due to its administrative requirements and the need for monitoring whether the limits are obeyed in reality. The effectiveness of the monitoring seems to be an issue. This is because, the price limits do not remove the incentive to bargain for higher price and quality in the private market as the consumers who are willing to pay higher price for higher quality are still there, and the dual practitioners might feel inclined to contract with them.

There is also an open question about what happens if the health authority audits the reports of a dual practitioner who provides quality below the lowest acceptable quality q^m in the public health care sector. If there are no punishments in the public sector regardless of the quality of treatment, it cannot be excluded that the profit-maximising physicians (dual practitioners) provide quality below the minimum acceptable quality q^m in any case. It also seems interesting to try to define what indeed is the lowest quality q^m and what it would take to go below it in the provision of health care. One can, for example, pose the question whether providing a treatment of the quality below the acceptable minimum would not impose malpractice costs on the dual practitioner.

Eventually, if the analysis allowed also for the option that there are also consumers who completely opt out of the public health care system and seek health care directly in the private market, it is possible that the dual practitioners' income from the private practice does not depend exclusively on the self-referrals. According to the theory of labour economics, in such circumstances, a physician would be involved in dual practice only if the better paid job, out of the two jobs he/she holds, does not alone produce a satisfactory level of income. Consequently, if the private practice is taken to be the better paid job and dual practice is still prevalent under such circumstances, it means that there must already exist a (natural) income ceiling in the private market. In such a setting the need for an introduction of a price regulation (or additional income ceiling) becomes less obvious.

5.1.1. More on dual practice, consumer welfare and quality of public health care

Biglaiser & Ma (2007) compare a purely public health care system with a mixed public-private one, where the private health care is provided exclusively by dual practitioners. Delfgaauw (2007) analyses the effects of dual practice for the public patients' wealth comparing a public-private system with and without dual practitioners. The author argues that allowing dual practice decreases welfare of public sector consumers as compared to a public-private system without dual practitioners. The model by Delfgaauw (2007) presents a health care system where patients do not face any costs of the public health care at the point of consumption, but they have to pay for the private health care services. There are no waiting times in the public sector. There is a number M of patients. A fraction of patients are poor patients who cannot afford private health care. There are also two sets of physicians, N altruistic physicians and $P-N$ regular physicians. The altruistic physicians always prefer to work in the public sector because they value the patients' benefits high and also value the opportunity to provide treatment to poor patients who cannot afford costly

private health care. The regular physicians value patients' benefits less than the altruistic physicians and thus, provide lower quality of health care in the public sector than the altruistic physicians. Consequently, a public sector patient's benefit is maximised when he/she receives a treatment from one of the altruistic physicians. The author defines the probability that a patient is treated by one of the altruistic physicians as a ratio of the number of altruistic physicians N to the number of patients in the public sector M , i.e. N/M .

Some of the regular physicians work in the private health care market where they attract some of the patients by offering a higher quality than the maximum health care quality in the public sector. As a consequence, relatively rich patients opt out of the public sector. The author argues that the existence of the private health care market, alongside the public one benefits the poor public sector patients because, when some rich consumers seek health care directly in the private market, the consumers remaining in the public sector have an increased probability of receiving a treatment of high quality from one of the altruistic physicians. Regarding dual practice, the author concludes that it decreases the benefits of the private market because fewer patients opt out of the public sector than when different physicians work either in the public or the private markets. The reason is that rich patients know that they can be matched with a dual practitioner in the public sector and bargain on a referral in case they are not satisfied with the public health care quality. The increase in the number of patients seeking health care in the public sector is interpreted to decrease the probability of poor patients to be matched with one of the altruistic physicians. There is, however, one major flaw of the model: defining the probability that a patient is matched with one of altruistic physician as N/M is not correct. Indeed, since the number of altruistic physicians in the public sector does not vary in the model, the probability that a patient is matched with one of the altruistic physicians does not vary either, regardless of dual practice.

Bir & Eggleston (2003) discuss the hypothesis that self-referrals of patients from the dual practitioner's public practice to the private practice might be in fact a form of demand inducement. There exists anecdotal evidence supporting the hypothesis (Gruen et al., 2002; WHO, 2000). The authors underline that consumers with low educational level might be especially vulnerable to demand inducement. Indeed, if a consumer knows little or nothing about what he/she might expect in the public sector in terms of the treatment's quality, a self-referral would not be a Nash bargaining solution described in the model by Biglaiser and Ma (2007). Bir & Eggleston (2003) use a dataset on the choice of public versus private primary health care providers by patients in Indonesia. The authors conclude that the choice of the private provider cannot be well explained by the difference in quality between the public and the private providers, especially for the poor consumers. It should be, however, underlined that the results do not yield any information concerning to the self-referrals, but only concerning to the consumers' initial decision where to seek health care in case of illness.⁸

The model by Barros & Ollivella (2005) sheds more light on the type of patients who accept self-referrals with regard to the severity of their condition. It is often argued that dual practitioners refer only the low severity patients (cream-skim patients) to their private practice. The authors show that

⁸ The study by Mitchell and Sass (1995) is not considered here as evidence in support of the supplier-induced demand hypothesis with regard to the self-referrals. This is because the study analyses physical therapy provision in physician-owned clinics, where the physicians do not provide the physical therapy themselves and thus their financial interests in referring patients to a particular clinic remain undisclosed to the patient.

this is not always the case. Whether dual practitioners are able to refer only low severity patients to their private practice depends on the rationing policy in the public sector. Here, rationing is understood to mean that only patients whose severity exceeds a certain threshold (for a given illness) are admitted to the public sector waiting list. This constrains the choice of the severities from among which dual practitioners can choose. The authors argue that, when the rationing policy sets the severity thresholds high, only the most severe patients are on the waiting lists. As the number of high severity cases is limited in the population of patients, such a rationing policy means that the waiting times are relatively short. For short waiting times, the least severe patients on the waiting lists are less tempted to opt out of the public health care system and buy expensive private services instead. The patients in the model weigh disutility from waiting for treatment against the fee for the private health care. Lower severity patients incur lower costs from waiting and thus, are less likely to accept self-referral the shorter are the waiting times. The authors conclude that in such a setting dual practitioners would have to restore to referring also the more severe cases from the waiting list. The outcome is defined by the authors as partial cream-skimming. The partial cream-skimming is different from full cream-skimming where dual practitioners refer only the mildest cases from the public sector waiting lists to the private practice. Similarly, partial cream-skimming occurs also when the rationing policy is very lax, i.e. when even patients with very low severity are admitted to the public sector waiting lists. This makes waiting times longer but the additional patients on the waiting lists are those who experience lower disutility from waiting and are less eager to pay for the private health care. In short, the model by Barros & Olivella (2005) can be summarised as indicating that dual practitioners might find it difficult to cream-skin the low-severity patients from the public sector waiting lists because, these patients face lower costs of waiting and find it less attractive to pay for the private health care.

Coming back to the argument that dual practitioners might provide quality below the minimum quality q^m in the public sector, physician absenteeism in the public practice is most commonly named to be the real life example of such a situation (Ferrinho, 2004, Ensor & Duran-Moreno, 2002; WHO, 2000). Indeed, when working schedules of the public and the private practice are conflicting, an opportunity cost of presence in the public sector might lead to absenteeism. Ferrinho et al. (2004) reports that due to a widespread absenteeism among dual practitioners, the access to the public health care is affected negatively. The studies cited in support of this argument, however, indicated absenteeism among public sector doctors as a general trend in low-income and middle-income countries. In other words, absenteeism seems to be prevailing regardless of dual practice. The strongest evidence for absenteeism regards physicians in Latin America but the studies identify a climate of blatant impunity as a main factor behind absenteeism, not dual practice (Alcázar & Andrade, 2001; Jaén & Paravisini, 2001; Giedion et al., 2001). Moreover, the studies do not allow for disentangling the dual practice impact from the impact of poor management, unsatisfactory working conditions, and low public salaries. Besides, in such a context, lack of commitment and low productivity are said to be expected. Poor management and lack of financial resources create dissatisfaction among the public health care professionals, and this leads to poor quality of public service. Moreover, in many low income countries both, governments and professional bodies representing the medical profession fail to develop standards for service provision against which quality could be measured (WHO, 2000).

Regarding high income countries, absenteeism among dual practitioners during their public practice hours is stated to be a common practice in the UK (Ferrinho et al., 2004; Ensor & Duran-Moreno, 2002). Yet, the supporting evidence is in fact a single observation by *The Sunday Times* that the health authority had employed detectives to investigate whether its consultants carried out private work during the public practice hours (The Sunday Times, January 15, 1995). Suspicion that dual practitioners in the UK manipulate their public practice to the benefit of the private one was strengthened by evidence on some surgeons' apparent under-involvement in NHS work (Humphrey & Russell, 2004). There is, however, no evidence that such an under-involvement is directly associated with the prevalence of dual practice. In fact there is evidence suggesting that dual practitioners perform more treatments in the NHS than physicians who are not involved in dual practice: A study of English consultant surgeons' performance in the NHS for five surgical specialties in a two-year period (1998-2000) revealed that the surgeons with a part-time contract had significantly higher absolute activity rates according to NHS Hospital Episode Statistics than the consultants with full-time contracts (Bloor et al., 2004). The results from a multi-level model adjusting for case-mix, indicate that consultant surgeons with a part-time contract typically finished 129 more consultant episodes per year. Since the part-time contract status usually means that a physician is engaged in dual practice, the evidence may suggest that dual practitioners belong to a group of more active/efficient physicians. A corresponding opinion about a proactive attitude of dual practitioners is revealed in an Italian study (De Pietro, 2006).

Eventually, in a discussion on dual practitioners' commitment to the public service, other types of dual job holding such as a combination of the public clinical practice with teaching or research activities or non-health related jobs is assessed as neutral to the public service or even beneficial (Ferrinho et al., 2004). It should be, however, noticed that the problem of competition for time and effort might equally arise between the schedules of public service and any other job.

5.2. Dual practice and the public sector waiting lists

Morga and Xavier (2001) point out that long waiting times in the English NHS appear to affect the purchasing of private health insurance. The authors state that the waiting lists in the public hospitals are owned and managed by the hospital specialists who are directly responsible for treating the patients on their lists. In such circumstances, the possibility to provide private health care services may create a perverse incentive to treat fewer patients in the NHS so as to increase the demand for the private health care services. In other words, dual practitioners might face an incentive to minimise the number of patients treated in the public sector. The authors recognise, however, that this argument has little support, as the NHS physicians receive an exogenously fixed salary and do not face any explicit incentive to treat a high number of patients, regardless of the involvement in dual practice. Still, the authors refer to the policy debate in which it has been stressed that the current salary system in the NHS is not efficient and that it should be reformed to include an explicit performance related element(s). With the aim to contribute to the debate the authors build a model of a performance-based payment system in the NHS and analyse possible consequences of dual practice for such a system. The conclusions are that if a performance-based payment system is to be introduced and dual practice is allowed, the authorities should take into account the level of income the dual practitioner might receive from private practice. If the performance-based bonus in the NHS does not outweigh the private practice profits, dual practitioners might not feel inclined to

increase the number of patients treated in the NHS. This is because they would still prefer long waiting times so that patients opt out of the NHS and buy the private health care instead. In other words, the prevalence of dual practice might make the performance-based system more expensive. Regarding the above argumentation, the major issue is the elasticity of the demand for the private health care with regard to the public sector waiting times. If this elasticity is low, dual practitioners might not feel inclined to manipulate the waiting times, as the benefits from such manipulations might be relatively small. The question about the elasticity requires further research. The existing evidence, provided by McAviney & Yannopoulos (1994) indicates that the elasticity is very low. The paper uses, however, data from 1980s.

It seems also interesting to recall from section 2, the English study by Humphrey and Russell (2004), which reports that several of the interviewed dual practitioners acknowledged the charge that physicians allow their waiting lists to grow in order to increase demand for the private health care services. These respondents refuted the argument, however, arguing in turn that due to increased managerial surveillance of the waiting lists such a manipulation would be impossible to pass undetected. Respondents underlined that there exist peers-monitoring and that one needs to stand up to the expectations in order to avoid censure from colleagues. Eventually, the study by Bloor et al. (2004), recalled in the subsection 5.1.1., also appears to undermine the charges that dual practitioners are interested in decreasing the number of treatments in the public sector.

A different approach to the analysis of the dual practice effect for the public sector waiting times is offered in the model by Iversen (1997). The model focuses on the analysis of the effect of a private sector on the public sector waiting times. It provides also insights into the dual practice influence on public sector waiting times in a system where the admissions to the public sector waiting lists are rationed. In the model the consumers choose between the public health care, which is free at the point of consumption and the private health care, for which they have to pay. The choice of the public or the private health care is determined by a consumer's income, public sector waiting time, and the price of the private health care, if the admissions to the waiting lists are not rationed. When rationing of admissions takes place, the demand for the private health care depends additionally on the number of patients admitted to the waiting lists. The number of patients who demand health care (public or private) is exogenously given.⁹ The public sector treatments are provided by a representative hospital. The authorities (local sponsor), which own the hospital, allocates a fixed budget to the hospital. The hospital has a capacity, measured by, e.g. a number of staffed beds. The relation between the hospital's capacity, the number of patients and the waiting times is such that when the number of patients increases, an increase in the capacity is required to maintain the waiting times constant (for a given technology of production). It is concluded that, if the admissions to the waiting lists are not rationed the demand the hospital is facing is exogenously determined. Thus, for a given capacity, the authorities determine also waiting times by the decision on the hospital's budget. In other words, the hospital does not have any degrees of freedom left with regard to the waiting times. Yet, when the authorities decide to ration waiting-list admissions, the rationing is in fact performed by the hospital physicians, which gives them some discretionary powers with regard to the number of patients. The discretionary powers of the physicians are such that, for a given capacity the waiting times can be influenced by changing the number of patients

⁹ It is determined by the number of referrals from general practitioners, who are assumed to act as perfect agents for the patients.

being admitted to the waiting lists. If the physicians who ration the waiting-list admissions are dual practitioners, they will be in favour of a higher number of patients admitted as it increases the public sector waiting times. The dual practitioners' income in the private practice depends positively on the number of patients demanding private health care. Hence, dual practitioners have an interest in increasing the number of patients opting out of the public system due to long waiting times. The changes in the number of patients admitted to the public sector waiting lists do not influence the public salary of dual practitioners, which is exogenously fixed. Of course, an increase in the number of patients admitted to the waiting lists must decrease the demand for the private health care, which is contrary to the dual practitioners preferred outcome. Yet, the decrease in the demand for the private health care due to the increase in the number of admissions to the public sector waiting lists is stated to be offset by the increase in the private demand achieved due to longer waiting times.

The model by Iversen (1997) illustrates how dual practice might matter for the public sector when the admissions to the public waiting lists are rationed. The observation that dual practitioners might use their discretionary powers in the area of rationing of the public sector waiting-list admissions, in order to maximise their private practice profits, introduces an important new aspect into the discussion of the dual practice effects for the public health care provision. The crucial aspect seems to remain whether the decrease in the demand for the private health care, due to the increase in the number of patients admitted to the public sector waiting lists, is always over-compensated by an increase in the private demand due to the longer waiting times. If it is not, dual practitioners would rather favour decrease in the number of waiting-lists admissions, which, according to the model, would lead to shorter waiting times in the public sector. The issue requires further investigation.

5.3. “Physician dual practice.” (Eggleston & Bir, 2006)

Eggleston & Bir (2006) develop a model based on total compensation theory and contracting limitations. The authors argue that a useful approach to analysing dual practice is to think of a government as an employer offering a total compensation package to physicians, which includes both a salary and a non-financial benefit in form of right to engage in dual practice. The authors stress that the argument that dual practice leads to a decrease in the public health care quality ignores the possibility that dual practice might help to retain the most skilled physicians in the public sector. Disallowing dual practice would put the government at a competitive disadvantage for attracting the most skilled physicians. In low-income countries allowing dual practice might help governments to recruit medical staff to assure not only quality, but simply patient access to the public health care services. A government decision on dual practice might be seen such that the government seeks to maintain a given quality in the public health care sector, while minimising the total cost of physicians' salaries and the social costs of dual practice. Put differently, governments allow dual practice when the opportunity costs of raising public sector salaries are high compared to the social costs of dual practice. The authors observe also that the social trade-off between the benefits and the costs of dual practice hinges on contracting and monitoring capacity of the government.

The total compensation approach to the problem of dual practice is of relevance not only for the low-income countries. The authors underline that salaries of the public physicians usually do not differ with the level of skills or reputation. The most skilled and well-known physicians might have

high value in the private market in any country. Hence, the public sector physicians in developed countries might simply compare their public salaries to earnings they could have in the private sector. Dual practice allows them to take at least some advantage of their market value.

The approach proposed by Eggleston & Bir (2006) seems to find support in practice. Jan et al. (2005) indicate that allowing dual practice helps to keep valuable human resources in countries where the public sector is seriously underfinanced. The authors provide an example from South Africa, where private practitioners are offered part-time public contract to provide public health care services in rural areas. Similarly, allowing dual practice was assessed to improve the retention of staff in public hospitals in Bahrain (WHO, 2000). In Greece, when between 1983 and 2002 dual practice was banned, many senior level physicians decided to leave the public sector (García-Prado & González, 2007).

To add to the discussion on the total compensation approach, one can argue that the regulation of dual practice can entail considerable costs. Hence, even without substantial benefits from allowing dual practice, the potential social costs of dual practice have to be set against the potential social costs of implementing and enforcing the regulation.

5.3.1. “Private medical services in the Italian public hospitals: the case for improving HRM.” (De Pietro, 2006)

De Pietro (2006), using the example of the Italian NHS, discusses how public hospitals can benefit in terms of human resource management from allowing their physician to provide private health care services in the public facilities, outside the standard working hours. When dual practitioners are allowed to provide private services within the public facilities, their performance can be observed, which provides information about the physicians’ skills. Similarly to Eggleston & Bir (2006), the author argues that the public hospitals could use dual practice as a tool for competing in the labour market for outstanding professionals or health professionals who are in short supply. It is underlined that the pay schemes in the public hospitals in Italy are rigid due to national collective bargaining and can only partially accommodate the differences in the market value of different professionals.

Private practice run within public hospital facilities is also described as introducing into the public sector a context in which physicians can enjoy greater autonomy, choose the preferred mix of activities, and develop new competencies. In general, an involvement in private practice alongside the public one can offer not only financial rewards but also extra opportunities to realise individual aspirations.

De Pietro (2006) also points out that managing private services provided by the employees of the public hospital might prove difficult due to administrative complications such as a need to run a separate accounting system for the private services. Besides, the borders between the time schedules of the public and the private work might be blurred. Nevertheless, private practice run within the public facilities means that dual practitioners do not charge their private patients directly, as this is managed by the public hospital administration. Such an arrangement allows for control of the extent of the private practice workload, income and level of prices. Moreover, the public hospital is entitled to deduct its own costs incurred due to the provision of treatment to the private patients. In this way, misappropriation of the public supplies can possibly be eliminated. On top of

that, the hospital can generate extra income from renting out its facilities outside the normal working hours.

6. Regulatory response to dual practice

The regulation of dual practice is very often seen as unavoidable, and the choice of regulatory measure constitutes a major concern of many health authorities (García-Prado & González, 2007). A review of different regulatory measures in force is provided by García-Prado & González (2007). Complete bans on dual practice among public sector physicians are reported to be in force in Canada, China, and in some states of India. A ban exists also in Ghana (WHO, 2000). In countries such as Kenya and Zambia only senior physicians are allowed to engage in dual practice. In Indonesia, public sector physicians who decide additionally to open a private practice are conscripted into the public service for at least 3 years. In other countries, the regulation of dual practice takes the form of encouraging public sector physicians to run the private practice within the public sector facilities, outside the normal working hours – Germany and Ireland. As described above, De Pietro (2006) reports such an arrangement to exist in Italy. In France a similar possibility is offered to physicians employed in university hospitals and to physicians with certain specialities in all public hospitals. In Bahrain public hospital specialists can provide limited ambulatory care in public facilities (WHO, 2000). In Austria and Italy dual practitioners face a constraint with regard to the maximum number of services that they can provide in the private practice. Portugal, Spain, and Thailand are examples of countries where public sector physicians who agree not to engage in dual practice receive a monetary bonus on top of their public sector salary. Dual practice is also limited through the imposition of an income ceiling on the private practice profits – France and the UK. In France the income from a dual practitioner's private practice cannot exceed 30 percent of the physician's total income. Regarding the UK, Morris et al. (2008) report that until October 2003 consultants working in the NHS in England with a full-time contract were allowed to undertake private practice but were limited to earning a private income no higher than 10 percent of their NHS income – the so-called “10 percent rule”. Consultants employed on part-time contracts – up to ten-elevenths of a full-time contract – did not face such a ceiling on their private practice income. According to a consultant contract introduced in November 2003, there is no restriction on the income from the private practice, but undertaking private practice at the expense of NHS work can affect pay progression. Moreover, consultants are required to disclose details of their private practice.

In the Italian NHS, physicians' dual practice can take two forms – a combination of full-time employment in the NHS with a private practice within public hospitals facilities or a combination of part-time employment with an external autonomous private practice. In the latter case, NHS physicians who are engaged in the external private practice cannot reach leading positions in the public hospital units/departments. Public hospitals are required by law to reserve 5-10 percent of total beds for private activities¹⁰. In case, the public hospitals cannot accommodate the required percentage of private beds, the law requires them to contract private external facilities (De Pietro, 2006).

¹⁰ In 2001 the percentage of the public hospitals beds used for pay beds was, however, less than 1.5 percent.

As already mentioned in section 5.1., in regulation of dual practice the enforceability of the regulatory measures is a crucial issue. International experience shows that bans are rarely enforceable and dual practice still exists outside the regulatory jurisdiction of the governments. Dual practice is observable despite of the ban in Ghana, India, and Zambia. In low income countries such bans are often not enforced due to a low regulatory capacity resulting from the resource constraint in the public sector (Jan et al., 2005). In Portugal and Greece, where dual practice was banned until 1993 and 2002, respectively, dual practice was common even before the ban was lifted (García-Prado & González, 2007; WHO, 2000). Morris et al. (2008) use data derived from tax returns of over 92 percent (24,407) of all NHS consultants in England for the financial year 2003/4 and calculate the ratio of private practice income to NHS income, and examine this with respect to the 10 percent rule: The ratio of mean private income to the NHS income was 0.26 for consultants on full-time contracts, which was greater than the level stipulated by the 10 percent rule. It should be noticed, however, that the authors state that the data on private income could include self-employed income unrelated to the provision of medical services. Still, the authors refer to the earlier evidence which supports the conclusion that NHS consultants on full-time were breaking the terms of their contracts by earning more than allowed in the private practice.

The issue of regulatory response to dual practice is particularly complex in low-income countries. It is underlined in the literature that crude regulatory responses such as bans on dual practice are not only unenforceable, but simply inappropriate. This is due to complexities such as the need to retain staff in the public sector, where salaries are often below the costs of living (Jumpa, 2007; Jan et al., 2005; Macq et al., 2001). The regulations, when enforced, may also lead to problems such as brain drain of physician to other countries, the growth of the informal health care sector, and an increase in under-the-table payments. Indeed, the bans do not prevent these practices, but instead take them outside the regulatory system. Where it is not realistic to expect the physicians to devote their time exclusively to the public sector job due to low salaries, it should be openly acknowledged. In contrast to the regulatory measures, the value of recognising and allowing dual practice is that it enables some degree of control over the quality and safety of the privately provided health care. Under ban or restrictions at least some private services would be provided in a black market where quality and safety assurances are more difficult to impose. Through the dual practitioners' involvement in the public sector; the public educational programmes have influence on the quality of at least part of the private health care services. Moreover, dual practice provides the policy makers with an option of transmitting public health care initiatives, such as vaccination programmes, into the private health care market (Jan et al., 2005; Ferrinho et al., 2004; Macq et al., 2001).

Ferrinho et al. (2004) also notice that a mere elimination of dual practice, will not automatically guarantee a good code of conduct among public sector physicians, if the working environment is not supportive due to, e.g. lack of career prospects and/or prevalence of favouritism and arbitrariness, which are very prevalent in the public sector of these countries. Hence, it is also stressed that while regulatory measures cannot be enforced in some of the countries due to their costliness, other steps could be taken. These could be actions aimed at improving working conditions and career prospects in the public health care sector by removing such obstacles to equal career prospects as clientelism and arbitrary judgements (Macq et al., 2001).

Moreover, the issues discussed in connection to dual practice seem often to be more about medical practice in general rather than dual practice as an isolated activity. In particular in the countries where the public health care system is heavily underfinanced, dual practice reflects various wider systemic characteristics such as limited financial resources, lack of public service quality standards, limited regulatory capacity or over-supply of physicians. Jumba et al. (2007) underline the importance of understanding the broader macroeconomic and regulatory context in which dual practice occurs. If dual practice or the physicians' behaviour is a consequence of system-level issues, the regulatory response most probably would have to be as far-fetching.

In the contributions discussing the regulatory response to dual practice it is underlined that the incentive to shift effort from the public to the private practice is resulting from the differences in the remuneration schemes in the two sectors – fixed salary with little incentive to increase output in the public sector and fee-for-service in the private one. Jan et al. (2005) propose that one approach could be to change the public sector remuneration system into a fee-for-service. This solution, however, viewed in the light of the broader health economics literature, might cause many new problems such as over-provision of health care services, possibly at the expense of their quality (McGuire, 2000).

In sum, restrictions on dual practice do not remove incentives to provide private health care services and thus, might be inefficient. Regulations such as limits on the private practice income or the number of privately provided services impose administrative costs and have to be coupled with a system of punishments in order to be enforceable. These might pose a challenge not only to the governments in the low-income countries. The introduction of exclusive contracts for the public sector physicians, with a financial compensation for the foregone private practice profits, appears to be an expensive solution. From among the theoretical contributions, presented in the earlier sections, a model by González (2004) proposes the introduction of exclusive contracts. In the model, however, the costs of exclusive contracts are calculated under the assumption that all public sector physicians are dual practitioners. In reality not all public sector physicians engage in dual practice, but the exclusive contracts have to apply to all. Consequently, the costs of exclusive contracts increase by the amounts paid to physicians would not engage in dual practice anyway.

Regulation in form of private services being provided inside the public facilities outside the normal working hours of the public sector seems to be worth more attention. It seems that this form of dual practice may benefit the public hospital due to the opportunity to generate income from using the facilities outside the normal working hours. It also allows for controlling the private practice workload and income. It involves, on the other hand, administrative costs and might lead to misuse of the public supplies and equipment. It might also prove difficult to set clear cut boundaries between the public working time and the time for providing the private health care services. The magnitude of the latter, however, could be limited depending on the general organisational principles in the public hospital.

Eventually, a non-financial motivation behind dual practice cannot be excluded. It might well be that dual practitioners self-impose restrictions on their participation in the private market in order to be able to support the public service. In such circumstances, the imposition of additional legal restrictions on dual practice sends a message of distrust to the physicians, which could have negative implications for their commitment to the public service.

7. Concluding remarks

The economic theory of dual practice is relatively limited and recent. The different contributions to the literature on physician dual practice concentrate often on very distinctive separate issues. In general, there is a need for further analysis and much more hard evidence than is currently available. The literature can be summarised as rather inconclusive. In particular, studies on the subject of dual practice implications for the costs of the public health care provision are inconclusive. It appears, however, that arguments about negative aspects of physician dual practice dominate in the literature. Yet, the negative predictions often depend on strict assumptions adopted in the analysis. The conclusions about negative effects of dual practice for the physicians' labour supply cannot be taken as general, as they do not take into account such characteristics of the real world as limited income of consumers or the inflexibility of physicians' employment contracts in the public sector. The conclusions about the negative effects of dual practice for the quality of the public health care could be biased by the assumption that the physicians working exclusively in the public sector are altruistic as opposed to the dual practitioners who are said to be always pure profit-maximisers. Where both, dual practitioners and physicians not involved in dual practice, are assumed to be rational profit-maximisers, it is not recognised that the single jobholding physicians must have different preferences with regard to the combination of income (working time) and leisure. Meanwhile, the latter has implications for the analysis of labour market participation constraint and marginal cost of effort. Moreover, studies cited in support of arguments about the dual practice negative effects for the quality of the public health care, do not allow for disentangling the possible dual practice effects from the effects of unsatisfactory working conditions, low salaries, or lack of quality standards in the public sector.

Implications from the literature can also be seen as contradictory, e.g. dual practitioners are expected to care about their reputation in the public sector in order to increase demand for their privately offered services but, at the same time, they are also suspected of purposefully providing poor service quality in the public sector, also in order to increase the demand for the private health care services. In the latter case, however, one may ask whether the patients who opt out of the public system due to its poor quality would really like to buy the private services from the same physician. It could be that the two opposing incentives i.e. the strive for good reputation and the incentive to lower the quality, balance each other. Hence, it would be interesting to investigate the matter more closely. Moreover, some of the issues raised in connection to the prevalence of dual practice have already been studied with regard to the physicians in general, e.g. supplier induced demand, preferences for long waiting times, the tendency to over-provide health care, or cream-skimming of profitable patients. Thus, there is a need for more studies comparing the behaviour of dual practitioners with physicians not involved in dual practice, in uniform settings and under the assumption of uniform objectives of the physicians in the two groups.

Further, little is known about the motivation behind dual practice. Concretely, the assumption about rational profit-maximisation as the exclusive motive behind dual practice seems to be derived from the evidence that the private practice offers lucrative financial opportunities. Still, while it seems that dual practitioners take up private practice predominantly to increase their income, it does not automatically imply that dual practice as a whole is a profit-maximising combination. Dual

practitioners seem to face promising opportunities in the private market, but they spend relatively little time in the private practice and predominantly supply labour to the lower paid public sector job. Moreover, labour economics studies indicate the existence of motives other than financial behind multiple job holding, especially among high income, highly skilled professionals. Indeed, with scarce information about the motivation behind the division of labour between the public and the private practice, much about the phenomenon remains unknown. If physicians engage in dual practice for reasons other than financial ones the arguments against dual practice could not be sustained easily.

Finally, implementing a ban or restrictions on dual practice is not free of side-effects. The interests of the physicians and the public employers or the physicians and the patients cannot be aligned costlessly. The challenge is to minimize this shortfall to manageable proportions at reasonable costs. The question whether this is better done by introducing a ban, or restrictions on dual practice, or some less drastic solutions like a disclosure of physicians' financial interests in self-referrals, or no regulation at all remains to be answered. A crucial issue is the effectiveness of such regulatory measures, especially in the markets where incentives to establish a private practice are strong. The current knowledge leaves too many questions open for supporting without a doubt the complicated exercise of regulation of dual practice.

It also appears that a different approach should be taken in the analysis of impact of dual practice on the public health care provision in high-income and low-income countries, as in the latter the public health care sector might simply be inexistent without physicians being able to supplement the extremely low public salaries with income from the private practice. In low-income countries, issues discussed in connection to dual practice are about medical practice in general rather than dual practice as an isolated activity. The regulation of dual practice might prove disadvantageous or inefficient without the recognition of the economic and regulatory factors which might simply force physicians into dual or multiple job holding.

Acknowledgements

The author would like to thank Mickael Bech, Terkel Christiansen and Dennis D. Dietrich for advice and critical comments received during the preparation of this paper.

References

- Alcázar L, Andrade R. Induced demand and absenteeism in Peruvian hospitals. In: Di Tella R, Savedoff WD, editors. *Diagnosis Corruption. Fraud in Latin America's Public Hospitals*. Inter-American Development Bank; 2001. p.123-162.
- Amirault T. Characteristics of multiple jobholders, 1995. *Monthly Labor Review* 1997;120:9-15.
- Averett SL. Moonlighting: multiple motives and gender differences. *Applied Economics* 2001;33:1391-1410.
- Barros PP, Olivella P. Waiting lists and patient selection. *Journal of Economics and Management Strategy* 2005;14:623-46.
- Bennett N, Phillips Carson P, Carson KD, Blum TC. A comparison of traditional and atypical workers: demographic, behavioral and attitudinal differences. *Journal of Business and Psychology* 1994;8:467-474.
- Biglaiser G, Ma CA. Moonlighting: public service and private practice. *RAND Journal of Economics* 2007;38:1113-1133.
- Bir A, Eggleston K. Physician dual practice: access enhancement or demand inducement? Tufts University, Working Paper; 2003.
- Blomqvist Å. The doctor as a double agent: information asymmetry, health insurance, and medical care. *Journal of Health Economics* 1991;10:411-432.
- Bloor K, Maynard A, Freemantle N. Variation in activity rates of consultant surgeons and the influence of reward structures in the English NHS. *Journal of Health Services Research and Policy* 2004;9:76-84.
- Bosworth D, Dawkins P, Stromback T, editors. *The economics of labour market*. Longman; 1996.
- Brekke K, Sørsgard L. Public versus private health care in a national health service. *Health Economics* 2007;16:579-601.
- Conway K, Kimmel J. Male labor supply estimates and the decision to moonlight. *Labour Economics* 1998;5:135-166.
- Cullis J, Jones P. National health service waiting lists. *Journal of Health Economics* 1985;4:119-135.
- Cullis J, Jones P. Waiting lists and medical treatment: analysis and policies. In: Culyer A, Newhouse J, editors. *Handbook of Health Economics*, Vol. 1. Elsevier; 2000. p.1201-1250.
- De Pietro C. private medical services in the Italian hospitals: The case for improving HRM. *Health Policy* 2006;78:56-69.
- Delfgaauw J. Dedicated doctors: Public and private provision of health care with altruistic physicians. Tinbergen Institute Discussion Paper TI 2007-010/1; 2007.
- Eggleston K, Bir A. Physician dual practice. *Health Policy* 2006;78:157-66.
- Ellis RP, McGuire TG. Optimal payment system for health services. *Journal of Health Economics* 1990;9:375-396.
- Ellis RP, McGuire TG. Provider behavior under prospective reimbursement: cost-sharing and supply. *Journal of Health Economics* 1986;5:129-151.
- Ellis RP, McGuire TG. Supply-side and demand-side cost sharing in health care. *Journal of Economic Perspectives* 1993;7:135-151.

- Ellis RP. Creaming, skimping and dumping: provider competition on the intensive and extensive margins. *Journal of Health Economics* 1998;17:537-555.
- Ensor T, Duran-Moreno A. Corruption as a challenge in effective regulation in the health sector. In: Saltman RB, Busse R, Mossialos E, editors. *Regulating entrepreneurial behaviour in European health care systems*. Open University Press; 2002. p.106-124.
- Ferrinho P, Van Lerberghe W, Fronteira I, Hipólito F, Biscaia A. Dual practice in the health care sector: review of evidence. *Human Resources for Health* 2004;2.
- Gal-Or E. Optimal reimbursement and malpractice sharing rules in health care markets. *Journal of Regulatory Economics* 1999;16:237-265.
- García-Prado A, González P. Policy and regulatory responses to dual practice in the health sector. *Health Policy* 2007;84:142-152.
- Giedion U, Morales LG, Acosta OL. The impact of health reforms on irregularities in Bogota hospitals. In: Di Tella R, Savedoff WD, editors. *Diagnosis Corruption. Fraud in Latin America's Public Hospitals*. Inter-American Development Bank; 2001. p.163-198.
- González P. On a policy of transferring public patients to private practice. *Health Economics* 2005;14:513-27.
- González P. Should physician dual practice be limited? An incentive approach. *Health Economics* 2004;13:505-524.
- Gruen R, Anwar R, Begum T, Killingsworth JR, Normand C. Dual job holding practitioners in Bangladesh: an exploration. *Social Science & Medicine* 2002;54:267-279.
- Humphrey C, Russell J. Motivation and values of hospital consultants in south-east England who work in the National Health Service and do private practice. *Social Science & Medicine* 2004;59:1241-1250.
- Iversen T. The effect of a private sector on the waiting time in a National Health Service. *Journal of Health Economics* 1997;16:381-396.
- Jaén MH, Paravisini D. Wages, capture and penalties in Venezuela's public hospitals. In: Di Tella R, Savedoff WD, editors. *Diagnosis Corruption. Fraud in Latin America's Public Hospitals*. Inter-American Development Bank; 2001. p.57-94.
- Jamal M, Baba VV, Rivière R. Job stress and well-being of moonlighters: the perspective of deprivation or aspiration revisited. *Stress Medicine* 1998;14:195-202.
- Jamal M, Crawford RL. Consequences of extended work hours: a comparison of moonlighters, overtimers, and modal employees. *Human Resource Management* 1981;20:18-23.
- Jamal M. How much do we really know about moonlighters? *Public Personnel Management* 1992;21:65-73.
- Jan S, Bian Y, Jumpa M, Meng Q, Nyazema N, Prakongsai P, Mills A. Dual job holding by public sector health professionals in highly resource-constrained settings: problem or solution? *Bulletin of the World Health Organization* 2005;83:771-776.
- Jumpa M, Jan S, Mills A. The role of regulation in influencing income-generating activities among public sector doctors in Peru. *Human Resources for Health* 2007;5.
- Kimmel J, Conway KS. Who moonlights and why? Evidence from SIPP. WE Upjohn Institute for Employment Research, Working Paper 95-40; 1995.
- Lundborg P. Job amenity and the incidence of double work. *Journal of Economic Behavior and Organization* 1995;26:273-287.

- Mabry BD. Economics of manpower and the labor market. Intex Press; 1973.
- Macq J, Ferrinho P, De Brouwere V, Van Lerberghe W. Managing health services in developing countries: Between ethics of the civil servant and the need for moonlighting: Managing and moonlighting. *Human Resources for Health Development Journal* 2001;5:17-24.
- Martimort D. Exclusive dealing, common agency, and multiprincipals incentive theory. *RAND Journal of Economics* 1996;27:1-31.
- Martinussen PE, Hagen TP. Reimbursement systems, organisational forms and patient selection: evidence from day surgery in Norway. *Health Economics, Policy and Law* 2009;4:139-158.
- McAvinchey ID, Yannopoulos A. A cost function approach to the choice between private and public acute care provision. *Scottish Journal of Political Economy* 1994;41:194-211.
- McGuire A, Hughes D. The Economics of the hospital: issues of asymmetry and uncertainty as they affect hospital reimbursement. In: Scott A, Maynard A, Elliot R, editors. *Advances in Health Economics*. Wiley; 2003. p.121-141.
- McGuire T. Physician agency. In: Culyer AJ, Newhouse JP, editors. *Handbook of Health Economics*, Vol. 1. Elsevier; 2000. p.461-536.
- Mezzetti C. Common agency with horizontally differentiated principals. *RAND Journal of Economics* 1997;28:323-345.
- Mitchell J, Sass T. Physician ownership of ancillary services: indirect demand inducement or quality assurance? *Journal of Health Economics* 1995;14:263-89.
- Morga A, Xavier A. Hospital specialists' private practice and its impact on the number of NHS patients treated and on the delay for elective surgery. *The University of York, Discussion Papers in Economics* 2001/01; 2001.
- Morris S, Elliott B, Ma A, McConnachie A, Rice N, Skatun D, Sutton M. Analysis of consultants NHS and private incomes in England in 2003/4. *Journal of the Royal Society of Medicine* 2008;101:372-380.
- Newhouse J. Health plans and health providers: efficiency in production versus selection. *Journal of Economic Literature* 1996;34:1236-1263.
- Perlman R. Observations on overtime and moonlighting. *Southern Economic Journal* 1966;33:237-244.
- Pratt JW, Zeckhauser RJ. *Principals and agents: the structure of business*. Harvard Business School Press; 1991.
- Rickman N, McGuire A. Regulating providers' reimbursement in a mixed market for health care. *Scottish Journal of Political Economy* 1999;46.1:53-71.
- Sæther ME. Physicians' labour supply: the wage impact on hours and practice combination. *Labour* 2005;19:673-703.
- Shisko R, Rostker B. The economics of multiple job holding. *The American Economic Review* 1976;66:298-308.
- The Sunday Times. NHS hires private detectives to spy on doctors. Rogers L, Lightfoot L. (London) 1995;January 15.
- World Health Organization. *The effects of economic and policy incentives on provider practice*. Geneva. 2000.

Studies in Health Economics present the results of health economics research at Institute for Public Health, Health Economics, University of Southern Denmark.

Professor Mickael Bech is editor of the series. He is professor of health economics and head of the department of Health Economics University of Southern Denmark.

Further information
Institute of Public Health
Department of Health Economics
University of Southern Denmark
J.B. Winsløvsvej 9, 1
DK-5000 Odense C
Denmark
Telephone: +45 6550 3081
Fax: +45 6550 3880
email: hmj@sam.sdu.dk
ISBN nr.:978-87-89021-76-8