Research Paper 11/05

The Effect of an Increase in the Rate of Payment on General Practitioners' Intrinsic and Extrinsic Motivation

Yan Feng, Shelley Farrar, Matt Sutton and Ada Ma

November 2011

Yan Feng[a], Shelley Farrar[b], Matt Sutton[c] and Ada Ma[b]

[a]Office of Health Economics, [b]Health Economics Research Unit, University of Aberdeen
[c]Health Methodology Research Group, School of Community Based Medicine, University of Manchester

OHE Research Paper 11/05

November 2011

Corresponding author

Yan Feng
Economist
Office of Health Economics
7th Floor, Southside, 105 Victoria Street
London SW1E 6QT
Email: yfeng@ohe.org
Phone +44 (0)207 747 8863

Acknowledgements

Abstract

This paper investigates how the increased rate of Quality and Outcomes Framework (QOF) payments implemented on 1 April 2005 affects Scottish general practitioners' (GPs) intrinsic, extrinsic and overall motivation. A first difference method is used to model GPs' intrinsic and overall motivation. GPs' extrinsic motivation is modelled using a probit model and Mundlak approach. The main finding is that the increased QOF payment effectively motivated GPs' health care supply, but it also crowded out GPs' intrinsic motivation. The results suggests that using strong financial incentives to further motivate already well-motivated health care professionals may have unintended effects on their performance.

Contents

1.   Introduction

A key question addressed in health economics is how to effectively motivate health care professionals to supply the optimum amount and type of care. In economics, the design of incentives for health care supply is commonly discussed under principal-agent theory. This theory can be used to analyse the ways in which principals (e.g. health authorities) attempt to design 'optimal' contracts which will induce the efforts of their agents (e.g. GPs and hospital staff) and maximise the principals' own utility.

Principal-agent theory predicts that the relation between 'high powered' compensation methods and employees' productivity or outputs is positive. This conclusion is based on the assumption that individuals are rational (Von Neumann and Morgenstern, 1944), driven to maximise their utility and are essentially selfish. The assumption posits that only factors driven by self-interest should be included in the utility function. For instance, for health professionals acting as suppliers in the health care market, these two assumptions (the drive to maximise utility and the dominance of selfishness) are consistent with profit-maximising behaviour. Any behaviour that generates costs but not income will not be included in the supplier's utility function.

Doctors and financial incentives

The existing literature examines the effect of changes in financial incentives on doctors' behaviour. In a review examining the effects of different methods of remunerating UK GPs, Donaldson and Gerard (1989) concluded that the fee-for-service remuneration method can lead to an increased and unwarranted demand for fee-yielding services from patients on the recommendation of their doctors. This inflates health care costs but has little or no effect on health itself. Donaldson and Gerard based their conclusions on evidence gathered from three different lines of research. Firstly, there is evidence that the rate of surgery for common operations in Canada and the USA is higher than in the UK. One explanation for this is that the USA doctors are remunerated for these procedures on a fee-for-service basis (Vayda, et.al., 1984, Vayda, Mindell and Rutkow, 1982, Vayda, 1973). Secondly, there is evidence linking an increase in the number of doctors within a geographic area and increased use of services. Doctors may provide more services in order to maintain their income in the face of greater competition (Cromwell and Mitchell, 1985, Phelps, 1986, Fuchs, 1978). Thirdly, there is evidence of a negative correlation between the reimbursement rates for medical services and medical service intensity (Rice, 1983, Evans, 1974). This evidence suggests that doctors adjust how they treat patients in order to maintain their target income.

Gosden, et al. (2001) conducted an extensive literature review of the impact of payment systems on the behaviour of primary care physicians. They concluded that the fee-for-service system results in a higher quantity of service in primary care compared to the capitation system. They cite, among others, the work of Davidson, et al. (1992) and Krasnik, et al. (1990). Davidson, et al. (1992) provided empirical evidence regarding the effects of alternative approaches to paying physicians caring for children in the Medicare programme in the US. Their results suggest that the child visit per year rate is higher among physicians in the fee-for-service group than among those paid by capitation. Krasnik, et al. (1990) explored the effects on GPs' activities of changing their remuneration from a capitation-based system to a mixed fee-per-item and capitation-based system in Copenhagen city. They found that the GPs who changed from being paid by capitation to mixed capitation and fee-for-service payment increased their number of consultations (both face-to-face and by telephone) and range of diagnostic and curative services. Gosden, et al. (2001) also found that fee-for-service resulted in more patient visits, greater continuity of care and higher compliance with the recommended

number of visits than salary payment. Gosden, et al. (2001) based their work on a study conducted by Hickson, et al. (1987) on the performance of paediatric residents in Vanderbilt Community Clinic (US) between September 1983 and June 1984. The subjects of this study were ten second-year and eight third-year residents. They were placed into nine pairs. In each pair, one individual received $2 per patient visit (the fee-for-service group) and the other received $20 per month (the salary group). Hickson's study found that the overall numbers of completed and scheduled visits were lower among salaried physicians than among fee-for-service physicians.

In addition to physicians in primary care, financial incentives are also found to have a positive effect on the supply of health care for other health care professionals (e.g. consultants in hospitals and dentists). In the UK National Health Services (NHS), hospital doctors are mainly paid by salary for their NHS work. The salary is not directly determined by the quantity or quality of consultants' work, but by the salary band. Their performance related payment is not their NHS work, but the completion of private work outside the NHS. The pay from private work is a function of the number of patients they treated. Morris, et al. (2008) found that the ratio of mean private income to mean NHS income was 26% for NHS consultants in England in 2003/4. However, full time NHS consultants are allowed to earn private income equivalent to no more than 10% of their NHS income. The result suggests that NHS consultants provide more private work than they are allowed to do and is evidence that NHS consultants are strongly extrinsically motivated for their private work.

In Scotland, the payment for self-employed dentists includes a capitation payment for each patient registered and a fee for each treatment administered. This creates a financial incentive for self-employed dentists to increase their service intensity. Chalkley and Tilley (2006) found that self-employed dentists in Scotland are more likely to treat patients who are exempt from payment than their salaried counterparts. This result provides evidence that the type of remuneration can have an effect on the distribution of treatments and hence the supply decisions of health care professionals.

Doctors and non-financial motivation

One of the potential problems with using principal-agent theory to understand and predict how the supply of health care responds to incentives is that it does not take into account the non-monetary motivation of doctors. Elements other than monetary factors may make up a doctor's utility function. For example, doctors could be driven by personal interest, concern for others or social conscience. As the utility derived by doctors from non-monetary motivated activities has not yet been appropriately integrated into their utility function, how non-monetary motivation responds to changes in financial incentives is unknown. It is possible that financial incentives have a negative impact on doctors' non-monetary motivation, and therefore their overall motivation (Frey, 1997). It is this hypothesis that motivates this paper. We attempt to investigate theoretically and empirically the following question: Does using financial incentives that aim to increase health care supply undermine health care providers' intrinsic motivation?

One of the first works to introduce the idea of intrinsic motivation into economics literature was Bruno Frey's paper *Tertium Datur: Pricing, Regulating and Intrinsic Motivation*, published in 1992. His book *Not Just for the Money*, published in 1997, proposed the integration of intrinsic motivation into the utility function and provided empirical evidence to support Frey's 'crowding' theory. Frey's theory retained the concept of the price effect, already widely used by economists, to explain individuals' behaviour in the market: individuals increase their supply of a service in line with financial rewards. In addition to the price effect, crowding theory assumes that the marginal utility derived from intrinsic motivation is affected by financial incentives: financial incentives may 'crowd out' intrinsic motivation, offsetting the price effect on supply. This is the major contribution of crowding theory.

In fact, Frey was not the first economist to question the wisdom of using financial incentives to motivate people. He was inspired by Titmuss's work, published in 1970, on the effect of monetary compensation on blood donation. Titmuss (1970) compared the blood collecting systems in Britain (voluntary) and the US (mainly paid) and concluded that paying for blood adversely affected individuals' altruistic motivation to donate.

Another important figure in the development and application of intrinsic motivation in economics is Julian Le Grand. His book *Motivation, Agency and Public Policy: Of Knights and Knaves, Pawns and Queens,* published in 2003, raised the profile of intrinsic motivation in health economics. Le Grand argued that the introduction of financial incentives may turn health care providers from knights (intrinsically motivated) to knaves (extrinsically motivated). In other words, strong external incentives may affect the behaviour and motivation of health care professionals in potentially harmful ways.

Specific to the health care sector, it is suggested that income is not the only influential factor in health care providers' utility functions. Marshall and Harrison (2005) argue that financial incentives play an important role in motivating NHS health care professionals. However, beyond this, 'internal moral motivation' also drives professionals' behaviours. Dolea and Adams (2005) suggest that health care providers are motivated not just by payment but also by the work itself, their relationships at work, workplace conditions, opportunities for professional development, how they are managed, and organisational policies. Scott (2002) argues that health care professionals are motivated by a combination of self-interest, altruism and other factors. A doctor may justify a treatment on clinical grounds as being of likely benefit to the patient, while also receiving remuneration or an enhanced reputation. Scott (2000, 1997) argues that in addition to income, other issues also should be factored into the utility function of health care providers, such as patients' welfare, interests (health and other undefined interests) and economic well-being; the interests of society; control of supplier-induced demand; doctors' reputation, status, intellectual satisfaction and autonomy; and group norms. Gothill (1998) suggests that doctors are motivated to help others in an attempt to develop and maintain their positive identities. These identities are in part shaped by their desire to be perceived as good doctors by their patients and co-workers.

Over the last two decades, economists have found that neoclassical microeconomic theory is not always able to explain how employees respond to incentives. Introducing intrinsic motivation to create a richer model of human behaviour might help researchers better explain the behaviours observed (Falk and Kosfeld, 2006, Fehr and Gachter, 2000, Frey, 1997).

Intrinsic motivation and extrinsic motivation: Identifying a workable definition for economics

Intrinsic motivation is not well defined in economics. It is a concept taken from psychology. Deci (1975) defined intrinsically motivated behaviour as: *'behaviors which a person engages in to feel competent and self-determining'.* Ryan and Deci (2000a) defined it as '*the doing of an activity for its inherent satisfactions rather than for some separable consequence'.* Similar definitions can be found in recent psychology literature. Psychologists have a clear, consistent and narrow definition of intrinsic motivation. When health economists talk about the intrinsic motivation of health care professionals, they usually refer to professionalism and altruism, subjective terms that do not appear in the definitions formulated by psychologists.

From the economists' point of view, the important distinction between extrinsic and intrinsic motivation is whether the motivation is the result of financial incentives. The psychologists' definition of intrinsic motivation is too narrow for the purposes of economic study as it includes only

one aspect of non-extrinsic motivation. For psychologists, behaviours that result from non-extrinsic motivation[1] are not incentivised by tangible rewards such as payment. It could be argued that the non-extrinsic motivations defined by psychologists may be generalised by economists as intrinsic motivation.

Intrinsic motivation could be redefined for economists as motivation that incentivises an individual to perform without external intervention. Gregg, et al. (2008) and Bell and Hart (1999) have used unpaid work as a proxy for public sector motivation. This paper uses individuals' supply of non-QOF-incentivised work as the proxy for intrinsic motivation in the theoretical and empirical analyses.

The concept of extrinsic motivation in psychology is quite similar to that applied by economists. The assumption made in microeconomics that individuals are essentially rational and self-interested is equivalent to the psychologists' assumption that individuals are extrinsically motivated. Extrinsically motivated behaviours derive utility from the external rewards linked to the behaviours: '*It is a construct that pertains whenever an activity is done in order to attain some separable outcomes*' (Ryan and Deci, 2000a).

Intrinsic motivation and financial incentives

Changes in remuneration can affect health care providers' supply of intrinsically motivated work in two ways: by affecting the utility they derive from intrinsically motivated work and the income they forgo by working unpaid. Frey's crowding theory (1997) is used to support the integration of utility derived from unpaid work, while Becker's time allocation theory (1965) provides support for the inclusion of forgone income. This section briefly discusses how they might be applied to develop a theoretical model of unpaid labour supply in health care.

Frey (1997) suggested that individuals' intrinsic motivation will be 'crowded in' when they feel their work has been recognised by the incentives. In contrast, he listed four psychological conditions under which incentives have a negative effect on agents' intrinsic motivation. First, both commands and rewards could 'crowd out' individuals' intrinsic motivation, as they make individuals feel restricted. Second, performance-contingent rewards may make individuals feel the locus of control has been shifted from intrinsic to extrinsic incentives. Third, how individuals perceive the external intervention depends on their level of intrinsic motivation. The intervention might actually undermine highly intrinsically motivated individuals if they see those with low intrinsic motivation being similarly rewarded. Fourth, intrinsic motivation might be crowded out if the external intervention does not imply acknowledgement of agents' intrinsic motivation.

Becker (1965) argues that indirect cost should be treated in the same way as direct cost. Indirect cost is defined as the market value of the time forgone while direct cost refers to market price of goods. The time spent on activities that do not bring monetary benefit to the participant is defined as unpaid hours. In his model, individuals maximise their utility by consuming a set of commodities within given budget and time constraints. Ideally, marginal utility from the last unit consumed should equal marginal cost, which is the sum of the goods' market price (direct cost) and the value of

---

[1] According to self-determining theory, an individual's motivation can be categorised into one of five classes (Deci and Ryan, 2000). Depending on the degree to which motivation has been internalised and integrated into one's personal values, the five types of motivation are: external motivation, introjected motivation, identified motivation, integrated motivation and intrinsic motivation (Ryan and Deci, 2000a,b).

time spent (indirect cost). The value of time spent is the forgone income which could have been earned by undertaking paid work rather than unpaid work. Applying this idea to analyse health care professionals' intrinsic motivation, the resources spent on services that bring no additional income can be reconstructed by adding together the material cost and the forgone income. Within this framework, increasing the price paid per unit of paid work will undermine the amount of unpaid work supplied because the value of forgone income has also increased.

This paper integrates the intrinsic motivation of GPs' into their utility function. The developed theoretical model makes the link between GPs' incomes with their extrinsic and intrinsic motivation. The empirical analysis will test the hypotheses that are drawn from the theoretical model developed.

2. Quality and Outcomes Framework

GPs play an important role in the UK National Health Service (NHS) system. With the exception of accidents and emergency care, GPs are usually the first point of contact for patients for all medical services. As gatekeepers, GPs are responsible for referring patients from the primary care sector to the secondary care sector. Around three per cent of the NHS workforce in Scotland work as General Medical Practitioners[2].

In the UK, GPs work in practices either single-handed or in a group with other GPs. On average, practice list size is around 5,300 with 4.4 GPs[3] per practice (Gravelle, et al., 2010). As independent contractors, they provide primary care services to registered patients. They bear the cost of care such as the cost for premises and nursing staff. Practices are run as for-profit businesses. Thus, the major driving force behind the provision of these health care services is the generation of income. However, GPs may also be intrinsically motivated to provide excellent care to their patients and may work overtime to achieve that. It is common to observe employees in the health care sector working beyond their contracted hours or their contracted level (Wise, 2005).

The GPs' contract is the mechanism through which government attempts to incentivise GPs. Since the 1990s GPs' main income has been from capitation payment supplemented by additional fees for specific services (e.g. payments for maternity care), target payments (e.g. target for child vaccination and cervical cytology screening service) and allowances (e.g. payments for seniority and postgraduate education activities) (Weller and Maynard, 2004). One of the most significant changes in GPs' contracts over the past two decades was the introduction of the pay-for-performance system in the form of the Quality and Outcomes Framework (QOF) (Roland, 2004). Pay-for-performance is an incentive tool widely used in public agencies and organisations (Dixit, 2002, Dixit, 1997). It creates a direct connection between payments to individuals and the quality and/or quantity of the work they produce. The new General Medical Service (GMS) contract was introduced in the NHS on 1st April 2004. As part of the new contract, QOF made performance payments an important part of GPs' income, though capitation remains the main form of payment to GPs. The QOF measures a combination of process and outcome, and therefore can be used as a proxy for the quality and quantity of care provided by GPs.

2.1 QOF payment mechanism

QOF performance data is centrally held in a national IT system called the Quality and Outcomes Framework Management and Analysis System (QMAS). QMAS gives general practices and NHS Boards objective evidence and feedback on the quality of care delivered to patients. It supports the QOF element of the GP contract and has been in operation since 2004. This dataset, maintained by QMAS, makes is possible to evaluate how the increase in QOF payments that was introduced in April 2005 has affected GPs' performance; performance data from before and after the policies were implemented is part of the QMAS dataset.

This paper focuses on Scottish GPs' performance in the clinical domain indicators as the majority of these indicators have maximum and minimum payment thresholds, not the case for most of the non-clinical indicators. Thresholds are the key variables used to calculate GPs' extrinsic and intrinsic

---

[2] The total number of NHS Scotland staff was 165,551 as at 30 September 2008, 4,916 of whom worked as General Medical Practitioners (ISD [2010b]).
[3] The dataset used in this paper shows that the average number of GPs in a practice is 4.48.

motivation. The QOF payment for the clinical domain is the sum of the payments achieved from all the clinical indicators. The payment is calculated in the same way for all indicators: achieved points multiplied by the value of adjusted pounds per point.

First, the achieved points are determined by the practice's performance as measured by the performance index (α_i). This is a nonlinear function of the proportion of patients treated. No points are awarded if the performance index is lower than the minimum payment threshold (α_l). Full points are given if the performance index reaches or is above the maximum payment threshold (α_u). Achieved points are linearly correlated to the performance index somewhere between the maximum and minimum payment thresholds. The performance index for indicator i at year t in practice g is calculated by using the number of patients treated (d_{igt}) out of the number of eligible patients (D_{igt}). It is explained by equation (1). The relation between the performance index and achieved points for one clinical indicator in QOF is shown by Figure 1.

$$\alpha_{igt} = \frac{d_{igt}}{D_{igt}} \tag{1}$$

Second, the raw value of pounds-per-point is adjusted by the Contractor Population Index (CPI)[4]. This adjustment gives practices with large lists a high pounds-per-point value. The value also is adjusted by the Adjusted Disease Prevalence Factor (ADPF) before it is used to calculate payments. This makes QOF payments high for practices with high prevalence rates. After these two adjustments have been made, each general practice will have a unique pounds-per-point value for every disease.
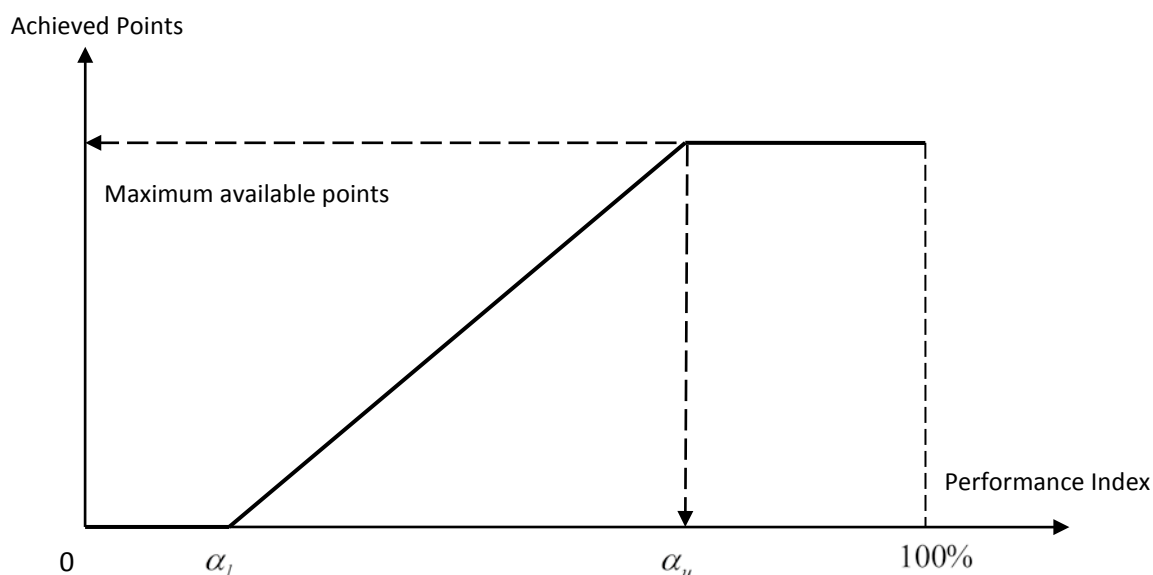
Figure 1. Performance index and achieved points for one clinical indicator in QOF

---

[4] The CPI is a national average practice list size; an individual practice's actual list size is used to calculate a 'population factor' for payment purposes. CPI is calculated by the number of registered patients in a practice divided by the average list size of the practice in Scotland. In this paper, the average list size takes 5,095.

2.2   QOF payments in 2004/5 and 2005/6

The areas of QOF payment, allocation of QOF indicators, points and the value of pounds-per-point in 2004/5 and 2005/6 are summarised in Table 1.

This paper focuses on the performance of Scottish practices in the clinical domain, which contained ten disease areas in 2004/5 and 2005/6. Each disease area has a number of indicators with a number of points allocated to each indicator. In each financial year, there were 76 indicators and 550 points allocated to the clinical domain. The pounds-per-point value was £75 in 2004/5, increasing to £124.64 after 1 April 2005. The effect of the change in pounds-per-point value under the QOF scheme is evaluated in the empirical part of this paper.

Table 1
QOF payment areas, allocation of QOF indicators, points and pounds-per-point values in 2004/5 and 2005/6

	2004/5 and 2005/6		Pounds-per-point in 2004/5	Pounds-per-point in 2005/6
	Number of indicators	Total points available		
Clinical domain	76	550	variable[1]	variable[1]
Organisational domain	56	184	£75	£124.64
Patient experience domain	4	100	£75	£124.64
Additional services domain	10	36	variable[2]	variable[2]
Additional payment points				
Holistic care	-	100	£75	£124.64
Quality practice	-	30	£75	£124.64
Access bonus	-	50	£75	£124.64
Total		1050		

1: Within the clinical domain, the baseline payment per point is adjusted up or down for each practice according to a 'prevalence rate' derived from the QOF register applicable to each individual indicator. A QOF prevalence rate is simply the total number of patients on the register, expressed as a proportion or percentage of the total number of patients registered with the practice.
2: Within the additional services domain, the baseline payment per point is adjusted up or down for each practice according to the number of patients within the target population for each additional service type, relative to the national average target population size for that additional service.
3: Source: http://www.isdscotlandarchive.scot.nhs.uk/isd/3305.html [Accessed 1 November 2011].

3. Theoretical model

This section lays out the theoretical framework for the empirical analysis. The developed theoretical model aims to explore whether the increased QOF income in 2005 crowded out GPs' intrinsic motivation. It also aims to explore the effect of increased income on the structure of GPs' working motivation under QOF. The list of symbols is explained in Table 2.

Table 2. List of symbols

Symbols	Definitions
α_l	Lower payment threshold
α_u	Upper payment threshold
α_i	Performance index
t_u	Time input to produce one unit of unpaid QOF work
MC_u	Marginal material cost to produce unpaid QOF work
MC_p	Marginal material cost of produce paid QOF work
$\bar{w}$	Average unit time earnings from the time spent on paid QOF work
λ	Marginal utility of money
MU_u	Marginal utility derived from unpaid QOF work
MU_p	Marginal utility derived from paid QOF work
$(MU_u)_{\bar{w}}$	Effect of changes in income on the marginal utility derived from unpaid QOF work
$(MU_u)_{z_u}$	Effect of unpaid work outputs on the marginal utility derived from unpaid QOF work
Z_u^*	Optimal amount of unpaid QOF work in 2004/5

3.1 The measurement of GPs' extrinsic and intrinsic motivation

Technology is assumed to be constant. Practices are assumed to be motivated by money only when the performance index α_i is no higher than the upper threshold, i.e. $\alpha_i \in [0, \alpha_u]$. Practices with a performance index located in $(\alpha_u, 1]$ (i.e. beyond the maximum payment threshold) are assumed to be both intrinsically and extrinsically motivated; at this point, GPs can make no further income from QOF. The level of intrinsic motivation can be measured by looking at outputs from the number of patients treated over and above the upper threshold, i.e. $\alpha_i - \alpha_u$.

The QOF services supplied over the maximum payment threshold are incentivised by their intrinsic motivation. Strictly speaking, they are not unpaid. GPs have contracted responsibilities to provide health care services to all the registered patients. They receive payments, such as capitation, at least for all the services supplied. The services which described as unpaid QOF service are in fact those that are not incentivised by the QOF scheme. This description applies to the rest of this paper for all services referred to as 'unpaid'.

3.2 Effect of increased QOF income on GPs' unpaid and paid QOF work

GPs may stop supplying services before the maximum payment threshold is achieved. Equation (2) describes this condition.

$$MU_p = \lambda \times (MC_p) \tag{2}$$

Equation (2) suggests that the equilibrium condition for GPs' performance in 2004/5 was achieved when their supply of QOF services reached the point where marginal cost was equal to marginal income. MU_p is the marginal utility derived from QOF payment for paid QOF services. MU_p decreases as outputs increase because of the greater effort required. λ is the marginal utility of money. As practices operate as businesses, GPs also bear the material costs of the care provided. MC_p is the marginal material cost to produce one unit of paid QOF work.

Where GPs go beyond the full payment threshold and supply unpaid QOF services, the equilibrium condition is achieved when the marginal utility derived from unpaid work is equal to the marginal material cost and forgone income. Equation (3) describes this condition.

$$MU_u = \lambda \times (\overline{w}t_u + MC_u) \tag{3}$$

MU_u is the marginal utility from supplying unremunerated work under QOF. MU_u decreases as outputs go up as a result of the increased effort required. $\overline{w}t_u$ is the value of time or forgone income to provide one unit of unpaid work under QOF. $\overline{w}$ is the average unit time earnings from the time spent on paid work. t_u is the units of time input to produce one unit of unpaid work. MC_u is the marginal material cost to produce one unit of unpaid QOF work.

Equation (3) suggests that the equilibrium condition is achieved when the marginal utility from unpaid work equals the marginal costs. These costs include both opportunity cost and material cost.

Where GPs' performance surpasses the maximum payment threshold for some indicators but not others, the equilibrium condition in these practices may be described as follows:

$$MU_u = \lambda \times (\overline{w}t_u + MC_u) = MU_p = \lambda \times (MC_p) \tag{4}$$

Equation (4) suggests that where practices do not achieve the maximum available QOF points in the clinical domain, GPs must choose a level of production that maximises utility from both paid and unpaid work. The equilibrium condition for this is when the marginal utility from unpaid work is equal to the marginal utility from paid work.

The equilibrium conditions (2), (3) and (4) are the results of maximising the utility of QOF work whether or not it attracts additional payment, under the constraints of working time and QOF payment function. Practices which have achieved full points in all clinical indicators can choose how to spread their unpaid QOF work among the various indicators. Practices that have not yet to achieve full payment from any clinical indicators must choose how to spread their paid work across all clinical indicators. Practices that only have achieved full points in some QOF clinical indicators must choose between doing paid work for the indicators for which they have not yet achieved full points and doing unpaid work for the indicators for which they have already achieved full points. If

they have achieved full points in more than one indicator, they must choose to one to which they will dedicate more of their unpaid hours.

The policy change in 2005/6 (to increase the value of $\overline{w}$) increased the value of marginal income for paid QOF work in 2005/6. The marginal material costs to produce paid QOF work did not change over time.

$$MU_p > \lambda \times MC_p \qquad (5)$$

Equation (5) suggests that the policy change in 2005 provides an incentive for GPs to treat more patients where their performance index fell below the maximum threshold in 2004/5.

Using equations (3) and (4), the following section explores how the 2005 increase in QOF income affected GPs' optimal intrinsically motivated outputs. Taking the first order condition of $\overline{w}$ on both sides of equation (3) (it is equivalent to the intrinsic motivation part of equation [4]), we get equation (6). The effect of the increased income on GPs' intrinsic motivation could be specified as follows:

$$\frac{\partial Z_u^*}{\partial \overline{w}} = \frac{\lambda \times t_u - (MU_u)_{\overline{w}}}{(MU_u)_{Z_u}} \qquad (6)$$

Z_u^* is the optimal amount of intrinsically motivated outputs in 2004/5. The effect of increased income on the marginal forgone income is positive $(\lambda \times t_u > 0)$. As suggested in Becker (1965), when forgone income is unobserved, it can be treated as the value of paid income for the same services. Practices receive no further remuneration for treating patients over the maximum threshold in QOF, so the value of forgone income for all intrinsically motivated outputs increased in 2005 following the increase in cash payment per point. $(MU_u)_{Z_u}$ is the effect of unpaid work outputs on the marginal utility derived from unpaid work. It is negative because the marginal effort that is required to produce unpaid work increases in line with increases in unpaid outputs. The effect of increased income on the marginal utility of supplying unpaid work is unknown, as it depends on how GPs perceive the increased payment. Where the value of $(MU_u)_{\overline{w}}$ is positive, it suggests that GPs' perception of the increased income is positive.

If the effect of the increased QOF income on the marginal utility from supplying unpaid work is positive and the effect is not big enough to offset the positive effect on marginal forgone income $[\lambda \times t_u > (MU_u)_{\overline{w}} > 0]$, then the increase in QOF income has a negative effect on unpaid work $(\frac{\partial Z_u^*}{\partial \overline{w}} < 0)$;

If the effect of the increased QOF income on the marginal utility from supplying unpaid work is positive and the effect is big enough to offset its positive effect on marginal forgone income $[(MU_u)_{\overline{w}} > \lambda_1 \times t_u > 0]$, then the increase in QOF income has a positive effect on unpaid work $(\frac{\partial Z_u^*}{\partial \overline{w}} > 0)$;

If the effect of the increased QOF income on the marginal utility from supplying unpaid work is negative $[\lambda_1 \times t_u > 0 > (MU_u)_{\bar{w}}]$, then the increase in QOF income has a negative effect on unpaid work $(\frac{\partial Z_u^*}{\partial w} < 0)$.

Hypothesis 1: High income fosters GPs' intrinsic motivation, as having more disposable income encourages GPs to supply more unpaid work.

Hypothesis 2: High income will undermine GPs' intrinsic motivation as the associated opportunity cost is high. Equation (6) states that the monetary value of unit of time $\bar{w}$ will increase as income increases. Forgone income $\bar{w}t_u$ will rise accordingly, causing the marginal forgone income of supplying intrinsically motivated services to go up $(\lambda \times t_u > 0)$.

Hypothesis 3: The high marginal utility derived from unpaid work will foster GPs' intrinsic motivation. However, how this utility relates to income is unknown. Therefore, the effect of increased income on the marginal utility of supplying unpaid work also is unknown.

Hypothesis 4: High income fosters GPs' extrinsic motivation as the marginal utility from paid work increases. Referring to equation (5), it is suggested that increased income results in increased marginal utility MU_p. Therefore, high income has a positive effect on GPs' extrinsic motivation.

4.   Data

This paper explores the change in the performance of Scottish GPs under QOF using data from financial years 2004/5 and 2005/6. Practices' performance data and QOF policy data was downloaded from the Information Services Division (ISD)[5]. Data on GPs' characteristics comes from the GP Contractor Database held at ISD. Only observations from GMS contract practices are used in the empirical analysis.

Thirty nine indicators are eligible for this research. The names and definitions of the indicators are listed in Appendix A. The criteria for eligibility are described below. First, the eligible indicators must have maximum and minimum payment thresholds. They are used to distinguish GPs' intrinsic and extrinsic motivations. Second, eligible indicators should cover the same period of fifteen months[6].

The sample used for the analyses includes 63,574 observations collected from 869 practices, 31,787 observations for each financial year. The unit of observation is the information submitted by practices in a particular financial year for one clinical indicator. The summary of these statistics are presented in Tables 3 and 4. Information about the maximum and minimum payment thresholds, available points and practices performance for the 39 selected clinical indicators under eight diseases domains in 2004/5 and 2005/6 is presented in Appendices B and C.

Table 3
Summary of statistics for observations in 2004/5

Year	2004/5			
	Mean	SD	Min	Max
Performance index	95.49	12.70	0.97	100
Marginal income	8.19	13.61	0.98	300.86
Threshold for full QOF achievement	114.35	101.14	3	838
Thousands of patients per GP	1.31	0.48	0.08	3.80
List size	5551.61	3176.90	153	23324
Observation	31787			

Table 4
Summary of statistics for observations in 2005/6

Year	2005/6			
	Mean	SD	Min	Max
Performance index	98.86	5.70	6.48	100
Marginal income	13.12	19.96	0.67	329.15
Threshold for full QOF achievement	118.43	103.04	3	830
Thousands of patients per GP	1.31	0.47	0.07	3.80
List size	5564.42	3181.11	146	23525
Observation	31787			

---

5 http://www.isdscotlandarchive.scot.nhs.uk/isd/3305.html (ISD [2010b]). [Accessed 1 November 2011].
6 There are two types of clinical indicators: 1) indicators with a nine month period and 2) indicators with a fifteen month period. Clinical indicators under type 1) are excluded from our analysis. The performance on indicators that cover different time periods are not comparable. It is likely that the changes in performance in indicators with a long time period are greater than those with a short time period. Furthermore, rate of change may different between the first nine months and the second six months for indicators with a fifteen month period.

5.   Variables

This paper models three types of GP motivation: intrinsic, extrinsic and overall motivation.

5.1 Dependent variables

Intrinsic motivation: It is measured by the number of patients treated in excess of the number of patients that need to be treated to get the maximum QOF payment divided by the number of patients could be treated without QOF payment. Practices do not receive additional remuneration from the QOF scheme for treatments beyond the maximum payment threshold. We suggest in the introduction that intrinsic motivation can be measured by quantifying the extent to which activity goes beyond the contracted level. Applying this principle to the QOF, intrinsic motivation can be quantified as that part of practices' output that is produced over and above the maximum payment threshold.

The motives behind the amount of unpaid QOF work undertaken can be complicated. For instance, it may reflect that GPs' desire to provide high quality care to encourage more patients to register with their practices. Or they could be motivated by altruism and professionalism, based on genuine concern about the welfare of patients. Intrinsic motivation also might be motivated by social norms. For instance, because GPs care about their reputation, they perform better than the assigned performance level. The measurement of intrinsic motivation that is used in this paper (the supply of unpaid work) cannot distinguish among the motives for unpaid QOF work.

Extrinsic motivation: It is measured by a dichotomous variable that indicates whether a practice reaches the maximum payment threshold. It is equal to one if the practice reaches the maximum payment threshold; otherwise it is given a zero. As discussed in the introduction, extrinsic motivation can be defined as that part of motivation that is incentivised by payment. Achieving the maximum payment threshold is a sign that GPs are extrinsically motivated.

Overall motivation: It is measured by the number of patients treated out of the number of eligible patients. Assuming GPs work from a combination of extrinsic and intrinsic motivation, their overall motivation should be indicated by their overall performance, including both paid and unpaid activities.

5.2 Independent variables

The independent variables include the *monetary value of QOF services*, *GPs' characteristics* and *patients' characteristics*. In total, there are 17 independent variables.

Three out of the 17 variables are time variant. They are the *marginal monetary value of QOF services supplied by GPs, thousands of patients per GP* and *number of patients needed to treat to get the full QOF payments.* The remaining 14 variables are time invariant.

Marginal monetary value of QOF services supplied by GPs: The fundamental question addressed in this paper is how financial incentives affect GPs' work motivation. The coefficient of the monetary value of the QOF services supplied by GPs could help answer this. The list of symbols is explained in Table 5.

Table 5. List of symbols

Symbols	Definitions
v_t	Unadjusted national value of Pounds per QOF point
$ADPF_{igt}$	Adjusted Disease Prevalence Factor for indicator i practice g at year t
CPI_{gt}	Contractor Population Index
π_i^{max}	Full available QOF points
D_{igt}	Number of eligible patients
$\alpha_u D_{igt}$	Number of patients required to treat to achieve upper threshold
$\dfrac{v_t ADPF_{igt} CPI_{gt} \pi_i^{max}}{\alpha_{iu} D_{igt}}$	Marginal income forgone

When a practice's performance index is below the maximum payment threshold, this variable is the proxy for income that the general practice receives for treating every patient in the QOF scheme. When the performance index is higher than the maximum payment threshold, this variable is used as a measurement of foregone income from every patient treated for free. As discussed in the introduction, when foregone income is unobserved it can be treated as the value of paid income for the same services (Becker, 1965). The calculation of this variable is as follows:

$$\text{Marginal QOF income} = \frac{v_t ADPF_{igt} CPI_{gt} \pi_i^{max}}{\alpha_{iu} D_{igt}}$$

$$\text{Where } \alpha_l < \alpha_i < \alpha_u \tag{7}$$

Equation (7) refers to the marginal value of QOF income when the performance index α_i locates in $\alpha_l < \alpha_i < \alpha_u$. It refers to the marginal value of forgone income when the performance index α_i locates in $\alpha_u \leq \alpha_i \leq 1$.

See Appendix D for an example of calculation.

The above calculation is arguable in two ways. First, in observations where the performance index falls below the maximum payment threshold, equation (7) overestimates the income from every patient treated. It is because practices cannot achieve the full QOF payment under this condition. The value of QOF income from every patient treated or the forgone income from every patient treated without QOF payment should be calculated as follows:

$$\text{Marginal QOF income} = \frac{v_t ADPF_{igt} CPI_{gt} \pi_i^{max} \times \dfrac{\alpha_i - \alpha_l}{\alpha_u - \alpha_l}}{\alpha_{iu} D_{igt}}$$

$$\textit{Where } \alpha_l < \alpha_i < \alpha_u \tag{8}$$

The difference between equation (7) and (8) is that the nominator in equation (8) is only part of the total available income in equation (7). It is the case that the calculation in equation (6) could bias the marginal QOF income upward, and that this bias may have an impact on the results of subsequent estimations. However, as the objective is to find a consistent way of calculating the marginal value of QOF income, whether or not the performance index is below or above the maximum payment threshold, this calculation will be retained in the empirical study.

Second, the denominator in equation (7) should use $(\alpha_u - \alpha_l)$ instead of α_u as practices are not paid any QOF payments until they have treated the minimum number of patients as stipulated by the minimum payment threshold α_l. The reason for using α_u rather than $(\alpha_u - \alpha_l)$ is that QOF payments are contingent on practices' overall performance. This performance is measured by index α_i rather than $(\alpha_i - \alpha_l)$. Although only performance beyond the minimum payment threshold is taken into account in the calculation of QOF income, the income is based on total performance.

5.3 GPs' characteristics

Thousands of patients per GP: This variable is calculated by the list size of the practice divided by the number of GPs. It is expected to have a negative effect on GPs' intrinsic motivation. The marginal effort that is required to supply unpaid work might be higher for practices with large list sizes than small ones. Also, the income for practices with high value of the thousands of patients per GP is expected to be high. Therefore, they might be less motivated to supply unpaid QOF work to attract patients than those practices with small numbers of patients per GP.

It is also expected that the number of patients per GPs has a negative effect on GPs' extrinsic motivation. This is because the marginal effort that is required to reach the maximum payment threshold for practices with large list sizes is high.

Number of patients need to treat to get the full QOF payments: This variable is calculated by the maximum payment threshold multiplied by the eligible number of patients. It is expected to have a negative effect on the supply of extrinsically and intrinsically motivated QOF work. The higher the number of patients that need to be treat to get the full QOF payment, the greater the marginal effort required.

Ex-fund-holding status: The fund-holding programme for GPs was introduced in the UK in 1991 and withdrawn in 1999. Practices with fund-holding status were asked to use a given budget to purchase primary and secondary care for their patients, pay for the drugs prescribed and pay for services provided by the non-medical staff (Ellwood, 1997). It is expected that practices with ex-fund-holding status are more extrinsically motivated than non-ex-fund holders. This payment scheme encouraged practices to utilise their budgets as efficiently as possible, as they were permitted to put any surplus back into their practices to provide better care for patients. The surplus funds easily could be diverted to the income and for the comfort of GPs (Hausman and Le Grand, 1999). Extrinsically motivated GPs may well have opted to participate in the fund-holding programme for this reason. In other words, choosing to be a fund-holder could also be a sign of being financially motivated.

Dispensing status in 2004: This variable assigns a value of one for practices that are eligible to dispense drugs and appliances and zero otherwise. Dispensing practices usually are located in remote areas and therefore must have a small stock of drugs and medical appliances. The variable

inverse of population density, measured by hectares per person*, is included to control for rural status.

Training status in 2004: This is a dummy variable indicating whether practices are eligible to provide training to qualified doctors who want to become GPs.

Practice herfindahl index: This index indicates the intensity of competition among practices. A low herfindahl index suggests a more competitive environment.

Other GP characteristics include: *number of GPs in practice, percentage of female GPs, percentage of qualified GPs who were not born in the U and average age of GPs*.

5.4 Patients' characteristics

Standardised Illness Ratio (SIR) is a measurement of patients' health status with 100 as the baseline. It measures the proportion of people who assess themselves as having a limiting long-term illness. It is standardised by age. Higher ratios are associated with worse health and therefore higher material costs for GPs providing care for these patients.

The cost of treating patients also can be affected by their age. The proxies for this are the proportion of patients aged above 75 and proportion of patients aged below 15. Patients aged over 75 are likely to require a high level of primary care, incurring high costs for GPs. However, practices with a high proportion of patients under 15 are better able to achieve the maximum payment threshold. Young patients are not always covered by the QOF indicators, e.g. all the indicators relating to epilepsy target patients aged 16 and over. This suggests that practices with a high proportion of patients under 15 have a lower workload under QOF.

The Scottish Index of Multiple Deprivation 2004 (SIMD) is used as the measurement of patients' deprivation status. This measurement is derived from six individual deprivation domains: education, access, income, house, health and employment. A high SIMD score suggests patients have high deprivation status.

Patients' characteristics also include the proportion of patients with ethnic minority status.

6 Methods

This paper uses balanced panel data to test the effect of increased QOF payment on practices' working motivation. GPs' intrinsic motivation is modelled by means of a first difference model. A probit estimation using Mundlak approach is used to model GPs' extrinsic motivation. Finally, a first difference model is used to model GPs' overall motivation.

6.1 Intrinsic motivation

In the empirical literature, the before-after estimator is the strategy most commonly used to evaluate the effect of a programme (Heckman, et al., 1999). This is the method implemented in this paper to evaluate the effect of the 2005 increase in QOF income on GPs' intrinsic motivation. An important advantage of the first difference method is that the effects of unobserved time invariant characteristics (of indicators, diseases, GPs or patients) are excluded. The changes of intrinsic motivated output of GPs between 2004/5 and 2005/6 can be modelled using the following equation:

$$\Delta y'_{igt} = y'_{ig\,2005/6} - y'_{ig\,2004/5} = \beta_1 + (X_{ig(2005/6)} - X_{ig(2004/5)})\beta_2 + (\varepsilon_{ig\,2005/6} - \varepsilon_{ig\,2004/5}) \tag{9}$$

The dependent variable $\Delta y'_{igt}$ is the changes of GPs' intrinsic motivation between 2004/5 and 2005/6. y'_{igt} is the percentage of patients treated without QOF payment out of the number of patients who could be treated for free in year t by practice g for indicator i. X_{igt} is a $n \times 3$ matrix of the independent variables. The three independent variables that included by matrix X_{igt} are *marginal monetary value of QOF services provided by GPs, thousands of patients per GP* and *the number of patients need to treat to get the full QOF payments*. ε_{igt} is the error term, which includes the unobserved variables and random error.

6.2 Extrinsic motivation

GPs' extrinsic motivation is modelled using a probit model and Mundlak approach (Mundlak, 1978) in order to control for the correlation between unobserved individual effects.

The specification of GPs' extrinsic motivation takes the form:

$$y''_{igt} = \beta_1 + X_{igt}\beta_2 + X_g\beta_3 + X_p\beta_4 + Z_{igt}\beta_5 + \varepsilon_{igt}, \qquad \varepsilon_{igt} \sim N(0, \sigma_\varepsilon^2) \tag{10}$$

$$Z_{igt}\beta_5 = \bar{X}_{igt}\beta_6 + \beta_7 w_{ig}, \qquad w_{ig} \sim N(0, \sigma_w^2) \tag{11}$$

where y''_{igt} is the dichotomous variable for whether the performance index was above or below the maximum payment threshold in year t by practice g for indicator i. X_g is the characteristics of general practices. X_P is the characteristics of patients. Z_{igt} is the proxy for individual effect. $\bar{X}_{igt}$ is the mean of X_{igt} for two observations from the same practice and indicator over the two financial years. w_{ig} is the random individual effect with constant variance σ_w^2 and zero mean.

6.3 Overall motivation

The specification for GPs' overall motivation may be given as:

$$\Delta y_{igt}^{'''} = y_{ig\,2005/6}^{'''} - y_{ig\,2004/5}^{'''} = \beta_1 + (X_{ig(2005/6)} - X_{ig(2004/5)})\beta_2 + (\varepsilon_{ig\,2005/6} - \varepsilon_{ig\,2004/5}) \tag{12}$$

The dependent variable $\Delta y_{igt}^{'''}$ is the changes of GPs' overall motivation between 2004/5 and 2005/6. $y_{igt}^{'''}$ is the percentage of patients actually treated out of the number of eligible patients in year t by practice g for indicator i.

7   Results

The results of using the first difference method and balanced panel data to model GPs' intrinsic motivation are reported in Table 6. The results show the effect of the 2005 increase in QOF payments. It suggests that each one pound increase in income under the new policy reduced GPs' intrinsic motivation by 0.116%. They also show that GPs' intrinsic motivation was adversely affected by the increase in the number of patients they were required to treat to reach the upper threshold. The marginal effect of this change was to significantly reduce (by 0.3%) the number of patients treated for free. The marginal effect of increasing the number of patients per GP was to reduce the number of patients treated for free by 10.27%. This result is only significant at the 10% level. The distribution of the changes in intrinsic motivation over the two financial years is reported in Fig. 2.

Table 6. Determinants of intrinsic motivation by using the first difference model for financial years 2004/5 and 2005/6

Dependent variable	GPs' intrinsic motivation	t value
Marginal forgone income	-0.00116**	-4.70
Threshold for full QOF achievement	-0.00300**	-19.5
Thousands of patients per GP	-0.10269*	-1.86
Constant	0.14116**	57.21
Observations	31787	

* significant at 10% ** significant at 5%
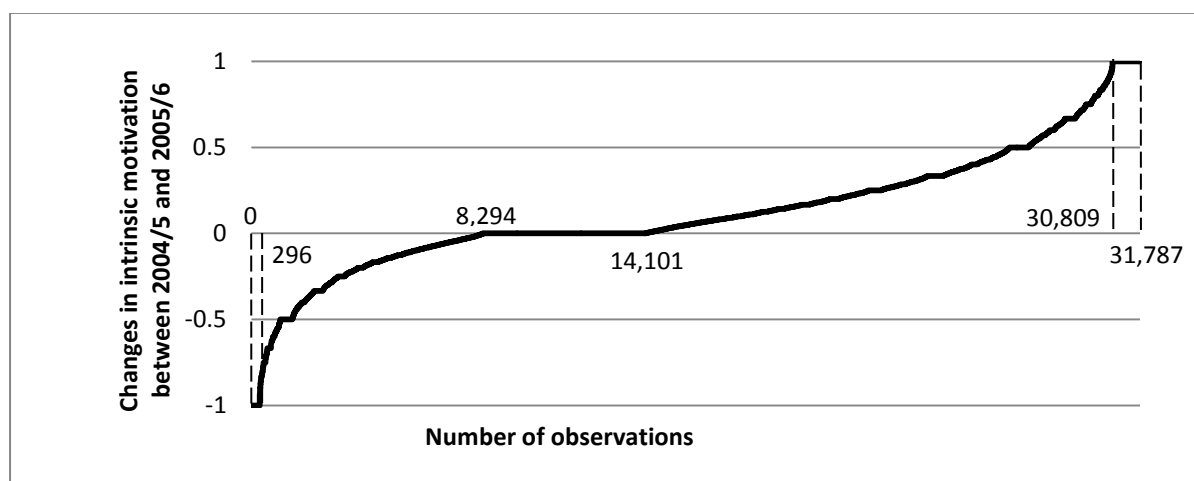
Figure 2. Changes in intrinsic motivation between 2004/5 and 2005/6

The estimated results from the separate regressions for the eight disease domains are reported in Table 7. Marginal forgone income was found to have negative and significant impact on the number of extra patients treated in three of the eight diseases: chronic heart disease, diabetes mellitus and stroke. In the cases of epilepsy and COPD, however, marginal foregone income was found to have a significant and positive effect on GPs' intrinsic motivation.

Table 7. Determinants of intrinsic motivation by using the first difference model for financial years 2004/5 and 2005/6

	Marginal forgone income	t value	Threshold for full QOF payment	t value	Thousands of patients per GP	t value	Constant	t value	Number of observations
Overall	-0.00116**	4.70	-0.0030**	-19.50	-0.1027*	-1.86	0.1412**	57.21	31787
Asthma	0.00013	0.21	-0.0025**	-6.40	0.0221	0.14	0.1053**	13.43	3310
CHD	-0.00670**	-4.96	-0.0045**	-15.56	-0.1960*	-1.72	0.1761**	28.70	5979
COPD	0.00717**	2.90	-0.0069	-7.62	0.2898	0.90	0.0235	1.21	1562
DM	-0.00156*	-1.68	-0.0021**	-8.37	-0.0474	-0.58	0.1360**	30.40	11942
Epilepsy	0.00419**	2.78	-0.0066**	-2.63	-0.3506	-1.28	0.0858**	5.10	2282
MH	-0.00060	-0.62	-0.0082**	-2.83	0.9960	1.57	0.0961**	2.36	846
Stroke	-0.00313*	-1.88	-0.0035**	-5.90	-0.2801**	-2.11	0.1756**	22.20	5006
Thyroid	0.00768	0.46	0.0018**	2.15	-0.3179	-0.98	0.0404	1.05	860

* significant at 10% ** significant at 5%

The results presented in Table 8 suggest that the increased marginal income in 2005/6 had a significant positive effect on GPs' extrinsic motivation; the probability of achieving full payment increased by 0.145%.

Table 8. Determinants of extrinsic motivation by using the probit model with Mundlak approach for financial years 2004/5 and 2005/6

Dependent variable	Probability of achieving the maximum payment threshold	Z value
Marginal income	0.00145**	7.54
Threshold for full QOF achievement	-0.00021**	-7.43
Thousands of patients per GP	0.05394	0.61
Mean of marginal income	-0.00026	-1.18
Mean of threshold for full QOF achievement	0.00039**	7.78
Mean of marginal thousands of patients per GP	-0.07048	-0.79
Number of GPs	-0.00869**	-5.29
Age	0.00007	0.11
Percentage of female GPs	0.01771	1.47
Proportion of GPs not born in UK	-0.05839**	-4.10
Ex-fund-holding practice	0.01402**	2.58
Practice Herfinahsl index	-0.04004**	-2.85
Training 2004	0.02278**	3.09
Dispensing 2004	-0.03056**	-2.54
Inverse of practice population density	-0.00041	-2.85
SIR	-0.00076**	-5.35
SIMD	0.00048*	1.95
Proportion of patients with ethnicity	0.47210**	3.57
Prop 75	-0.05924	-0.44
Prop 15	0.37450**	3.70
Observations	63574	

* significant at 10% ** significant at 5%

The results from modelling the overall motivation are presented in Table 9. It suggests that the rise in marginal income increased GPs' overall working motivation by 0.04%. GPs' overall motivation was negatively affected by an increase in the threshold for full QOF achievement. The marginal effect of

this change was to significantly reduce the number of patients treated by 0.098%. The marginal effect of the number of patients per GP was to reduce the overall patients treated by 3.07%.

Table 9. Determinants of overall motivation by using the first difference model for financial years 2004/5 and 2005/6

Dependent variable	The first difference of GPs' overall motivation	t value
Marginal income	0.00040**	5.93
Threshold for full QOF achievement	-0.00098**	-22.57
Thousands of patients per GP	-0.03070**	-2.03
Constant	0.04373**	64.57
Observations	31787	

*significant at 10% ** significant at 5%

8 Discussion

When the policy is evaluated using the balanced panel data, the results suggest that the April 2005 increase the rate of QOF payment had a negative effect on GPs' supply of unpaid work. This finding is supported by the separated regressions at disease domain level. At the same time, the model of overall motivation suggests that the policy raised GPs' performance under QOF. However, the positive effect is small. These findings provide evidence for the conclusion that the QOF policy is effective, but that it has a negative effect on GPs' intrinsic motivation. This conclusion has important implications for policy makers as it shows that using strong financial incentives to further motivate already well motivated professionals may not be an efficient way to influence behaviour and performance. The main results are summarised in Table 10.

Our results are in line with the predictions of the theoretical economic model (*Hypotheses 2 and 3*). Although higher income allows GPs to spend more time and money on unpaid work, negative effects also are associated with the increase in forgone income and changes in the utility GPs derive from doing unpaid work. Unfortunately, the empirical result presented here cannot explain how the three effects interact because of lack of information. The result shows that increased QOF income in 2005/6 had a positive effect on GPs' extrinsic motivation, providing supporting evidence for *Hypothesis 4*. The overall effect of increasing QOF payment was to increase GPs' overall motivation. This suggests that although the policy adjustment in 2005 was a qualified success, it undermined GPs' intrinsic motivation.

Table 10. The effect of financial incentives on GPs' extrinsic, intrinsic and overall motivation in 2004/5 and 2005/6

Motivation	Panel data analysis
Intrinsic motivation	-
Extrinsic motivation	+
Overall motivation	+

(+) indicates that the financial incentive has positive effect
(-) indicates that the financial incentive has negative effect

Limitations

The findings in this study are subject to the following limitations.

First, the study contains no information on GPs' attitude towards risk. GPs may have supplied unpaid work as a form of insurance that they would achieve the upper threshold. Appendix E reports where GPs are assumed to be risk averse. Although the results cannot entirely eliminate the effect of GPs' risk aversion on their provision of unpaid work, they do demonstrate that being risk averse is not the only reason that GPs do unpaid work.

Second, it is arguable whether the way foregone income is calculated here (as the best alternative use of the time GPs spend supplying free services under QOF) reflects real practice. Again, this information is not available. Appendix F reports what happens when an alternative method of foregone income calculation is applied. The results show that the negative effect of foregone income on GPs' intrinsic motivation does not change.

Third, QOF payment is remunerated at practice level, but how money is distributed within practices is unknown. Campbell, McDonald and Lester (2008) suggest that most doctors and nurses acknowledge that nurses have become the primary provider of health care for patients with chronic conditions. The new contract provides a significant incentive for practices to make changes in employment patterns. These changes or planned changes include employing more nurses, instituting nurse-led chronic disease management, and increasing the triage and management of minor illnesses by nurses (Ronald, et al., 2006). However, as partners, GPs award most of the QOF income to themselves; in many practices, the increased input of nurses has gone unrewarded. In other words, there is a mismatch between who receives the QOF payment and who does the work. The QOF is part of the GMS contract, and was introduced to elicit more effort from GPs, but the actual main performers are nurses. It would be interesting to know how QOF income and workloads are allocated within practices and whether the two things are correlated. Conversely, it might be argued that GPs are getting paid more under QOF because their practices are better managed, and that the problem with mismatching is merely a red herring.

Fourth, the number of practitioners is determined by a simple headcount rather than by specifying the number of full time staff or its equivalent. One of the determinants for GPs' outputs is working time. The number of working hours should have been taken into account when modelling the possibility that maximum payment threshold would be achieved and GPs' provision of unpaid work under QOF. However, the study lacks any information about the number of hours actually worked by GPs.

Fifth, the analysis does not take into account additional practice income from other clinical and non-clinical work. This might have a positive effect on GPs' provision of unpaid work as more money is available to spend on unpaid work. On the other hand, the effect may be negative as practitioners' time becomes more valuable. A trade off must be made between time allocated to unpaid QOF work, other clinical paid work and non-clinical paid work. Consequently, the opportunity costs involved in providing unpaid QOF work are higher for GPs with high alternative incomes. Sutton and Mclean (2006) provide evidence for this, suggesting that practices earning a high income from other sources exhibit lower quality scores.

Sixth, productivity is not taken into account in this analysis. The relative levels of productivity achieved in paid and unpaid work may affect the provision of unpaid work. Productivity data on the provision of care directly and not directly incentivised by money is needed to address this problem.

Seventh, we assumed that practices are only extrinsically motivated when they are below the full payment threshold. It is arguable that this assumption underestimates the intrinsic motivation of GPs, as they could also be intrinsically motivated when their performance indices locate below the maximum payment threshold. GPs run practices as businesses; therefore, extrinsic motivation is the major driving force within the GMS contract for GPs to treat patients. The QOF is an optional scheme; payment is contingent on GPs' performance as long as the performance index locates below the maximum payment threshold. It is therefore reasonable to assume that practices that choose to join the QOF scheme and have a performance index below the maximum payment threshold are on the whole extrinsically motivated.

Eighth, following their examination of data gathered from Scottish general practices in 2004/5, Gravelle, et al. (2010) suggested that the QOF payment system creates incentives for GPs to manipulate their reporting of prevalence rates and exception rates. The evidence from their paper may call into question the way extrinsic and intrinsic motivations are measured. Similar evidence found in Doran, et al. (2006) shows that a small number of general practices in England achieved high scores under QOF in 2004/5 by excluding large numbers of patients in exception reports. This

suggests that we cannot exclude the possibility that GPs exploit the QOF payment system to increase income. However, this limitation will not invalidate our results. More likely, it will enhance the evidence of a crowding out effect on intrinsic motivation. The practices could have had lower achievement rates by accepting few patients or diagnosing more new cases, thus lowering the measure of intrinsic motivation. The effect of intrinsic motivation will show up even stronger if the gaming effects can be isolated.

Ninth, the unobserved time variant of patients' characteristics and GPs' characteristics may lead to a biased estimation. For instance, patients' characteristics may have an effect on GPs' supply of QOF services through costs of treatment. However, all the patients' characteristics variables in our dataset are time invariant and so all are dropped from the panel data analysis.

9 Conclusion

This paper explores how GPs respond to financial incentives under the QOF. The results suggest that the policy change of 1 April 2005 increased income from QOF, enhancing GPs' extrinsic motivation but crowding out their intrinsic motivation. It concludes that increasing income may not be a completely efficient way of motivating GPs as intrinsic motivation is an important feature in their utility functions. This paper provides both theoretical and empirical evidence of the presence and importance of GPs' intrinsic motivation and how it influences the ways that GPs respond to financial incentives.

This result has important implications for policy makers as it shows that using strong financial incentives to further motivate already well-motivated professionals may have unintended effects on their behaviour and performance.

Appendix A. Descriptions of the 39 indicators in 2004/5 and 2005/6

Indicator names	Definitions
CHD3	The percentage of patients with coronary heart disease whose notes record smoking status in the past 15 months, except those who have never smoked where smoking status need be recorded only once.
CHD4	The percentage of patients with coronary heart disease who smoke, whose notes contain a record that smoking cessation advice or referral to a specialist service, where available, has been offered within the last 15 months.
CHD5	The percentage of patients with coronary heart disease whose notes have a record of blood pressure in the previous 15 months.
CHD6	The percentage of patients with coronary heart disease in whom the last blood pressure reading (measured in the last 15 months) is 150/90 or less.
CHD7	The percentage of patients with coronary heart disease whose notes have a record of total cholesterol in the previous 15 months.
CHD8	The percentage of patients with coronary heart disease whose last measured total cholesterol (measured in last 15 months) is 5 mmol/l or less.
CHD9	The percentage of patients with coronary heart disease with a record in the last 15 months that aspirin, an alternative anti-platelet therapy, or an anti-coagulant is being taken (unless a contraindication or side-effects are recorded).
STROKE3	The percentage of patients with TIA or stroke who have a record of smoking status in the last 15 months, except those who have never smoked where smoking status should be recorded at least once since diagnosis.
STROKE4	The percentage of patients with a history of TIA or stroke who smoke and whose notes contain a record that smoking cessation advice or referral to a specialist service, if available, has been offered in the last 15 months.
STROKE5	The percentage of patients with TIA or stroke who have a record of blood pressure in the notes in the preceding 15 months.
STROKE6	The percentage of patients with a history of TIA or stroke in whom the last blood pressure reading (measured in last 15 months) is 150/90 or less.
STROKE7	The percentage of patients with TIA or stroke who have a record of total cholesterol in the last 15 months.
STROKE8	The percentage of patients with TIA or stroke whose last measured total cholesterol (measured in last 15 months) is 5 mmol/l or less.
DM2	The percentage of patients with diabetes whose notes record BMI in the previous 15 months.
DM3	The percentage of patients with diabetes in whom there is a record of smoking status in the previous 15 months, except those who have never smoked where smoking status should be recorded once.
DM4	The percentage of patients with diabetes who smoke and whose notes contain a record that smoking cessation advice or referral to a specialist service, where available, has been offered in the last 15 months.
DM5	The percentage of diabetic patients who have a record of HbA1c or equivalent in the previous 15 months.
DM6	The percentage of patients with diabetes in whom the last HbA1C is 7.4 or less (or equivalent test/reference range depending on local laboratory) in last 15 months.
DM7	The percentage of patients with diabetes in whom the last HbA1C is 10 or less (or equivalent test/reference range depending on local laboratory) in last 15 months.
DM8	The percentage of patients with diabetes who have a record of retinal screening in the previous 15 months.
DM9	The percentage of patients with diabetes with a record of the presence or absence of peripheral pulses in the previous 15 months.
DM10	The percentage of patients with diabetes with a record of neuropathy testing in the previous 15 months.

DM11	The percentage of patients with diabetes who have a record of the blood pressure in the past 15 months.
DM13	The percentage of patients with diabetes who have a record of micro-albuminuria testing in the previous 15 months (exception reporting for patients with proteinuria).
DM14	The percentage of patients with diabetes who have a record of serum creatinine testing in the previous 15 months.
DM16	The percentage of patients with diabetes who have a record of total cholesterol in the previous 15 months.
DM17	The percentage of patients with diabetes whose last measured total cholesterol within the previous 15 months is 5mmol/l or less.
COPD4	The percentage of patients with COPD in whom there is a record of smoking status in the previous 15 months.
COPD5	The percentage of patients with COPD who smoke, whose notes contain a record that smoking cessation advice or referral to a specialist service, where available, has been offered in the past 15 months.
EPILEPSY2	The percentage of patients aged 16 and over on drug treatment for epilepsy who have a record of seizure frequency in the previous 15 months.
EPILEPSY3	The percentage of patients aged 16 and over on drug treatment for epilepsy who have a record of medication review in the previous 15 months
EPILEPSY4	The percentage of patients aged 16 and over on drug treatment for epilepsy who have been seizure free for the last 12 months recorded in the last 15 months.
THYROI2	The percentage of patients with hypothyroidism with thyroid function tests recorded in the previous 15 months.
MH2	The percentage of patients with severe long-term mental health problems with a review recorded in the preceding 15 months. This review includes a check on the accuracy of prescribed medication, a review of physical health and a review of co-ordination arrangements with secondary care.
MH4	The percentage of patients on lithium therapy with a record of serum creatinine and TSH in the preceding 15 months.
ASTHMA3	The percentage of patients with asthma between the ages of 14 and 19 in whom there is a record of smoking status in the previous 15 months.
ASTHMA4	The percentage of patients aged 20 and over with asthma whose notes record smoking status in the past 15 months, except those who have never smoked where smoking status should be recorded at least once.
ASTHMA5	The percentage of patients with asthma who smoke, and whose notes contain a record that smoking cessation advice or referral to a specialist service, where available, has been offered within the last 15 months.
ASTHMA6	The percentage of patients with asthma who have had an asthma review in the last 15 months.

Appendix B. Clinical indicators, maximum and minimum payment thresholds, maximum available points and GP performance in 2004/5

Indicator name	Minimum payment threshold	Maximum payment threshold	Maximum available points	Mean of achieved points (SD)	Mean percentage of patients treated for free (SD)	Mean percentage of patients treated (SD)
Secondary prevention in Coronary Heart Disease (CHD)						
CHD3	25%	90%	7	6.91 (0.47)	0.63 (0.33)	0.96(0.06)
CHD4	25%	70%	4	3.87 (0.37)	0.76 (0.25)	0.93 (0.08)
CHD5	25%	90%	7	6.92 (0.32)	0.62 (0.32)	0.96(0.05)
CHD6	25%	70%	19	18.86 (0.81)	0.53 (0.25)	0.86 (0.08)
CHD7	25%	90%	7	6.58 (0.86)	0.31 (0.33)	0.89 (0.10)
CHD8	25%	60%	16	15.20 (2.32)	0.31 (0.25)	0.71 (0.13)
CHD9	25%	90%	7	6.74 (0.44)	0.23 (0.29)	0.90 (0.06)
Stroke and Transient Ischaemic Attacks (STROKE)						
STROKE3	25%	90%	3	2.94 (0.24)	0.55 (0.37)	0.94 (0.07)
STROKE4	25%	70%	2	1.98 (0.11)	0.70 (0.33)	0.91 (0.10)
STROKE5	25%	90%	2	1.97 (0.11)	0.57 (0.36)	0.95 (0.06)
STROKE6	25%	70%	5	4.91 (0.34)	0.47(0.28)	0.84 (0.10)
STROKE7	25%	90%	2	1.78 (0.35)	0.23 (0.32)	0.85 (0.13)
STROKE8	25%	60%	5	4.48 (0.99)	0.23 (0.24)	0.66 (0.15)
Diabetes Mellitus (Diabetes)						
DM2	25%	90%	3	2.95 (0.18)	0.51 (0.35)	0.94 (0.06)
DM3	25%	90%	3	2.99 (0.11)	0.77 (0.26)	0.98 (0.04)
DM4	25%	90%	5	4.84 (0.48)	0.54 (0.42)	0.94 (0.09)
DM5	25%	90%	3	2.99 (0.09)	0.68 (0.31)	0.97 (0.04)
DM6	25%	50%	16	15.19 (2.05)	0.17 (0.19)	0.58 (0.11)
DM7	25%	85%	11	10.85 (0.60)	0.40 (0.28)	0.90 (0.06)
DM8	25%	90%	5	4.56 (0.70)	0.27 (0.35)	0.87 (0.11)
DM9	25%	90%	3	2.70 (0.46)	0.22 (0.30)	0.86 (0.12)
DM10	25%	90%	3	2.64 (0.52)	0.19 (0.29)	0.84 (0.13)
DM11	25%	90%	3	3.00 (0.04)	0.80 (0.24)	0.98 (0.03)
DM13	25%	90%	3	2.46 (0.66)	0.16 (0.29)	0.80 (0.16)
DM14	25%	90%	3	2.97 (0.13)	0.57 (0.35)	0.95 (0.05)
DM16	25%	90%	3	2.97 (0.13)	0.54 (0.34)	0.95 (0.05)
DM17	25%	60%	6	5.87 (0.54)	0.39(0.26)	0.75(0.12)
Chronic Obstructive Pulmonary Disease (COPD)						
COPD4	25%	90%	6	5.88 (0.47)	0.65 (0.37)	0.95 (0.07)
COPD5	25%	90%	6	5.83 (0.52)	0.57 (0.42)	0.94 (0.08)
Epilepsy						
EPILEPSY2	25%	90%	4	3.87 (0.42)	0.54 (0.42)	0.94 (0.09)
EPILEPSY3	25%	90%	4	3.80 (0.56)	0.50 (0.42)	0.92 (0.12)
EPILEPSY4	25%	70%	6	4.65 (1.57)	0.11 (0.23)	0.64(0.16)
Hypothyroid						
THYROI2	25%	90%	6	5.94 (0.33)	0.65(0.33)	0.96 (0.06)
Mental Health (MH)						
MH2	25%	90%	23	22.30 (2.62)	0.60 (0.43)	0.94 (0.10)
MH4	25%	90%	3	2.91 (0.26)	0.48 (0.49)	0.94(0.08)
Asthma						
ASTHMA3	25%	70%	6	5.67 (0.89)	0.42(0.35)	0.81 (0.15)
ASTHMA4	25%	70%	6	5.97(0.31)	0.75 (0.22)	0.92 (0.08)
ASTHMA5	25%	70%	6	5.92 (0.44)	0.62 (0.31)	0.88 (0.11)
ASTHMA6	25%	70%	20	17.50 (4.69)	0.29 (0.31)	0.73 (0.17)

Appendix C. Clinical indicators, maximum and minimum payment thresholds, maximum available points and GP performance in 2005/6

Indicator name	Minimum payment threshold	Maximum payment threshold	Maximum available points	Mean of achieved points (SD)	Mean percentage of patients treated for free (SD)	Mean percentage of patients treated (SD)
Secondary Prevention in Coronary Heart Disease (CHD)						
HD3	25%	90%	7	6.99 (0.14)	0.77 (0.24)	0.98 (0.03)
CHD4	25%	70%	4	3.97 (0.18)	0.62 (0.37)	0.96 (0.05)
CHD5	25%	90%	7	6.99 (0.11)	0.77 (0.23)	0.98 (0.03)
CHD6	25%	70%	19	18.98 (0.42)	0.66 (0.19)	0.90 (0.06)
CHD7	25%	90%	7	6.92 (0.33)	0.52 (0.32)	0.95 (0.05)
CHD8	25%	60%	16	15.90 (0.74)	0.52 (0.22)	0.81 (0.10)
CHD9	25%	90%	7	6.98 (0.21)	0.55 (0.30)	0.95 (0.04)
Stroke and Transient Ischaemic Attacks (STROKE)						
STROKE3	25%	90%	3	2.99 (.09)	0.68 (0.31)	0.97 (0.04)
STROKE4	25%	70%	2	2.00 (0.05)	0.78 (0.26)	0.94 (0.07)
STROKE5	25%	90%	2	1.99 (0.05)	0.73(0.27)	0.97 (0.03)
STROKE6	25%	70%	5	4.99 (0.18)	0.61 (0.22)	0.89 (0.07)
STROKE7	25%	90%	2	1.96 (0.14)	0.42 (0.35)	0.93 (0.07)
STROKE8	25%	60%	5	4.95 (0.30)	0.44 (0.24)	0.77 (0.10)
Diabetes Mellitus (Diabetes)						
DM2	25%	90%	3	2.99 (0.07)	0.63 (0.30)	0.96 (0.04)
DM3	25%	90%	3	3.00 (0.03)	0.86 (0.19)	0.99 (0.02)
DM4	25%	90%	5	4.95 (0.22)	0.64 (0.38)	0.96 (0.05)
DM5	25%	90%	3	3.00 (0.02)	0.79 (0.22)	0.98 (0.02)
DM6	25%	50%	16	15.54 (1.52)	0.22 (0.21)	0.21 (0.11)
DM7	25%	85%	11	10.97 (0.20)	0.49 (0.25)	0.92 (0.04)
DM8	25%	90%	5	4.86 (0.37)	0.48 (0.40)	0.93 (0.08)
DM9	25%	90%	3	2.89 (0.26)	0.34 (0.33)	0.91 (0.08)
DM10	25%	90%	3	2.87 (0.29)	0.31 (0.33)	0.90 (0.08)
DM11	25%	90%	3	3.00 (0.03)	0.87 (0.16)	0.99 (0.02)
DM13	25%	90%	3	2.80 (0.37)	0.25 (0.31)	0.88 (0.10)
DM14	25%	90%	3	2.99 (0.05)	0.71 (0.26)	0.97 (0.03)
DM16	25%	90%	3	3.00 (0.03)	0.69 (0.26)	0.97 (0.03)
DM17	25%	60%	6	5.99 (0.21)	0.54 (0.20)	0.82 (0.08)
Chronic Obstructive Pulmonary Disease (COPD)						
COPD4	25%	90%	6	5.97 (0.23)	0.73 (0.29)	0.97 (0.04)
COPD5	25%	90%	6	5.95 (0.26)	0.63 (0.37)	0.96 (.052)
Epilepsy						
EPILEPSY2	25%	90%	4	3.98 (0.15)	0.68 (0.36)	0.97 (0.05)
EPILEPSY3	25%	90%	4	3.96 (0.25)	0.64 (0.37)	0.96 (0.06)
EPILEPSY4	25%	70%	6	5.58 (0.91)	0.23 (0.28)	0.75 (0.13)
Hypothyroid						
THYROI2	25%	90%	6	5.99 (0.07)	0.72 (0.28)	0.97 (0.03)
Mental Health (MH)						
MH2	25%	90%	23	22.92 (0.46)	0.62 (0.40)	0.96 (0.042)
MH4	25%	90%	3	2.97 (0.12)	0.68 (0.46)	0.97 (0.06)
Asthma						
ASTHMA3	25%	70%	6	5.94 (0.37)	0.54 (0.32)	0.87(0.10)
ASTHMA4	25%	70%	6	5.99 (0.14)	0.84 (0.14)	0.95 (0.04)
ASTHMA5	25%	70%	6	5.98 (0.16)	0.68 (0.27)	0.91 (0.08)
ASTHMA6	25%	70%	20	19.69 (1.49)	0.45 (0.30)	0.830.10)

Appendix D. Example of the calculation of marginal QOF income

Take the performance of practice 10002 in financial year 2004/5 under indicator CHD5 as an example. The indicator's minimum payment threshold is 0.25 and the maximum payment threshold is 0.9. The practice had 7,103 patients on its list in 2004/5. The average list size in Scotland is 5,095. 235 patients are treated for this indicator out of 245 eligible patients. The maximum payment threshold is 90%. GPs' performance in this indicator surpassed the full payment threshold $(235 > 245 \times 90\%)$. Thus, the practice got the full available seven points. In 2004/5, each point was worth £75. The value of ADPF was 0.9.

Based on the information above, the QOF income from every patient treated can be calculated as follows:

$$\text{Marginal QOF income} = \frac{\text{national value of pounds per point} \times \text{ADPF} \times \text{CPI} \times \text{QOF points}}{\text{number of eligible patients} \times \alpha_u}$$

$$= \frac{£75 \times 0.9 \times \dfrac{7103}{5095} \times 7}{245 \times 0.9} = \frac{£658.72}{245 \times 0.9} = \frac{£658.72}{220.5} \approx \frac{£658.72}{221} = £2.98 \qquad \text{(D.1)}$$

The highest achievable QOF income from this indicator is £658.72. The denominator 221 is the number of patients this practice must treat to get full points for this indicator. The accurate value of the denominator is 220.5, but the practice will not get the maximum available points if they treat fewer than 221 patients. As a result, marginal income from treating the first 221 patients is £2.98. The marginal income forgone from treating the 222nd patient is £ 2.98.

Appendix E: The effect of marginal income on GPs' intrinsic motivation when they are assumed to be risk averse

Four models were estimated to establish whether the number of patients treated for free is the result of GPs' attitudes toward risk as well as their intrinsic motivation. If risk aversion is an important factor in GPs' working motivation, they may treat a few more patients for free than is actually needed in order to make sure that the maximum payment threshold is achieved at the end of the financial year.

The mean value of the number of patients treated for free is 10.13 (SD=15.76) for the first year and 13.51 (SD=18.47) for the second year in the pooled dataset. The number of patients treated for free was arbitrarily reduced by two in the first regression. The mean value of the number of patients treated for free after this reduction is 9.65 (SD=15.97) in the first year and 13.11 (SD=18.80) in the second year. The number of patients treated for free was arbitrarily reduced by five in the second regression. The mean value of the number of patients treated for free is then 9.27 (SD=16.36) in the first year and 12.96 (SD=19.35) in the second year. In the third regression, the maximum payment threshold was arbitrarily increased by 2%. After this adjustment, the number of patients treated for free on average was 8.15 (SD= 14.17) in the first year and 11.06 (SD= 16.86) in the second year. In the fourth regression, the maximum payment threshold was increased by 5%. The mean of number of patients treated for free is 5.64 (SD=12.03) in the first year and 7.84 (SD=14.69) in the second year. The specifications of the four regressions are identical to the ones discussed in Section 6.1.

Coefficients of the marginal forgone income from the four regressions show that they collectively have a significant negative effect on GPs' intrinsic motivation. This suggests that even if GPs are risk averse, the negative correlation between GPs' QOF income and their intrinsic motivation holds.

The disadvantage of this method is that the same level of risk aversion is imposed on all practices. In the absence of information about individual GPs' attitude to risk, it is not easy to estimate its actual importance. Given this limitation, the above findings can do no more than suggest that risk aversion is not a main driving force behind GPs' provision of unpaid work. However, it cannot be excluded from the list of possible explanations why GPs provide free services. On the other hand, it could be argued that there is no reason to believe that GPs' attitudes toward risk are correlated with their intrinsic motivation. In this case, even given the fact that GPs may treat a few extra patients in order to guarantee their full QOF income, the results still supports the conclusion that there is a negative relationship between income and GPs' intrinsic motivation.

Table E.1
The effect of marginal income on GPs' intrinsic motivation when GPs are assumed to be risk averse

Dependent variable	Difference in marginal income between 2004/5 and 2005/6	t value	Number of observations
Reducing number of patients treated for free by two[1]	-0.0072**	-31.17	28622
Reducing number of patients treated for free by five[2]	-0.0084**	-27.11	23937
Increasing the maximum payment threshold by 2%[3]	-0.0042**	-24.15	31333
Increasing the maximum payment threshold by 5%[4]	-0.0038**	-18.34	30364

* significant at 10% ** significant at 5%

Appendix F: Alternative method in the calculation of marginal QOF income

The following supplementary estimations test whether the findings reported in the main context would change if forgone income is calculated using other methods.

The first regression models a subsample of practices that did not achieve full points. The number of observations is 3,091. The forgone income is calculated as the difference between the full available QOF income and what was achieved in practice. The total forgone income is calculated from all thirty nine indicators in the study year. Forgone income is not calculated at indicator level as it is not easy to convert the forgone income from one group of indicators to another. Also it is wrong to take the average value of marginal forgone income to reflect the marginal forgone income for all indicators. Because the efforts and time input are different for different indicators. The results of the panel data analysis are reported in column two, Table F.1.

The second regression models the subsample of practices which achieved full points in all clinical indicators. There are 9,870 observations for 2004/5 and 2005/6. The results of the panel data analysis are reported in column four, Table F.1.

The results reported here suggest that the results reported in the main text are robust.

Table F.1
The effect of forgone income on GPs' intrinsic motivation using an alternative method to calculate forgone income

Dependent variable	Changes in Intrinsic motivation	t value	Changes in Intrinsic motivation	t value
Forgone income	-0.0001**	-24.01	-0.0006	-1.07
Threshold for full QOF achievement	-0.0032 **	-12.04	-0.0022**	-4.09
Thousands of patients per GP	-0.0128	-0.13	0.2580**	2.10
Constant	0.0906**	22.08	0.0211**	3.31
Observations	9870		3091	

* significant at 10% ** significant at 5%

References

Becker, G.S. (1976) *The economic approach of human beha*vior. Chicago and London: University of Chicago Press.

Bell, D.N.F. and Hart, R.A. (1999) Unpaid work*. Economica*. 66(262), 271-290.

Campbell, S.M., McDonald, R. and Lester, H. (2008) The experience of pay for performance in English family practice: A qualitative study. *Annals of Family Medicine*. 6(3), 228-234.

Chalkley, M. and Tilley, C. (2006) Treatment intensity and provider remuneration: Dentists in the British National Health Service. *Health Economics.* 15(9), 933-946.

Cromwell, J. and Mitchell, J.B. (1985) Physician-induced demand for surgery. *Journal of Health Economics*. 5(4), 293-313.

Davidson, S.M., et al. (1992) Prepayment with office-based physicians in publicly funded programs: Results from the children's Medicaid program*. Pediatrics.* 89(4), 761-767.

Deci, E.L. (1975) *Intrinsic Motivation.* New York: Plenum Press.

Deci, E.L. and Ryan, R.M. (2000) The 'what' and 'why' of goal pursuits: Human needs and the self-determination of behaviour. *Psychological Inquiry.* 11(4), 227-268.

Dixit, A. (2002) Incentives and organizations in the public sector: An interpretative review. *Journal of Human Resources.* 37(4), 696-727.

Dixit, A. (1997) Power of incentives in private versus public organizations. *American Economic Review.* 87(2), 378-382.

Dolea, C. and Adams, O. (2005) Motivation of health care workers – review of theories and empirical evidence. *Cahiers de Sociologie et de Démographie Médicales*. 45(1), 135-161.

Donaldson, C. and Gerard, K. (1989) Paying general practitioners: Shedding light on the review of health services. *Journal of the Royal College of General Practitioners*.39 (320), 114-117.

Doran, T. (2006) Pay for-performance programs in family practices in the United Kingdom. *New England Journal of Medicine*. 355(4), 375-384.

Ellwood, S.M. (1997) The response of fundholding family doctors to price signals. *Financial Accountability & Management.* 13(4), 345-365.

Evans, R.G. (1974) Modelling the economic objectives of the physician. *In* Fraser, R.D. ed. H*ealth economics symposium*. Kingston, Canada: Industrial Relations Centre, Queen's University. 33-45.

Francois, P. (2000) 'Public service motivation' as an argument for government provision. *Journal of Public Economics*. 78(3), 275-299.

Frey, B.S. (1997*) Not just for the money: An economic theory of personal motivation*. Cheltenham: Edward Elgar Publishing.

Frey, B.S. (1992) Tertium datur: Pricing, regulation and intrinsic motivation. *Kyklos.* 45(2), 161-184.

Fuchs, V.R. (1978) The supply of surgeons and the demand for surgical operations. *Journal of Human Resources* .13(SUPPL), 35-56.

Gosden, T., et al. (2001) Impact of payment method on behaviour of primary care physicians: a systematic review. *Journal of Health Services Research and Policy* .6 (1), 44-55.

Gothill, M. (1998) What do doctors want? Altruism and satisfaction in general practice. *Family Practice* 15 (SUPPL 1), S36-39.

Gravelle, H., Sutton, M. and Ma, A. (2010) Doctor behaviour under a pay for performance contract: Treating, cheating, and case finding. *Economic Journal.* 120(542), 129-156.

Gregg, P., Grout, A.P., Ratcliffe, A., Smith, S. and Windmeijer, F. (2008) *How important is pro-social behaviour in the delivery of public services?* Centre for Market and Public Organisation. Working paper. Bristol: University of Bristol.

Hausman, D. and Le Grand, J. (1999) Incentives and health policy: Primary and secondary care in the British National Health Service. *Social Science & Medicine*. 49(10), 1299-1307.

Heckman, J.J., Lalonde, R.J. and Smith, J.A. (1999) The economics and econometrics of active labor market programs. *In* Ashenfelter, O. and Layard, R. eds. *Handbook of labor economics*. Amsterdam: North Holland. 1865-2097.

Hickson, G.B., Altemeier, W.A. and Perrin, J.M. (1987) Physician reimbursement by salary or fee-for-service: Effect on physician practice behavior in a randomized prospective study. *Pediatrics*. 80(3), 344-350.

Information Services Division Scotland (ISD). (2010a) *General practice -- Quality & Outcomes Framework.* Available online at http://www.isdscotlandarchive.scot.nhs.uk/isd/3305.html. [Accessed 1 November 2011].

Information Services Division Scotland (ISD). (2010b) *Overall staff and turnover.* Available online at http://www.isdscotland.org/Health-Topics/Workforce/Overall-Staff-and-Turnover/. [Accessed 1 November 2011].

Krasnik, A., et al.(1990) Changing remuneration systems: Effects on activity in general practice. *British Medical Journal.* 300(6741), 1698-1701.

Le Grand, J. (2003) *Motivation, agency and public policy: Of knights and knaves, pawns and queens*. Oxford: Oxford University Press.

Marshall, M. and Harrison, S. (2005) It's about more than money: Financial incentives and internal motivation. *Quality and Safety in Health Care*. 14(1), 4-5.

Morris, S., et al.(2008) Analysis of consultants' NHS and private incomes in England in 2003/4. *Journal of the Royal Society of Medicine.* 101(7), 372-380.

Mundlak, Y. (1978) On the pooling of time series and cross section data. *Econometrica.*46(1), 69-85.

Phelps, C.E. (1986) Induced demand--can we ever know its extent? *Journal of Health Economics*. 6(4), 375-376.

Rice, T.H. (1983) The impact of changing Medicare reimbursement rates on physician-induced demand. *Medical Care.* 21(8), 803-815.

Roland, M. (2004) Linking physicians' pay to the quality of care- - a major experiment in the United Kingdom. *New England Journal of Medicine*. 351(14),1448-1454.

Ryan, R.M. and Deci, E.L. (2000a) Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*. 25(1), 54-67.

Ryan, R.M. and Deci, E.L. (2000b) Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist.* 55(1), 68-78.

Scott, A. (2002) For love or money: What makes us tick and leads to better health care. *Journal of Health Services Research and Policy.* 7(3), 184-185.

Scott, A. (2000) Economics of general practice. *In* Culyer, A.J. and Newhouse, J.P. eds. *Handbook of health economics.* Amsterdam: Newhouse Elsevier. 1175-1200.

Scott, A. (1997) *Designing incentives for GPs: A review of the literature on their preferences for pecuniary and non-pecuniary characteristics.* Health Economics Research Unit. Discussion Paper. Aberdeen: University of Aberdeen.

Sutton, M. and McLean, G. (2006) Determinants of primary medical care quality measured under the new UK contract: Cross sectional study. *British Medical Journal.* 332(7538), 389-390.

Titmuss, R.M. (1970) *The gift relationship: From human blood to social policy*. London: George Allen and Unwin Ltd.

Vayda, E.A. (1973) Comparison of surgical rates in Canada and in England and Wales. *New England Journal of Medicine.* 289(23), 1224-1229.

Vayda, E.A., Barnsley, J.M., Mindell, W.R. and Cardillo, B. (1984) Five-year study of surgical rates in Ontario's counties. *Canadian Medical Association Journal.*131(2), 111-115.

Vayda, E.A., Mindell, W.R. and Rutkow, I.M. (1982) A decade of surgery in Canada, England and Wales, and the United States. *Archives of Surgery*. 117(2), 846-853.

Von Neumann, J. and Morgenstern, O. (1944) *Theory of games and economic behavior*. Princeton: Princeton University Press.

Weller, D.P. and Maynard, A. (2004) How general practice is funded in the United Kingdom. *Medical Journal of Australia.* 181(2), 109-110.

Wise, S. (2005) *Professionalism and long working hours in NHS nursing*. Paper prepared for the 23rd Annual International Labour Process Conference. Strathclyde University. 21-23 March.