WORK IN PROGRESS: PLEASE DO NOT CITE WITHOUT PERMISSION

ARE DOCTORS CLOCK-WATCHING? THE EFFECT OF HOURS HARMONISATION ON HOSPITAL DOCTOR OUTPUTS

Matt Sutton
Eduardo Fé-Rodríguez

Health Economics
Health Methodology Research Group
Community Based Medicine
University of Manchester
Contact: matt.sutton@manchester.ac.uk

Abstract

Previous studies have demonstrated very substantial variations in the activity rates of hospital consultants in the UK. The new contract for UK hospital consultants introduced from April 2004 involved greater harmonisation of hours worked across consultants. This was intended to increase overall productivity but has been dubbed 'clock-watching'. We show that increased productivity could have been achieved by: increasing mean input; a mean-preserving reduction in input variation; and/or targeting of increases in input on individuals with higher marginal products. Using linked data from surveys undertaken in 2001 and 2006 and hospital administrative records from 2002/3 to 2005/6, we use fixed-effects negative binomial regression models to estimate the effect of hours worked on the number of patients treated. The resulting estimated individual marginal products can be used to judge whether the change in the distribution of hours worked increased aggregate output. We find that levels of output are highly variable across individual doctors. Hours devoted to clinical work are a significant determinant of the number of patients treated. Across individual doctors, there is a weak positive correlation between the changes in hours worked and the marginal product of hours spent on clinical care. The results suggest that, in isolation, the harmonisation of hours produced by the new consultant contract led to an increase of 16 patients treated per consultant per year, which equates to 2% of the mean. However, wide variation in marginal products across individual consultants suggests there is considerable scope for further increases in productivity.

November 2008

Acknowledgements

We are grateful to David Bailey for providing data and to other members of the project team that collected the doctors' survey data: Fiona French, Claudia Geue, Divine Ikenwilo, Gillian Needham, Catriona Rooke and Diane Skåtun. Preparation of the paper benefited from comments at seminars in Econometrics and Applied Economics at the University of Manchester and at the Centre for Health Economics at the University of York.

Introduction

New contracts were introduced for senior doctors across the UK in 2003/4. These contracts involved substantial pay rises and were intended to increase productivity. The new contract for family doctors (General Practitioners) introduced explicit output incentives. The new contract for senior hospital doctors (Consultants) introduced explicit input management. Hospital doctors have recently claimed that these input controls have led them to adopt a "clockwatching attitude" to their work (National Audit Office, 2007). In this paper we aim to test whether these input controls led to changes in aggregate output using a unique linked dataset containing doctors' work commitments and patient episodes.

Senior hospital doctors in the UK are salaried employees. Until 2003, they could be employed on one of four main contracts: Ordinary Full-Time (under which private income was restricted to no more than 10% of their full NHS salary); Maximum Part-Time (under which consultants received 10/11ths of a full NHS salary but faced no cap on their private income); Ordinary Part-Time (session-based payment); Honorary (primarily academic).

Williams and Buchan (2006) reported that there was widespread and wide-ranging concerns over the lack of accountability under this contract. There was no formal process for monitoring hours and many doctors reported working over their contracted hours while others were reported to not be fulfilling their NHS commitments. There was little link between pay and output and concern that doctors substituted private work for NHS work by maintaining long waiting lists.

Consequently, there was scope for considerable variation between doctors in both their inputs and their outputs. Bloor et al (2004) analysed variations in activity rates between consultants in five surgical specialties in 1998/9 and 1999/2000. They found substantial variations in activity between surgeons that were not confined to outliers. Depending on the specialty, the ratio of the upper to the lower quartile varied between 1.60 and 1.85.

Consultants could transfer to a new contract from October 2003. This aimed to link reward to 'NHS contribution' and created greater managerial control over consultants' work. Annual job plans are negotiated and agreed. This includes the quantum of hours and its composition across four categories of work: direct clinical care; supporting professional activities; additional NHS responsibilities; and external duties. Pay progression is determined by satisfactory performance against this job plan and the offer of the first four hours of overtime to the NHS. Around the same time the European Working Time Directive was introduced which capped working hours at 48-hours per week unless the employee signed a waiver.

In what may be seen as compensation for this loss of freedom, the contract reform substantially increased doctors' pay. Government figures (re-produced by Williams and Buchan, 2006) suggest basic pay increased by 49% between 2001 and 2005. In our data, consultants reported a mean 25% increase in income *after* tax from £5,170 per month in 2001/2 to £6,440 per month in 2005/6.

The increased management of consultant inputs was expected to lead to increases in consultant productivity. The then Scottish Executive included increases in consultant productivity as one of two 'time-releasing savings' targets for the health sector:

> "NHS health boards are expected to demonstrate an increase in consultant productivity by 1% per annum over a 3 year period from 1 April 2005 to 31 March 2008."

This was to be facilitated by the introduction of the new regime for managing inputs: "[the] focus on increasing consultant productivity will be underpinned by the new consultant contract". Meeting this productivity target was projected to secure 'time-releasing savings' worth £74M by 2007/8, but this target was subsequently judged to have been missed, was redefined and was then dropped altogether.

Setting a formal efficiency target in this area was high risk because remarkably little is known about consultant productivity, much less about the influences on it.[1] Moreover, there is little belief in the links between contracted and actual hours and between hours and outputs. The National Audit Office (2007) reported that many consultants saw the high numbers of Programmed Activities negotiated in the first year of the contract as a reward for the actual hours they worked. However, many trusts had subsequently reduced the number of Programmed Activities, so that many consultants no longer believed that their current contract reflected their working hours. Most importantly, outputs are likely to influenced by many factors and the target, and the judgement about whether it was achieved, did not take account of these factors.

In this paper we seek to isolate the contribution of the new contract to the change over time in the outputs of consultants. We examine how the introduction of the new contract influenced the distribution of contracted hours, whether this resulted in changes in the distribution of actual hours reported by consultants and whether the changes in the hours distribution resulted in changes in aggregate output.

We begin by highlighting how difficult the change in aggregate output is to predict based on aggregate statistics. The changes in output depend on the changes in hours, the shape of the production function and on individual heterogeneity in marginal products. For example, one possible explanation for how the harmonisation of hours could increase mean output is that output is concave in hours worked. Since there are diminishing output returns to hours, a mean-preserving reduction in the spread of hours will increase mean output.

Having identified how hours harmonisation may affect output, we use a unique dataset to test these mechanisms empirically. The dataset involves linkage of doctors' responses to two postal surveys with records from administrative NHS patient records. We examine activity in the two years before the introduction of the new contract and the two years after its introduction. We compare the ability of two input measures (self-reported and contracted hours of clinical work) to explain changes in output and provide results that allow for unobservable heterogeneity between individual consultants.

Methods

Inputs and aggregate production

We illustrate the difficulty of predicting the consequences for aggregate output of changes in the distribution of hours with just two workers. We consider scenarios in which the planner can increase mean inputs or induce a mean-preserving harmonisation of hours, and in which the production function is proportional or concave in input. Figures 1 and 2 provide a graphical exposition of the main arguments.

[1] Ongoing studies by Bloor and colleagues are rare exceptions.

The total output (Q) from these two workers is given by the sum of their individual outputs (q), which are individual-specific functions (f) of their hours of input (h):

$$Q = q_1 + q_2 = f_1(h_1) + f_2(h_2) \qquad [1]$$

If both workers are equally productive and output is proportional to input, so that $f_1(h_1) = \beta h_1$ and $f_2(h_2) = \beta h_2$, it is obvious that output can be increased by increasing the hours worked by either worker but a mean-preserving harmonisation of hours will have no impact on total output because $dq_1/dh_1 = dQ/dh_1 = \beta = dQ/dh_2 = dq_2/dh_2$.

If the production function for both workers is equal ($f_1(.) \equiv f(.) \equiv f_2(.)$) but not proportional, then a mean-preserving harmonisation of workers' hours can have an impact on aggregate output (see Figure 1). If the production function is concave, harmonisation of workers' hours will increase aggregate output because:

$$f(h_1 + \Delta) + f(h_2 - \Delta) > 0 \quad \text{if} \quad h_2 > h_1 \quad , \quad df/dh < 0 \qquad [2]$$

However, it is likely that the production function differs across workers. Once we allow *f(.)* to differ across workers the effect of hours harmonisation on aggregate output becomes harder to predict because individual 2's marginal product may still exceed individual 1's marginal product even if individual 2 is working more hours than individual 1 (Figure 2 provides an example).

In the case of a mean-preserving harmonisation of hours, what matters is whether the redistribution of hours is towards productive and away from unproductive workers. With *i* total workers, the overall effect on output is determined by the joint distribution of the changes in hours and each individual's marginal product:

$$\Delta Q = \sum_i \frac{dq_i}{dh_i} \Delta h_i \qquad [3]$$

Data

The information on doctors' input levels are from postal surveys of all doctors working for the NHS in Scotland in 2001 and in 2006 collected by NHS Education for Scotland/Health Economics Research Unit at the University of Aberdeen. The response rates to the surveys were 57% in 2001 and 52% in 2006. From the survey data we use sessional (old contract) and Programmed Activity (new contract) commitments, the number of hours worked per week, and personal and family characteristics.

A subset of respondents to the surveys consented to linkage of their data to administrative data held by ISDScotland. The linkage of the datasets was restricted to this sample. The survey responses were linked to activity data for each year between 2002/3[2] and 2005/6. This activity dataset covers patient episodes in 'acute' hospitals only, excluding long-stay specialties (such as psychiatry and continuing care of older people) and maternity services. Accordingly, we focus our analysis on consultants in acute hospital specialties only.

Records of each hospital episode were aggregated by ISDScotland by responsible consultant. The administrative data contain counts of inpatient episodes (Finished Consultant Episodes) and outpatient attendances for each consultant in each year, and the following mean characteristics of the inpatient episodes:

- the mean Healthcare Resource Group (HRG) based complexity index,
- the proportion of patients that are female,
- the average age of patients treated,
- the average deprivation score of patients treated,
- the proportion of cases treated as daycases,
- the proportions of cases admitted as emergencies and transfers, and
- the proportion of patients dying within 30 days of their admission.

[2] There was a change in the method by which patient activity was attributed to individual consultants between 2001/2 and 2002/3 such that figures before 2002/3 are not comparable.

The dataset also contains the number of months within each year in which the consultant has a patient discharge. This is used to correct the output models for doctors that are only active for part of the year.[3]

Through the linkage of these data sources we have data on 338 consultants working in acute hospitals in Scotland. We concentrate on consultants' clinical work. From the surveys in 2001 and 2006 we have two estimates of clinical time input: reported clinical hours worked per week and contracted clinical sessions/PAs. For these consultants we have four years' output data: two before the new contract (2002/3 and 2003/4); and two after the new contract (2004/5 and 2005/6). There are data on two output measures: numbers of Finished Consultant Episodes; and numbers of outpatients seen. As a proxy measure of quality we have the mortality rate for the patients of each consultant in each year.

Production function

We assume that the output (*q*) of individual *i* at time *t* is generated by an individual specific function (*f*) of hours (h) and other time-varying factors (X, including case-mix). We assume Cobb-Douglas production and measure individual-specific heterogeneity using time-invariant fixed effects:

$$q_{it} = \alpha_i h_{it}^{\beta} X_{it}^{\gamma} \quad [4]$$

The marginal product of individual *i* at time *t* is then:[4]

$$\left(\frac{\partial q_{it}}{\partial h_{it}}\right) = \beta\left(\frac{q_{it}}{h_{it}}\right) \quad [5]$$

The aggregate effect of the change in the distribution of hours can be estimated using the sum of the products of the individual changes in input and the individual marginal product estimates:

[3] We also experimented with health and sickness absence measures collected in the survey in case the doctor had reduced output because of a reduced numbers of days worked during the year. None of these was significant.
[4] Note that if β=1, the marginal product of individual *i* is the individual-specific rate of patients treated per hour.

$$\sum_i \Delta\hat{q}_i = \sum_i \left[\left(\frac{\partial q}{\partial h} \right)_i \Delta h_i \right] \quad [6]$$

In our application we calculate:

$$\sum_i \Delta\hat{q}_i = \sum_i \left[\hat{\beta} \left(\frac{q_{i,t=0}}{h_{i,t=0}} \right) (h_{i,t=1} - h_{i,t=0}) \right] \quad [7]$$

<u>Econometric estimation</u>

In order to capture the discrete nature of the output variable we adopt a Negative Binomial model. Given the panel structure of the data, we make use of the models described by Hausman, Hall and Griliches (1984).

There is an additional complication in the model. Measurement errors in *h* will influence the estimate of β, the denominator of the second term and the change over time represented by the third term in [7]. In particular, errors in $h_{t=0}$ will induce positive correlation between the estimated marginal products and the changes in hours. In our study the number of NHS clinical hours reported by each consultant is likely to be an endogenous covariate due to two different sources of correlation with unobservables:

1. Clinical hours are likely to be correlated with unobserved time invariant heterogeneity; thus, the total number of clinical hours may reflect a consultant's inherent skill.
2. Clinical hours are likely to be subject to measurement error. Thus, the reported value is likely to follow the scheme $clinical_{it} = clinical_{it}^{*} + \varepsilon_{it}$ where ε_{it} is a time varying error component.

Either case may be framed as a measurement error problem so that $E(y_{it} \mid x_{0,it}, x_{e,it}, x_{u,it}, u_i) = \tilde{\lambda}_{it} = \lambda_{it}\alpha_i$, where $\lambda_{it} = \exp(x_{0,it}\beta_1 + x_{e,i}\beta_2 + x_{u,it}\beta_3)$, $x_{o,it}$, $x_{e,it}$ and $x_{u,it}$ are k_1, k_2 and k_3 vectors of exogenous, endogenous and unobservable variables respectively, $\alpha_i = \exp(u_i)$, with u_i capturing unobserved time invariant heterogeneity, so that $y_{it} = \tilde{\lambda}_{it}\alpha_i + v_{it}$ for some i.i.d. v_{it} uncorrelated with the nonlinear part of the model, and β_i are conformable vectors of parameters which need to be estimated.

In the present context of panel data, it is possible to tackle both sources of endogeneity at once, by using HHG's Fixed Effect count data models. To see this, we note, firstly, that the conditional maximum likelihood formulation of the model eliminates the individual effects from the (conditional) likelihood function. Thus, expression (2.4.) in HHG shows how

$$P(y_{i1},...,y_{iT} \mid \sum_{t=1}^{T} y_{it}; x_{e,it}, x_{o,it}, x_{u,it}) = \frac{\left(\sum_t y_{it}\right)!}{\prod_t y_{it}!} \prod_t \left(\frac{\tilde{\lambda}_{it}}{\sum_t \tilde{\lambda}_{it}} \right)^{y_{it}}$$

However, time invariance of α_i implies that

$$P(y_{i1},...,y_{iT} \mid \sum_{t=1}^{T} y_{it}) = \frac{\left(\sum_t y_{it}\right)!}{\prod_t y_{it}!} \prod_t \left(\frac{\lambda_{it}}{\sum_t \lambda_{it}} \right)^{y_{it}}$$

and this does not depend on the unobserved time invariant heterogeneity. Unfortunately, λ_{it} still contains a set of endogenous variables, where the endogeneity arises via neglecting the correlated information set in the time varying regressors $x_{u,it}$. However, at this stage, Hausman's (1978) Two Stage Residual Inclusion (2SRI) may be used to obtain consistent estimates of the parameters β_j in the model (see also Terza et al, 2008). Given a set of informative instrumental variables, w_{it}, the residuals from the ancillary (nonlinear) regression $x_{e,it} = \rho(w_{it}'\gamma) + x_{u,it}$ (after consistent estimation) may be included as an explanatory variable in λ_{it}, so that one considers the conditional likelihood,

$$P(y_{i1},...,y_{iT} \mid \sum_{t=1}^{T} y_{it}) = \frac{\left(\sum_t y_{it}\right)!}{\prod_t y_{it}!} \prod_t \left(\frac{\lambda^*_{it}}{\sum_t \lambda^*_{it}} \right)^{y_{it}}$$

where now, $\lambda_{it} = \exp(x_{0,it}\theta_1 + x_{e,i}\theta_2 + (x_{e,it} - \rho(w_{it}'\hat{\gamma}))\theta_3)$. Intuitively, given that the estimator $\hat{\gamma}_n \to \gamma$ in probability, the maximizer of the log-likelihood function will be such that $|\theta_j - \beta_j| \to 0$ in probability, which justifies the method.

Results

Table 1 shows the trends in mean output per doctor over the four-year period. Mean finished consultant episodes per doctor is 694 in the first year 2002/3. This rises to 707 in 2003/4 and then declines in the first two years of the new contract.[5] The full distributions of FCEs in 2002/3 and 2005/6 are shown in Figure 3. Throughout the period, the mean number of outpatients seen per doctor per year falls from 716 in 2003/4 to 643 in 2005/6. The mortality rate rises from a mean of 2.33% in 2002/3 to 2.49% in 2005/6.

Descriptive statistics for the other key variables in the first year and the last year are shown in Table 2. There is an increase of 2% in the mean complexity of episodes from £1591 to £1626. There is a slight increase in the proportion of female patients and small declines in the average area income deprivation scores of admitted patients and the average length of stay. Under 5% of doctors are categorised as part-year workers. There are reductions in the proportions of doctors reporting significant commitments to non-clinical activities.

The mean number of reported clinical hours is 34 in both years but, as further illustrated in Figure 4, there is a substantial reduction in the extent of variation. The mean number of contracted clinical hours rises by just under 5 hours per week between the pre-contract and post-contract period and equals the reported clinical hours at the mean in the post-contract period. As illustrated in Figure 5, there is a marked shift in the contracted clinical hours distribution.

We begin the regression analysis with models of reported clinical hours regressed on contracted clinical hours and other determinants of input. The relationships between reported and contracted clinical hours are shown in Figure 6. The results using random-effects and fixed-effects specifications are shown in Table 3. Contracted clinical hours is a strong predictor of reported clinical hours and, as appears to be evident in Figure 6, this mapping changes significantly over time in the random-effects models. In both years the line of best-fit slopes upwards but does not lie on the 45^0 line and shows a better correspondence in 2006. In the random-effects models doctors with a spouse/partner work provide up to 4 hours fewer clinical hours per week than those without a spouse/partner. Male doctors with at least one

[5] Presumably why the efficiency target was dropped.

child under the age of five report up to two hours more clinical hours of work per week and females with children under five report fewer hours. Doctors with significant administrative, research or medical politics commitments work fewer clinical hours per week. Overall just under 40% of the variation in reported clinical hours can be explained. A joint test of the significance of these instruments is F(8, 564) = 23.12, which exceeds the Staiger-Stock (1997) recommended critical value of 10. The fixed-effects models indicate that only the change in contracted hours and the reduction in respondents reporting significant administrative work can explain the change in reported clinical hours.

In Table 4 we compare the results of two pooled output models with the fixed-effects specification.[6] The first pooled model does not contain specialty effects whereas the second contains 32 dummy variables representing the main specialty in which the doctor produces output. In the pooled models, the year dummies are not significant but 2005/6 is associated with significantly less output than 2002/3 in the panel data specification. In all models there is complementarity between outpatient and inpatient outputs. The mortality rate is positively related to the number of inpatient admissions in the pooled models but this is shown, in the panel data model, to be confounded by consultant heterogeneity. Each of the case-mix indicators is significant in the pooled models but in the fixed-effects specification only the HRG complexity is significantly associated with output and this is negative as expected. Clinical hours are significantly associated with output levels in all models, though the coefficient is substantially reduced as we allow first for specialty differences and then individual differences in productivity between doctors. Overall, the differences in the coefficients in the pooled and panel models suggest that ‘low productivity doctors’ have lower mortality rates, more complex case-mix, fewer patients from deprived areas, longer mean length-of-stay and shorter working hours. Females produce less output than males, though this is exaggerated if we do not allow for differences between specialties. The relationship between doctor age and output is an inverted-U function which reaches a maximum at the age of 45 years.

Table 5 contains the results from three panel data models. The first is reproduced from Table 4 and shows the effect of reported clinical hours on output. The second additionally contains the residuals from the model in Table 3. It shows that those consultants that reported more

[6] The included measures of case-mix are those that are ever significant in any models. The other available measures were not significant.

than expected increases in clinical hours had lower than expected increases in output. We interpret this as an indication that reported clinical hours is endogenous because of measurement error. Correcting for this endogeneity results in a substantially larger estimated effect of a change in clinical input on output. The third model uses contracted clinical hours as the measure of input. It also shows that input has a significant impact on output.

The distributions of the marginal products estimated using the two alternative input measures are shown in Figure 6. There is considerable variation using both measures. The mean (SD) values are 2.1 (1.9) and 2.9 (2.2) patients per annum using reported and contracted hours, respectively. The models in Table 5 are estimated on slightly different samples. Estimated on the same samples, the model log-likelihoods using the reported and contracted hours measures are, respectively, -5017.6 and -5014.5. The contracted hours measure therefore provides a marginally better fit to the data.[7]

Figure 7 shows a scatterplot for the two components of equation [6]: the individual changes in contracted clinical hours and the individual marginal product estimates. There is a weak positive relationship as shown by the upward sloping lowess line of best fit. The correlation between the change in contracted clinical hours and the estimated marginal products is 0.136 ($p=0.018$). Finally we obtain estimates of the aggregate changes in Q as the sum of the products of each individual's change in hours and their marginal products (equation [7]). The mean predicted change in output using contracted hours is 15.6 patients per year, which is a little over 2% of the mean.

Discussion

In this paper we have shown that, assuming output is a concave function of input, an increase in mean output could be generated by an increase in mean input, a mean-preserving reduction in the spread of inputs and/or a positive correlation between marginal products and changes in input.

[7] The best model includes both measures and each is positive and significant at $p<0.05$, with the contracted hours having a larger coefficient and z-statistic.

The dataset we have used indicates that the new contract for hospital doctors achieved greater harmonisation of hours. The hours reported by doctors show a mean-preserving reduction in the extent of variation. The relationship between reported and contracted hours becomes more consistent following the introduction of the new contract and the changes in contracted hours are significantly associated with the changes in reported hours. In this sense, there is more 'clock-watching' by consultants following the introduction of the new contract.

Whether this 'clock-watching' represents time spent being idle depends on its links with outputs. Our analysis of variations in output across doctors and over time shows that changes in input were significantly associated with changes in output. However, this production function exhibits considerable heterogeneity across individual doctors and there is evidence that changes in reported hours contain significant measurement error. The best input predictors of the changes in output are the changes in contracted clinical hours. Using this input measure we find that the changes in input are positively correlated with the estimated marginal products and, since the mean value of this input measure increases by 5 hours per week, we find that the new contract may have generated a small increase in mean output equal to 16 additional patients per consultant per year (equivalent to 2% of the mean).

Our study is one of very few that has analysed variations in the output of individual hospital doctors. The dataset we have assembled is unique but has a number of limiting features including the relatively small sample size, the limited proxy measure of quality and the lack of information on the associated health care teams.

Our analysis contains a number of key steps and we have not yet systematically tested the robustness of the findings to alternative modelling strategies. In particular, we have not bootstrapped the entire procedure to illustrate the uncertainty around our headline finding that the new contract induced, in isolation, a 2% increase in aggregate output. Moreover, we have not investigated whether the sample of consultants that participated in both rounds of the postal survey are representative of the entire population of consultants.

Nevertheless, our analysis illustrates the complexity of attempting to answer the question of whether the greater input management introduced with the new contract led to improvements in productivity. With such complex linkages between variables, such heterogeneity between individual doctors and such a range of influences on outputs, it is likely to always remain risky to set targets for productivity increases and difficult to evaluate the consequences of initiatives aiming to change productivity.

References

Bloor K, Maynard A, Freemantle N. Variation in activity rates of consultant surgeons and the influence of reward structures in the English NHS. *Journal of Health Services Research and Policy* 2004; 9(2): 76–84.

Hausman J. Specification Tests in Econometrics. *Econometrica* 1978; 46: 1251-1277.

Hausman J, Hall B, Griliches Z. Econometric Models for Count Data with an Application to the Patents-R & D Relationship. *Econometrica* 1984; 52(4): 909-938.

National Audit Office. *Pay Modernisation: A New Contract for NHS Consultants in England*. Report by the Comptroller and Auditor General. HC 335 Session 2006-2007. 19 April 2007.

Scottish Executive. *Building a Better Scotland: Spending Proposals 2005-2008: Enterprise, Opportunity, Fairness*. Edinburgh 2004.

Staiger D, Stock J.H. Instrumental variables regression with weak instruments. *Econometrica* 1997; 65: 557–586.

Terza JV, Basu A, Rathouz P. Two-Stage Residual Inclusion Estimation: Addressing Endogeneity in Health Econometric Modeling. *Journal of Health Economics* 2008; 27: 531-543.

Williams S, Buchan J. *Assessing the new NHS consultant contract: a something for something deal?* King's Fund 2006.

Table 1. Trends in mean output per doctor

Year	Finished episodes	Outpatients	Mortality Rate
2002/3	694.4	715.6	2.33%
2003/4	707.0	691.5	2.48%
2004/5	694.1	671.8	2.39%
2005/6	662.7	642.6	2.49%

Table 2 Descriptive statistics

Year		2002			2005	
Variables	N	Mean	SD	N	Mean	SD
Finished Episodes	338	694.4	516.0	338	662.7	527.5
Outpatients seen	338	715.6	694.3	338	642.6	587.8
Mortality rate	338	2.33%	2.57%	338	2.49%	2.62%
Complexity	338	1591.2	740.9	338	1626.2	731.9
Female patients	338	0.492	0.130	338	0.500	0.133
Deprivation	338	15.72	4.43	338	15.54	4.44
Length of stay	337	3.45	4.88	337	3.43	6.20
Part-year worker	338	4.7%	-	338	4.4%	-
Reported clinical hours	329	34.2	11.4	315	34.1	9.3
Contracted clinical hours	318	29.4	6.6	320	34.2	8.0
Partner/spouse	338	92.3%	-	338	93.5%	-
Child under 5 years	338	21.3%	-	338	11.2%	-
>4 hours admin per week	338	45.0%	-	338	25.1%	-
>2 hours research per week	338	28.4%	-	338	19.2%	-
>4 hours medical politics per week	338	10.7%	-	338	6.2%	-

Table 3 Models for reported clinical hours

Model	(1)		(2)		(3)		(4)		(5)		(6)	
	RE		RE		RE		RE		FE		FE	
Variable	**Coeff.**	**z**	**Coeff.**	**z**	**Coeff.**	**z**	**Coeff.**	**z**	**Coeff.**	**t**	**Coeff.**	**t**
2005[§]			-1.940	-0.7	-1.838	-0.6	-1.679	-0.6	1.144	0.3	0.242	0.1
Contracted hours	0.479	10.3	0.578	7.8	0.535	7.5	0.466	6.3	0.298	3.1	0.275	2.8
2005*Contracted hours			-0.023	-0.3	-0.051	-0.6	-0.038	-0.4	-0.065	-0.6	-0.053	-0.5
Has spouse/partner					-2.426	-1.7	-3.859	-2.6			-1.910	-0.7
Has child <5 years					1.888	1.8	1.848	1.8			-0.060	0.0
Female*Has child <5yrs					-4.110	-1.9	-3.685	-1.7			-0.770	-0.2
>4 hours admin p.w.					-3.051	-4.2	-2.532	-3.4			-2.981	-2.9
>2 hours research p.w.					-5.009	-5.7	-3.901	-4.2			-1.197	-0.7
>4 hrs med politics p.w.					-2.492	-2.0	-2.822	-2.2			-1.033	-0.6
NT	615		615		615		613		615		613	
N	333		333		333		333		333		333	
R^2	0.212		0.236		0.322		0.398					
R^2 (within)									0.051		0.096	
ρ	0.386		0.350		0.304		0.263		0.585		0.588	
Test year and year*contract interactions = 0	-	-	chi2(2)=17.18, p<0.001		chi2(2)=28.00, p<0.001		chi2(2)=18.59, p<0.001		F(2,279)=1.03, p=0.358		F(2,264)=1.84, p=0.161	
Contains output determinants?	No		No		No		Yes		No		Yes	

[§] Reference year is 2002.

Table 4 Comparison of pooled and panel output models

Model	Pooled (without specialties)		Pooled (with specialties)		Fixed effects	
Variable	Coeff.	z	Coeff.	z	Coeff.	z
2003§	0.046	0.9	0.037	0.9	0.002	0.2
2004§	0.024	0.5	0.021	0.5	-0.018	-1.2
2005§	-0.031	-0.6	-0.010	-0.3	-0.052	-3.5
Outpatients seen	0.060	4.7	0.099	5.7	0.132	8.6
Mortality Rate	0.203	15.8	0.062	3.4	0.006	0.7
Complexity index	-0.651	-15.4	-0.892	-15.0	-0.398	-6.5
Female patients	-0.225	-1.4	0.369	2.4	0.018	0.1
Income deprivation	0.014	3.3	0.016	4.3	-0.005	-0.9
Length of stay	-0.056	-15.2	-0.033	-6.1	-0.011	-1.8
Clinical time input	0.384	7.9	0.240	5.7	0.095	3.7
Female doctor	-0.183	-3.4	-0.090	-1.9	-	-
Doctor age	24.229	3.0	37.249	5.7	-	-
Doctor age^2	-3.182	-3.1	-4.888	-5.7	-	-

All models contain a dummy variable for part-year workers. The pooled model with specialties contains 32 speciality dummies. § Reference year is 2002.

Table 5 Panel output models

Measure of clinical input	Reported hours		2SRI		Contracted hours	
Variable	Coeff.	z	Coeff.	z	Coeff.	z
2003§	0.002	0.2	0.010	0.7	0.007	0.5
2004§	-0.018	-1.2	-0.035	-2.2	-0.036	-2.3
2005§	-0.052	-3.5	-0.050	-3.2	-0.066	-4.2
Outpatient attendances	0.132	8.6	0.123	7.8	0.124	8.3
Mortality Rate	0.006	0.7	-0.005	-0.6	-0.006	-0.7
Complexity index	-0.398	-6.5	-0.342	-5.5	-0.341	-5.7
Female patients	0.018	0.1	0.039	0.3	0.020	0.2
Income deprivation	-0.005	-0.9	-0.008	-1.3	-0.008	-1.3
Length of stay	-0.011	-1.8	-0.011	-1.9	-0.013	-2.2
Clinical time input	0.095	3.7	0.279	4.0	0.118	3.5
Residual			-0.215	-3.2		
Constant	5.380	11.8	4.399	8.4	4.993	11.0
NT	1280		1222		1268	
N	333		332		336	
Initial Log-L	-11645.1		-10550.8		-11491.1	
Model Log-L	-5367.9		-5012.5		-5250.8	

All models contain a dummy variable for part-year workers. § Reference year is 2002.

Figure 1 Hours harmonisation with common production function

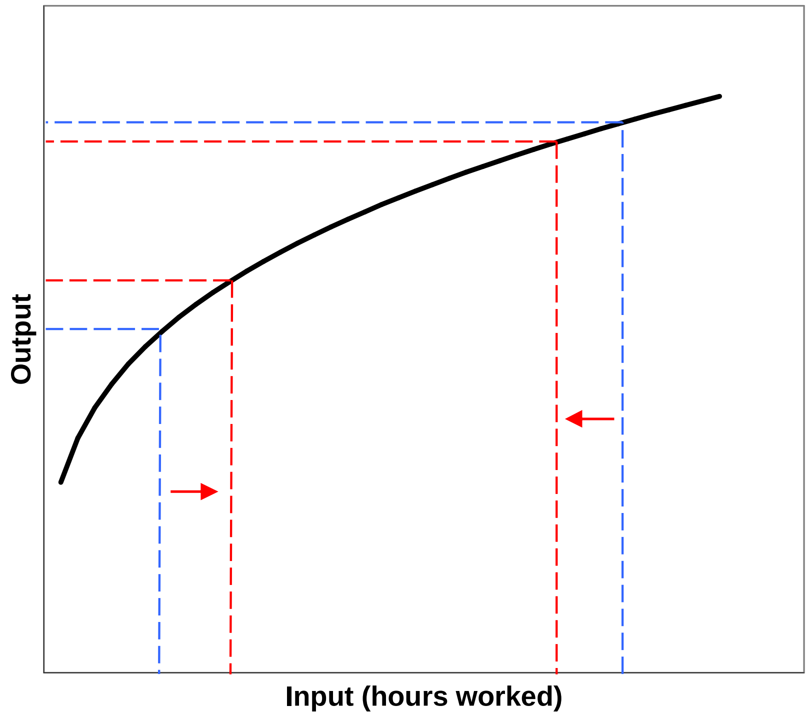


Figure 2 Hours harmonisation with variable production functions

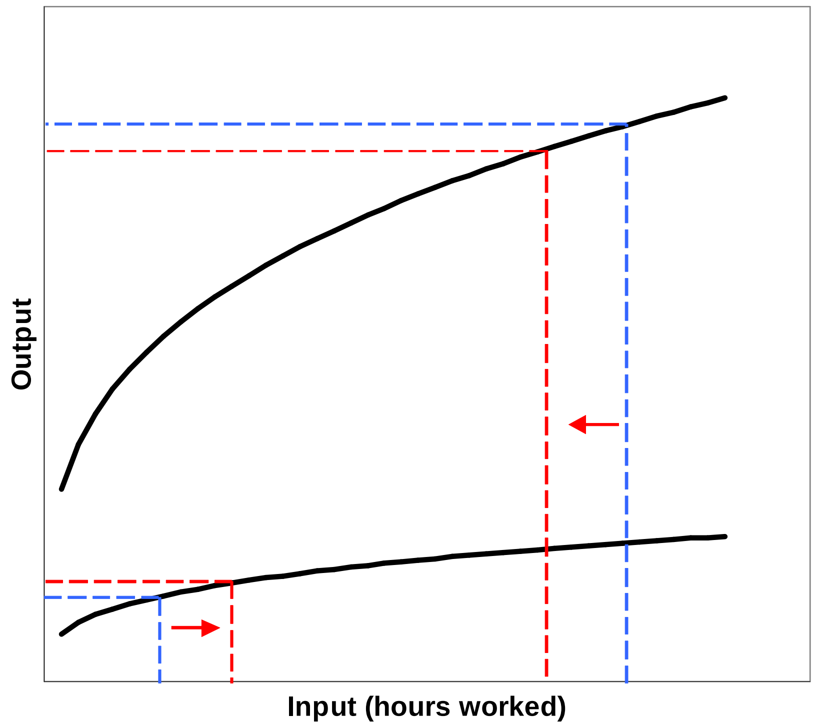
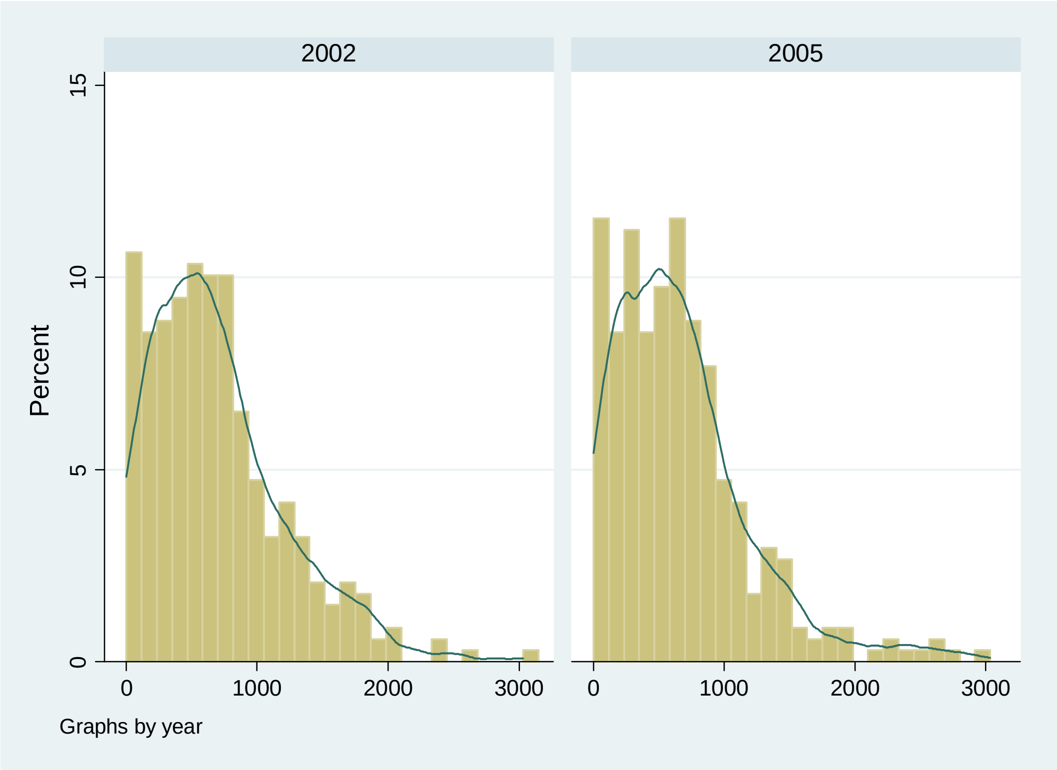
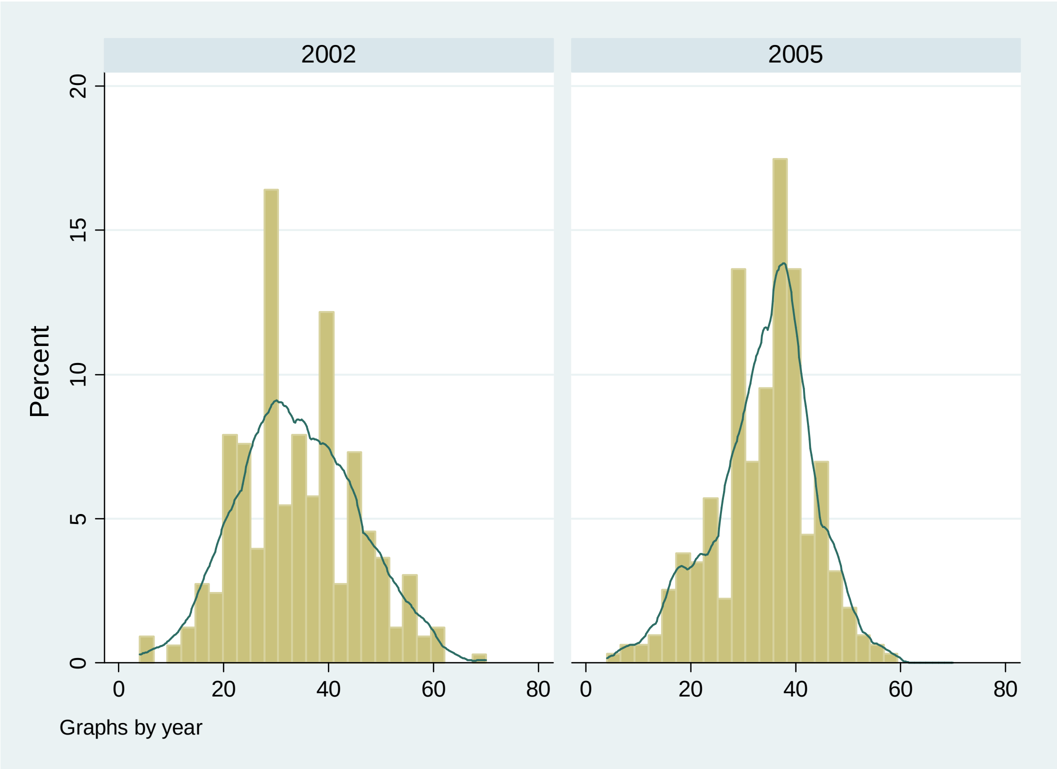


Figure 3 Distributions of Finished Consultant Episodes in 2002 and 2005



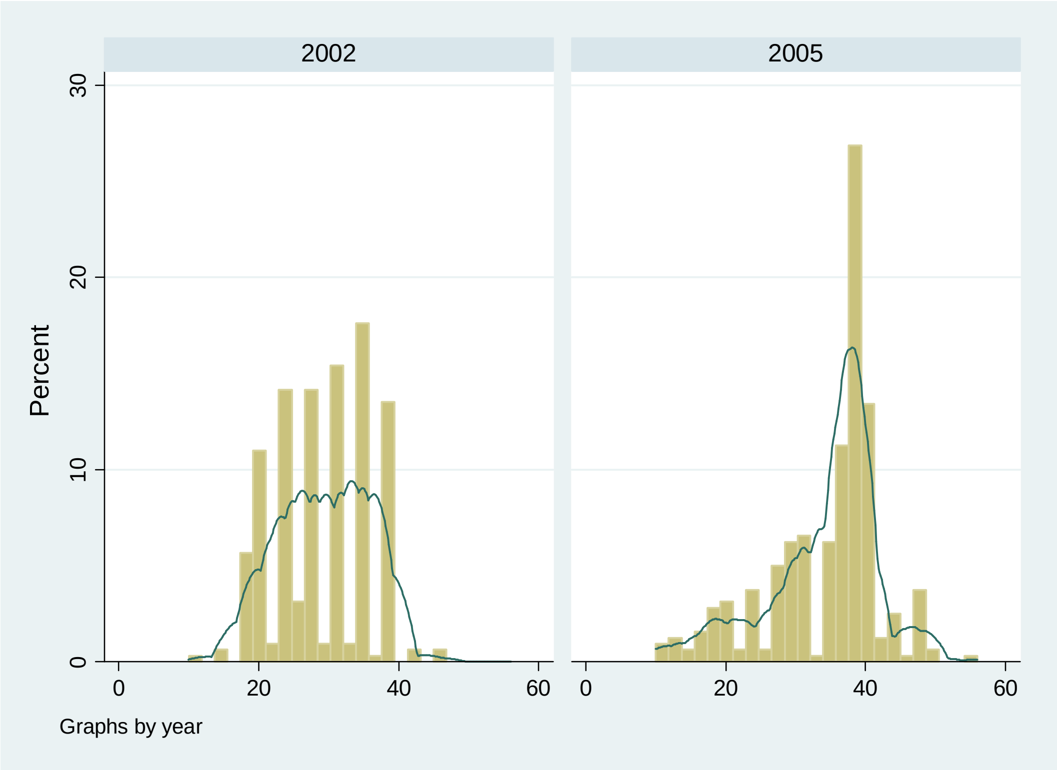
Smooth lines are kernel density functions.

Figure 4 Distributions of reported clinical hours worked per week



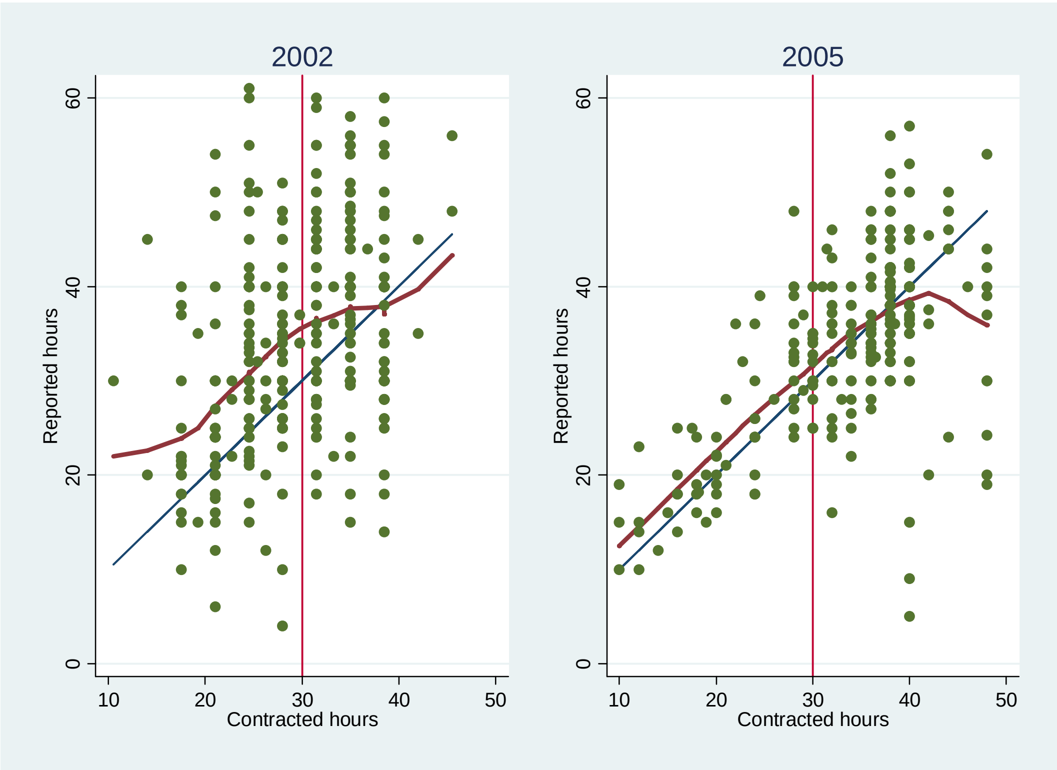
Smooth lines are kernel density functions.

Figure 5 Distributions of contracted clinical hours per week



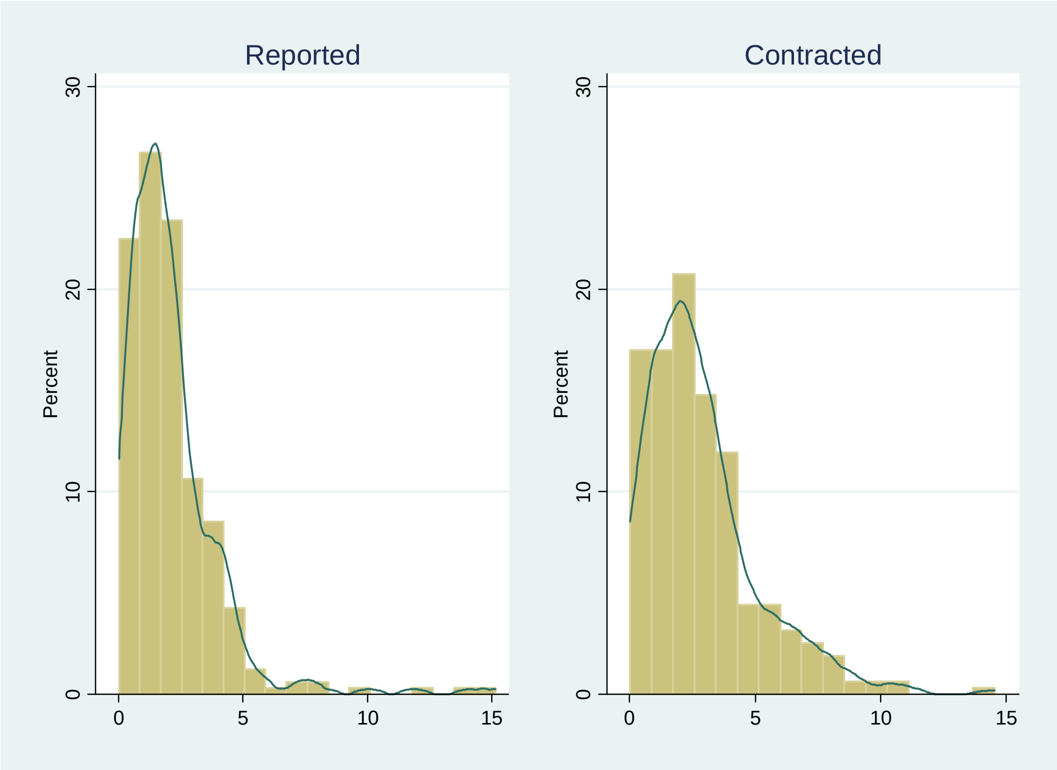
Smooth lines are kernel density functions.

Figure 6 Scatterplots of reported hours on contracted hours by year



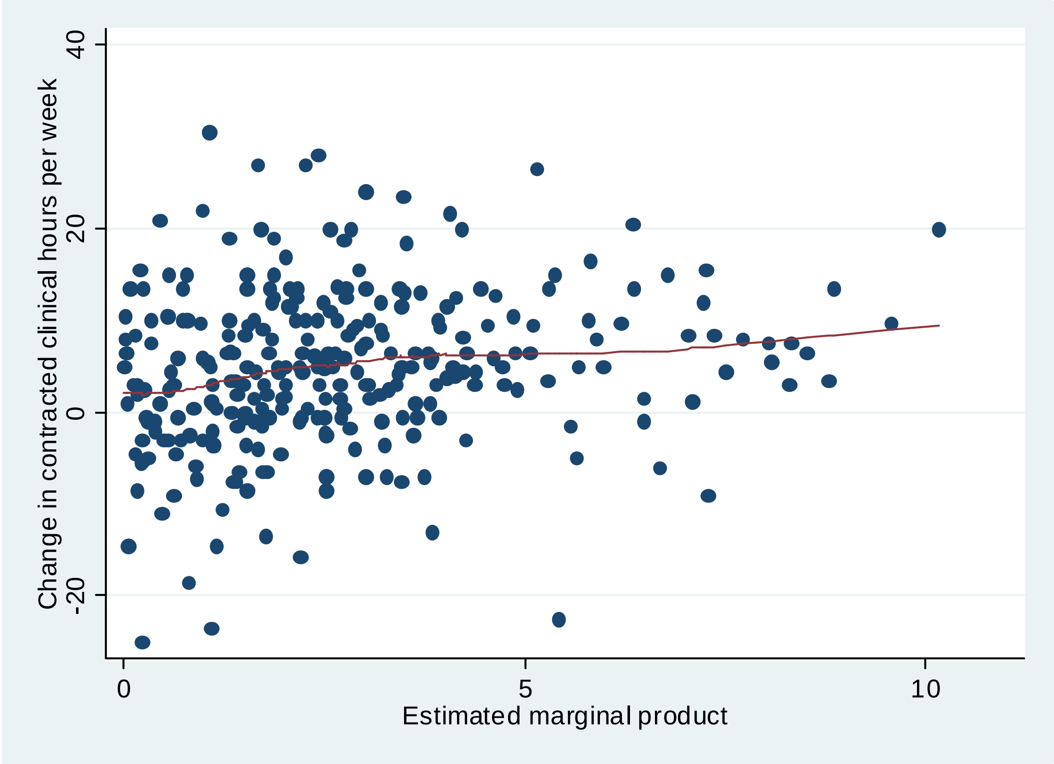
Straight line is 45^0-line. Smoothed lines of best fit are *lowess* (bandwidth = 0.8)

Figure 7 Distributions of estimated marginal products using alternative input measures



Smooth lines are kernel density functions.

Figure 8 Scatterplot of changes in contracted clinical hours on marginal products



Smoothed line of best fit is *lowess* (bandwidth = 0.8).