THE
ECONOMIC
JOURNAL

*The Economic Journal*, **122** (*June*), 555–589. Doi: 10.1111/j.1468-0297.2011.02472.x. © 2011 The Author(s). The Economic Journal © 2011 Royal Economic Society. Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

# THE EFFECTS OF PROMOTIONS ON HEART DISEASE: EVIDENCE FROM WHITEHALL*

*Michael Anderson and Michael Marmot*

The positive relationship between SES and health is well documented but limited evidence exists regarding the effect of an exogenous manipulation of SES on health. This article estimates the effect of promotions on heart disease using data on British civil servants from the Whitehall II study. Differences in promotion rates across departments and cohorts generate plausibly exogenous variation in promotion opportunities. The results suggest that promotions may reduce the probability of developing heart disease by 2.6–12.8 percentage points over a 15-year period. These estimates appear robust and are several times larger than cross-sectional estimates.

A long-standing debate exists regarding the relationship between socioeconomic status (SES) and health. The positive cross-sectional correlation between SES and health is well established (Marmot, 2003). Nevertheless, two surveys of the literature by economists document the difficulty in measuring the causal effect of SES on health as defined by Rubin (1974) and summarise the empirical evidence as inconclusive (Smith, 1999; Deaton, 2003). Some research suggests that lagged SES predicts future health outcomes (Adda *et al.*, 2003) and a large number of studies examine the structural channels through which the observed health gradient may operate (Marmot *et al.*, 1997; Kuper and Marmot, 2003; Chandola *et al.*, 2005). However, there is little evidence on the effect of an experimental manipulation of SES on health outcomes (Mealli and Rubin, 2003).

One salient finding in this literature is that income differentials across developed countries are uncorrelated with life expectancy but that income differentials within developed countries are strongly related to life expectancy (Deaton, 2003). Deaton and Paxson (2001) present a framework for analysing these patterns. Within this framework, health is an increasing function of the difference between an individual's income and the average income of his or her reference group. As a result, the underlying independent variable in the model – the gap between observed income and average reference group income – is unobserved. Instead, a noisy measure of this variable is observed. Conventional linear regressions of health on income may therefore understate the health effects of increased income if individuals have very different reference points (as is likely in cross-country regressions). This model explains the divergence between cross-country and within-country health gradients

* Corresponding author: Michael Anderson, 207 Giannini Hall MC 3310, Berkeley, CA 94720-3310, USA. Email: mlanderson@berkeley.edu.

We are extremely grateful to Tarani Chandola and other members of the Whitehall II Study Group, for their help in accessing the data used in this study. We also thank Jerome Adda, Anna Aizer, Josh Angrist, David Autor, David Card, Michael Greenstone, Jonathan Gruber, Ted Miguel and members of the MIT Labor Lunch, UC Berkeley Health Economics Workshop and UIC Health Economics Seminar for their insightful comments. Funding from the National Institute on Aging, through Grant Number T32-AG00186 to the National Bureau of Economic Research, the George Schultz Fund and the Robert Wood Johnson Foundation is gratefully acknowledged. The authors bear sole responsibility for the contents of this article.

and it suggests a way to minimise the attenuation bias. Specifically, if we limit the study sample to a relatively homogeneous population that shares a similar reference group, then within that population a strong relationship between income and health may arise.

The Whitehall II data focus on a plausible 'ready-made' reference group: British civil servants working in Inner London. The original Whitehall study collected data on over 18,000 white-collar male civil servants in the vicinity of the Whitehall area of London (Reid *et al.*, 1974). Although it was not the original aim of the study, Marmot *et al.* (1978) documented a significant relationship between employment grade and coronary heart disease (CHD). This finding surprised many researchers, who did not expect to observe a large health gradient in a relatively homogenous population. The substantial body of research originating from the first Whitehall study aided the design of a second Whitehall study, devised specifically to explore the causal channels between employment grade and health. This study, known as Whitehall II, collected longitudinal data on over 10,000 white-collar civil servants, beginning in 1985. It reconfirmed the original Whitehall results and advanced the hypothesis that factors such as increased job control could underlie the relationship between employment grade and heart disease (Marmot *et al.*, 1991; Bosma *et al.*, 1997). However, it has proven difficult to produce a convincing estimate of the effect of an exogenous manipulation of employment grade on CHD.

To measure the effect of promotions on CHD, we use a plausibly exogenous source of variation in employment grade that allows us to address issues of selection bias and measurement error. Specifically, we exploit variation in promotion rates across major Civil Service departments and cohorts. For any given promotion slot, candidates are selected on the basis of merit (Stanley, 2004). However, within departments candidates can only be promoted if a slot is available, and departments cannot easily change the number of available promotion slots. Departmental promotion rates during the study period therefore have little relationship to average civil servant quality within a department after conditioning upon grade of entry. Instead, they are a complex function of relative cohort sizes, departmental grade composition and overall employee departure rates (HM Treasury Office, 1985). Our empirical strategy uses variation in promotions across departments and cohorts as instruments to identify the effects of promotion on CHD. Using different subsets of this variation, we estimate large effects of promotions on heart disease.

Of course, civil servants may select into departments based on expectations of future promotion rates and some cohorts may be promoted quickly because they contain exceptionally skilled workers. If these unobserved factors are correlated with actual promotion rates, our instrument may not be valid. We therefore analyse a range of observable health measures that should not be affected by promotion to establish whether pre-treatment health is correlated with the instrument. We find that the correlation between pre-treatment health characteristics and the instrument is insignificant and often in the opposite direction of our instrumental variables (IV) estimates for heart disease. Furthermore, the overall correlation between pre-treatment health characteristics and individual promotions is small in magnitude and statistically insignificant, suggesting that selection bias plays a modest role even in the simple differences estimates.

The results suggest that promotions reduce the probability of heart disease by 3–13 percentage points over a 15-year period. The IV estimates are larger than the first differences estimates, and both sets of results are larger than cross-sectional estimates in the previous Whitehall literature. Nevertheless, they are consistent with other research on the causal effects of SES on health. We discuss several possible explanations for the divergence between the IV and differences estimates, including rational expectations, mis-specified reference groups and positive external effects of promotions.

The article is organised as follows. Section 1 describes the data. Section 2 discusses a simple model that provides a framework for interpreting the empirical results. Section 3 presents the statistical models and estimates of the effects of promotions on heart disease and other health outcomes. Section 4 discusses and tests potential threats to the validity of the instruments. Section 5 summarises the results and discusses possible explanations for the observed pattern of effects. Section 6 concludes.

## 1. Data and Descriptive Statistics

The Whitehall II data set is noteworthy in that it represents a large longitudinal data set containing both objective health measures and detailed employment grade data. The data set and its predecessor (the original Whitehall data) have been influential in the epidemiology literature; results from both data sets are often cited in the context of SES and health. To the best of our knowledge, this research represents the first analysis of the Whitehall II data by a researcher outside the Whitehall Study Group. The Whitehall II sample contains 10,308 civil servants employed in Inner London from 1985–7. All males and females aged 35–55 from 20 Whitehall departments were eligible for inclusion. Overall response rates were approximately 76%. The initial data collection consisted of a medical screening and a lengthy questionnaire. Follow-up data have been collected over four subsequent 'phases' since the initial screening, with questionnaires in 1990, 1992, 1996 and 1998, and medical screenings in 1992 and 1998 (each phase takes places over a two- to three-year period).[1] Questionnaire data include information on age, gender, employment grade, tenure, marital status, parent and sibling health, education and self-rated health. Medical measures available for this study include weight, height, mortality and the presence of ischaemic heart disease. Ischaemic heart disease – the focus of this study – was diagnosed via electrocardiogram during study screenings in 1985, 1992, or 1998 or in standard medical screenings at any other time from 1985 to 1999. Attrition rates in the subsequent phases range from 16% to 24% (Marmot and Brunner, 2004) but for ischaemic heart disease and mortality there is no attrition because participants' medical records are flagged by the National Health Service (NHS). Nevertheless, the ischaemic heart disease measure may be affected by the fact that 15% of our analytic sample was not screened in 1998. Heart disease – including fatal heart disease – is identified either in the 1998 medical screening or because an individual was diagnosed with heart disease by a medical doctor. Individuals who miss the 1998 screening and have little contact with the medical system thus may be miscoded as not having heart disease. We explore the

[1] Two additional phases, in 2001 and 2004, have also occurred. However, the first author does not have access to these data.

implications of this miscoding in Section 4.2 and conclude that it cannot explain our results.

Employment grades in the British Civil Service – the key explanatory variable – were standardised across most departments in 1971 (Her Majesty's Stationery Office, 1971). For research purposes, employment grades were further condensed into six primary grade levels. Ranked from highest to lowest, they are: Unified Grades 1–6 (Administrative), Unified Grade 7 (Administrative), Senior Executive Officer, Higher Executive Officer, Executive Officer and Clerical/Support Staff.[2] In the existing Whitehall literature, these grades are generally labelled 1 through 6, with 1 being the 'highest' grade (Unified Grades 1–6) and 6 being the 'lowest' grade (Clerical/Support Staff). In this article, however, we reverse the numbering so that 1 corresponds to the 'lowest' grade and 6 corresponds to the 'highest' grade. Although it is inconsistent with the previous Whitehall work, it makes the interpretation of regression coefficients more straightforward.

Promotions to higher grades entail raises; at the grade levels common in our data, a promotion increases salary by an average of 23–48% (HM Treasury Office, 1989). Health care benefits – which are provided through the NHS – do not change. Promotions can change working conditions, however. Our data contain questions regarding social support at work, job demands and decision latitude.[3] There is no relationship between promotions and social support at work but a promotion of one grade level is associated with a 4.0% (0.11 SD) increase in the job demands index and a 4.4% (0.16 SD) increase in the decision latitude index. Both increases are statistically significant. Although we cannot observe all aspects of the work environment, we experiment with controlling for the aspects that we do observe. We find that including these aspects as controls does not change the coefficient on promotions. Nevertheless, our promotion estimates necessarily capture some combination of income and work environment effects.

The sampling frame of the Whitehall II data merits special attention. In an ideal experiment, each civil servant would enter the data set as soon as he or she joined the Civil Service and remain in it until time of death. In the Whitehall II data, however, civil servants only enter the data set if they are employed in the Civil Service between 1985 and 1987. Because average tenure exceeds 17 years, the sample is selected – individuals are more likely to enter the sample if they remain in the Civil Service for several decades. If promotions positively affect health, this sample selection may attenuate the correlation between employment grade and health.[4] On the other hand, failure to

[2] Although titles such as 'Higher Executive Officer' may sound impressive, they in fact refer to relatively low ranking positions.

[3] Each of these three indices is created from six to ten underlying survey items. Our data set only includes the indices, not the underlying items. Examples of items include: Do you have to work very fast? Do you have to work very intensively? I have a good deal of say in decisions about work. My working time can be flexible. How often do you get help and support from your colleagues? How often can you delegate work effectively to your juniors?

[4] For example, assume that employment grade is randomly assigned and that employees leave the Civil Service – due to sickness or death – if their health index falls below $c$. If promotions improve health, then a larger share of the low grade employees will fall below $c$ and leave the Civil Service prior to the study's start date. These leavers will also tend to be the sicker individuals, so the remaining pool of low grade employees will be drawn from a healthier group of individuals than the remaining high grade employees. The estimated positive effect of promotions on health will therefore be attenuated (this is the standard result that a truncated left-hand side variable can cause attenuation bias).

receive a promotion could encourage more capable employees to leave the Civil Service if they are not promoted, possibly inducing a bias in the opposite direction.[5] Signing the direction of the overall bias resulting from the sample selection procedure is infeasible.

To explore the sampling frame issue, we perform additional analyses using only employees who joined the Civil Service in 1980 or later. These employees have an average tenure of only 3.5 years upon entry to the sample. While it is possible that some sample selection issues remain, their effects should be reduced in comparison to those in the overall sample. We therefore view the 1980+ subsample as a reasonable approximation to the ideal sampling frame and use it for a robustness check.

Table 1 reports descriptive statistics. The Table is broken into three columns. The first column reports summary statistics for the entire Whitehall II sample. The second column reports summary statistics for employees that joined the Civil Service at the lowest two grade levels and have non-missing data in the 1990 follow-up survey. Restricting the sample in this manner avoids contamination by 'Fast Stream' employees, as discussed in Section 3.1, and this sample is the primary analytic sample for the differences and IV models. The last column reports summary statistics for employees that joined the Civil Service during the 1980s and entered at the lowest two grade levels. This subsample is small, but it provides a useful robustness check because it is least likely to be affected by the sampling frame issue.

In comparing the first two columns in Table 1, most measures are fairly close, but the average grade level of the primary analytic sample is lower because it excludes workers that joined the Civil Service at high grade levels. In comparing the second and third columns, females account for 37% of the primary analytic sample but 57% of the 1980+ subsample. This discrepancy occurs because females are concentrated in lower-grade and lower-tenure positions, and the 1980+ subsample contains primarily lower-grade and lower-tenure workers. There is little difference in age between the primary analytic sample and the 1980+ subsample, however, because both are limited to employees aged 35–55. If the overall sample did not select civil servants based on age, the average age of the 1980+ subsample would be lower than that of the primary analytic sample. Because of the differences between the primary analytic sample and the 1980+ subsample, we focus on whether the 1980+ subsample generates a significant result rather than making a direct comparison between coefficient magnitudes in the two samples.

## 2. Theoretical Framework and Empirical Specifications

We develop a simple theoretical model to understand some of the mechanisms through which promotions may affect heart disease. The basis for this model comes from Deaton and Paxson (2001), who present a model in which health is an increasing

[5] Suppose that promotions are randomly assigned and that good employees, whose opportunity cost of employment is greater, leave the Civil Service if they do not receive a promotion within the first five years. Poor employees, in contrast, stay regardless of whether they receive a promotion. The observed pool of promoted employees will therefore consist of a mix of good and poor employees, whereas the observed pool of non-promoted employees will consist largely of poor employees. If employee quality is positively correlated with health, the estimates could overstate the impact of promotions on health.

Table 1

*Summary Statistics*

| Variable | Full sample | Analytic sample | 1980+ Subsample |
|---|---|---|---|
| Grade level (1–6) | 3.23 | 1.81 | 1.30 |
| | (1.69) | (1.52) | (0.63) |
| Female | 0.331 | 0.372 | 0.569 |
| | (0.471) | (0.483) | (0.496) |
| Age | 44.4 | 44.3 | 44.5 |
| | (6.1) | (6.1) | (6.2) |
| Tenure | 17.6 | 18.4 | 3.5 |
| | (8.5) | (8.6) | (2.2) |
| College | 0.469 | 0.387 | 0.354 |
| | (0.499) | (0.487) | (0.479) |
| CHD in 1985 | 0.041 | 0.044 | 0.037 |
| | (0.198) | (0.204) | (0.189) |
| CHD in 1999 | 0.118 | 0.132 | 0.119 |
| | (0.322) | (0.338) | (0.327) |
| CHD in 1985 or 1999 | 0.133 | 0.148 | 0.133 |
| | (0.340) | (0.355) | (0.339) |
| Sample size | 10,308 | 4,677 | 649 |

*Notes.* Parentheses contain standard deviations. All variables are measured at baseline (1985) except CHD in 1999 and Any CHD. Analytic sample contains employees that joined the Civil Service at the first two grade levels and have non-missing data in the 1990 follow-up sample. 1980+ subsample contains employees who joined the Civil Service in 1980 or later at the first two grade levels. Most statistics in the first column contain <10,308 observations because some values are missing.

linear function of relative income (i.e. income relative to some unobserved reference point). A key implication of this model is that observed income is a noisy measure of relative income, so regressing health on observed income generates an attenuated estimate of relative income's effect on health.

We enrich this model along several dimensions in a micro panel data setting. First, we model health as a function of permanent income rather than current income; this change highlights the role that expectations play. Second, we allow absolute income to have an effect independent of relative income. Finally, we examine different ways in which individuals may form their reference points, $r_i$. To mirror our estimating equations, we change the outcome of interest from health ($h$) to sickness ($s$).

The basic model takes the form

$$\mathrm{E}(s_{it}|\alpha_i, \gamma_t, y_{it}^{\mathrm{p}}, r_{it}) = c + \alpha_i\gamma_t + \beta_1 y_{it}^{\mathrm{p}} + \beta_2(y_{it}^{\mathrm{p}} - r_{it}). \tag{1}$$

$$y_{it}^{\mathrm{p}} = y_{it} + \sum_{j=1}^{\infty} \delta^j \mathrm{E}(y_{it+j}), \quad r_{it} = r_i. \tag{2}$$

The coefficient $\beta_1$ captures the direct effect of permanent income on sickness, whereas the coefficient $\beta_2$ captures the effect of relative income on sickness. Both coefficients are presumed to be weakly negative. Introducing these two channels separately is useful in a panel data setting because the channels have different implications for different estimators. The reference point, $r_i$, is assumed constant over time (we relax this assumption later). The individual effect, $\alpha_i$, represents the stock of sickness with which an individual enters the sample; this stock may be a product of both genetics and environment. To parallel Grossman (1972), we assume an individual's stock of sickness grows over time (i.e. the stock of health decays over time). The time effect, $\gamma_t$, therefore increases over time: $\gamma_{t+1} \geq \gamma_t \geq 0$.

We cannot directly estimate the conditional expectation above for several reasons. First, both permanent income, $y^{\mathrm{P}}$, and the reference point, $r$, are unobserved. Instead, we observe employee grade level, which is effectively a measure of current income, $y_{it}$. Furthermore, the unobserved individual stock of sickness, $\alpha_i$, is likely correlated with current grade level. Because the stock grows over time ($\gamma_t$ is not constant), differences estimators should reduce but not necessarily eliminate the selection bias arising from the unobserved $\alpha_i$.

In the context of our model, changes in employee grade level can affect health through two channels. First, they increase the financial resources available for health-related investments by increasing permanent income – this channel is represented by $\beta_1$ in (1). Second, they increase an individual's relative income – this channel is represented by $\beta_2$. Relative income may matter in the context of promotions either because higher grade positions are inherently more prestigious (regardless of salary) or because employees care about income relative to some reference point. After normalising $y_{it}$ so that a 1-unit increase represents the income gain associated with a promotion, the expected effect of an unanticipated permanent promotion on sickness is $\beta_1 + \beta_2$.

Now consider a cross-sectional regression of $s_{it}$ (sickness) on $y_{it}$ (current grade level) using observations in period $t$. Under the model presented in (1), the probability limit of the coefficient on $y_{it}$ is:

$$\mathrm{plim}(\hat{\beta}_{\mathrm{ols}}) = (\beta_1 + \beta_2)\frac{\sigma_{y_t y_t^{\mathrm{P}}}}{\sigma^2_{y_t}} + \frac{\gamma_t \sigma_{y_t \alpha}}{\sigma^2_{y_t}} - \beta_2 \frac{\sigma_{y_t r}}{\sigma^2_{y_t}}, \tag{3}$$

where $\sigma_{y_t y_t^{\mathrm{P}}}$ is the covariance between current grade and permanent grade, $\sigma_{y_t}\alpha$ is the covariance between current grade and the unobserved individual effect, $\sigma_{y_t}r$ is the covariance between current grade and the unobserved reference point and $\sigma^2_{y_t}$ is the variance of current grade (see online Appendix A.3 for derivation). Plim$(\hat{\beta}_{\mathrm{ols}})$ may diverge from $\beta_1 + \beta_2$ for several reasons. The first term on the right side of (3) is attenuated relative to $\beta_1 + \beta_2$ because current grade is an imperfect measure of permanent grade.[6] The third term also generates upwards bias (towards zero) as long as an individual's reference point is positively correlated with his grade level. The second term, however, generates downwards bias (away from zero) if sicker individuals concentrate in lower grades. In that case, the covariance between grade level and the individual effect ($\alpha_i$) is negative. Net plim$(\hat{\beta}_{,\mathrm{ols}})$ could thus be larger or smaller than $\beta_1 + \beta_2$.

Next, consider a specification that relates changes in sickness to changes in grade. Under the model presented in (1), when regressing $\Delta s_{it}$ on $\Delta y_{it}$, the probability limit of the coefficient on $\Delta y_{it}$ is:

$$\mathrm{plim}(\hat{\beta}_{\mathrm{diff}}) = (\beta_1 + \beta_2)\frac{\sigma_{\Delta y_t \Delta y_t^{\mathrm{P}}}}{\sigma^2_{\Delta y_t}} + \frac{\Delta\gamma_t \sigma_{\Delta y_t \alpha}}{\sigma^2_{\Delta y_t}} - \beta_2 \frac{\sigma_{\Delta y_t \Delta r}}{\sigma^2_{\Delta y_t}}, \tag{4}$$

[6] In terms of the quantities in (3), the covariance between current and permanent grade is less than the variance of current grade.

where $\sigma_{\Delta y_t \Delta y_t^{\mathrm{p}}}$ is the covariance between changes in current grade and changes in permanent grade, $\sigma_{\Delta y_t}\alpha$ is the (cross-sectional) covariance between changes in current grade and the unobserved individual effect, $\sigma_{\Delta y_t}\Delta r$ is the covariance between changes in current grade and the unobserved reference point, and $\sigma^2_{\Delta y_t}$ is the variance of changes in current grade (see online Appendix A.4 for derivation).

How do $\mathrm{plim}(\hat{\beta}_{\mathrm{diff}})$ and $\mathrm{plim}(\hat{\beta}_{\mathrm{ols}})$ compare? We make two assumptions to conduct this comparison. First, based upon standard results in labour economics, we assume that differencing exacerbates measurement error in permanent grade level.[7] Second, we assume that an individual's reference point, $r_{it}$, responds 'slowly' to changes in grade (we discuss what this assumption implies in practical terms in footnote 8). Under these assumptions, the first term in $\mathrm{plim}(\hat{\beta}_{\mathrm{diff}})$ – the measurement error term – is more attenuated than its counterpart in $\mathrm{plim}(\hat{\beta}_{\mathrm{ols}})$. The attenuation is particularly strong if promotions are well anticipated; in the extreme case in which changes in $y_{it}$ are perfectly anticipated, there is no correlation between $\Delta y_{it}$ and $\Delta y_{it}^{\mathrm{p}}$. The second term in $\mathrm{plim}(\hat{\beta}_{\mathrm{diff}})$ is closer to zero than its counterpart in $\mathrm{plim}(\hat{\beta}_{\mathrm{ols}})$ if differencing reduces the degree of selection bias. The third term in $\mathrm{plim}(\hat{\beta}_{\mathrm{diff}})$ is a positive bias term. It is smaller (i.e. generates *less* bias towards zero) than its counterpart in $\mathrm{plim}(\hat{\beta}_{\mathrm{ols}})$ when $r_{it}$ responds slowly to changes in grade. Under our initial assumption that $r_{it}$ is fixed over time, the third term becomes zero – more generally, it should be smaller than its counterpart in (3) unless $r_{it}$ responds very quickly.[8] Overall, the first two terms (measurement error in permanent grade and selection bias) attenuate $\hat{\beta}_{\mathrm{diff}}$ relative to $\hat{\beta}_{\mathrm{ols}}$, whereas the third term (reference point bias) increases the magnitude of $\hat{\beta}_{\mathrm{diff}}$ relative to $\hat{\beta}_{\mathrm{ols}}$.

Finally consider an IV estimator using an instrument $z_i$ (see online Appendix A.5 for derivation). The probability limit of $\hat{\beta}_{\mathrm{iv}}$ is:

$$\mathrm{plim}(\hat{\beta}_{\mathrm{iv}}) = (\beta_1 + \beta_2)\frac{\sigma_{z\Delta y_t^{\mathrm{p}}}}{\sigma_{z\Delta y_t}} - \beta_2 \frac{\sigma_{z\Delta r}}{\sigma_{z\Delta y_t}}. \tag{5}$$

How do $\mathrm{plim}(\hat{\beta}_{\mathrm{iv}})$ and $\mathrm{plim}(\hat{\beta}_{\mathrm{diff}})$ compare? The first term in $\mathrm{plim}(\hat{\beta}_{\mathrm{iv}})$ is less attenuated than its counterpart in $\mathrm{plim}(\hat{\beta}_{\mathrm{diff}})$ if promotions generated by an exogenous instrument are less anticipated than endogenous promotions.[9] The second term in $\mathrm{plim}(\hat{\beta}_{\mathrm{iv}})$ is either larger or smaller than its counterpart in $\mathrm{plim}(\hat{\beta}_{\mathrm{diff}})$ depending on the instrument and the specification of the reference point $r_{it}$. Our instruments are department and cohort level promotion rates. If reference points are formed based on lagged grade level or the average grade level of colleagues of similar skill, then the second term is larger in our IV estimator than in the differences estimator. If reference

[7] Formally, we assume that the correlation between changes in current income (grade) and changes in permanent income (grade) is generally lower than the correlation between current income (grade) and permanent income (grade) (Bound and Krueger, 1991).

[8] Suppose that reference points adjust quickly, so that $r_{it} = y_{it-1}$. The reference point term will be larger in $\hat{\beta}_{\mathrm{ols}}$ than $\hat{\beta}_{\mathrm{diff}}$ if $\sigma_{y_t y_t-1}/\sigma^2_{y_t} > \sigma_{\Delta y_t \Delta y_{t-1}}/\sigma^2_{\Delta y_t} = (\sigma_{y_t y_{t-1}} + \sigma_{y_{t-1} y_{t-2}} - \sigma^2_{y_{t-1}} - \sigma_{y_t y_{t-2}})/\sigma^2_{\Delta y_t}$. Using approximate values from the data, we estimate that the left side of the inequality is several times larger than the right side of the inequality, even when adjustment takes only one period. Only when adjustment is instantaneous, i.e. $r_{it} = y_{it}$, are the two sides of the inequality equal.

[9] If exogenously generated promotions are less likely to be anticipated, then they are more likely to represent actual increases in permanent grade (the variable that is being measured with error). This case is a specific example of the standard result that IV estimators can eliminate measurement error bias.

points are formed based on the average grade level of an employee's department or cohort, then the second term is smaller in our IV estimator than in the differences estimator.[10] The selection term containing $\alpha_i$ disappears as the instrument is presumed uncorrelated with selection factors. Overall, the first term probably increases the magnitude of $\hat{\beta}_{\text{iv}}$ relative to $\hat{\beta}_{\text{diff}}$, the second term has an ambiguous effect, and the absence of a selection term reduces the magnitude of $\hat{\beta}_{\text{iv}}$ relative to $\hat{\beta}_{\text{diff}}$.

The theoretical framework generates ambiguous predictions but highlights several interesting possibilities. First, cross-sectional and panel estimates of the effect of promotions on heart disease need not be biased away from zero. Attenuation bias can arise due to measurement error in permanent grade level and reference points. Second, panel estimates need not be smaller in magnitude than cross-sectional estimates. Although greater measurement error and reduced selection bias may reduce the magnitude of panel estimates, reduced bias from omitted reference points can increase the magnitude of panel estimates. Third, the relationship between IV estimates and panel or cross-sectional estimates can depend heavily on how individuals form expectations, both in relation to permanent grade level and reference points. Finally, the ordering of the estimates becomes somewhat clearer if relative position does not matter (i.e. $\beta_2 = 0$). In that case, panel estimates are likely smaller than cross-sectional estimates, but the relationship between IV estimates and panel estimates remains ambiguous.[11]

Ideally we could estimate both $\beta_1$ and $\beta_2$ consistently. However, this is not feasible since we only have variation in promotion rates and cannot observe reference points. Furthermore, even exogenous variation in permanent grade level need not estimate $\beta_1 + \beta_2$ consistently since reference points may adjust in response to promotions. We thus expect our IV coefficients to be attenuated relative to $\beta_1 + \beta_2$, even if they consistently estimate the effect of a promotion. Note, however, that the combined effect of $\beta_1 + \beta_2$, while interesting from a theoretical standpoint, is not an effect that we would observe in real data, as it corresponds to manipulating an individual's permanent grade level while forcing her reference point to remain unchanged. The 'reduced form' effect of a promotion – i.e. the effect that includes any adjustments to the reference point – should be smaller than $\beta_1 + \beta_2$. As we discuss in Section 6, we interpret the panel and IV estimates as plausible lower and upper bounds (in magnitude), respectively, on the 'reduced form' effect of a promotion on health.

[10] Formally, if $r_{it}$ is fixed over time then this term is zero in both the differences and IV cases (see online Appendix A.6 for derivation). If $r_{it} = y_{it-1}$, then this term should be larger for the differences estimator than for IV (see online Appendix A.7 for derivation). If $r_{it} = \bar{y}_{cdt}$, where $\bar{y}_{cdt}$ is the average grade level in an employee's department-by-cohort, then this term should be smaller for the differences estimator than for IV since our instrument leverages department-by-cohort variation in promotions (see online Appendix A.8 for derivation). Finally, if $r_{it} = \bar{y}_{kt}$, where $\bar{y}_{kt}$ is the average grade level of employees with skill level $k$, then this term should be larger for the differences estimator than for IV because the instrument compares individuals of equivalent skill level (see online Appendix A.9 for derivation) – see Section 5 for further discussion of this case.

[11] Panel estimates are likely smaller than cross-sectional estimates because the only term that increases the magnitude of the panel estimates relative to the cross-sectional estimates is the reference point term. The relationship between IV estimates and panel estimates remains ambiguous because IV estimates may have less measurement error (increasing their magnitude) and less selection bias (decreasing their magnitude).

## 3. Results

### 3.1. *Cross-sectional Results*

The primary health outcome in the Whitehall studies is CHD. Cross-sectional OLS results for the entire Whitehall II sample reveal that employment grade is strongly correlated with the presence of CHD. Table 2 presents results for the regression:

$$\text{CHD}_{id} = \beta\text{Grade}_{id} + \mathbf{X}_{id}\boldsymbol{\delta} + \epsilon_{id}. \tag{6}$$

The dependent variable is the presence of CHD, $\text{Grade}_{id}$ is the worker's employment grade at sample entry, and $\mathbf{X}_{id}$ is a set of controls. Subscript $i$ refers to an individual, and subscript $d$ refers to a department. Table 2 reports coefficients for the presence of any CHD (CHD that was present upon entering the sample or that occurred between 1985 and 1999) regressed upon grade level. Column (1) includes no controls, whereas column (2) controls for gender, quadratics in age and tenure, year of entry into the Civil Service and college education (this is the preferred specification).[12] In both regressions, an increase of one grade level is associated with a statistically significant reduction of approximately one percentage point in the prevalence of heart disease. The sample size changes from column (1) to column (2) because some covariates are missing values for some observations, but estimating the model in column (1) using the sample in column (2) generates similar results.

Column (3) estimates the same regression as in column (1), but only for individuals that have non-missing data in the 1990 follow-up survey. An increase of one grade level is associated with a statistically significant reduction of 1.2 percentage points in the prevalence of heart disease. Column (4) further restricts the sample to include only employees that entered the Civil Service at Grades 1 or 2 (Clerical/Support Staff and

Table 2

*OLS Regressions of Coronary Heart Disease (CHD) on Employment Grade*

| | Dependent variable: any CHD | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Grade level | −0.011 | −0.010 | −0.012 | −0.008 | −0.004 | −0.031 |
| | (0.002) | (0.002) | (0.004) | (0.004) | (0.004) | (0.014) |
| Sample restrictions | None | None | Non-missing in 1990 | Non-missing in 1990 Enter Grades 1–2 | Non-missing in 1990 Enter Grades 1–2 | Enter 1980+ Enter Grades 1–2 |
| Covariates | No | Yes | Yes | No | Yes | Yes |
| $R^2$ | 0.003 | 0.021 | 0.021 | 0.001 | 0.022 | 0.027 |
| *N* | 10,307 | 7,541 | 5,966 | 4,677 | 4,677 | 649 |

*Notes.* Models with covariates control for gender, quadratics in age and tenure, year of entry to Civil Service and college attendance. 'Non-missing in 1990' denotes that the sample is limited to employees that have non-missing data in the 1990 follow-up sample. 'Enter Grades 1–2' denotes that the sample is limited to employees that joined the Civil Service at the first two grade levels. 'Enter 1980+' denotes that the sample is limited to employees who joined the Civil Service in 1980 or later. Parentheses contain standard errors clustered at the department level.

[12] Tenure is not perfectly colinear with age and year of entry into the Civil Service because employees are surveyed at different points throughout a two to three-year window.

Executive Officers); this sample is the primary analytic sample. This restriction eliminates a group of employees known as 'Fast Stream' employees. These employees are hired with an explicit expectation that they will quickly advance through the ranks, but they always begin at Grade 3 (Higher Executive Officer) or above (Price, 2006). In this sample, an increase of one grade level (without controlling for covariates) is associated with a marginally significant reduction of 0.8 percentage points in the prevalence of heart disease. Controlling for covariates in this sample, presented in column (5), reduces the coefficient to −0.4 percentage points and eliminates any statistical significance. This result implies that among employees joining the Civil Service at the lowest grade levels, there is little cross-sectional relationship between current grade level and prevalence of heart disease.

The coefficients in columns (4) and (5) are similar in magnitude to previously published Whitehall results. For example, Marmot *et al.* (1997) report an average reduction in CHD incidence of 0.7 percentage points per grade level when adjusting only for age and gender (i.e. more covariates than column (4) but fewer covariates than column (5)). An equivalent specification in our data generates a coefficient of −0.5 percentage points; the small discrepancy arises because we do not observe the CHD measure at the same point as Marmot *et al.* Marmot *et al.* (1997) also find that the cross-sectional CHD gradient becomes statistically insignificant when adjusting for a broad range of covariates (although some of their covariates, such as job control, are factors that are endogenous to employment grade).

Column (6) presents results from the same specification as in column (5) but limits the sample to include only employees entering the Civil Service from 1980 onwards at the two lowest grades. Within this subsample, which approximates the idealised sampling frame in which we capture and retain all employees after Civil Service entry, an increase of one grade level is associated with a statistically significant 3.1 percentage point reduction in the prevalence of heart disease. The coefficient from the 1980+ subsample is almost eight times larger than the coefficient from the primary analytic sample, suggesting that the sampling frame issue discussed in Section 1 attenuates the observed grade level–CHD relationship in the primary sample. Such attenuation would be consistent with the attrition pattern from 1985 to 1999 – poor health is highly predictive of sample attrition. The difference between the coefficients, however, is at the margin of statistical significance.

Table 3 presents coefficients for a regression of the presence of CHD on five grade level dummies (the omitted category is the lowest grade level) and a full set of controls. This specification allows the health gradient to vary across grade levels. The first column presents results for the entire sample with complete data on covariates; the second column presents results for individuals that joined the Civil Service at the lowest two grade levels and have non-missing data in the 1990 follow-up survey (the primary analytic sample). In both columns, there appears to be a discrete jump at Grade 3 (Higher Executive Officer). However, the individual grade level dummies do not explain significantly more of the variation in CHD than the single grade level variable.[13]

[13] We find F-statistics of 0.77 and 1.77 respectively for the first and second columns when testing whether the individual grade level dummies explain more of the variation than the single grade level variable. Neither is statistically significant.

Table 3

*OLS Regressions of Coronary Heart Disease (CHD) on Grade Dummies*

| | Dependent variable: any CHD | |
|---|---|---|
| | (1) | (2) |
| Grade 6 | −0.045 | −0.017 |
| | (0.012) | (0.031) |
| Grade 5 | −0.035 | −0.028 |
| | (0.013) | (0.023) |
| Grade 4 | −0.022 | −0.004 |
| | (0.014) | (0.027) |
| Grade 3 | −0.024 | −0.038 |
| | (0.017) | (0.023) |
| Grade 2 | 0.004 | 0.001 |
| | (0.015) | (0.020) |
| Sample restrictions | None | Non-missing in 1990<br>Enter Grades 1–2 |
| $R^2$ | 0.021 | 0.024 |
| *N* | 7,541 | 4,677 |

*Notes.* All models control for gender, quadratics in age and tenure, year of entry to Civil Service and college attendance. Results in column (1) are estimated on the full sample; results in column (2) are estimated on the analytic sample (employees that joined the Civil Service at the first two grade levels and have non-missing data in the 1990 follow-up sample). Parentheses contain standard errors clustered at the department level.

We therefore parameterise grade level as a single variable rather than as multiple dummies for the remainder of the analysis.

### 3.2. *Promotions and Heart Disease*

To estimate the relationship between promotions and heart disease, we regress changes in the presence of heart disease on changes in grade level. In our data, the presence of heart disease is measured in 1985 and 1999, whereas employment grade is measured in 1985 and 1990. We thus test whether changes in employment grade from 1985 to 1990 predict changes in heart disease from 1985 to 1999. Table 4 reports coefficients for the model:

$$\text{Change in CHD}_{id} = \beta \text{ Change in Grade}_{id} + \mathbf{X}_{id}\boldsymbol{\delta} + \epsilon_{id}. \tag{7}$$

Change in $\text{CHD}_{id}$ equals 1 if the worker develops CHD after 1985 and 0 otherwise (or, in a small number of cases in which a worker loses heart disease, −1). Change in $\text{Grade}_{id}$ is the number of grade levels that an employee has been promoted (or demoted) between 1985 and 1990. Eighty-three per cent of the primary analytic sample had no change in grade level over this period, 16% had a promotion of one grade level, and 1% each had a promotion of two grade levels or a demotion of one grade level. Promotion rates from each initial grade level are as follows: 11.2% from Grade 1, 17.5% from Grade 2, 20.8% from Grade 3, 21.6% from Grade 4 and 16.8% from Grade 5. $\mathbf{X}_{id}$ is defined as before. The first column in Table 4 reports the results for a regression of changes in heart disease on changes in grade level with no covariates. The sample includes individuals that joined the Civil Service at the lowest two grade levels (i.e. that are not Fast Stream entries) and have non-missing data in the 1990 follow-up survey.

Table 4

*OLS Regressions of Changes in Coronary Heart Disease (CHD) on Promotions*

| | Dependent variable: change in CHD | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Change in grade level | −0.035 (0.014) | −0.026 (0.015) | −0.025 (0.014) | −0.025 (0.013) | −0.057 (0.014) |
| Change in job demands | | | −0.006 (0.004) | | |
| Change in social support at work | | | −0.014 (0.004) | | |
| Change in decision latitude | | | 0.002 (0.006) | | |
| Sample restrictions | Non-missing in 1990 Enter Grades 1–2 | Non-missing in 1990 Enter Grades 1–2 | Non-missing in 1990 Enter Grades 1–2 | Non-missing in 1990 | Enter 1980+ Enter Grades 1–2 |
| Covariates | No | Yes | Yes | Yes | Yes |
| $R^2$ | 0.002 | 0.009 | 0.011 | 0.010 | 0.012 |
| *N* | 4,677 | 4,677 | 4,442 | 5,966 | 649 |

*Notes*. Models with covariates control for gender, quadratics in age and tenure, year of entry to Civil Service and college attendance. Coefficients on the job demands, social support and decision latitude indices correspond to an 1-SD change in any of these indices. 'Non-missing in 1990' denotes that the sample is limited to employees that have non-missing data in the 1990 follow-up sample. 'Enter Grades 1–2' denotes that the sample is limited to employees that joined the Civil Service at the first two grade levels. 'Enter 1980+' denotes that the sample is limited to employees who joined the Civil Service in 1980 or later. Parentheses contain standard errors clustered at the department level.

A promotion of one grade level predicts a statistically significant 3.5 percentage point decrease in the probability of developing heart disease. Adding a full set of covariates (gender, age, tenure and college education) to the specification in column (2) reduces the coefficient to −2.6 percentage points but it remains marginally significant.

Column (3) presents results from the same specification as in column (2) but controls for changes in job demands, social support and decision latitude. Including these controls does not impact the coefficient on promotions. An 1-SD increase in social support at work predicts a statistically significant 1.4 percentage point decrease in the probability of developing heart disease but there is no significant relationship between heart disease and job demands or decision latitude. Column (4) presents results from the same specification as in column (2) but includes employees that entered the Civil Service at any grade level. This sample includes potential Fast Stream employees. Expanding the sample has minimal impact on our estimates – the coefficient changes to −2.5 percentage points and remains marginally significant. Column (5) presents results from a specification that regresses change in CHD on change in grade since first hire; the sample includes only employees that entered the Civil Service from 1980 onwards at the lowest two grades. Within this 1980+ subsample, which approximates the idealised sampling frame in which we capture and retain all employees upon entry to the Civil Service, a promotion of one grade level is associated with a statistically significant 5.7 percentage point reduction in the prevalence of heart disease. However, this estimate is unique to this sample – applying the new specification to the original sample or the original specification to the 1980+ subsample generates insignificant estimates. The magnitude of the coefficient is due in part to the distribution of promotions. Employees starting at Grade 1 have a promotion rate of 8.7% in this subsample, whereas employees starting at Grade 2 have a promotion rate of 38.3%, and the estimates in Table 3 suggest that promotions have the biggest impact when moving from Grades 2 to 3.

In all cases, the relationship between changes in grade level and changes in heart disease (Table 4) is larger in magnitude than the cross-sectional relationship between grade level and heart disease (Table 2). This fact suggests that grade level relative to a reference point could affect health (i.e. $\beta_2 \neq 0$ in (1)). If relative grade had no effect on health ($\beta_2 = 0$ in (1)), then we would expect $|\hat{\beta}_{\text{diff}}| < |\hat{\beta}_{\text{ols}}|$ because of increased attenuation from measurement error and reduced selection bias. Instead, the magnitude of $\hat{\beta}_{\text{diff}}$ increases relative to $\hat{\beta}_{\text{ols}}$, suggesting that relative grade may play a role and that reference points are endogenous. This possibility is consistent with the literature on subjective well-being and relative income, which concludes that relative income is an important determinant of happiness (Ferrer-i-Carbonell, 2005; Luttmer, 2005).

The differences estimates diverge from previously published cross-sectional Whitehall results in two respects. First, the magnitudes are larger; the −2.6 percentage point coefficient reported in column (2) of Table 4 is almost four times the analogous estimate from Marmot *et al.* (1997). Second, unlike in Marmot *et al.* (1997), controlling for job demands, social support at work and decision latitude has only a modest impact on the grade level coefficient; the coefficient is of similar magnitude in both columns (2) and (3). This result suggests that promotions affect health through channels beyond changes in the work environment. These comparisons should nevertheless be interpreted subject to the caveats that the standard errors in Table 4 are relatively large and that the work environment variables are endogenously determined.

### 3.3. *Instrumental Variables Results*

Even in the differences model, the issue of confounding is a primary concern. The treatment of interest, promotion, is not randomly assigned and it is likely that promoted and non-promoted individuals differ in important ways that are not caused by the treatment itself (as represented by the term $\alpha_i$ in (1)). One possibility that has received attention in the literature is 'health selection', or the possibility that the causal channels run from health to employment grade because healthy people are more likely to be selected for higher grade positions (Marmot and Davey Smith, 1997; Adda *et al.*, 2003; Chandola *et al.*, 2003). However, many other possibilities exist. Prior to treatment, promoted individuals may differ from non-promoted individuals in terms of family background, psychological disposition or living environment. All of these factors could affect both the initial level and the subsequent trajectory of CHD, confounding interpretation of the results. In addition, measurement error in promotions may attenuate the estimated coefficient, as CHD is likely a function of the entire history of promotions – past, present and future – rather than employment grade at a single point in time. Even future differences in grade level could affect current health if individuals have rational expectations.

To address these concerns, we estimate an IV model that leverages exogenous variation in promotion rates. To form an instrument, we exploit a plausibly exogenous source of variation in employment grade: promotion rates across major Civil Service departments. Examples of major departments include the Department of Trade and Industry, the Department for Transport and the Home Office.[14] Since the Northcote-Trevelyan report in 1854, promotion within Civil Service departments has officially been on the basis of merit (Stanley, 2004). However, the distribution of promotion opportunities across departments has little relationship to merit[15] and it is difficult for employees to transfer between departments after entering the Civil Service.[16] The *Civil Service Statistics* state that 'vacancies [within departments] arise through retirements, resignations, promotions to yet higher grades, or through the creation of new posts, offset by any posts that have been lost…There are marked differences between individual departments due to variations in relative grade sizes and in levels of wastage [worker departures]' (HM Treasury Office, 1985). Furthermore, much of the variation in promotion rates during the period in question arises from the large expansion of hiring that occurred during World War II. This expansion had a differential effect on departments and caused a wave of retirements that occurred from the late 1970s through the late 1980s (HM Treasury Office, 1989). Figure 1, reproduced from *Civil Service Statistics 1989*, demonstrates the substantial change in the age distribution that

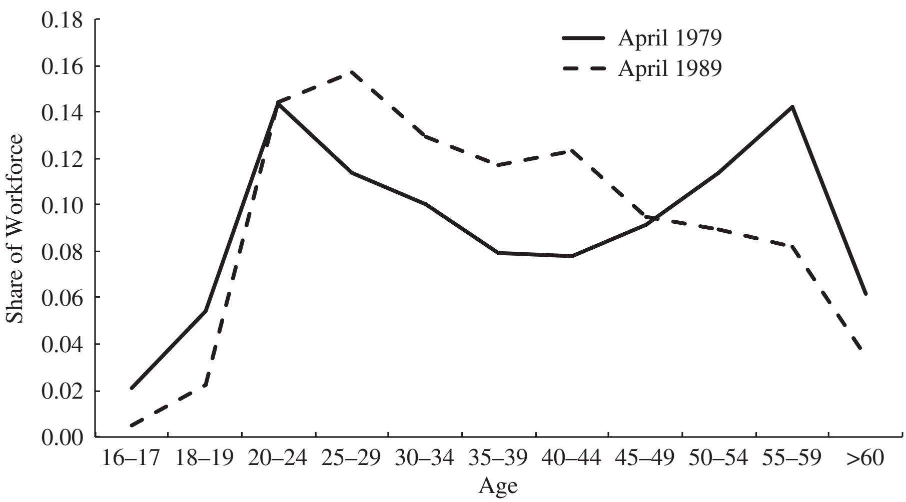


Fig. 1. *Changes in the Civil Service Age Distribution From 1979 to 1989*
*Note*. Reproduced from *Civil Service Statistics 1989*.

[14] The results do not seem very sensitive to the omission of various departments (see online Appendix A.1).

[15] The distinction between promotion opportunities and average grade level is important. Average departmental grade level will generally be correlated with employee quality, because departments with a greater number of high-grade positions will directly recruit higher quality employees into those slots. However, conditional on initial grade level, there is no reason to believe that departmental promotion rates are correlated with employee quality. The one exception to this rule are Fast Stream employees (see Section 3.1). We eliminate Fast Stream employees by limiting our sample to employees entering at Grades 1–2.

[16] An employee who wants to transfer departments needs approval from his or her current department managers and must pass specific recruitment exercises (Civil Service, 2011). In general, between 0% and 3% of employees transferred out of their initial department during the first five years of data collection.

occurred during this period as World War II cohorts reached the mandatory retirement age of 60 (Mein *et al.*, 2003). The age distribution in 1979 (the solid line) has much more mass concentrated among older employees than the age distribution in 1989 (the dashed line).

Ideally, we would use the pattern of wastage within departments and cohorts in London as the instrument for departmental promotion rates. However, data of this detail are not available from the Civil Service Statistics Office. Instead, we use the observed departmental promotion rate as an instrument for individual promotions. We predict individual promotions by using the average observed departmental promotion rate for employees between 1985 (the start of the study) and 1990 (the follow-up survey). In our primary specifications, we allow this rate to vary by five-year cohorts because promotion opportunities across departments should affect different cohorts differentially (conceptually similar instruments have been used in other contexts, such as von Wachter and Bender (2006)). In particular, an increase in the departmental promotion opportunities should have the strongest effect on the cohort that is situated directly 'below' the preponderance of the opening positions.[17] Our five-year cohorts begin in 1945 and run through to 1985; there are thus eight cohorts spaced at five-year intervals. With 18 departments and eight cohorts, we could have up to 143 department-by-cohort indicators (including department and cohort main effects). In practice, we have 114 department-by-cohort indicators as a few cohorts do not appear in all departments.

Figure 2 visually summarises our identification strategy and core results. It plots the average change in heart disease against the average promotion rate for each department-by-cohort. Department–cohorts with high promotion rates experience smaller increases in heart disease rates relative to department–cohorts with low promotion rates. To implement the identification strategy in a regression framework, we use a set of department dummies interacted with cohort dummies as our instruments. Since we have multiple instruments, we estimate coefficients using two-stage least squares (2SLS). As in the previous Section, the dependent variable is an individual's change in heart disease over the sample period. The first-stage regression is:

$$\text{Change in Grade}_{idc} = \textbf{Dept-Cohort}_{dc}\gamma + \mathbf{X}_{idc}\boldsymbol{\theta} + v_{idc}. \tag{8}$$

Variables are as previously defined, except for $\mathbf{Dept}_{dc}$, which represents a set of department-by-cohort dummies (the $c$ subscript denotes cohort). First-stage results indicate a highly significant relationship between promotion odds and department assignment; the F-statistic on the department–cohort dummies is 155.6 ($p = 0.000$).

[17] Subsequent cohorts should also experience additional promotions as positions are vacated by the newly promoted cohort. Nevertheless, there are two reasons why the cohort situated directly below the promotion opportunities experiences the largest immediate number of promotions. First, there is a lag in filling positions, so it can take time for the promotions to 'trickle down' to subsequent cohorts. Second, some positions are filled by outside hires rather than promotions from within the Civil Service, and this fraction increases at lower grade levels. During the period in question, 31% of promotion opportunities at Grade 2 were filled by outside hires, 3% of promotion opportunities at Grade 3 were filled by outside hires and 0% of promotion opportunities at Grade 4 were filled by outside hires (HM Treasury Office, 1989). These figures imply that the number of promotions that occurs from a given number of promotion opportunities decays in subsequent cohorts. For example, suppose that 100 Grade 4 employees retire. This retirement wave might translate into 100 promotions for Grade 3 employees, 97 promotions for Grade 2 employees, and 67 promotions for Grade 1 employees.

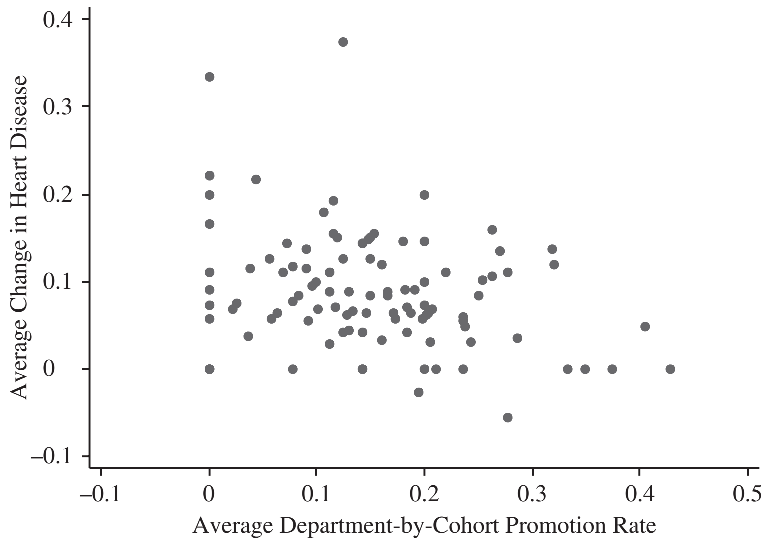


Fig 2. *Change in Heart Disease Versus Department-by-Cohort Promotion Rate*
*Notes.* This Figure plots the average change in heart disease against the average promotion rate for each department-by-cohort. For precision only department–cohorts with more than five workers are included in the figure (these department–cohorts account for over 99% of the sample).

Of primary concern is that the quality of employees may vary across departments or across cohorts within a department. Even if there is selection at the department level, cross-department variation in promotion rates can still be a valid source of identification as long as the selection is not systematically correlated with the future openings of departmental vacancies. If department managers construct some cohorts to be more skilled and other cohorts to be less skilled within the same department, however, then the high-skill cohort may advance more quickly than the low-skill cohort. We test for this possibility by regressing past promotion rates for a given department-by-cohort on current promotion rates for that department-by-cohort. In this regression, a one grade level increase in current department-by-cohort promotion rates (our instrument) predicts a $-0.65$ grade level increase in past promotion rates ($t = -1.2$). Thus current department-by-cohort promotion rates do not arise from systematic differences in skill across cohorts within a given department; were that the case, the more skilled cohorts would advance faster in both the previous period and the current period.

Using department-by-cohort level variation also allows us to estimate models that include both department fixed effects and cohort fixed effects. Department fixed effects can control for the possibility that some departments may have consistently high promotion rates and attract a select group of employees. Cohort fixed effects can control for the possibility that different cohorts may experience different outside opportunities due to time-varying labour market conditions. The inclusion of both sets of fixed effects does not qualitatively change our conclusions. We further analyse selection bias and other potential threats to validity in depth in Section 4. Our analysis reveals that the instrument does not appear to be correlated with other factors that could independently affect health.

We estimate the effect of promotions on changes in heart disease using the second-stage regression:

$$\text{Change in CHD}_{idc} = \beta \widehat{\text{Change in}}\ \text{Grade}_{idc} + \mathbf{X}_{idc}\boldsymbol{\delta} + \epsilon_{idc}. \quad (9)$$

Variables are as previously defined, except for $\widehat{\text{Change in}}\ \text{Grade}_{idc}$, which is the fitted value for promotions from the first stage. Table 5 reports results from 2SLS regressions. The first column includes no controls. The coefficient estimate is statistically significant and implies that a promotion reduces the prevalence of heart disease by 13.3 percentage points (t = −2.4).[18] The second column controls for gender, quadratics in age and tenure, year of entry into the Civil Service and college attendance. The coefficient changes slightly to −12.8 percentage points and remains statistically significant (t = −2.8). The third column includes all of the controls in column (2) and adds cohort fixed effects as additional controls – the identification now comes from within-cohort variation in promotion rates across departments. The coefficient changes to −11.0 percentage points but remains statistically significant (t = −2.3). The fourth column includes all of the controls in column (2) and adds department fixed effects as additional controls – the identification now comes from within-department variation in promotion rates across cohorts. The coefficient increases to −14.4 percentage points and remains statistically significant (t = −2.8). The fifth column includes all of the controls in column (2) and adds both cohort and department fixed effects as additional controls. The coefficient changes to −11.9 percentage points and remains statistically significant (t = −2.3).

The sixth column includes all of the controls in column (2) but changes the set of instruments from department-by-cohort dummies to department-by-grade level dummies. If a wave of departures leads to promotions for employees at a particular grade level, then the department-by-grade level dummies may be valid instruments for promotions.[19] Using department-by-grade level dummies as instruments generates a statistically significant coefficient of −13.8 percentage points (t = −2.4). This estimate is of similar magnitude to the estimate in column (2), suggesting that either specification is valid.

All promotion coefficients in Table 5 are of similar magnitude and generally significant. The estimates for the preferred specification (the second column) imply that a promotion of one grade level reduces the probability of developing heart disease during a 15-year period by 13 percentage points. This value is over half the average rate of CHD in departments with the most heart disease and corresponds to an effect size of 0.38. The coefficient is also five times larger than the comparable differences coefficient. Nevertheless, the standard errors are large; the differences estimate lies at the edge of the confidence interval for the 2SLS estimate.

Table 6 explores the sensitivity of the 2SLS regression coefficient to different sample selection criteria. All regressions in Table 6 control for a full set of covariates (gender, quadratics in age and tenure, year of entry into the Civil Service and college

[18] As in previous specifications, we limit the sample to employees that joined the Civil Service at the lowest two grade levels and have non-missing data in the 1990 follow-up survey. We expand the estimation sample below and find that our conclusions do not qualitatively change.

[19] The department-by-grade level dummies have good predictive power; the first-stage F-statistic is 78.6.

Table 5

*2SLS Regressions of Changes in Coronary Heart Disease (CHD) on Promotions*

| | Dependent variable: change in CHD | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Change in grade level | −0.133 (0.055) | −0.128 (0.046) | −0.110 (0.047) | −0.144 (0.052) | −0.119 (0.051) | −0.138 (0.057) |
| Covariates | No | Yes | Yes | Yes | Yes | Yes |
| Fixed effects | | | Cohort | Dept | Cohort and Dept | |
| Instruments | Dept-by-Coh | Dept-by-Coh | Dept-by-Coh | Dept-by-Coh | Dept-by-Coh | Dept-by-Grade |
| *N* | 4,677 | 4,677 | 4,677 | 4,677 | 4,677 | 4,677 |

*Notes.* Models with covariates control for gender, quadratics in age and tenure, year of entry to Civil Service and college attendance. In all columns, the sample is limited to employees that joined the Civil Service at the first two grade levels and have non-missing data in the 1990 follow-up sample. Parentheses contain two-stage least squares (2SLS) standard errors clustered at the department level.

Table 6

*2SLS Regressions Across Different Subsamples*

| | Dependent variable: change in coronary heart disease | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Change in grade level | −0.128 (0.046) | −0.122 (0.057) | −0.107 (0.047) | −0.099 (0.046) | −0.179 (0.064) |
| Joined Civil Service | 1944+ | 1944+ | 1944+ | 1944+ | 1980+ |
| Grade in 1985 | Any | Any | 1–5 | 1–4 | Any |
| Civil Service entry grade | 1–2 | Any | 1–2 | 1–2 | 1–2 |
| *N* | 4,677 | 5,966 | 4,432 | 3,872 | 649 |

*Notes.* All models control for gender, quadratics in age and tenure, year of entry to Civil Service and college attendance. Parentheses contain two-stage least squares (2SLS) standard errors clustered at the department level.

education). The magnitude of the coefficient varies across samples, but the sign remains consistently negative, and all of the estimates achieve statistical significance. The first column reports results for the primary analytic subsample – it replicates column (2) of Table 5. The second column reports results from a sample that includes employees that joined the Civil Service at any grade level (i.e. it includes Fast Stream employees). The coefficient is similar in magnitude to the estimate from the primary analytic subsample (−0.122 *versus* −0.128) and remains statistically significant.

The third column of Table 6 reports results from a sample restricted to employees that joined the Civil Service at the lowest two grade levels (i.e. non-Fast Stream employees) and were below the top grade levels when sampling began (1985). The latter restriction drops employees that have no possibility of further promotion. The coefficient changes to −10.7 percentage points but remains statistically significant. The fourth column reports results from a sample restricted to employees that joined the Civil Service at the lowest two grade levels and were below the top two grade levels in 1985. The latter restriction drops employees that have a very low probability of achieving a promotion.[20] The coefficient is smaller in magnitude than the estimate from the primary analytic subsample ($-0.099$ *versus* $-0.128$) but remains statistically significant.

Column (5) of Table 6 presents results from the same specification as in column (1) but limits the sample to include only employees that entered the Civil Service from 1980 onwards at the lowest two grade levels. Within this 1980+ sample, which approximates the idealised sampling frame, the 2SLS coefficient increases in magnitude – a promotion of one grade level reduces the prevalence of heart disease by 17.9 percentage points ($t = -2.8$). The magnitude and statistical significance of this coefficient suggest that the sampling frame issues discussed in Section 1 are not generating the significant relationships reported in Table 5.

### 3.4. *Other Health Outcomes*

We examine two other available health outcomes – self-reported health and mortality rates – to explore whether promotions have a broader impact on health beyond reducing heart disease. Self-reported health is measured on a scale of 1–5; 1 corresponds to excellent health, and 5 corresponds to poor health. We thus refer to the measure as 'self-reported ill health'. Mortality measures whether an individual is deceased by 1999.

The first column of Table 7 reports results from regressions of each health outcome on grade level in 1985 and the full set of controls. There are negative, statistically significant cross-sectional relationships between grade level and self-reported ill health and mortality. The second column of Table 7 reports results from regressing the change in each outcome (relative to the value at sample entry) on the change in grade level from 1985 to 1990 and the full set of controls. In the case of self-reported health, we control for its value at sample entry rather than differencing because the format of the question changes over time. With the exception of self-reported ill health in 1990, all of the regression coefficients are statistically insignificant. The third column reports results from the 2SLS analogue of the model in the second column with the department-by-cohort promotion rate as the instrument. There is a negative, statistically significant relationship between grade level and self-reported ill health in 1995 and 1998. There is no significant 2SLS relationship between mortality and grade level, although the standard errors are too large to generate meaningful conclusions.

Overall, there is a strong cross-sectional relationship between other health measures and grade level and a strong relationship between changes in self-reported health and the department-by-cohort promotion rate. The relationship between changes in self-reported health and the department-by-cohort promotion rate suggests that promotions impact other dimensions of health beyond heart disease; controlling for heart disease at sample baseline and in 1999 has no impact on the coefficients or standard errors in these regressions. However, the modest correlation between individual

[20] An employee in the second highest grade level can only be promoted to the top grade level. This promotion is difficult to achieve – Civil Service documents describe employees falling in our top grade level to be analogous to military officers with the ranks of Brigadier General up to Field Marshal.

Table 7

*Regressions of Other Health Outcomes on Grade Level/Promotions*

| | OLS | Differences | 2SLS | *N* |
|---|---|---|---|---|
| Self-reported ill health in 1990 | −0.095 | −0.050 | −0.147 | 4,652 |
| | (0.018) | (0.023) | (0.188) | |
| Self-reported ill health in 1992 | −0.076 | 0.035 | −0.401 | 4,148 |
| | (0.009) | (0.024) | (0.220) | |
| Self-reported ill health in 1995 | −0.114 | 0.002 | −0.277 | 3,981 |
| | (0.014) | (0.015) | (0.115) | |
| Self-reported ill health in 1998 | −0.090 | 0.000 | −0.327 | 3,571 |
| | (0.010) | (0.025) | (0.120) | |
| Deceased by 1999 | −0.005 | −0.006 | −0.039 | 4,671 |
| | (0.002) | (0.006) | (0.030) | |

*Notes*. All models control for gender, quadratics in age and tenure, year of entry to Civil Service and college attendance. Sample size decreases in self-reported health regressions over time due to sample attrition. Parentheses contain standard errors clustered at the department level.

promotions and changes in self-reported health suggests that measurement error and expectations may play a role in the relationship between grade level and health. We discuss these issues in depth in Section 5.

## 4. Potential Threats to Validity

Several potential objections exist to interpreting the 2SLS estimates as causal effects of promotions on CHD. All of these issues focus on the possibility of a positive correlation between department-by-cohort promotion rates and employee quality. First, employees in different departments or cohorts may select (or be selected) into departments based on future expectations of department promotion rates. Second, attrition from the Civil Service after sampling begins may be non-random. Third, it is possible that departments directly affect health in ways other than through promotions. Finally, the fact that managers have discretion to fill promotion opportunities through outside hires raises concerns about finite sample bias (particularly given the large number of instruments employed). We analyse all of these issues and conclude that the results do not appear to be driven by any of them. We present the analyses of the first two issues in this Section and the analyses of the last two issues in online Appendices A.1 and A.2.

### 4.1. *Selection Prior to Sample Entry*

If individuals in different cohorts select into their departments based upon expectations of future promotion odds, a significant relationship between employee quality and department-by-cohort promotion rates may arise. The results in Table 5 suggest that selection into departments with high promotion rates is not driving the 2SLS coefficient because the inclusion of department fixed effects does not decrease the coefficient's magnitude. However, department fixed effects are insufficient to control for selection at the department-by-cohort level. For example, a department with an impending mid-level retirement wave might attract a cohort of recruits that anticipates a large number of promotion opportunities. To test whether selection is contaminating the instrument, we examine a variety of observable characteristics that are correlated with health but should not be affected by promotions. The results indicate that there is no systematic relationship between these characteristics and department promotion rates.

We examine the relationship between the average department-by-cohort promotion rate and pre-treatment health characteristics in Table 8 (in this context the term 'pre-treatment' refers to outcomes which should not be affected by grade level). If selection into department cohorts is driving our IV results, then we should observe a positive relationship between pre-treatment health and department-by-cohort promotion rates. To implement these tests, we place each pre-treatment outcome on the left-hand side of our 2SLS regression and test whether the coefficient on grade level is significant. The first column of Table 8 reports results from estimating (9), with each row substituting a different pre-treatment outcome for Change in $\text{CHD}_{idc}$. We also test whether each pre-treatment outcome is correlated with individual promotions by estimating an

Table 8

*Falsification Tests – 2SLS and OLS Regressions of Other Outcomes on Promotions*

| | 2SLS (Dept-by-cohort IVs) | OLS (individual promotions) | Dep Var mean (SD) | *N* |
|---|---|---|---|---|
| *Family history of:* | | | | |
| Diabetes | 0.016 | 0.004 | 0.100 | 4,459 |
| | (0.054) | (0.008) | (0.300) | |
| Heart disease | 0.129 | 0.007 | 0.256 | 4,511 |
| | (0.088) | (0.012) | (0.436) | |
| High blood pressure | −0.009 | −0.018 | 0.368 | 4,529 |
| | (0.151) | (0.015) | (0.482) | |
| Stroke | −0.106 | −0.012 | 0.164 | 4,470 |
| | (0.065) | (0.011) | (0.370) | |
| *Health behaviours* | | | | |
| Ever smoked | 0.066 | 0.020 | 0.508 | 4,653 |
| | (0.124) | (0.017) | (0.500) | |
| Hours of exercise per week | 0.695 | 0.252 | 11.73 | 4,462 |
| | (2.057) | (0.357) | (9.33) | |
| *'Pre-treatment' characteristics* | | | | |
| Ever heart trouble | −0.018 | 0.005 | 0.069 | 4,662 |
| | (0.044) | (0.008) | (0.253) | |
| CHD at sample entry | 0.006 | 0.007 | 0.044 | 4,677 |
| | (0.042) | (0.009) | (0.204) | |
| Height (inches) | 0.47 | 0.19 | 67.34 | 4,673 |
| | (0.38) | (0.11) | (3.78) | |
| BMI | −0.42 | −0.08 | 24.64 | 4,670 |
| | (0.81) | (0.09) | (3.52) | |
| Chronic illness | 0.034 | −0.020 | 0.319 | 4,652 |
| | (0.098) | (0.013) | (0.466) | |
| CHD risk index (1998) | −0.002 | 0.001 | 0.118 | 4,677 |
| | (0.022) | (0.004) | (0.120) | |
| CHD risk index (change in CHD) | −0.008 | −0.006 | 0.075 | 4,677 |
| | (0.026) | (0.005) | (0.104) | |

*Notes.* CHD, coronary heart disease. Each row reports results from a regression of the indicated dependent variable in a 2SLS model (first column) or in an OLS model (second column). In all cases, the reported coefficient corresponds to the coefficient on the promotion variable. All models control for gender, quadratics in age and tenure, year of entry to Civil Service and college attendance. The sample contains employees that joined the Civil Service at the first two grade levels and have non-missing data in the 1990 follow-up sample. Parentheses contain 2SLS or OLS standard errors clustered at the department level.

OLS regression of each outcome on observed promotions. The second column of Table 8 reports results from estimating (7), with each row substituting a different pre-treatment outcome for Change in $CHD_{id}$. We estimate both specifications using the primary analytic sample.

We analyse three sets of outcomes: family medical history, individual health behaviour and individual health characteristics that should not be immediately affected by promotions. The family medical history variables measure whether an employee reported that a parent experienced a given condition (e.g. high blood pressure). The conditions, reported in the first set of rows in Table 8, include diabetes, heart attacks, high blood pressure or stroke. The results demonstrate that there is no significant relationship between the instrument and parental conditions. No coefficient is statistically significant and the signs on the coefficients go in both directions. If selection into departments were driving the 2SLS results in Section 3.3, we would expect a negative relationship between promotion rates and family conditions. The results in the OLS column demonstrate that there is also no significant relationship between individual promotions and parental conditions – the coefficients in this column are small, precisely estimated and statistically insignificant.

The second set of rows in Table 8 reports results for health behaviour. Such behaviour includes smoking (measured as whether the employee ever smoked before entering the sample) and hours of activity (exercise) per week at sample entry. The first column demonstrates that there is no significant relationship between the instrument and smoking or exercise; the second column demonstrates that there is no significant relationship between individual promotions and smoking or exercise.

The third set of rows in Table 8 reports results for 'pre-treatment' health characteristics. These characteristics include a history of heart trouble, the presence of CHD at sample entry, height, body mass index at sample entry and the presence of a chronic illness at sample entry. The first column demonstrates that there is no significant relationship between the instrument and any pre-treatment health characteristic, and two of the five coefficients imply an adverse relationship between pre-treatment health and promotion rates. The second column indicates that individuals that are promoted are taller than individuals that are not promoted. This result is marginally significant ($t = 1.7$) and is consistent with findings in the existing literature (Case and Paxson, 2008). There is no significant relationship between individual promotions and any other pre-treatment health characteristic.

The final row in Table 8 implements a summary test that combines all 12 outcomes into a single measure. Testing a summary index can increase statistical power and correct for multiple inference (Anderson, 2008). The first CHD summary risk index contains the fitted values from a regression of CHD status in 1998 on all 12 risk measures (family medical history, health behaviours and pre-treatment health characteristics). This index summarises all of the risk measures, weighting them in relation to their correlation with CHD, and provides a more powerful test than examining each risk measure individually. If there is any systematic relationship between pre-treatment outcomes and the instrument (or promotions, in the case of the second column), it is more likely to be detected in the summary index. We construct a second version of this index in the same manner but substitute the change in CHD status for CHD status in 1998.

The 1998 CHD summary risk index coefficient for the 2SLS specification implies that a one grade level increase in the department-by-cohort promotion rate decreases (i.e. makes better) the risk index by a statistically insignificant 0.2 percentage points (t = −0.1). Although the standard error is large (2.2 percentage points), a 95% confidence interval has a lower bound of −4.5 percentage points. It thus does not cover the coefficient estimate of −12.8 from the analogous 2SLS specification in Table 5. The summary risk index coefficient on promotions (the second column) indicates that a promotion is associated with a statistically insignificant 0.2 percentage point increase in the CHD risk index (t = 0.4). A 95% confidence interval has a lower bound of −0.8 percentage points and does not cover the coefficient estimate of −2.6 from the analogous differences specification in Table 4. The change in CHD summary risk index coefficients is also statistically insignificant. The lower bounds on 95% confidence intervals are −6.0 percentage points and −1.6 percentage points, respectively, for the 2SLS and OLS specifications.

To compare the magnitude of the individual 'pre-treatment' coefficient estimates and the estimated treatment effects, we normalise all outcomes by their standard deviations. Figures 3 and 4 plot the distribution of the estimated pre-treatment effect sizes (light grey) and the estimated treatment effect size (dark grey). Figure 3 plots the distribution of the 2SLS coefficients in the first column of Table 8, after these coefficients have been normalised by the standard deviations of the respective dependent variables. This distribution (light grey) is compared against the coefficient from the CHD 2SLS regression, after it is normalised by the standard deviation of CHD (dark grey). Figure 4 plots the distribution of the OLS coefficients in the second column of Table 8, after these coefficients have been normalised by the standard deviations of the respective dependent variables. This distribution (light grey) is compared against the coefficient from the CHD regression in column (2) of Table 4, after it is normalised by the standard deviation of CHD (dark grey).

Figure 3 reveals that the relationship between instrumented promotions and heart disease is stronger than the relationship between instrumented promotions and any of the 11 pre-treatment outcomes. This suggests that selection bias alone is unlikely to explain the 2SLS coefficient; at a minimum, the selection bias for the CHD outcome needs to be stronger than the selection bias for any of the pre-treatment outcomes.[21] Figure 4 reveals that the relationship between observed promotions and heart disease is stronger than the relationship between observed promotions and any of the eleven pre-treatment outcomes. This suggests that selection bias is not even the primary determinant of the differences coefficient.

Overall, there is little evidence of selection into departments or cohorts. There is no significant relationship between any pre-treatment health characteristic and the instrument, and the summary risk index is uncorrelated with the instrument as well. Including the pre-treatment health characteristics as control variables in the main 2SLS regression has little impact on the grade level coefficient; the estimated effect of promotions on heart disease changes from −12.8 percentage points to −12.6 percentage

[21] One clear difference between the CHD outcome and the pre-treatment outcomes is that the heart disease outcome is specified in changes (i.e. new cases of heart disease), whereas the pre-treatment outcomes are specified in raw levels. However, specifying the CHD outcome in levels (measured at the end of the sampling period) does not change the magnitude of the standardised 2SLS coefficient.

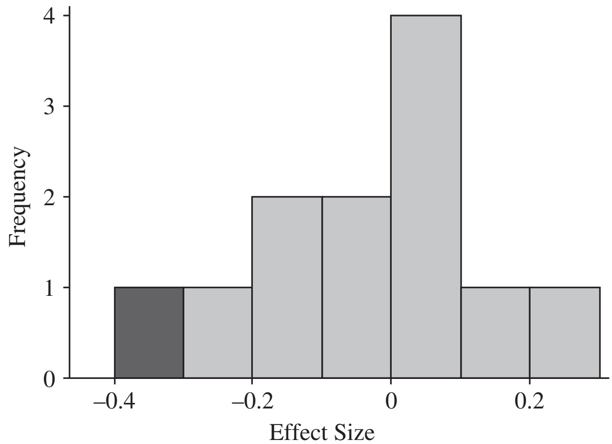


Fig. 3. *Two-stage Least Squares (2SLS) Effect Sizes for Pre-treatment Outcomes*
*Notes*. This Figure plots the distribution of effect sizes for pre-treatment outcomes (light grey) and coronary heart disease (dark grey). Each effect size is computed by dividing an outcome (actual or pre-treatment) by its standard deviation and then using the standardised outcome as the dependent variable in a 2SLS regression in which department-by-cohort promotion rates serve as instruments for observed promotions.

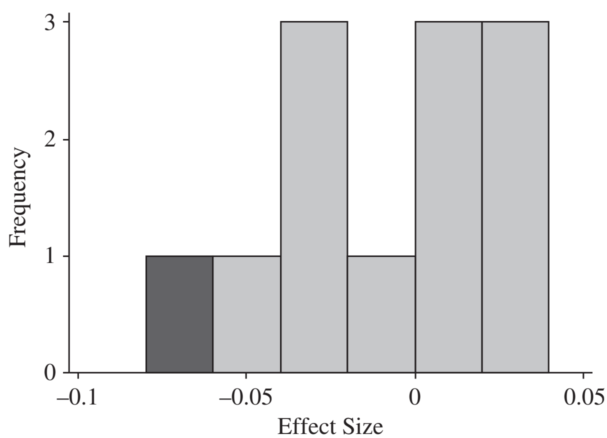


Fig. 4. *OLS Effect Sizes for Pre-treatment Outcomes*
*Notes*. This Figure plots the distribution of effect sizes for pre-treatment outcomes (light grey) and coronary heart disease (dark grey). Each effect size is computed by dividing an outcome (actual or pre-treatment) by its standard deviation and then using the standardised outcome as the dependent variable in a regression on observed promotions.

points ($t = -3.0$). In comparison, there is limited evidence of selection into promotions at the individual level. Height appears to be correlated with promotions, and controlling for pre-treatment health characteristics in the main differences regression reduces the coefficient on promotions from $-2.6$ percentage points to $-2.2$ percentage points ($t = -1.6$). Nevertheless, there is no significant relationship between the summary risk index and individual promotions, suggesting that selection bias may be modest even in the differences specifications. These conclusions are unchanged if we control for department and cohort fixed effects in all specifications.

### 4.2. *Selective Attrition After Sample Entry*

Two types of attrition after sample entry can affect our estimates. First, there is attrition in the explanatory variable – 22% of employees in our sample leave the Civil Service between sample entry and the 1990 follow-up survey. For these employees, we do not know their grade level in 1990. Second, there is partial attrition in the dependent variable – 24% of individuals in our sample do not attend their 1998 medical screenings. For these individuals, the measurement of heart disease is less accurate than for individuals who attend their 1998 screenings. Both types of attrition could cause bias in either direction.

To test whether attrition in the promotion variable can affect our estimates, we estimate a reduced-form version of our main 2SLS models without dropping observations with missing promotion data. These estimates are feasible because we have complete outcome data for employees that leave the Civil Service (up to the medical screening issue mentioned above). We first estimate the relationship between CHD and department-by-cohort promotion rates using observations with complete promotion data. We then assign employees that leave the Civil Service the average promotion rate of their department-by-cohort. Finally, we regress CHD on the average department-by-cohort promotion rate and covariates.

Table 9 presents results for models that include employees with missing promotion data. For comparison purposes, the first column reports the coefficient from a reduced form regression of changes in heart disease on department-by-cohort promotion rates estimated on the primary analytic sample (−0.124). The second column reports the coefficient from the same specification run on a sample that includes employees with missing promotion data. Each employee, including those with missing promotion data, is assigned the average promotion rate of his or her department-by-cohort. Including employees with missing promotion data changes the coefficient from −0.124 to

Table 9

*Reduced Form Regressions of Changes in Coronary Heart Disease on Promotion Rates While Including Workers that Leave Civil Service*

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Dept-by-cohort promotion rate | −0.124 (0.045) | −0.101 (0.035) | −0.079 (0.038) | −0.128 (0.042) | −0.103 (0.047) |
| Sample restrictions | Non-missing in 1990 | | | | |
| Fixed effects | None | None | Cohort | Dept | Cohort and Dept |
| *N* | 4,677 | 5,981 | 5,981 | 5,981 | 5,981 |

*Notes.* All models regress the change in heart disease on the average department-by-cohort promotion rate and control for gender, quadratics in age and tenure, year of entry to Civil Service and college attendance. 'Non-missing in 1990' denotes that the sample is limited to employees that have non-missing data in the 1990 follow-up sample. When employees with missing 1990 data are included in the sample, their promotion rates are imputed as the average promotion rate for their department-by-cohort. All samples are limited to employees that joined the Civil Service at the first two grade levels. Parentheses contain standard errors clustered at the department level.

−0.101, but the estimate remains statistically significant (t = −2.9). The third column replicates the second column but adds cohort fixed effects. The coefficient changes to −0.079 but remains marginally significant (t = −2.1). The fourth column replicates the second column but adds department fixed effects. The coefficient changes to −0.128 and remains statistically significant (t = −3.0). The last column replicates the second column but adds both department and cohort fixed effects. The coefficient is approximately unchanged relative to the second column (−0.103) and remains statistically significant (t = −2.2). Overall, including employees that left the Civil Service does not substantially affect our conclusions.

The estimates in Table 9 do not fully rule out the possibility of bias from missing promotion data, however. When estimating the models in Table 9, we must assume that Civil Service leavers would have been promoted at the same rate as other employees in their department–cohorts. This assumption is violated if leaving is endogenous to promotions or department health. We first explore the sensitivity of our estimates to leaving being endogenous to promotions. If everyone who left would not have been promoted, then the coefficient in column (2) changes from −0.101 to −0.107 but remains statistically significant (t = −2.2). If everyone who left would have been promoted, then the coefficient in column (2) changes from −0.101 to −0.088 but remains statistically significant (t = −2.3). We also experiment with varying the average promotion rate for leavers in single percentage point increments from 1% to 99%. Across these values, the smallest coefficient estimate is −0.089, whereas the largest coefficient estimate is −0.122. Our results are therefore robust to leaving being endogenous to promotions in a manner that is constant across departments.

A more extreme possibility is that leaving varies by department in a manner that is related to both promotion rates and health. In particular, if everyone leaving a healthy department were promoted at the average rate for their department whereas everyone leaving a sick department were promoted at an above average rate for their department, our estimates would be too large. We test the sensitivity of our results to this possibility by assigning leavers from healthy department–cohorts the average promotion rate for their department–cohort, and leavers from unhealthy department–cohorts a promotion rate that is above average for their department–cohort. We define healthy department–cohorts as department–cohorts with heart disease incidence rates below the mean heart disease incidence rate and unhealthy department–cohorts as department–cohorts with heart disease incidence rates above the mean heart disease incidence rate. Our estimates remain statistically significant up to the point at which the promotion rate among leavers in an unhealthy department–cohort is 28% higher than the average promotion rate for their department–cohort (the coefficient at this point is −0.070.) A relationship of this strength would imply that the interaction between leaving and department–cohort health is a stronger predictor of promotions than any of our independent variables, including gender, tenure or college attendance. Our results thus appear robust to a high degree of endogeneity in leaving the Civil Service.

The other form of attrition that could affect our estimates is attrition in the outcome variable. The heart disease measure is available for every employee in our analytic sample, but 24% of employees did not attend their 1998 medical screenings. Heart disease for these individuals is still identified if they are diagnosed with heart disease by a medical doctor. Nevertheless, an individual who misses the 1998 screening and has minimal contact with the medical system could be incorrectly coded as not having heart disease. If promotions reduce the probability of attending the 1998 screening, our coefficient estimates could be too large.

We test the sensitivity of our results to attrition in the outcome variable by calculating reasonable bounds on the impact of the missing medical screenings. To calculate these bounds, we assume that any employee that missed the 1998 medical screening and was not diagnosed with heart disease elsewhere would have developed heart disease at double the average incidence rate in our sample (18%). Under this assumption, our 2SLS coefficient estimate on promotions changes from $-0.128$ to $-0.116$ but remains statistically significant ($t = 2.8$). The results are insensitive to attrition in medical screenings in large part because there is no statistically significant relationship between department–cohort promotion rates and medical screening attrition.

## 5. Discussion

Differences estimates suggest that a promotion reduces the probability of developing heart disease by 2.6 percentage points over a 15-year period, while 2SLS results imply a larger reduction of 12.8 percentage points. A number of objections to interpretation of the 2SLS estimates as causal effects exist, including employee selection into departments or cohorts, selective attrition, independent effects of departments on health and finite sample bias. However, these biases do not appear to drive the results.

We cannot isolate how much of the reduction in heart disease works through increased income and how much works through status (relative position) or work environment effects. The point estimates are large enough, however, to suggest that absolute income is not the only causal channel. The increased income associated with a promotion has some permanence; employees promoted one grade level by our instrument remain 0.55 grade levels ($t = 2.0$) above their peers 8–12 years later.[22] If the entire effect were assumed to run through the channel of (current) income, the elasticity of heart disease with respect to income would be around $-2$ (in comparison, the differences estimates imply an elasticity of heart disease with respect to income in the range of $-0.4$ to $-0.8$). This elasticity is an order of magnitude larger than the elasticity of mortality with respect to income that Deaton and Paxson (2001) estimate, the elasticity of health-related symptoms with respect to income that Lindahl (2005) reports, or the elasticities of several child health measures with respect to income that Case *et al.* (2002) find. It is three to four times larger than the elasticity of mortality with respect to six-year averages of income that Sullivan and von Wachter (2009) report.

One clear pattern that emerges from the results is that the 2SLS estimates generally exceed the magnitude of the differences estimates. This may seem surprising, as one would expect selection effects to bias differences estimates towards overstating the true causal effect. However, we should note that the standard errors on the 2SLS coefficients are relatively large; a true effect located near the differences coefficient could plausibly generate the observed data even with no bias in either estimator. Furthermore, the 2SLS estimates implicitly weight distinct segments of the Civil Service population

[22] The sample for which we have grade level information in the latter years, however, is less than half the size of our primary analytic sample.

differently from the differences estimates. Figures 5 and 6 plot the share of promoted individuals (grey bars) and the 2SLS weight share (black bars) by cohort and initial grade level respectively.[23] These Figures demonstrate that, relative to the differences estimates, the 2SLS estimates place more weight on employees from very recent or very early cohorts and on employees who are in the lowest grade level when sampling begins. Re-estimating the differences model using the 2SLS weights increases the magnitude of the differences coefficient in the heart disease regressions from $-0.026$ to $-0.048$. Nevertheless, a substantial gap remains between the 2SLS and differences coefficients.

The model presented in (1) and its implications for the statistical estimators highlight two reasons why the 2SLS coefficients could converge to larger values than the differences coefficients. First, using current grade level as a proxy for permanent income may attenuate the differences coefficient relative to the 2SLS coefficient. Second, how employees define their reference groups can have different implications for the differences and 2SLS estimators.[24]

Employment grade is measured at only a few points in time. However, it is presumably the entire history of promotions, rather than employment grade at a single point in time, which determines the likelihood of CHD. The independent variable of interest is therefore measured with error, generating attenuation bias as shown in (3) and (4). Solon (1992) and Zimmerman (1992), and more recently Mazumder (2005), demonstrate that this type of measurement error generates substantial attenuation bias when estimating the intergenerational elasticity of income. Sullivan and von Wachter (2009) find that using a six-year average of earnings instead of a single year of earnings increases the observed relationship between earnings and mortality by 70%. The same measurement error phenomenon probably attenuates the differences estimates in the Whitehall II data. The 2SLS estimator, by contrast, remains consistent in the presence of classical measurement error and is therefore typically of larger magnitude.[25]

Rational expectations may further increase the magnitude of the 2SLS estimator relative to the differences estimator. Most employees can likely predict their chances of promotion in comparison to their co-workers with reasonable accuracy. However,

[23] The 2SLS weight share is calculated as described in Angrist and Imbens (1995). The 2SLS weight for each individual is calculated separately and then the weights across all individuals in a given cohort or initial grade level are summed together.

[24] The possibility of attenuation bias in the differences coefficients relative to the 2SLS coefficients is supported by the existing quasi-experimental research on education – one of the most important determinants of SES – and health. Lleras-Muney (2005) uses changes in compulsory schooling laws to estimate the effect of education on mortality and concludes that one additional year of education reduces the probability of dying in the next decade by at least 3.6 percentage points. This estimate is approximately three times the magnitude of her corresponding least squares coefficient. Oreopoulos (2007) implements a similar research design and finds that schooling has large, beneficial effects on health. His 2SLS estimates range from 100% to 200% of the magnitude of his OLS estimates. MacInnis (2006) uses the Vietnam Draft Lottery to estimate the effect of college attendance on morbidity. She concludes that college graduation reduces smoking, obesity, mental distress and Type 2 diabetes by between 1.7 and 9.9 SD, depending on the disease. These 2SLS estimates range from 2 to 15 times the magnitude of her OLS estimates.

[25] Alternatively, if the entire history of promotions is condensed into a single binary variable, the results in Angrist and Imbens (1995) demonstrate that the coefficient estimated by IV may be overstated, although the sign will be correct. Although the promotion variable is not literally coded as binary, it only takes on a small number of discrete values, and the true effect may work though the entire history of promotions. This could cause the 2SLS estimate to exceed the differences estimate.

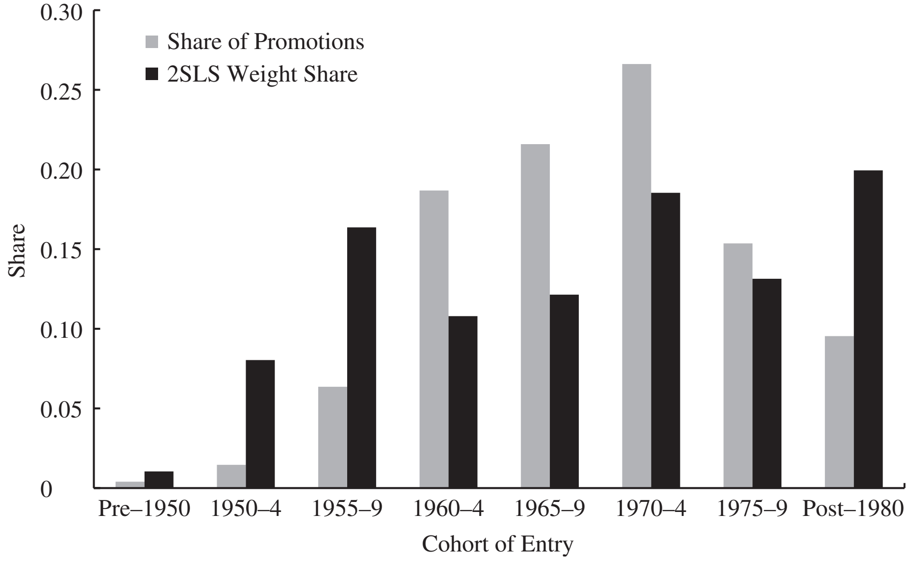


Fig. 5. *Two-stage Least Squares (2SLS) Weights by Cohort of Entry*
*Notes*. This Figure plots the weight that each cohort contributes to the differences estimate (grey bars) and the 2SLS estimate (black bars). Weight shares have been normalised so that they sum to one across all cohorts.

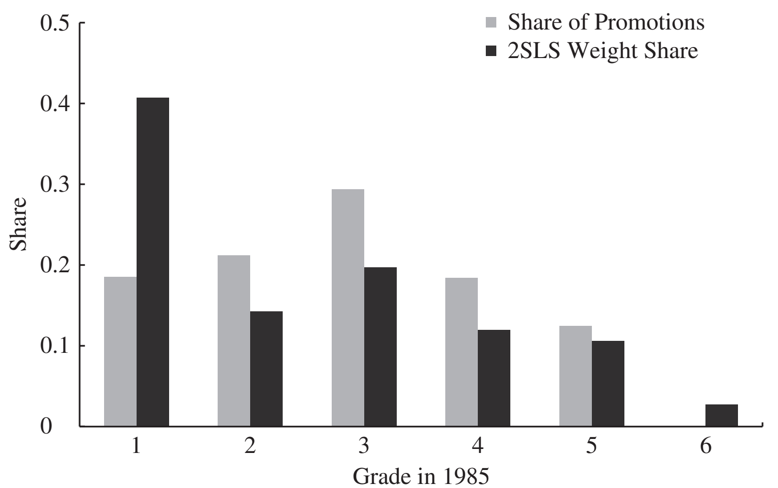


Fig. 6. *Two-stage Least Squares (2SLS) Weights by Grade in 1985*
*Notes*. This Figure plots the weight that each grade level contributes to the differences estimate (grey bars) and the 2SLS estimate (black bars). Weight shares have been normalised so that they sum to one across all cohorts.

variation in department or cohort level promotion rates affects the promotion of the marginal employee within a given department or cohort. By definition, this employee cannot accurately forecast whether he will be promoted with a high level of confidence. Therefore, the promotions utilised by the differences estimator should be more likely to be anticipated, whereas the promotions utilised by the 2SLS estimator should be less likely to be anticipated. As we define permanent income as $y_{it}^{\mathrm{p}} = y_{it} + \sum_{j=1}^{\infty} \delta^{j}\mathrm{E}(y_{it+j})$, a perfectly anticipated increase in grade level should have no effect on permanent income. This implies that changes in both absolute income and relative income may be mismeasured, particularly in the differences specification.

How employees construct their reference groups can also influence the relative magnitudes of the differences and 2SLS estimators. An IV regression estimates the Local Average Treatment Effect (LATE), or the average treatment effect for 'compliers', i.e. individuals whose treatment status is affected by the instrument (Angrist *et al.*, 1996). In the context of our Whitehall instrument, the LATE compliers are employees who are skilled enough to merit promotion in high promotion rate scenarios but not skilled enough to merit promotion in low promotion rate scenarios. All of the compliers should thus have relatively similar skill levels. Highly skilled employees, who are always promoted, and highly unskilled employees, who are never promoted, will not be compliers. If workers construct their reference groups based in part on skill, then all LATE compliers, whether promoted or not, will tend to share similar reference points and observed income will be a reasonable proxy of relative income for the 2SLS estimator.[26] Furthermore, if the compliers know that they are skilled enough to be promoted in a higher promotion rate department, those that are denied promotions may become particularly frustrated, causing deleterious health effects. The differences estimator, by contrast, leverages variation in promotions between 'always-takers' – employees who are skilled enough to be promoted even in low promotion rate departments – and 'never-takers' – employees who are not skilled enough to be promoted even in high promotion rate departments (Angrist *et al.*, 1996). These workers are likely to have substantially different reference points, and so observed income will be a poor proxy of relative income for the differences estimator.

Existing empirical evidence supports the hypothesis that individuals form reference groups based on skill. Clark and Senik (2010) find that Europeans most frequently cite colleagues as the reference group against whom they compare themselves, and several quasi-experimental studies that match individuals with roughly equivalent skills demonstrate large effects of relative status on health. For example, Redelmeier and Singh (2001) demonstrate that Oscar winners live 3.9 years longer on average than Oscar nominees, and Rablen and Oswald (2008) find a similar pattern when comparing Nobel Prize winners from 1901 to 1950 with matched Prize nominees who did not win. In complementary research, Becker *et al.* (2008) find that baseball players that receive just enough votes to enter the Hall of Fame have significantly longer life expectancy than players that fall just short of entering the Hall of Fame. Taken as a whole, these studies, combined with our Whitehall findings, suggest that reference groups may be highly localised and defined in part by skill.

[26] The notion that reference groups may be defined in part by skill seems compelling. For example, academic faculty are more likely to compare their credentials in reference to their peers at other institutions rather than to compare their credentials relative to the custodial staff in their own department.

In contrast to the findings in populations with sharply defined reference groups, Boyce and Oswald (forthcoming) find a weaker relationship between promotions and health measures in the British Household Panel Survey. Individuals promoted to manager (the highest job classification in the survey) report a significant drop in doctors' visits relative to individuals who are not promoted. The magnitude of this drop is two to three times larger than the cross-sectional difference in doctors' visits between managers and non-managers. However, promoted individuals display minimal improvements in self-reported health and they report significant increases in mental strain. The impacts of promotions to supervisory positions (the middle job classification in the survey) are generally insignificant. Overall, the results suggest that the relationship between health and promotions may be weaker when rankings or reference groups are less sharply defined.

In addition to the factors highlighted by the model, there is another reason why the 2SLS coefficients could exceed the differences coefficients. If promotions have positive external effects on the co-workers of employees who are promoted, then the 2SLS coefficients will be inflated. In statistical terms, this would be a violation of the non-interference portion of the 'stable unit treatment value assumption' (Rubin, 1980). For example, workers in departments with high promotion rates may be content because they believe that they too will soon receive a promotion and general upward mobility may improve the work environment for everyone. In that case, the differences estimate would tend to understate the true overall effect of promotions, because some non-promoted employees would receive a beneficial treatment when their co-workers were promoted. The 2SLS estimate, by contrast, would tend to overstate the effect for a treated individual, because it assumes that the entire effect is operating only through individuals that are promoted. The 2SLS coefficient would therefore be estimating the net internal and external effects of a promotion on heart disease, rather than simply an internal effect.

## 6. Conclusion

We use department-by-cohort promotion rates as plausibly exogenous sources of variation in employment grade to estimate the effect of promotions on CHD. Our estimates are sizable, implying that a promotion from the lowest grade level reduces the probability of heart disease by 2.6 to 12.8 percentage points. These estimates do not appear to be driven by employee selection or endogeneity of the instruments, and they are consistent with other estimates of the causal effects of SES on health. Nevertheless, the range of potential effects is large, as the IV estimates have wide confidence intervals and are of substantially greater magnitude than the differences estimates. Given the rich literature on mismeasurement of long run earnings, we interpret the differences estimates as a lower bound (in magnitude) on the effect of an unanticipated promotion. The likelihood of positive external effects suggests that the IV estimates represent an upper bound (in magnitude) on the effect of an unanticipated promotion. The true impact of an unanticipated promotion is thus likely to lie between the IV and differences estimators.

The finding that favourable shocks can positively affect health complements recent research showing that adverse shocks can negatively affect health (Eliason and Storrie, 2009; Sullivan and von Wachter, 2009). Whether the results generalise to populations beyond British civil servants depends upon several factors. On the one hand, like most British workers the Whitehall employees are engaged in white-collar work and earn incomes that place them near the centre of the British income distribution.[27] Their health care, provided by the NHS, is similar to that received by almost all Britons. Along these dimensions they are representative of a typical British worker. However, unlike some private employers, the Civil Service has clearly defined employment grades. If the clear delineation of employment grades enhances the effects of promotions on health, then the effects of promotions for the greater population may be smaller than the effects estimated here. Nevertheless, the results suggest that a significant health gradient can appear even among individuals whose differences in SES appear small on a global scale.

*University of California*
*University College London*

*Submitted: 11 June 2010*
*Accepted: 20 June 2011*

Additional Supporting information may be found in the online version of this article:

**Appendix S1.** Department Checks on Health.
**Appendix S2.** Finite Sample Bias.
**Appendix S3.** OLS Estimand.
**Appendix S4.** Differences estimand.
**Appendix S5.** Instrumental Variables Estimand.
**Appendix S6.** Differences and IV Estimands with Fixed Reference Points.
**Appendix S7.** Differences and IV Estimands with Lagged Reference Points.
**Appendix S8.** Differences and IV Estimands with Cohort-based Reference Points.
**Appendix S9.** Differences and IV Estimands with Skill-based Reference Points.

Please note: The RES and Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the authors of the article.

## References

Adda, J., Chandola, T. and Marmot, M. (2003). 'Socio-economic status and health: causality and pathways', *Journal of Econometrics*, vol. 112(1), pp. 57–63.

Anderson, M. (2008). 'Multiple inference and gender difference in the effects of early intervention: a reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects', *Journal of the American Statistical Association*, vol. 103(484), pp. 1481–95.

Angrist, J. and Imbens, G. (1995). 'Two-stage least squares estimation of average causal effects in models with variable treatment intensity', *Journal of the American Statistical Association*, vol. 90(430), pp. 431–42.

Angrist, J., Imbens, G. and Rubin, D. (1996). 'Identification of causal effects using instrumental variables', *Journal of the American Statistical Association*, vol. 91(434), pp. 444–55.

[27] The 25th, 50th and 75th percentiles of the 1990 British income distribution were 6,182, 9,473 and 14,000 pounds respectively, whereas the salary bands for the median grade level in our data set range from a minimum of 7,286 pounds to a maximum of 12,621 pounds (HM Treasury Office, 1990; Goodman and Webb, 1994; O'Donoghue, 1998). Over 70% of the British labour force was employed in white-collar occupations in 2003 (Begum, 2004).

Becker, D., Chay, K. and Swaminathan, S. (2008). 'Mortality and the Baseball Hall of Fame: an investigation into the role of status in life expectancy', American Economics Association Conference Presentation, New Orleans, January 4.

Begum, N. (2004). 'Employment by occupation and industry', *Labour Market Trends*, vol. 112(6), pp. 227–34.

Bosma, H., Marmot, M., Hemingway, H., Nicholson, A., Brunner, E. and Stansfeld, S. (1997). 'Low job control and risk of coronary heart disease in Whitehall II (prospective cohort) Study', *British Medical Journal*, vol. 314(7080), pp. 558–65.

Bound, J. and Krueger, A. (1991). 'The extent of measurement error in longitudinal earnings data: do two wrongs make a right?', *Journal of Labor Economics*, vol. 9(1), pp. 1–24.

Boyce, C. and Oswald, A. (forthcoming). 'Do people become healthier after being promoted?', *Health Economics*.

Case, A., Lubotsky, D. and Paxson, C. (2002). 'Economic status and health in childhood: the origins of the gradient', *American Economic Review*, vol. 92(5), pp. 1308–34.

Case, A. and Paxson, C. (2008). 'Stature and status: height, ability, and labor market outcomes', *Journal of Political Economy*, vol. 116(3), pp. 499–532.

Chandola, T., Bartley, M., Sacker, A., Jenkinson, C. and Marmot, M. (2003). 'Health selection in the Whitehall II Study, UK', *Social Science and Medicine*, vol. 56(10), pp. 2059–72.

Chandola, T., Siegrist, J. and Marmot, M. (2005). 'Do changes in effort-reward imbalance at work contribute to an explanation of the social gradient in angina?', *Occupational and Environmental Medicine*, vol. 62(4), pp. 223–30.

Civil Service (2011). 'Jobs FAQs', http://www.civilservice.gov.uk/jobs/Background/Frequently-Asked-Questions/index.aspx# (last accessed 8 September 2011).

Clark, A. and Senik, C. (2010). 'Who compares to whom? The anatomy of income comparisons in Europe', Economic Journal, vol. 120(544), pp. 573–94.

Deaton, A. (2003). 'Health, inequality, and economic development', *Journal of Economic Literature*, vol. 41(1), pp. 113–58.

Deaton, A. and Paxson, C. (2001). 'Mortality, education, income and inequality among American cohorts', in (D. Wise, ed.), *Themes in the Economics of Aging*, pp. 123–65, Chicago: University of Chicago Press.

Eliason, M. and Storrie, D. (2009). 'Does job loss shorten life?', *Journal of Human Resources*, vol. 44(2), pp. 277–302.

Ferrer-i-Carbonell, A. (2005). 'Income and well-being: an empirical analysis of the comparison income effect', *Journal of Public Economics*, vol. 89(6), pp. 997–1019.

Goodman, A. and Webb, S. (1994). 'For richer, for poorer: the changing distribution of income in the United Kingdom, 1961–91', Commentary No. 42, Institute for Fiscal Studies, London.

Grossman, M. (1972). 'On the concept of health capital and the demand for health', *Journal of Political Economy*, vol. 80(2), pp. 223–55.

Her Majesty's Stationery Office (1971). *Civil Service Statistics 1971*, London: Her Majesty's Stationery Office.

HM Treasury Office (1985). *Civil Service Statistics 1985*, London: Her Majesty's Stationery Office.

HM Treasury Office (1989). *Civil Service Statistics 1989*, London: Her Majesty's Stationery Office.

HM Treasury Office (1990). *Civil Service Statistics 1989–90*, London: Her Majesty's Stationery Office.

Kuper, H. and Marmot, M. (2003). 'Job strain, job demands, decision latitude, and risk of coronary heart disease within the Whitehall II Study', *Journal of Epidemiology and Community Health*, vol. 57(2), pp. 147–53.

Lindahl, M. (2005). 'Estimating the effect of income on health and mortality using lottery prizes as an exogenous source of variation in income', *Journal of Human Resources*, vol. 40(1), pp. 144–68.

Lleras-Muney, A. (2005). 'The relationship between education and adult mortality in the United States', *Review of Economic Studies*, vol. 72(1), pp. 189–221.

Luttmer, E. (2005). 'Neighbors as negatives: relative earnings and well-being', *Quarterly Journal of Economics*, vol. 120(3), pp. 963–1002.

MacInnis, B. (2006). 'Does college education impact health? Evidence from the pre-lottery Vietnam Draft', Working Paper, U.C. Berkeley Department of Agricultural and Resource Economics.

Marmot, M. (2003). 'Understanding social inequalities in health', *Perspectives in Biology and Medicine*, vol. 46(3), pp. S9–S23.

Marmot, M., Bosma, H., Hemingway, H., Brunner, E. and Stansfeld, S. (1997). 'Contribution of job control and other risk factors to social variations in coronary heart disease incidence', *The Lancet*, vol. 350(9073), pp. 235–9.

Marmot, M. and Brunner, E. (2004). 'Cohort profile: the Whitehall II Study', *International Journal of Epidemiology*, vol. 34(2), pp. 251–6.

Marmot, M. and Davey Smith, G. (1997). 'Socio-economic differentials in health: the contribution of the Whitehall studies', *Journal of Health Psychology*, vol. 2(3), pp. 283–96.

Marmot, M., Davey Smith, G., Stansfeld, S., Patel, C., North, F., Head, J., White, I., Brunner, E. and Feeney, A. (1991). 'Health inequalities among British civil servants: the Whitehall II Study', *The Lancet*, vol. 337(8754), pp. 1387–93.

Marmot, M., Rose, G., Shipley, M. and Hamilton, P. (1978). 'Employment grade and coronary heart disease in British civil servants', *Journal of Epidemiology and Community Health*, vol. 32(4), pp. 244–9.

Mazumder, B. (2005). 'Fortunate sons: new estimates of intergenerational mobility in the United States using social security earnings data', *Review of Economics and Statistics*, vol. 87(2), pp. 235–55.

Mealli, F. and Rubin, D. (2003). 'Assumptions allowing the estimation of direct causal effects', *Journal of Econometrics*, vol. 112, pp. 79–87.

Mein, G., Martikainen, P., Hemingway, H., Stansfeld, S. and Marmot, M. (2003). 'Is retirement good or bad for mental and physical health functioning? Whitehall II Longitudinal Study of civil servants', *Journal of Epidemiology and Community Health*, vol. 57(1), pp. 46–49.

O'Donoghue, J. (1998). 'Harmonised index of consumer prices: historical estimates', *Economic Trends*, vol. 541, pp. 49–55.

Oreopoulos, P. (2007). 'Do dropouts drop out too soon? Wealth, health, and happiness from compulsory schooling', *Journal of Public Economics*, vol. 91, pp. 2213–29.

Rablen, M. and Oswald, A. (2008). 'Mortality and immortality: the Nobel Prize as an experiment into the effect of status on longevity', *Journal of Health Economics*, vol. 27(6), pp. 1462–71.

Redelmeier, D. and Singh, S. (2001). 'Survival in Academy Award-winning actors and actresses', *Annals of Internal Medicine*, vol. 134(10), pp. 955–62.

Reid, D., Brett, G., Hamilton, P., Jarrett, R., Keen, H. and Rose, G. (1974). 'Cardiorespiratory disease and diabetes among middle-aged male civil servants', *The Lancet*, vol. 1(7856), pp. 469–73.

Rubin, D (1974). 'Estimating causal effects of treatments in randomized and nonrandomized studies', *Journal of Educational Psychology*, vol. 66(5), pp. 688–701.

Rubin, D (1980). 'Randomization analysis of experimental data: the Fisher Randomization Test: comment', *Journal of the American Statistical Association*, vol. 75(371), pp. 591–3.

Smith, J. (1999). 'Healthy bodies and thick wallets: the dual relation between health and economic status', *Journal of Economic Perspectives*, vol. 13(2), pp. 145–66.

Solon, G. (1992). 'Intergenerational income mobility in the United States', *American Economic Review*, vol. 82(3), pp. 393–408.

Stanley, M. (2004). *How to Be a Civil Servant*, London: Politico's Publishing.

Sullivan, D. and von Wachter, T. (2009). 'Average earnings and long-term mortality: evidence from administrative data', *American Economic Review: Papers and Proceedings*, vol. 99(2), pp. 133–8.

von Wachter, T. and Bender, S. (2006). 'In the right place at the wrong time: the role of firms and luck in young workers' careers', *American Economic Review*, vol. 96(5), pp. 1679–705.

Zimmerman, D. (1992). 'Regression toward mediocrity in economic stature', *American Economic Review*, vol. 82(3), pp. 409–29.