Published in final edited form as:

Scott J Polit Econ. 2011 ; 58(3): 347–377.

MARRIAGE, BMI, AND WAGES: A DOUBLE SELECTION APPROACH

Heather Brown*

* University of Aberdeen

Abstract

Obesity rates have been rising over the past decade. As more people become obese, the social stigma of obesity may be reduced. Marriage has typically been used as a positive signal to employers. If obese individuals possess other characteristics that are valued in the labour market they may no longer face a wage penalty for their physical appearance. This paper investigates the relationship between marital status, body mass index (BMI), and wages by estimating a double selection model that controls for selection into the labour and marriage markets using waves 14 and 16 (2004 and 2006) of the British Household Panel Survey. Results suggest that unobserved characteristics related to marriage and labour market participation are causing an upward bias on the BMI coefficients. The BMI coefficient is positive and significant for married men only in the double selection model. The findings provide evidence that unobserved characteristics related to success in the marriage and labour market may influence the relationship between BMI and wages.

I Introduction

Obesity rates have continued to rise over the past few decades in much of the developed world. For example in the United Kingdom, in 1993, 13% of adult men and 16% of women were obese, but currently, these proportions have risen to 24% for both men and women (NHS: The Information Centre, 2008). Rising obesity rates have led to a flurry of research activity examining the consequences of obesity on various aspects of life from marriage to employment to happiness (Baum and Ford, 2004; Morris, 2006, 2007; Oswald and Powdthavee, 2007; Chiappori *et al.*, 2009).

The obesity and labour market outcomes literature (Sargent and Blanchflower, 1994; Baum and Ford, 2004; Cawley, 2004; Morris, 2006, 2007; Han *et al.*, 2009) have reported mixed effects of the impact of body mass index (BMI) on wages. For example, Cawley (2004) found a wage penalty for white women only, Morris (2006) and Han *et al.* (2009) find a significant wage penalty for all women. Sargent and Blanchflower (1994), Morris (2007), and Han *et al.* (2009) find no employment penalty for obese men but an employment penalty for obese women which is ethnicity dependent.

This paper develops an alternative marriage market hypothesis to better understand the counterintuitive finding in many studies (Cawley, 2004; Morris, 2007; Han *et al.*, 2009) of a positive effect of BMI on wages for men and mixed results of the effect of BMI on wages for women. If unobserved characteristics valued in both the marriage market and labour markets are more important than physical appearance for prospective partners/employers, this may partially explain the observed wage premium for married men. Changing social

norms regarding the 'ideal' body weight may mean that heavier individuals are no longer penalised in the marriage and/or labour market if they possess other characteristics that are valued in these markets. A national health service in the United Kingdom also means that there are not the extra medical costs for employers of individuals with a higher BMI. The mixed results of the impact of BMI on wages for women may be caused by heavier women being penalised in the marriage but not labour market. Single women with a higher BMI may then choose to invest more time and resources to paid employment leading to a wage premium.

To test how controlling for selection into the labour and marriage markets impacts on the effect of BMI on wages for married and single men and women waves 14 and 16 (2004–2006) of the British Household Panel Survey (BHPS)¹ are used. Various model specifications are estimated to examine how the inclusion of health in the double selection, single selection, and standard wage equation influence the impact of wages on BMI for the four sub-samples of: (1) married men; (2) single men; (3) married women; and (4) single women. To further test how marriage may impact on wages for these four groups a Blinder–Oaxaca decomposition is employed. Discrimination against married women or single men in the labour market may influence the impact of BMI on wages.

This paper is organised as follows. Section II discusses the motivation behind the econometric model. Section III introduces the econometric framework. The data is described in section IV. Section V presents the empirical results. Finally, section VI concludes.

II Background

There have been three studies (Baum and Ford, 2004; Bhattacharya and Bundorf, 2005; Oreffice and Quintana-Domeque, 2009) investigating the possible mechanisms through which obesity may impact on wages. Baum and Ford (2004) found significant effects of health, human capital differentials, and employer and consumer distaste for obese workers. Bhattacharya and Bundorf (2005) investigate how health care costs influence wages for obese individuals, finding a wage penalty only for those employed at firms providing health insurance. In the UK, there is no employer sponsored health care. There will be no direct costs of employing obese individuals eliminating this effect on the wage rate of obese individuals. Oreffice and Quintana-Domeque (2009) suggest a marriage market mechanism to explain why there may not be a wage penalty for obese men. They found that married men to compensate their spouse for their heavier weight, tended to work longer hours and subsequently earn higher wages compared with heavier unmarried men and heavier single and married women. Similarly, Chiappori *et al.* (2009) find that wages, BMI, and height matter for matching in the marriage market adding further evidence that characteristics valued in the marriage market may influence other observed outcomes.

An alternative explanation for a positive obesity effect on wages can be found in the social norms literature. Empirical evidence shows that social standards of physical appearance are important determinants of behaviour (see e.g. Garner *et al.*, 1980; Mazur, 1986). Burke and Heiland (2007) find that individuals attempt to maintain weight around a socially accepted norm. Oswald and Powdthavee (2007) develop a theory where the disutility of obesity decreases in a society with a higher prevalence rate of obesity, as utility will depend upon relative weight. If the 'ideal' weight has been increasing as obesity rates have risen, there may no longer be a wage penalty for having a higher BMI. This hypothesis may also explain differences in the impact of BMI on wages by race (Cawley, 2004; Han *et al.*, 2009).

¹These are currently the only two waves containing the height and weight information needed to calculate BMI (kilograms divided by metres squared).

Different cultural norms regarding acceptable weight rather than health or productivity factors could be driving the impact of BMI on wages.

This paper links the social norms and marriage market literature to offer an alternative hypothesis to explain the mixed results in the BMI and wage literature regarding the direction of the effect of BMI on wages. Changing social norms of acceptable body weight may have reduced the penalties associated with having a higher BMI in both the marriage and labour market. If this is the case, the expectation is that of wage premium rather than a wage penalty for heavier married men who possess other characteristics which are highly valued in both markets. However, there may be disequilibrium for women, where characteristics valued in the labour market may not have equal merit in the marriage market. In the marriage market women are more likely to be judged by their physical appearance than men. Single heavier women may not be penalised in the labour market because of their weight and could therefore choose to devote more time and resources to the labour market to compensate for failure in the marriage market. This suggests that heavier single women may not face a wage penalty in the labour market. The econometric model, by controlling for unobserved characteristics related to the marriage and labour market, will attempt to disentangle how marital status influences the impact of BMI on wages.

III Econometric Framework

A double selection model is used to account for how success in the marriage market impacts on the influence of BMI on wages. As was discussed in the previous section, it is expected that marriage will positively influence the wage rate for men. Thus, it is likely that there may be a wage premium for employed married men irrespective of their physical appearance. On the other hand, for women, success in the marriage market may not act as a positive signal to potential employers. Single heavier women who devote more time and resources to the labour market may receive a wage premium. Thus, if this double selection process is not accounted for in the estimation strategy, the coefficient on BMI in the wage equation will be biased upwards.

Following Wetzels and Zorlu (2003), to deal with these sample selection issues, the underlying decision process of employment and marriage is explicitly modelled using a bivariate probit model (Heckman, 1979; Maddala, 1983; Greene, 2003). It is assumed that marriage and employment are jointly determined in a double selection framework.

To illustrate this double selection process, there are four possible combinations of marriage and employment outcomes:

	Married	
Employed	0	1
0	Q_1	Q_2
1	Q_3	Q_4

Where Q_1 represents a state of not being in employment or marriage, Q_2 represents a state of not being in employment but being married, Q_3 represents a state of being in employment but not in marriage, and Q_4 represents a state of employment and marriage. This implies that we observe two possible sets of outcomes for each individual:

$$Y_i = \begin{cases} Y_{1it}=1 & Y_{2it}=1 \\ Y_{1it}=0 & Y_{2it}=0 \end{cases} \quad (1)$$

where $i = 1, 2, \dots, n$ and $t = 1, 2$.

Let Y_{1it} be an individual's observed employment status which equals one if an individual (i) is employed in time (t) and is zero otherwise and Y_{2it} is an individual's observed marital status which equals one if an individual (i) is married in time (t) and is zero otherwise.

This pair of decisions is presented in a single standard bivariate probit model. Becker's (1974) theory on marriage considers the gain to individuals from marriage. Individuals choose a spouse based upon personal characteristics that maximise the household production function. Firms choose employees based on characteristics related to human capital stock that will maximise the firm's output and profits. Therefore, individual characteristics such as education, age, and health are likely to impact on participation in the marriage and labour market. Thus, it is assumed that similar observed and unobserved factors influence the likelihood of success in the marriage and labour markets. However, these factors may impact on labour and marriage market outcomes differently:

$$\begin{aligned} Y_{1it} &= X_{it}\beta_1 + \varsigma_1 H_{1it} + \tau_1 B_{1it} + \varepsilon_{1it}, \\ Y_{2it} &= X_{it}\beta_2 + \varsigma_2 H_{2it} + \tau_2 B_{2it} + \varepsilon_{2it}, \end{aligned} \quad (2)$$

$$\varepsilon_{1it} = \alpha_{1i} + u_{1it} \quad \text{and} \quad \varepsilon_{2it} = \alpha_{2i} + u_{2it},$$

where $i = 1, 2, \dots, n$ and $t = 1, 2$.

Y_{1it} and Y_{2it} are defined above. X_{it} is a vector of individual characteristics with the associated parameter of coefficients, β_1 and β_2 for the employment and marriage outcomes respectively. H_{it} is a vector of health variables with the associated parameter of coefficients ς_1 and ς_2 for the employment and marriage outcomes respectively. B_{it} is a scalar for BMI with the associated coefficient of parameters τ_1 and τ_2 for the employment and marriage outcomes, respectively. The error terms ε_{1it} and ε_{2it} are comprised of individual effects α_i and a random error component, u_{it} . It is assumed that ε_{1it} and ε_{2it} follow a bivariate standard normal distribution $E[\varepsilon_{1it}] = E[\varepsilon_{2it}] = 0$, $\text{Var}[\varepsilon_{1it}] = \text{Var}[\varepsilon_{2it}] = 1$ with the two error terms being correlated $\text{Cov}[\varepsilon_{1it}, \varepsilon_{2it}] = \rho$.

There are several issues that need to be addressed in the estimation strategy to investigate the role of marriage on the impact of BMI on wages. Firstly, if the unobserved individual effects are correlated with the explanatory variables then the random effects bivariate probit specification may give inconsistent results because it violates the condition:

$\alpha_i | X_i, H_i, B_i \sim \text{IN}(0, \sigma_\alpha^2)$ in the participation and marriage equation respectively. This is likely to be the case as unobserved individual effects such as motivation, sociability, and friendliness are likely to influence the likelihood of being married and participating in the labour market. Mundlak (1978) proposed a method to overcome this problem, that accounts for the positive correlation between the explanatory variables X_i , H_i and B_i and the individual effects α_i by modelling this relationship explicitly. This approach takes the group means of the time varying explanatory variables in the participation and marriage equations.

The group means of the explanatory variables are then included as additional explanatory variables in the bivariate probit model, as a proxy for removing the time invariant individual effects, α_i . Modelling this dependence allows for unbiased estimation of β , τ , and ς , regardless of whether or not X_i , H_i , and B_i and α_i are independent in the marriage and participation equations (Ebbes *et al.*, 2004). The model is then estimated by a maximum likelihood function which is described in greater detail in Greene (2003).

The second issue in the estimation strategy is to correct for the bias on the BMI coefficient in the wage equation caused by the observed outcomes of Q_1 , Q_2 , and Q_3 . This bias caused by selection into the labour and marriage market is reduced by computing an inverse Mill's ratio (Heckman, 1976) using the theory of truncated normal distribution. The inverse Mill's ratios calculated from the bivariate probit models are then added as additional regressors to the wage function reducing the bias on the BMI coefficient from not accounting for selection into the marriage and labour market.

To calculate the inverse Mill's ratio, the conditional expectation $E[Y_{1it}|Y_{1it}=1, Y_{2it}|Y_{2it}=1, X_{it}, H_{it}, B_{it}]$ is taken from the bivariate probit equation in (2) when $Y_{1it}=1$ and $Y_{2it}=1$ to compute the unconditional expectation $E(Y_{1it}, Y_{2it}|X_{it}, H_{it}, B_{it})$.² As explained above, it is assumed that the explanatory variables may have different impacts on labour market and marriage outcomes, therefore, an inverse Mill ratio is calculated for both Y_{1it} and Y_{2it} :

$$\begin{aligned}\widehat{\lambda}_1 &= \frac{(-X_{1it}\beta_1 + H_{1it}\varsigma_1 + B_{1it}\tau_1)}{\Phi(X_{1it}\beta_1 + H_{1it}\varsigma_1 + B_{1it}\tau_1)} = \frac{\phi(X_{1it}\beta_1 + H_{1it}\varsigma_1 + B_{1it}\tau_1)}{\Phi(X_{1it}\beta_1 + H_{1it}\varsigma_1 + B_{1it}\tau_1)}, \\ \widehat{\lambda}_2 &= \frac{(-X_{2it}\beta_2 + H_{2it}\varsigma_2 + B_{2it}\tau_2)}{\Phi(X_{2it}\beta_2 + H_{2it}\varsigma_2 + B_{2it}\tau_2)} = \frac{\phi(X_{2it}\beta_2 + H_{2it}\varsigma_2 + B_{2it}\tau_2)}{\Phi(X_{2it}\beta_2 + H_{2it}\varsigma_2 + B_{2it}\tau_2)},\end{aligned}\quad (3)$$

where for $\widehat{\lambda}_1$, $X_{1it}\beta_1 + H_{1it}\varsigma_1 + B_{1it}\tau_1$ indicates how the explanatory variables impact on labour market participation for Y_{1it} . For $\widehat{\lambda}_2$, $X_{2it}\beta_2 + H_{2it}\varsigma_2 + B_{2it}\tau_2$ represents how the explanatory variables impact on marital status, Y_{2it} . ϕ and Φ are the cumulative density function and probability distribution function, respectively, for a standard normal distribution (Wooldridge, 2002). Therefore, the inverse Mill's ratio is defined as the ratio between the standard normal probability distribution function and standard normal cumulative distribution function evaluated at each $X_{it}\beta + H_{it}\varsigma + B_{it}\tau$ for Y_{1it} and Y_{2it} .³

III.1 Wage equation

To determine the impact of BMI on wages, a modified Mincerian wage equation (Mincer, 1974) is employed which estimates the statistical relationship between market wages, education, experience, BMI. The most generalised specification is

$$\begin{aligned}\ln W_{it} &= \kappa X_{it} + \gamma H_{it} + \xi L_{it} + \psi B_{it} + \eta_1 \lambda_{1it} + \eta_2 \lambda_{2it} + \varepsilon_{it}, \\ \varepsilon_{it} &= \alpha_i + u_{it},\end{aligned}\quad (4)$$

where $i = 1, 2, \dots, n$, $T = 1, 2$.

Let $\ln W$ be the log of hourly wage for individual i in period t . The vectors X_i , H_i , and B_i are as defined in equation (2). The vector L_i represents labour market variables such as firm size and type of occupation, with the associated parameter of coefficients, ξ . λ_{1it} is the inverse Mill's ratio calculated for labour market participation with the associated parameter, η_1 and λ_{2it} is the inverse Mill's ratio calculated for marital status with the associated parameter of

²In this case Y_{1it} and/or Y_{2it} can be equal to one or zero.

³The Mundlak method is used with bivariate probit models estimated to calculate the Inverse Mill's ratios for the fixed effects wage model. The standard bivariate probit is estimated to calculate the Inverse Mill's ratios for the random effects generalised least squares wage model.

coefficients η_2 . The error term ε_i is comprised of unobservable individual effects α_i and a random error term, u_{it} . The model is estimated by generalised least squares controlling for random effects.

To control for potential omitted variable bias caused by the error term ε_i being correlated with the explanatory variables, a fixed effects wage model is also estimated. In this specification the unobserved effects α_i are removed from the model by taking the panel level averages of the explanatory variables.

The double selection framework assumes that the error terms from these bivariate probit and wage equations are normally distributed and not independent. A necessary identification restriction for this framework is that at least one of the explanatory variables included in the bivariate probit equation is excluded from the wage equation (Billari and Borgoni, 2005). The reason for this exclusion restriction is that the inverse Mill's ratio is a non-linear function of the explanatory variables in the bivariate probit equation (X_i, H_i, B_i); thus, the second stage equation (wage function) is identified because of this non-linearity. However, the non-linearity of the inverse Mill's ratio is based upon the normality assumptions of the bivariate probit equation which is not normally tested or justified. Therefore, in order to make the source of identification clear, it is advisable to have an explanatory variable in the bivariate probit equation, which is not included in the wage equation (Greene, 2003). The explanatory variables that are only included in the bivariate probit to meet this exclusion restriction are having preschool and/or school age children at home as well as non-labour income. It is assumed that these variables will influence participation in the labour and marriage market rather than wages. The standard errors in the random effects and fixed effects wage equation are bootstrapped to correct the bias in the standard errors caused by the two stage estimation procedure.

III.2 Robustness checks

There are two robustness checks on the results from the double selection model. A wage equation which does not account for selection and a wage equation which only controls for selection into the labour market using a univariate probit model for labour market participation are also estimated. In the latter model specification, the associated inverse Mill's ratio calculated to correct for selection into labour is then added as the only inverse Mill's ratio in the wage equation.

III.3 Wage differentials

The Blinder–Oaxaca method is employed to test the role of marriage in explaining wage outcomes. This method decomposes the mean differences in log wages for the outcome groups of interest, in this case married and unmarried respondents (Blinder, 1973; Oaxaca, 1973). The wage differential between the two groups is divided into an ‘explained’ component which is group differences in productivity arising from individual characteristics such as education or job tenure, and an ‘unexplained’ component often referred to as discrimination. If unobserved characteristics related to marital status influences wage outcomes then not accounting for selection into marriage will bias the coefficients in the wage equation. Formally the Blinder–Oaxaca decomposition is

$$\ln \bar{W}_M - \ln \bar{W}_N = \widehat{\kappa}_M, \widehat{\gamma}_M, \widehat{\xi}_M, \widehat{\psi}_M \left(\bar{X}_M - \bar{X}_N, \bar{H}_M - \bar{H}_N, \bar{L}_M - \bar{L}_N, \bar{B}_M - \bar{B}_N \right) + \bar{X}_N, \bar{H}_N, \bar{L}_N, \bar{B}_N \left(\widehat{\kappa}_M - \widehat{\kappa}_N, \widehat{\gamma}_M - \widehat{\gamma}_N, \widehat{\xi}_M - \widehat{\xi}_N, \widehat{\psi}_M - \widehat{\psi}_N \right), \quad (5)$$

where the subscripts M and N refer to married and non-married individuals respectively, $\widehat{\kappa}$, $\widehat{\gamma}$, $\widehat{\xi}$, $\widehat{\psi}$ represent the estimated coefficients, $\bar{X}$, $\bar{H}$, $\bar{L}$, $\bar{B}$ show the mean characteristics. The

first-term on the right-hand side of the equation is the wage gap due to individual characteristics and the second-term measures the unexplained portion of the wage gap between married and single respondents.

IV Data

The empirical analysis uses waves 14 and 16 (2004–2006) of the British Household Panel Survey (BHPS). The BHPS is a longitudinal study of approximately 5000 nationally representative private households, where 10,000 individuals aged 16 or older are surveyed.⁴ The BHPS questionnaire covers a wide range of topics ranging from employment status, wages, various health measures, and education. The first wave of the survey was conducted between 1 September 1990 and 30th April 1991, and it has been administered annually since. The initial household selection for the survey was determined by using a two-stage stratified systematic sampling procedure designed to give each address an approximately equal probability of selection. The same individuals are re-interviewed in each wave. In Wave 9 (1999), two additional samples were recruited from Scotland and Wales allowing for independent analysis of the countries, and comparison with England. In Wave 11 (2001), an additional sample from Northern Ireland of approximately 2000 households was added to increase the representativeness of the sample across the UK. As mentioned previously, Waves 14 and 16 are currently the only two waves which contain information on respondent's height and weight allowing the calculation of BMI.

The sample is restricted to individuals between the ages of 18 and 65 years old which is assumed to be the typical age range when individuals would participate in the labour market. This is also the age range which has been validated for use of the WHO BMI classification system to grade obesity (WHO, 1995, 2000, 2004). All pregnant women are omitted from the sample as this condition may impact BMI and may affect labour market outcomes. An unbalanced panel is used.

IV.1 Labour market participation

Labour market participation is defined according to employment status. The working category comprises respondents who claim to be employed or self-employed. The unemployed category consists of individuals who claim to be unemployed and are actively looking for employment. Respondents who are long-term sick or disabled, on maternity leave, retired, students, as well as partaking in family care are excluded from the analysis.

IV.2 Marital status

Marital status is a binary variable which equals one if an individual is married or cohabiting and is zero if the respondent is never married. Couples in a civil partnership, respondents who are divorced, separated, and widowed are excluded from the analysis. Individuals who are divorced, separated, or widowed were at one time successful in the marriage market and may share similar characteristics to married individuals potentially impacting on their labour market outcomes.

IV.3 Covariates in the bivariate probit equations

The explanatory variables in the bivariate probit equation are used to determine how individual characteristics, in particular health and BMI impact on labour market participation and marital status. Other individual characteristics included in the analysis are age, age squared, children under five, children aged five to fifteen, highest educational

⁴From wave 9 youth data for household members aged 11–15 were included in the survey. Height and weight measures for youth respondents were provided in waves 14 and 16; however, in this analysis the youth data is not used.

qualification received, non-labour income, and region. Married individuals are more likely to have children than their single counterparts which may impact on their labour market participation. Individual characteristics such as age and education will influence the likelihood of being successful in the labour and marriage market. BMI measures are included in all model specifications. A self-assessed health variable is included in some model specifications.

Health and BMI may impact on labour market participation. Empirical evidence suggests (see Currie and Madrian, 1999; Brown *et al.*, 2010) that poor health will reduce the likelihood of participating in the labour market. If individuals with a higher BMI have worse health than their leaner counterparts this will impact on their labour market participation. Individuals with a higher BMI may also be discriminated against in the hiring process. Thus, models are estimated with and without a health variable to determine if BMI has a direct or indirect effect on labour market participation. Health and BMI also act as clear signals on the marriage market. For example, BMI and health may act as an observable signal for unobserved characteristics such as future health and potential life expectancy. BMI can also signal preferences for other lifestyle characteristics such as how one chooses to spend their leisure time.

BMI is calculated using self-reported height measured in feet and inches, metres and centimetres, as well as weight measured in stones and pounds, and kilograms. Overweight and obese are measured using the WHO (1995, 2000, 2004) BMI classification system. Overweight is defined as a BMI between 25 and 30 kg/m² and obese is defined as BMI of greater than or 30kg/m². All model specifications are estimated twice: firstly, with a continuous measure for BMI and secondly with the indicator variables for overweight and obese. In the model, specifications estimated with the indicator variables for overweight and obese, healthy weight individuals or those with a BMI between 18.5 and 25 kg/m² are the base category in the analysis. All models are estimated twice with these two measures of BMI as a robustness check on the results.

IV.4 Hourly wage

In order to analyse the relationship between wages, BMI, and health, the dependent variable used in the wage function is the hourly wage of the respondent. There is no hourly wage variable in the BHPS. Therefore, the hourly wage variable is calculated following the literature (Bardasi and Taylor, 2005; Brown and Taylor, 2005) using usual gross pay per month and number of hours normally worked per week. To construct the variable, usual monthly pay is divided by number of hours normally worked per week which is then multiplied by 12/52 to standardize the two terms:

$$Hwage = [(\text{MONTHLYPAY}) / (\text{HOURS})] \times (12/52).$$

In the wage models hourly wage is in logarithm form to normalise wages.

IV.5 Covariates in wage function

The wage equation includes the individual characteristics from the bivariate probit models with the exceptions described in section III.1. The wage model also includes job experience and experience squared to control for on-the-job human capital accumulation. A binary variable indicating if an individual is employed in the private or public sector is included in the analysis. On average, private sector workers command a higher hourly wage than those in the public sector. If workers with a higher BMI are discriminated in the private sector

they may self-select into the public sector which could impact on their observed hourly wage.

To control for the effect of occupation on wage a categorical variable which is divided into the five broad groups of (1) professional; (2) manager; (3) skilled; (4) semi-skilled; (5) unskilled and other, based upon the Standard Occupational Classification (2000) system, is included in the wage function. Firm size and if the respondent is employed full or part time is also controlled for in the wage equations.

As was the case with the participation equation, some models are estimated which do not include the health variable to determine how the relationship between health and BMI impacts on wages.

A full list of the variables included in the double selection framework are shown in Appendix A.

IV.6 Bias

There are a number of types of bias likely to arise in the empirical analysis. BMI, health, and wages are self-reported which may lead to measurement error and bias. The fixed effect wage models remove individual heterogeneity reducing the bias caused by measurement error. Firstly, focusing on potential bias in the BMI variable, values were compared with more objective measures of BMI from other UK surveys such as the Health Survey of England where height and weight are measured by a healthcare professional. Similar mean BMI measures are found in both studies indicating that self-reported BMI measures should not bias the analysis. For a US population, Kuckmarzi *et al.* (2001) found that a continuous self-reported BMI measure may be more accurate than indicator variables for overweight and obese if people tend to miss-report height and weight if they are on the boundaries of the three BMI categories. As a robustness check on the results, all models are estimated twice; once with the indicator variables for overweight and obese and secondly with the continuous BMI variable.

Next to ensure accuracy of the derived variable for hourly wage, this variable was compared with wage information from the Office of National Statistics (ONS) over the same time period. Similar mean hourly wages were found in the two data sets.⁵ Finally, to reduce bias on the health coefficient, some model specifications were estimated with specific health conditions that are more objective than self-reported health.

It is possible that individuals sharing similar characteristics may choose to not report their height and weight. A systematic relationship between BMI, health, marital status, and labour market outcomes may result in attrition bias. To test for attrition bias a simple test proposed by Verbeek and Nijman (1992) is estimated. Two test variables are used: (1) If the couple was present in both waves 14 and 16; (2) If the couple was present in wave 14 the likelihood that they would be present in wave 16. These test variables are regressed together with a set of conditioning variables controlling for marital status, labour market participation, and other individual characteristics on the male and female obese indicator variables using a random effects probit model. A Wald test is performed to determine if the test coefficients are equal to zero. Results indicate that attrition should not severely bias the results.

Finally, BMI and health may be endogenously related to labour market outcomes and marital status. The obesity literature has found mixed evidence on the endogeneity of BMI on labour market outcomes. It is partially a problem of weak instruments. For example,

⁵Comparison tables are available upon request.

Cawley (2004) using sibling weight as an instrument finds no evidence of the endogeneity of BMI on wages. Morris (2006, 2007) finds some evidence of the endogeneity of BMI on female labour market participation but no evidence of the endogeneity of BMI on wages for either gender using mean BMI and the prevalence of obesity in the local authority district where the respondent lives as instruments. The main focus of this paper is how unobserved characteristics related to labour market participation and marriage influence the impact of BMI on wages. To avoid extra complications, it is assumed that endogeneity should not severely bias the results. The health variable is also excluded from some models to reduce the potential endogeneity bias caused by this variable.

IV.7 Descriptive statistics

Descriptive statistics for the full list of explanatory variables included in the analysis are shown in Table 1. The sample is divided by marital status and gender. The single sample is younger than the married sample, has a lower average BMI, and less respondents in the obese BMI category. Married respondents have a higher mean hourly wage and more accumulated job experience in their current position. Thus, if marriage is associated with a higher BMI, higher levels of job specific human capital, and a higher mean hourly wage, not accounting for marital status in the wage equation may bias the coefficient on BMI upwards for both genders.

V Results and Discussion

The double selection framework is used as an estimation strategy to determine the role of participation in the marriage and labour market on the impact of BMI on wages. The analysis is divided by marital status and gender. In all of the models estimated; * indicates significance at the 10% level, ** indicates significance at the 5% level and *** indicates significance at the 1% level. All models are estimated with and without health coefficients to try and disentangle the relationship between health and BMI on wages. In the wage models, the first four columns show the models estimated with the health coefficients. Columns 1 and 2 are for single respondents and columns 3 and 4 are the results for married respondents. The last four columns (columns 5–8) show the models estimated without the health coefficients. Columns 5 and 6 show the results for single respondents and columns 7 and 8 show the findings for married respondents. Random and fixed effects models are estimated to test for the role of time invariant unobservables on the impact of BMI on wages. The random effects models are presented in the odd columns and the fixed effect models are shown in the even columns. The results generated from the models estimated with the dummy variables for overweight and obese are similar to those from the continuous BMI variable. As discussed in section IV.7, Kuckmarzi *et al.* (2001) found that a self-reported continuous BMI measure was more accurate than dummy variables for overweight and obese. Therefore, for ease of exposition and to reduce the potential bias from using self-reported measures, only models estimated with the continuous BMI variable are shown.⁶

The first model presented is a wage equation that does not control for selection into the marriage or labour market in Tables 2A and 2B for men and women respectively. Next, to determine how controlling for selection into the labour market effects the results from the wage equation a single selection model is estimated in Tables 3A and 3B for men Tables 4A and 4B for women. The double selection model controlling for selection into the labour and marriage market is shown in Tables 5A and 5B for men and Tables 6A and 6B for women. To further disentangle the influence of marital status on wages, results from a Blinder–Oaxaca decomposition are presented in Tables 7A and 7B for men and women, respectively.

⁶Results from the models with the indicator variables for overweight and obese are available upon request.

Starting with the results from the wage model that does not control for selection, Table 2A shows the results for men. BMI has a positive effect on wages for single and married men in the random effects specification in columns 1, 3, 5, and 7. These findings suggest that unobserved time invariant effects such as motivation or genetic endowment may influence the relationship between BMI and wages for men. Morris (2006) shows that under certain econometric specifications BMI has a positive and significant effect on male earnings in the United Kingdom.

The wage equation that does not control for selection into the labour and marriage market for women is shown in Table 2B. BMI has a small, positive, and significant effect on wages in the random effects specification in columns 1 and 5 for single women. These findings are consistent with the hypothesis that single women with a higher BMI who are unsuccessful in the marriage market may have a greater attachment to the labour market resulting in higher hourly wages. Oreffice and Quintana-Domeque (2009) also find no significant effect of BMI on the log of earnings for married women.

V.1 Controlling for selection into the labour market

To control for selection into the labour market a univariate probit for participation is estimated in Tables 3A and 4A for men and women, respectively. The tables show the marginal effects for the coefficients. The inverse Mill's ratio calculated from the probit equation is added as an additional explanatory variable to the wage equation for men in Table 3B and women in Table 4B.

Results from the participation equation in Tables 3A and 4A for men and women, respectively, indicate that better health has a positive and significant impact on the likelihood of participating in the labour market for both genders in the random effects and the Mundlak probit models (columns 1 and 2). These findings are consistent with the health and labour market literature (Currie and Madrian, 1999). In the participation models, the sign of the BMI coefficient changes from positive to negative when moving from the model specification which includes health to the models which only have BMI measures. However, the BMI coefficients are only significant for women in those models which do not include health, in columns 3 and 4 in Table 4A. This suggests that the combined effect of obesity and health may influence the participation decision for women. These results are consistent with Morris (2007) who found that health influences the impact of obesity on employment status for women.

The wage models that account for selection into the labour market are shown in Tables 3B and 4B for men and women, respectively. The inverse Mill's ratio in these models are negatively significantly in all model specifications for both genders in Tables 3B and 4B. This implies that there are unobserved characteristics which increase the probability of selection into the labour market and increase the probability of a lower than average score on the dependent variable (wages). For both genders, the magnitude of the BMI coefficients in Table 3B for men and Table 4B for women are smaller than those from the wage model that does not control for selection into the labour market in Tables 2A for men and Table 2B for women suggesting that not accounting for selection into the labour market leads to an upward bias on the BMI coefficient.

BMI has a positive and significant effect on wages in all random effects model specifications (columns 1, 3, 5, and 7) for both single and married men and in the fixed effect model specifications for married men (columns 4 and 8). These results suggest that after controlling for unobserved time invariant individual effects married men may possess other time varying unobserved characteristics that are valued in the labour market irrespective of their BMI. For single women in columns 1, and 5 in Table 5B, BMI has a

positive and significant effect on wages. These findings add further evidence to suggest that the wage premium for single women with a higher BMI is due to their greater attachment to the labour market resulting from less time and family commitments. For both genders, in the wage models controlling for selection into the labour market (Tables 3B and 4B), health does not have a significant impact on wages.

V.2 Controlling for selection into marriage and labour

The results from the bivariate probit model which jointly estimates participation in the labour and marriage market are presented in Tables 5A and 6A for men and women, respectively. The marginal effects for the coefficients are shown. A separate inverse Mill's ratio is calculated for the participation and marriage equation to control for participation in both these markets. The inverse Mill's ratios are then added as additional explanatory variables to the wage model in Tables 5B and 6B for men and women, respectively.

For both genders in the bivariate probit models in Tables 5A and 6A, better health has a positive and significant effect on participation in the labour market. The health coefficients do not have a significant impact on the likelihood of being married for either gender. In Table 5A for men, the BMI variable does not have a significant impact on labour market participation or marriage in all model specifications. This implies that BMI does not impact on selection into the marriage or labour market for men. For women, in Table 6A, BMI has a significant and negative impact on labour market participation in columns 5 and 6, which do not include the health variable. In columns 3, 4, 7, and 8 the continuous BMI variable has a negative and significant impact on the likelihood of being married. This suggests that physical appearance is an important signal on the marriage market for women. Thinner women appear to be more successful in the marriage market which is consistent with expectations. *Rho* is positive and significant for both genders indicating that married respondents are more likely to participate in the labour market. The wage equations that control for selection into the labour and marriage market are presented in Table 5B for men and Table 6B for women. Comparing the results from the wage models controlling for selection into the marriage and labour market to the wage equations that control for selection into the labour market only (Table 3B for men and Table 4B for women) and the wage models that do not control for any type of selection (Tables 2A and 2B for men and women, respectively) the coefficient on BMI is smaller in the wage model controlling for selection into the marriage and labour market for both genders. This suggests that not controlling for both selection into the labour market and marriage may result in an upward bias on the BMI coefficients. The inverse Mill's ratios for marital status and labour market participation are significant and negative for both genders in Tables 6B and 7B in columns 1, 2, 5, and 6.

The continuous BMI variable has a positive and significant impact on wages for married men only in the random effects specification in columns 3 and 7. These results differ from the wage equations in Tables 3A and 4B where BMI had a positive and significant effect on wages for both single and married men in the random effects specification. These findings suggest that unobserved characteristics related to marriage may influence wages for men. These results are similar to Oreffice and Quintana-Domeque (2009) who find a BMI wage premium for married men. These results confirm our hypothesis that married men may possess characteristics which are valued both in the labour and marriage market irrespective of physical appearance resulting in a higher wage for all married men.

For women, the sign of the coefficients on the BMI variables changed from positive for single women in columns 1 and 5 in Tables 3B and 6B to negative and insignificant in Table 7B. These findings are consistent with much of the obesity and wages literature that found a negative effect of BMI on wages for women (e.g. Sargent and Blanchflower, 1994; Baum and Ford, 2004; Morris, 2007). The results are inconsistent with the hypotheses of a

compensating effect by single women with a higher BMI. The positive effect of marriage on wage reflected by the inverse Mill's ratio in the female equations also suggest single women may not have an advantage in the labour market.

Health does not have a significant impact on wages for men in Table 5B. For married women in column 3 in Table 6B, being in good health has a positive and significant impact on wages.

V.3 Blinder–Oaxaca decomposition

The final step in investigating the role of marriage on the impact of BMI on wages is the Blinder–Oaxaca decomposition to investigate how unobserved characteristics influence the wage rate. The sign and significance of the unexplained component will indicate if single men and/or married women are discriminated against in the labour market influencing the impact of BMI on wages. The results from the decomposition are presented in Tables 7A and 7B for men and women, respectively.

For men in Table 7A, single men earn 0.16 less than married men because of discrimination in the double selection model. The magnitude of the unexplained component of health and BMI coefficients is smaller in the models controlling for selection into the labour market and selection into the marriage and labour market. This suggests that unobserved characteristics related to success in the labour and marriage market influences the impact of BMI on wages.

In Table 7B, for women, the total unexplained wage component is negative in the single and double selection equation compared with the equation which does not control for selection where the total unexplained coefficient is positive. This indicates that married women face a wage penalty. This may explain the wage premium for single heavier women observed in the base wage model (Table 2B) and the wage model that controls for only selection into labour (Table 4B).

Overall, these results imply that single men and married women face wage penalties. This suggests that not controlling for selection into the marriage and labour market may lead to biased coefficients on the BMI variables in the wage equation.

VI Conclusion

This paper develops a marriage market hypothesis to try and explain some of the mixed results concerning the direction of the effect of BMI on wages found in the earlier literature (Sargent and Blanchflower, 1994; Baum and Ford, 2004; Cawley, 2004; Morris, 2006, 2007; Han *et al.*, 2009). If changing social norms have reduced the stigma of obesity, then it is expected that a higher BMI may not be penalised in the labour and/or marriage market. Social norms regarding weight are likely to be gender dependent.

A double selection framework is estimated to determine how selection into the labour and marriage markets influence the impact of BMI on wages for a married and single sample. These results are compared with a standard wage equation and wage model only controlling for selection into the labour market. Results indicate that the role marriage plays on the effect of BMI on wages differs by gender. In the standard wage equation and the model controlling for only selection into the labour market there is a wage premium on increasing BMI for single men and women, as well as married men. After controlling for selection into the marriage and labour market there is a wage premium for only married men. These findings suggest that not controlling for unobserved characteristics related to marriage participation may lead to an upward bias on the BMI coefficients in the simple wage model

and the wage equation which only controls for selection into the labour market. Further evidence from a Blinder–Oaxaca decomposition indicates that there is discrimination against single men and married women in the labour market.

These findings indicate that as the mean BMI rates continue to rise the social stigma of having a higher BMI declines. This implies that the outcomes of individuals with a higher BMI will not be different from those with a lower BMI if they possess other characteristics that are valued in the market. Future work on the impact of obesity needs to account for changing social norms influencing the impact of BMI on the outcome of interest. Social norms will impact the effectiveness of public policies to reduce obesity rates. Individuals may be less likely to change their behaviour if there is no social penalty of having a higher BMI.

Acknowledgments

This work was funded by the MRC/ESRC/NIHR under grant G0802291.

Appendix

Appendix A:

Table A1

Variable Labels and Definitions

Variable name	Description	BHPS Code
<i>Dependent variable</i>		
<i>Employment Status</i>		
Jobstatus	0- Working (Employed and Self-Employed) 1-Unemployed	JBSTAT
<i>Hourly wage</i>		
Maritalstatus	0 = never married 1 = married/cohabiting	
Lhwage	Log of hourly wage: $\text{hourlywage} = \frac{\text{MONTHLYPAY}}{\text{HOURS}} \times (12/52)$	MONTHLYPAY = PAYGU HOURS = JBHRS
<i>Explanatory variables</i>		
Age	Age in years	AGE
Agesquared	Age squared	AGE
Sex	Gender	SEX
<i>Marital Status</i>		
<i>Pregnant</i>		
Preg	0 = Not Pregnant 1 = Pregnant	HLPREG
<i>Region</i>		
Area	1 = Greater London 2 = Southeast 3 = Southwest 4 = Midlands 5 = Northwest 6 = Yorks, Humberside, & Rest of North	REGION

Variable name	Description	BHPS Code
	7 = Wales	
	8 = Scotland	
	9 = Northern Ireland	
	<i>Education</i>	
Education	0 = No qualifications	QFACHI
	1 = CSE or O level	
	2 = HND, HNC, teaching, or A-level	
	3 = Higher or First Degree	
	<i>Children</i>	
Preschoolkids	0 = No children aged between zero and four	NCH02+NCH04
	1 = Has at least one child aged zero to four	
Schoolagekids	0 = Has no children 5–11 or children 16+ at home	NCH511+NCH1618
	1 = Has at least one child 5–11 or children 16+ at home	
Nonlabourincome	Non-labour income: HOUSEHOLD INCOME minus ANNUALLABOURINCOME	FIHHYR = ANNUAL HOUSEHOLD INCOME HHSIZE = HOUSEHOLD SIZE FIHHYR/HHSIZE = HOUSEHOLDINCOME FIYRL = ANNUAL LABOUR INCOME NONLABOURINCOME = HOUSEHOLDINCOME-FIYRL
	<i>Self-assessed health</i>	
Health	0 = Very Poor	HLSTAT
	1 = Poor	
	2 = Fair	
	3 = Good	
	4 = Excellent	
	<i>BMI</i>	
BMI	Body Mass Index measured in kilograms divided by meters squared	HLHTF+HLHTI = HEIGHT HLWTS+HLWTP = WEIGHT BMII = (WEIGHT/ HEIGHT^2) × 703
Healthyweight	1 = 18.5–25 kg/m ² , 0 = Otherwise	BMI
Overweight	1 = 25–30 kg/m ² , 0 = Otherwise	BMI
Obese	1 = 30+ kg/m ² , 0 = Otherwise	BMI
	<i>Job Tenure</i>	
Jobexp	Number of years at current job:	JBBGY4 (2007 minus year current job started)
Expsquared	Square of JOBEXP	JBBGY4^2
	<i>Job Sector</i>	
Sectpr	0 = Public Sector	JBSECT
	1 = Private Sector	
	<i>Firm Size</i>	
Jobsizes	1 = Between 1 and 25 employees	JBSIZE
	2 = Between 25 and 100 employees	

Variable name	Description	BHPS Code
	3 = Between 100 and 500 employees	
	4 = Between 500 and 1000 employees	
	5 = 1000 or more employees	
	<i>Full/PT Employment</i>	
Ftptemployment	0 = Part time	JBFT
	1 = Full time	
	<i>Occupation</i>	
Occupation	1 = Professional (employers large, employers small, professional, self-employed and professional employed)	JBSEG
	2 = Manager (managers large, managers small, int non-manual foreman, and foreman manual.)	
	3 = Skilled (skilled manual workers)	
	4 = Semi-skilled (int. non manual work, junior non-manual personal service worker, semi-skilled manual workers and own account workers.)	
	5 = Unskilled/Other (unskilled manual workers, farmers employer, agricultural workers, farmers on won account and members of armed forces.)	

References

- Bardasi, E.; Taylor, M. Wages and marriage. 2005. ISER Working Papers Number 2005-1
- Baum CL, Ford WF. The wage effects of obesity: a longitudinal study. *Health Economics*. 2004; 13:885–99. [PubMed: 15362180]
- Becker GS. A theory of marriage: part II. *The Journal of Political Economy*. 1974; 82:S11–26.
- Bhattacharya, J.; Bundorf, MK. The incidence of the health care costs of obesity. 2005. National Bureau of Economic Research Working Paper No. 11303
- Billari F, Borgoni R. Assessing the use of sample selection models in the estimation of fertility postponement effects. *Statistical Methods and Applications*. 2005; 14:389–402.
- Blinder AS. Wage discrimination: reduced form and structural estimates. *The Journal of Human Resources*. 1973; 8:436–55.
- Brown S, Roberts J, Taylor K. Reservation wages, labour market participation, and health. *Journal of the Royal Statistical Society Series A*. 2010; 173:501–29.
- Brown S, Taylor K. Wage growth, human capital and financial investment. *The Manchester School*. 2005; 73:686–708.
- Burke MA, Heiland F. Social dynamics of obesity. *Economic Inquiry*. 2007; 45:571–9.
- Cawley J. The impact of obesity on wages. *The Journal of Human Resources*. 2004; 39:451–74.
- Chiappori PA, Oreffice S, Quintana-Domeque C. Fatter attraction: anthropometric and socioeconomic characteristics in the labour market. *IZA DP*. 2009:4594.
- Currie, J.; Madrian, BC. Health, health insurance and the labour market. In: Ashenfelter, O.; Card, D., editors. *Handbook of Labour Economics*. Elsevier; Amsterdam: 1999.
- Ebbes P, Bockenholt U, Wedel M. Regressor and random-effects dependencies in multilevel models. *Statistica Neerlandica*. 2004; 58:161–78.
- Garner DM, Garfinkel PE, Schwartz D, Thompson M. Cultural expectations of thinness in women. *Psychological Reports*. 1980; 47:183–91.
- Greene, WH. *Econometric Analysis*. Prentice Hall; Upper Saddle River, NJ: 2003.
- Han E, Norton EC, Stearns SC. Weight and wages: fat versus lean paychecks. *Health Economics*. 2009; 18:535–48. [PubMed: 18677723]

- Heckman J. The common structure of statistical models of truncation, sample selection, and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement*. 1976; 5:475–92.
- Heckman J. Sample selection bias as a specification error. *Econometrica*. 1979; 47:153–61.
- Kuckmarzi RJ, Fanelli M, Najjar T. Effects of age on validity of self-reported height, weight and Body Mass Index: findings from the Third National Health and Nutrition Examination Survey, 1998–2004. *Journal of American Dietetic Association*. 2001; 101:28–34.
- Maddala, GS. *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge University Press; New York: 1983. *Econometric Society Monographs*
- Mazur A. U.S. trends in feminine beauty and overadaptation. *Journal of Sex Research*. 1986; 22:281–303.
- Mincer, J. *Schooling, Experience and Earnings*. NBER Press; New York: 1979.
- Morris S. Body Mass Index and occupational attainment. *Journal of Health Economics*. 2006; 25:347–64. [PubMed: 16290234]
- Morris S. The impact of obesity on employment. *Labour Economics*. 2007; 14:413–3.
- Mundlak Y. On the pooling of time series and cross-section data. *Econometrica*. 1978; 46:69–85.
- NHS: The Information Centre. [accessed 13 September 2009] Body Mass Index (BMI), by survey year, age, and sex. 2008. Available at <http://www.ic.nhs.uk/webfiles/publications/hseupdate08/HSE2008-TrendTables.xls>
- Oaxaca RL. Male–female wage differentials in urban labour markets. *International Economic Review*. 1973; 14:693–709.
- Oreffice, S.; Quintana-Domeque, C. Fatter attraction: marital status and the relationship between BMI and labour supply. 2009. Working Paper
- Oswald AJ, Powdthavee N. Obesity, unhappiness, and the challenge of affluence: theory and evidence. *Economic Journal*. 2007; 117:441–54.
- Sargent JD, Blanchflower DG. Obesity and stature in adolescence and earnings in young adulthood. *Archives of Paediatric and Adolescent Medicine*. 1994; 148:681–7.
- Verbeek M, Nijman T. Testing for selectivity bias in panel data models. *International Economic Review*. 1992; 33:681–703.
- Wetzels, C.; Zorlu, A. Wage effects of motherhood: a double selection approach. Núcleo de Investigação em Microeconomia Aplicada (NIMA). 2003. Universidade do Minho Working Paper No. 22
- Wooldridge, J. *Econometric Analysis of Cross-Section and Panel Data*. MIT Press; Cambridge, MA: 2002.
- World Health Organisation (WHO). [accessed 29 October 2008] BMI classification. 2004. 1995, 2000 Available at http://www.who.int/bmi/index.jsp?introPage=intro_3.html

Table 1

Descriptive Statistics

Variables	Men		Women	
	Single	Married	Single	Married
Job status: unemployed	0.14 (251)	0.04 (230)	0.09 (152)	0.03 (155)
Working	0.86 (1548)	0.96 (5138)	0.90 (1381)	0.97 (5208)
Hourly wage	8.55 (1756)	13.11 (5043)	8.11 (1742)	10.04 (5254)
Age	29.33 (2422)	44.03 (7307)	28.63 (2276)	43.49 (8282)
Region: Greater London	0.05 (167)	0.04 (372)	0.06 (183)	0.04 (378)
Southeast and East of England	0.13 (399)	0.14 (1248)	0.12 (388)	0.14 (1341)
Southwest	0.05 (145)	0.06 (533)	0.04 (134)	0.06 (566)
Midlands	0.09 (265)	0.11 (901)	0.08 (247)	0.10 (941)
Northwest	0.06 (171)	0.06 (534)	0.06 (186)	0.06 (565)
Yorks, Humberside, and rest of North	0.09 (302)	0.08 (805)	0.07 (224)	0.08 (857)
Wales	0.15 (47)	0.17 (1533)	0.16 (491)	0.17 (1639)
Scotland	0.18 (544)	0.17 (1501)	0.16 (503)	0.17 (1600)
Northern Ireland	0.19 (567)	0.16 (1384)	0.23 (699)	0.16 (1517)
Education: No qualifications	0.18 (543)	0.27 (2405)	0.16 (498)	0.29 (2687)
CSE or O-Level	0.37 (1123)	0.27 (2370)	0.39 (1171)	0.32 (3022)
HND, HNC, teaching or A-level	0.32 (974)	0.30 (2593)	0.31 (930)	0.25 (2312)
First or Higher degree	0.13 (398)	0.16 (1406)	0.14 (431)	0.15 (1387)
Children: Preschool kids	0.02 (3090)	0.15 (8926)	0.09 (3080)	0.14 (9524)
School age kids	0.18 (3090)	0.25 (8926)	0.27 (3080)	0.27 (9524)
Non-labour income	4672.59 (3090)	3270.18 (8924)	5063.95 (3080)	5672.39 (9523)
Job Experience (years)	4.86 (1821)	7.52 (5124)	4.20 (1804)	7.19 (5321)
Sector: Public	0.16 (285)	0.20 (1030)	0.22 (407)	0.40 (2140)
Private	0.84 (1543)	0.80 (4116)	0.78 (1407)	0.60 (3205)
Job Size: Between 1 and 25 employees	0.22 (401)	0.18 (907)	0.24 (422)	0.22 (1161)
Between 25 and 100 employees	0.42 (764)	0.38 (1962)	0.46 (827)	0.43 (2227)
Between 100 and 500 employees	0.19 (346)	0.25 (1286)	0.16 (297)	0.18 (967)
Between 500 and 1000 employees	0.09 (157)	0.08 (430)	0.06 (100)	0.07 (359)
1000 or more employees	0.09 (155)	0.11 (551)	0.09 (168)	0.11 (573)
Hours Worked: Part time	0.20 (394)	0.07 (449)	0.37 (692)	0.41 (2328)
Full time	0.80 (1615)	0.93 (5863)	0.63 (1179)	0.60 (3407)
Occupation: Professional	0.05 (99)	0.12 (731)	0.04 (69)	0.05 (299)
Manager	0.153 (271)	0.27 (1679)	0.12 (233)	0.18 (1038)
Skilled	0.15 (306)	0.16 (1032)	0.01 (13)	0.01 (67)
Semi-skilled	0.59 (1187)	0.41 (2566)	0.80 (1516)	0.72 (4136)
Unskilled/Other	0.07 (154)	0.05 (302)	0.03 (54)	0.03 (198)
Health: Very Poor	0.01 (29)	0.02 (145)	0.02 (50)	0.02 (183)
Poor	0.04 (131)	0.06 (549)	0.06 (179)	0.08 (751)
Fair	0.17 (529)	0.19 (1726)	0.18 (538)	0.19 (1961)

Variables	Men		Women	
	Single	Married	Single	Married
Good	0.48 (1478)	0.48 (4237)	0.48 (1469)	0.46 (4410)
Excellent	0.30 (920)	0.25 (2269)	0.27 (842)	0.24 (2215)
BMI	24.80 (2528)	26.99 (7956)	24.55 (2112)	26.70 (7250)
Healthy weight	0.56 (2528)	0.34 (7956)	0.56 (2112)	0.40 (7250)
Overweight	0.29 (2528)	0.46 (7956)	0.25 (2112)	0.35 (7250)
Obese	0.12 (2528)	0.20 (7956)	0.13 (2112)	0.23 (7250)
<i>n</i> (marital status)	3090	8926	3080	9524

Note: () show sample size. With the exception of age, job experience, hourly wage, non-labour income, BMI, and the three weight groups all variables shown are percentages.

Table 2A

Wages (Men)

Wages	Single (RE)	Single (FE)	Married (RE)	Married (FE)	Single (RE)	Single (FE)	Married (RE)	Married (FE)
Health states	Column 1	Column 2	Column 3	Column 4	Column 5	Column 6	Column 7	Column 8
Poor	0.18 (0.19)	0.12 (0.42)	0.06 (0.09)	-0.03 (0.11)				
Fair	0.18 (0.17)	0.07 (0.40)	0.07 (0.08)	0.01 (0.10)				
Good	0.15 (0.17)	0.03 (0.40)	0.12 (0.08)	0.02 (0.10)				
Excellent	0.11 (0.17)	0.001 (0.40)	0.15 (0.08)	0.02 (0.10)				
BMI	0.02 *** (0.02)	0.01 (0.01)	0.01 *** (0.002)	0.01 (0.03)	0.01 *** (0.002)	0.01 (0.01)	0.001 *** (0.002)	0.006 (0.003)
<i>n</i>	1176	1176	4152	4152	1176	1176	4152	4152

Note: The other control variables in the equation are job experience, experience squared, job size, region, occupation, sector, and full or part-time employment.

Standard errors are in parentheses.

Significant at 1% level.

**

Significant at 5% level.

*

Significant at 10% level.

Table 2B

Wages (Women)

Wages	Single (RE)	Single (FE)	Married (RE)	Married (FE)	Single (RE)	Single (FE)	Married (RE)	Married (FE)
Health states	Column 1	Column 2	Column 3	Column 4	Column 5	Column 6	Column 7	Column 8
Poor	-0.06 (0.16)	0.04 (0.32)	0.001 (0.06)	0.02 (0.08)				
Fair	0.02 (0.16)	0.05 (0.31)	0.02 (0.06)	0.01 (0.08)				
Good	0.02 (0.06)	0.03 (0.30)	0.02 (0.06)	0.04 (0.07)				
Excellent	0.03 (0.16)	0.13 (0.31)	0.03 (0.06)	0.02 (0.08)				
BMI	0.01 *** (0.001)	-0.003 (0.01)	0.002 (0.002)	0.001 (0.005)	0.01 *** (0.002)	-0.003 (0.01)	0.002 (0.001)	0.001 (0.004)
<i>n</i>	891	891	3643	3643	891	891	3643	3643

Note: The other control variables in the equation are job experience, experience squared, job size, region, occupation, sector, and full or part-time employment.

Standard errors are in parentheses.

Significant at 1% level.

**

Significant at 5% level.

*

Significant at 10% level.

Table 3A**Labour Market Participation (Men)**

Participation	RE Probit	Mundlak Probit	RE Probitmfx (Column 3)	Mundlak Probit
Health states	mfx (Column 1)	mfx (Column 2)	mfx (Column 2)	mfx (Column 4)
Poor	0.16 *** (0.03)	0.07 *** (0.04)		
Fair	0.38 *** (0.03)	0.36 *** (0.04)		
Good	0.67 *** (0.04)	0.64 *** (0.04)		
Excellent	0.53 *** (0.06)	0.56 *** (0.08)		
BMI	0.01 (0.01)	0.001 (0.01)	−0.02 (0.01)	−0.02 (0.02)
Log likelihood	−1066.8513	−1066.8513	−1722.7649	−1722.7649

Note: The other control variables in the equation are regional dummies, age, age squared, education, non-labour incomes, and dummy variables for preschool and school age children.

Standard errors are in parentheses.

Significant at 1% level.

**
Significant at 5% level.

*
Significant at 10% level.

Table 3B

Wage Equation Accounting for Selection into Labour (Men)

Wages	Single (RE)	Single (FE)	Married (RE)	Married (FE)	Single (RE)	Single (FE)	Married (RE)	Married (FE)
Health states	Column 1	Column 2	Column 3	Column 4	Column 5	Column 6	Column 7	Column 8
Poor	0.15 (0.15)	0.16 (0.18)	0.09 (0.12)	0.04 (0.12)				
Fair	0.12 (0.16)	0.01 (0.22)	0.13 (0.14)	0.01 (0.14)				
Good	0.08 (0.17)	0.03 (0.16)	0.18 (0.14)	0.02 (0.14)				
Excellent	0.05 (0.17)	0.03 (0.16)	0.21 (0.14)	0.02 (0.14)				
BMI	0.01 *** (0.003)	-0.002 (0.002)	0.01 *** (0.002)	0.01 *** (0.003)	0.01 *** (0.003)	-0.004 (0.01)	0.01 *** (0.002)	0.01 * (0.005)
Inv1	-0.11 *** (0.02)	-0.08 ** (0.02)	-0.14 *** (0.03)	-0.09 ** (0.03)	-0.17 *** (0.05)	-0.17 ** (0.05)	-0.15 *** (0.06)	-0.14 ** (0.06)
<i>n</i>	1430	1430	4296	4296	1430	1430	4296	4296

Note: The other control variables in the equation are job experience, experience squared, job size, region, occupation, sector, and full or part-time employment.

Standard errors are in parentheses.

*** Significant at 1% level.

** Significant at 5% level.

* Significant at 10% level.

Table 4A**Labour Market Participation (Women)**

Participation	RE Probit	Mundlak Probit	RE Probit	Mundlak Probit
Health States	mfx (Column 1)	mfx (Column 2)	mfx (Column 3)	mfx (Column 4)
Poor	0.08 [*] (0.03)	0.08 ^{**} (0.04)		
Fair	0.34 ^{***} (0.08)	0.39 ^{***} (0.04)		
Good	0.30 ^{***} (0.08)	0.24 ^{***} (0.04)		
Excellent	0.66 ^{***} (0.04)	0.68 ^{***} (0.04)		
BMI	−0.01 (0.01)	−0.001 (0.01)	−0.04 ^{**} (0.01)	−0.04 ^{**} (0.01)
Log likelihood	−716.13503	−716.13503	−712.5607	−712.5607

Note: The other control variables in the equation are regional dummies, age, age squared, education, non-labour incomes, and dummy variables for preschool and school age children.

Standard errors are in parentheses.

Significant at 1% level.

**
Significant at 5% level.

*
Significant at 10% level.

Table 4B

Wage Equation Accounting for Selection into Labour (Women)

Wages	Single (RE)	Single (FE)	Married (RE)	Married (FE)	Single (RE)	Single (FE)	Married (RE)	Married (FE)
Health states	Column 1	Column 2	Column 3	Column 4	Column 5	Column 6	Column 7	Column 8
Poor	-0.03 (0.11)	0.02 (0.14)	0.01 (0.07)	0.05 (0.04)				
Fair	0.11 (0.09)	0.04 (0.10)	0.004 (0.07)	0.001 (0.08)				
Good	0.09 (0.11)	0.03 (0.08)	0.003 (0.07)	0.001 (0.08)				
Excellent	0.14 (0.11)	0.04 (0.13)	0.05 (0.08)	0.02 (0.09)				
BMI	0.01 ** (0.003)	0.004 (0.01)	0.001 (0.002)	0.001 (0.004)	0.001 * (0.003)	0.004 (0.004)	0.001 (0.001)	0.001 (0.005)
Inv1	-0.07 *** (0.03)	-0.06 *** (0.03)	-0.05 *** (0.02)	-0.01 ** (0.04)	-0.05 *** (0.02)	-0.02 *** (0.01)	-0.02 *** (0.01)	-0.01 ** (0.001)
<i>n</i>	1182	1182	3802	3802	1182	1182	3803	3803

Note: The other control variables in the equation are job experience, experience squared, job size, region, occupation, sector, and full or part-time employment.

Standard errors are in parentheses.

*** Significant at 1% level.

** Significant at 5% level.

* Significant at 10% level.

Table 5A

Labour Market Participation and Marriage (Men)

Health states	RE Probit		Mundlak Probit		RE Probit		Mundlak Probit		RE Probit		Mundlak Probit	
	Participation	marriage	Participation	marriage	Participation	marriage	Participation	marriage	Participation	marriage	Participation	marriage
	mf (Column 1)	mf (Column 2)	mf (Column 3)	mf (Column 4)	mf (Column 5)	mf (Column 6)	mf (Column 7)	mf (Column 8)	mf (Column 9)	mf (Column 10)	mf (Column 11)	mf (Column 12)
Poor	0.05 (0.04)	0.08 (0.04)	0.03 (0.02)	0.05 (0.03)								
Fair	0.09 (0.04)	0.14 *** (0.04)	0.10 (0.06)	0.12 (0.08)								
Good	0.17 *** (0.04)	0.24 *** (0.04)	0.09 (0.02)	0.10 (0.02)								
Excellent	0.18 *** (0.04)	0.26 *** (0.04)	0.10 (0.02)	0.13 (0.02)								
BMI	0.003 (0.001)	0.003 (0.001)	0.02 (0.01)	0.02 (0.01)	0.001 (0.001)	0.002 (0.001)	0.001 (0.006)	0.001 (0.006)				
Log likelihood	-3568.793	-3568.793	-3568.793	-3568.793	-3786.9304	-3786.9304	-3786.9304	-3786.9304				
ρ	0.21 *** (0.04)	0.20 *** (0.04)	0.21 *** (0.04)	0.20 *** (0.04)	0.21 *** (0.04)	0.20 *** (0.04)	0.21 *** (0.04)	0.20 *** (0.04)				

Note: The other control variables in the equation are regional dummies, age, age squared, education, non-labour incomes, and dummy variables for preschool and school age children.

Standard errors are in parentheses.

*** Significant at 1% level.

** Significant at 5% level.

* Significant at 10% level.

Table 5B

Wage Equation Accounting for Selection into Labour and Marriage (Men)

Wages	Single (RE)	Single (FE)	Married (RE)	Married (FE)	Single (RE)	Single (FE)	Married (RE)	Married (FE)
Health states	Column 1	Column 2	Column 3	Column 4	Column 5	Column 6	Column 7	Column 8
Poor	0.11 (0.10)	0.08 (0.37)	0.12 (0.09)	0.19 (0.09)				
Fair	0.05 (0.11)	0.03 (0.35)	0.19 (0.09)	0.22 (0.12)				
Good	0.01 (0.12)	0.15 (0.35)	0.19 (0.09)	0.26 (0.13)				
Excellent	0.01 (0.13)	0.10 (0.36)	0.16 (0.09)	0.26 (0.13)				
BMI	0.003 (0.002)	-0.001 (0.001)	0.004** (0.002)	-0.01 (0.01)	0.002 (0.002)	-0.01 (0.01)	0.005** (0.001)	0.01 (0.006)
Inv1	-0.11* (0.02)	-0.42*** (0.11)	-0.31*** (0.08)	-0.46*** (0.11)	-0.10* (0.02)	-0.42*** (0.09)	-0.34*** (0.08)	-0.30*** (0.07)
Inv2	-0.29*** (0.04)	-0.56*** (0.11)	-	-	-0.28*** (0.04)	-0.60*** (0.11)	-	-
<i>n</i>	1430	1430	4296	4296	1430	1430	4296	4296

Note: The other control variables in the equation are job experience, experience squared, job size, region, occupation, sector, and full or part-time employment.

Standard errors are in parentheses.

Significant at 1% level.

**

Significant at 5% level.

*

Significant at 10% level.

Table 6A

Labour Market Participation and Marriage (Women)

Health states	RE Probit		Mundlak Probit		RE Probit		Mundlak Probit		RE Probit		Mundlak Probit	
	Participation	mfx (Column 1)	Participation	mfx (Column 2)	Marital Status	mfx (Column 3)	Marital Status	mfx (Column 4)	Participation	mfx (Column 5)	Participation	mfx (Column 6)
Poor	0.04 (0.04)		0.06 (0.04)		0.05 (0.07)		0.06 (0.07)					
Fair	0.14 *** (0.04)		0.18 *** (0.04)		0.09 (0.06)		0.09 (0.05)					
Good	0.24 *** (0.04)		0.29 *** (0.04)		0.14 (0.06)		0.16 (0.06)					
Excellent	0.27 *** (0.04)		0.32 *** (0.04)		0.13 (0.06)		0.11 (0.06)					
BMI	0.001 (0.001)		0.001 (0.001)		-0.01 ** (0.005)		-0.03 *** (0.001)		-0.001 ** (0.001)		-0.001 ** (0.01)	
Log likelihood	-2451.8364		-2451.8364		-2451.8364		-2451.8364		-2475.8923		-2475.8923	
ρ	0.36 *** (0.06)		0.36 *** (0.06)		0.36 *** (0.06)		0.36 *** (0.06)		0.36 *** (0.06)		0.36 *** (0.06)	

Note: The other control variables in the equation are regional dummies, age, age squared, education, non-labour incomes, and dummy variables for preschool and school age children.

Standard errors are in parentheses.

*** Significant at 1% level.

** Significant at 5% level.

* Significant at 10% level.

Table 6B

Wage Equation Accounting for Selection into Labour and Marriage (Women)

Wages	Single (RE)	Single (FE)	Married (RE)	Married (FE)	Single (RE)	Single (FE)	Married (RE)	Married (FE)
Health states	Column 1	Column 2	Column 3	Column 4	Column 5	Column 6	Column 7	Column 8
Poor	0.07 (0.13)	0.19 (0.27)	0.02 (0.06)	0.01 (0.09)				
Fair	0.10 (0.12)	0.13 (0.26)	0.08 (0.06)	0.02 (0.10)				
Good	0.18 (0.13)	0.08 (0.26)	0.13* (0.05)	0.08 (0.10)				
Excellent	0.16 (0.13)	0.001 (0.26)	0.09 (0.05)	-0.07 (0.11)				
BMI	-0.001 (0.002)	-0.005 (0.01)	-0.002 (0.003)	-0.001 (0.005)	-0.0002 (0.002)	-0.0001 (0.01)	-0.001 (0.001)	-0.004 (0.04)
Inv1	-0.35*** (0.10)	-0.69*** (0.12)	-0.20*** (0.04)	-0.20*** (0.06)	-0.39*** (0.09)	-0.77*** (0.08)	-0.16*** (0.04)	-0.22*** (0.09)
Inv2	-0.41*** (0.04)	-0.83*** (0.15)	-	-	-0.41*** (0.05)	-0.82*** (0.18)	-	-
<i>n</i>	1182	1182	3802	3802	1182	1182	3803	3803

Note: The other control variables in the equation are job experience, experience squared, job size, region, occupation, sector, and full or part-time employment.

Standard errors are in parentheses.

Significant at 1% level.

**

Significant at 5% level.

*

Significant at 10% level.

Table 7A

Blinder-Oaxaca Decomposition (Men)

	No selection		Single selection		Double selection	
	Explained	Unexplained	Explained	Unexplained	Explained	Unexplained
Health	-0.002 (0.001)	0.39* (0.15)	-0.004 (0.001)	0.25 (0.18)	0.002 (0.002)	0.14 (0.18)
BMI	0.02*** (0.003)	-0.22** (0.07)	0.01*** (0.003)	-0.08 (0.07)	0.004 (0.002)	-0.02 (0.07)
Total	0.11*** (0.01)	0.29*** (0.01)	0.06*** (0.01)	0.21*** (0.02)	0.04*** (0.01)	0.16*** (0.04)
<i>n</i>	5726	5726	5926	5926	5926	5926

Note: The other control variables in the equation are job experience, experience squared, job size, region, occupation, sector, and full or part-time employment. In the single selection equation an inverse Mill's ratio estimated from the probit model in Table 4A is included to control for selection into labour. In the double selection equation and inverse Mill's ratio estimated from the bivariate probit model in Table 6A is included as an explanatory variable in the wage equation to control for selection into marriage and labour market participation.

Standard errors are in parentheses.

Significant at 1% level.

**

Significant at 5% level.

*

Significant at 10% level.

Table 7B

Blinder-Oaxaca Decomposition (Women)

	No selection		Single selection		Double selection	
	Explained	Unexplained	Explained	Unexplained	Explained	Unexplained
Health	-0.001 (0.001)	0.19 (0.14)	-0.004 (0.008)	1.15*** (0.23)	0.004 (0.001)	0.58*** (0.13)
BMI	0.01** (0.002)	-0.22 (0.07)	0.003 (0.002)	-0.17* (0.06)	0.002 (0.002)	0.01 (0.06)
Total	0.02 (0.01)	0.15*** (0.01)	-0.04*** (0.01)	0.11*** (0.02)	-0.05*** (0.01)	-0.20*** (0.04)
<i>n</i>	4535	4535	4920	4920	4920	4920

Note: The other control variables in the equation are job experience, experience squared, job size, region, occupation, sector, and full or part-time employment. In the single selection equation an inverse Mill's ratio estimated from the probit model in Table 5A is included to control for selection into labour. In the double selection equation and inverse Mill's ratio estimated from the bivariate probit model in Table 7A is included as an explanatory variable in the wage equation to control for selection into marriage and labour market participation.

Standard errors are in parentheses.

Significant at 1% level.

**

Significant at 5% level.

*

Significant at 10% level.