CEPS INSTEAD

WORKING PAPERS

Body size and wages in Europe: A semi-parametric analysis

Vincent HILDEBRAND[1]
Philippe VAN KERM[2]

[1] Glendon College, York University - Department of Economics , Canada and CEPS/INSTEAD, Luxembourg
[2] CEPS/INSTEAD, Luxembourg

Working Paper No 2010-09
June 2010

Body size and wages in Europe: A semi-parametric analysis*

Vincent A. Hildebrand†
Department of Economics, Glendon College, York University, Canada
and CEPS/INSTEAD, Luxembourg

Philippe Van Kerm‡
CEPS/INSTEAD, Luxembourg

May 2010

Abstract

Evidence of the association between wages and body size –typically measured by the body mass index– appears to be sensitive to estimation methods and samples, and varies across gender and ethnic groups. One factor that may contribute to this sensitivity is the non-linearity of the relationship. This paper analyzes data from the European Community Household Panel survey and uses semi-parametric techniques to avoid functional form assumptions and assess the relevance of standard models. If a linear model for women and a quadratic model for men fit the data relatively well, they are not entirely satisfactory and are statistically rejected in favour of semi-parametric models which identify patterns that none of the parametric specifications capture. Furthermore, when we use height and weight in the models directly, rather than equating body size with the body mass index, the semi-parametric models reveal a more complex picture with height having additional effects on wages. We interpret our results as consistent with the existence of a wage premium for physical attractiveness rather than a penalty for unhealthy weight.

Keywords: Body Mass Index ; obesity ; wages ; partial linear models ; ECHP

JEL classification codes: C14 ; J31 ; J71

*This research is part of the MeDIM project (*Advances in the Measurement of Discrimination, Inequality and Mobility*) supported by the Luxembourg 'Fonds National de la Recherche' (contract FNR/06/15/08) and by core funding for CEPS/INSTEAD from the Ministry of Culture, Higher Education and Research of Luxembourg. The authors would like to thank seminar participants at the Australian National University and Alessio Fusco, Eric Kam and Don Williams for useful comments.

†Vincent Hildebrand, Department of Economics, York Hall 363, Glendon College, York University, 2275 Bayview ave., Toronto, Ontario, M4N 3M6 Canada. E-mail: vincent@econ.yorku.ca.

‡Philippe Van Kerm, CEPS/INSTEAD, B.P.48, L-4501 Differdange, Luxembourg. E-mail: philippe.vankerm@ceps.lu.

1 Introduction

Since the seminal contribution of Register and Williams (1990), evidence of a wage penalty associated to excessive body weight has been repeatedly reported – in Denmark (Greve, 2008), England (Morris, 2006), Germany (Cawley et al., 2005), Sweden (Lundborg et al., 2010), the United States (Averett and Korenman, 1996; Pagan and Davila, 1997; Baum II and Ford, 2004; Cawley, 2004; Mocan and Tekin, 2009) and continental Europe as a whole (Brunello and d'Hombres, 2007; Atella et al., 2008; Villar and Quintana-Domeque, 2009). Other contributions have however failed to confirm the robustness of this negative association in Australia (Kortt and Leigh, 2010), the United States (Hamermesh and Biddle, 1994; Norton and Han, 2008) and across European countries (Fahr, 2006; Atella et al., 2008). Overall, findings from the literature suggest that the significance of this association differs widely across gender and ethnic groups (Averett and Korenman, 1996; Cawley, 2004; Han et al., 2009; Mocan and Tekin, 2009) and is sensitive to the choice of estimation methods and data considered.[1]

One factor that may account for the lack of robustness of these findings is the likely non-linearity of the relationship between body size and wages, which if not modeled appropriately may lead to attenuated or misleading associations (as if, e.g., one were fitting a linear regression through an inverted U-shape relationship). A number of pathways explaining how obesity may translate into lower wages (such as strict productivity arguments or personal prejudice factors) have been hypothesized and empirically tested (Baum II and Ford, 2004; Han et al., 2009; Lundborg et al., 2010), but the literature does not provide clear guidelines regarding the shape of this association. Most studies have relied on simple parametric wage models with the Body Mass Index (BMI) — defined as a person's weight (in kilograms) divided by her height (in meters) squared — included as a regressor along with human capital and job-related characteristics. The two most common specifications either assume a linear association between BMI and (the logarithm of) wages or allow for a non-linear relationship by categorizing BMI using conventional clinical thresholds to capture whether individual respondents are obese, overweight or underweight. The latter specification is intuitively appealing as it allows for differential wage effects of body mass for deviations above or below clinically recommended values. However, it may still suffer important shortcomings. The discretization of the BMI

[1]For instance, Han et al. (2009) recently reported a significant negative association for females in the US using 13 years of data from the NLSY79 which includes respondents aged 18–43 over the 1991–2000 period. In contrast, Norton and Han (2008) did not find any significant effects using the third wave of the NLSAH which covers respondents aged 18–26 in the 2001–2002 period. Furthermore, the significance of the association may also be sensitive to the measure of body weight considered (Wada and Tekin, 2007; Burkhauser and Cawley, 2008).

variable is somewhat arbitrary as there is no *a priori* guarantee that conventional clinical thresholds are adequate to pick up levels where obesity starts affecting wages. For instance, it is not unreasonable to believe that moderate deviations of BMI around a central value or within a socially accepted range (possibly outside clinically recommended ranges) do not trigger an immediate wage penalty. Mis-locating 'turning points' in the relationship between BMI and wages will lead to an attenuation of the estimated impact of obesity on wages. In addition, a piecewise constant specification does not identify differential wage effects of body size *within* BMI categories.

This paper addresses these concerns by estimating partial linear regression models that allow us to examine the shape of the BMI-wage relationship without imposing functional form assumptions. Two recent studies have adopted this approach to examine this association in China (Shimokawa, 2008) and in the United States (Gregory and Ruhm, 2009). Our analysis revisits in a similar fashion the association between BMI and wages in Europe, and assesses the suitability of standard parametric models. To preview our results, we find that a linear model describes the association for women reasonably well while it is better captured by a quadratic model for men. The fit of parametric models is however far from fully satisfactory and the semi-parametric models identify patterns that none of the parametric specifications can capture. In particular, we observe, especially among Northern European men, that body size has no effect on wages over a broad, median BMI range (which does not coincide with classic BMI classifications), but bites strongly outside of this range.

The discussion of functional form specifications can be taken further, however. It often goes unnoticed that using the BMI as a measure of body size in effect imposes a specific relationship between height, weight and wages. There has been little concern about the validity of this approach.[2] A further contribution of this paper is to take advantage of the semi-parametric approach to examine how height and weight relate to wages *without* using the BMI functional form. By using a bivariate extension of the partial linear model adopted in Shimokawa (2008) and Gregory and Ruhm (2009), we are able to test whether a non-parametric function associating freely weight and height to wages is significantly different from a non-parametric function associating BMI to wages. This allows us to check whether the BMI functional form reliably captures the relationship between body size and wage. Perhaps unsurprizingly given recent evidence of a height-related wage premium –see, *inter alia*, Case and Paxson (2008)–, we find a significant independent effect of height after controlling for BMI. This suggests that considering BMI alone is too restrictive to fully capture the complexity of the association between body size and

[2]Kan and Lee (2009) is a recent exception. They adopted a similar approach than the one used in this paper to re-estimate the wage effects of weight among US white females on the sample of Cawley (2004).

wages. This, in turn, may also account for the limited robustness of the empirical findings on BMI and wages.

Parametric and semi-parametric models of body size and wages are detailed and discussed in Section 2. Our sample from the European Community Household Panel is described in Section 3 along with a review of existing evidence. Estimation results are shown in Section 4. Section 5 concludes with a discussion of our results.

2 Modeling body size and wage

2.1 Model specifications

Virtually all studies of the wage effects of body size estimate a wage equation which can be embedded in the following model:

$$y_i = X_i\beta + g(h_i, w_i) + e_i \tag{1}$$

where y_{it} is the logarithm of individual i hourly wage, X_i is a vector of individual attributes affecting wage (such as education, work experience), and e_i is a residual term. The bivariate function $g(h_i, w_i)$ captures the effect of body size on wage where body size is a function of height (h_i) and weight (w_i). The vast majority of studies summarize body size from height and weight using the body mass index, so the bivariate $g(h_i, w_i)$ can be reduced to the univariate function $f(w_i/h_i^2) \equiv f(\text{BMI}_i)$, and equation (1) becomes:[3]

$$y_i = X_i\beta + f(\text{BMI}_i) + e_i. \tag{2}$$

At this point, studies differ in the specification of f. Most of them make further

[3]See the early contributions of Register and Williams (1990), Averett and Korenman (1996) or Pagan and Davila (1997) and more recently Baum II and Ford (2004), Cawley (2004), Cawley et al. (2005), Conley and Glauber (2005), Morris (2006), Brunello and d'Hombres (2007) and Atella et al. (2008) among many others. Only some recent studies have considered more complex measures of body mass. Wada and Tekin (2007) used alternative measures of body composition from bioelectrical impedance analysis (BIA) to measure body fat, and Burkhauser and Cawley (2008) discussed the appropriateness of the BMI as measure of body fat.

parametric assumptions. They either assume that BMI enters the wage equation linearly,[4]

$$f^{\mathrm{I}}(\mathrm{BMI}_i) = \gamma_1 \mathrm{BMI}_i$$

or adopt a piecewise constant specification using categories for being underweight, overweight or obese according to conventional clinical classification,[5]

$$f^{\mathrm{II}}(\mathrm{BMI}_i) = \alpha_1 \mathbf{1}(\mathrm{BMI}_i \leq 18.5) + \alpha_2 \mathbf{1}(25 \leq \mathrm{BMI}_i < 30) + \alpha_3 \mathbf{1}(\mathrm{BMI}_i \geq 30),$$

where $\mathbf{1}(\bullet)$ evaluates to 1 if the expression in brackets is true and 0 otherwise. This second specification is somewhat more flexible since it allows for non-linearity in the wage–body mass relationship. However, there is no guarantee that BMI categories derived from medical evidence on increased morbidity are meaningfully related to the way wage is associated to body size. Two recent studies on Danish and US data allowed non-linear wage effects of BMI by assuming a quadratic functional form (Greve, 2008; Wada and Tekin, 2007),

$$f^{\mathrm{III}}(\mathrm{BMI}_i) = \gamma_1 \mathrm{BMI}_i + \gamma_2 \mathrm{BMI}_i^2.$$

These last two specifications allow for a potential wage penalty for either low (underweight penalty) or high BMI (overweight penalty) or both. Nevertheless, these parametric assumptions remain relatively strict and violations may yield mis-specification bias.

A more flexible semi-parametric approach in which the wage effects of BMI enters the wage equation non-parametrically has recently been considered by Shimokawa (2008) and Gregory and Ruhm (2009). In this case,

$$f^{\mathrm{IV}}(\mathrm{BMI}_i) = \hat{f}(\mathrm{BMI}_i)$$

where $\hat{f}$ is an unknown, smooth function estimated along with the β parameter vector. Unlike parametric models, this specification does not constrain the shape of the association between BMI and wage. It allows for different penalties for underweight, overweight and obesity, it does not impose single-peakedness (so that no wage difference may be seen for broad ranges of BMI, for example), and it does not rely on any pre-determined thresholds to define underweight, overweight or obesity.

[4]As in Cawley (2004); Cawley et al. (2005); Morris (2006); Brunello and d'Hombres (2007); Norton and Han (2008).

[5]Recent studies which have adopted this specification include Cawley (2004), Cawley et al. (2005) and Fahr (2006). Others have restricted their focus on overweight and obese respondents (Norton and Han, 2008; Han et al., 2009) or have limited the scope of their study to the obese (Baum II and Ford, 2004).

While this last specification is flexible, it still rests on the modeling assumption that combining height and weight into the body mass index and including the latter in a wage regression adequately describes the relationship between wage and body size. However, there are reasons to conjecture that height alone has an independent effect on wages since height tends to be associated with factors such as physical attractiveness, strength or cognitive ability that affect wages (Steckel, 1995; Case and Paxson, 2008; Cinnirella and Winter, 2009). As in Kan and Lee (2009), we question this fundamental assumption by considering a model in which both height and weight enter the wage equation non-parametrically and not through the body mass index; that is, we specify directly

$$g(h_i, w_i) = \hat{g}(h_i, w_i)$$

in equation (1), where $\hat{g}$ is an unknown smooth, bivariate function which is estimated along with the β parameters. This allows us to identify potential mis-specification in the use of the BMI index in the wage equation.

2.2 Estimation and specification tests

While estimation of the fully parametric models is standard, the flexible specifications for $\hat{f}$ and $\hat{g}$ require semi-parametric estimators. Alternative estimators of such 'partially linear' models can be chosen from. We adopt Yatchew's (1997) differencing estimator (see also Yatchew, 2003). Popular alternatives is the more computationally intensive 'double residual' estimator (Robinson, 1988) or estimation based on smoothing splines (Ruppert et al., 2003; Wand, 2003). In our large sample application, these three estimators resulted in almost identical results and our choice was eventually guided by ease of implementation.

Differencing estimation of equation (1) is a two-step procedure. The first stage involves estimation of the β parameters net of the effect of w_i and h_i:

$$\hat{\beta}^D = (DX)^{-T}(DX)^T Dy$$

where $X = (X_1, X_2, \ldots, X_n)^T$, $y = (y_1, y_2, \ldots, y_n)^T$, and D is an 'optimal' differencing matrix (defined in Yatchew (1997)) applied after ordering the data according to BMI (univariate model) or h_i and w_i (bivariate model). The covariance matrix of $\hat{\beta}^D$ is given by

$$\hat{V}(\hat{\beta}^D) = \left(1 + \frac{1}{2m}\right) s^2_{diff}(DX)^{-T}$$

where $s^2_{diff} = n^{-1}\hat{v}'\hat{v}$ is the residual variance of the differenced regression, $\hat{v} = Dy - DX\hat{\beta}^D$ and m is the order of differencing.[6] The second stage involves estimation of the non-parametric component ($\hat{f}$ or $\hat{g}$) by regressing non-parametrically the first-stage residuals $\hat{v}_i$ on BMI_i or on w_i and h_i using, e.g., local polynomial regression (or any standard non-parametric regression). See Yatchew and No (2001) for an application of this technique.

The differencing approach offers a straightforward way to test parametric specifications against flexible non-parametric estimates. Let $\phi(z;\theta)$ be a parametric function with parameters θ (such as f^{I}, f^{II}, or f^{III} defined above). Under the null hypothesis that $f(z) = \phi(z;\theta)$, the test statistic

$$(mn)^{0.5}\frac{\left(s^2_{res} - s^2_{diff}\right)}{s^2_{diff}} \xrightarrow{D} N(0,1) \tag{3}$$

where s^2_{diff} is defined above and is obtained using optimal mth order differencing weights, and $s^2_{res} = n^{-1}\hat{w}^T\hat{w}$ is the residual variance in the parametric model, $\hat{w} = y - X\hat{\beta} - \phi(z;\hat{\theta})$ (Yatchew, 2003, p.63). Note that, because it relies on the differencing principle, computation of the test statistic does not depend on estimation of the unknown f function but only of the fully parametric model and of the $\hat{\beta}^D$ parameters of the linear components in the semi-parametric model. This specification test allows us to formally test our various specifications against each other.

3 Data

3.1 BMI and wages in the European Community Household Panel survey

Our study exploits longitudinal data extracted from the European Community Household Panel survey (ECHP).[7] The ECHP survey is a large-scale, general-purpose panel

[6]The variance expression is valid provided 'optimal' differencing weights are used to construct D. See Yatchew (1997, 2003) for details. Heteroscedastic-consistent, 'robust' standard errors can be estimated using the classic 'sandwich' formula. See Yatchew (2003, p.72) and StataCorp (2007). Our estimates are based on optimal differencing weights at the order 100, with robust standard errors.

[7]The public-use ECHP database was created, maintained and centrally distributed by Eurostat. See EUROSTAT (2003) or Lehmann and Wirtz (2003) for more information on the database, and Peracchi (2002) for an independent critical review. All our results are based on the final release (April 2004) of the ECHP Users' Database.

survey run in fifteen EU countries over the period 1994–2001. The database contains a wide range of household- and individual-level information on income and living conditions, employment, education, health, demographic characteristics. In the last four waves (1998–2001), the ECHP included information on respondents' height and weight.

Several studies have recently used the ECHP to document the wage effects of body mass in Europe. In a regression model where BMI enters a log-wage equation linearly, Brunello and d'Hombres (2007) found existence of a significant European wide wage penalty to obesity –of greater magnitude for men. In contrast, relying on a piecewise constant specification in BMI capturing whether a respondent is underweight, overweight or obese according to clinical thresholds, Atella et al. (2008) suggested that this European wide wage penalty only affects overweight and obese females. Fahr (2006) further investigated non-linearities in this association with a model which allows to disentangle the independent wage effects of deviations from both socially accepted body mass and medically recommended thresholds. His results suggest that deviations from medically recommended BMI are more hurtful to female earnings than deviations from social norms.[8] The opposite observation seems to hold for men. This is broadly consistent with Atella et al.'s (2008) finding of a more significant negative wage penalty for overweight and obese female respondents since they defined BMI categories according to conventional clinical thresholds. It is also consistent with the claim that BMI score may capture more accurately excessive body fatness in females than in males (Wada and Tekin, 2007; Burkhauser and Cawley, 2008). Fahr (2006) still relied, however, on normative assumptions regarding what constitutes socially acceptable BMI scores or more generally, continued to rely on ad hoc assumptions regarding the location of potential turning points shaping the BMI wage association.

Differences in methodology and sample selection make it difficult to readily compare the results reported in Fahr (2006), Brunello and d'Hombres (2007) and Atella et al. (2008). However, the estimated wage effects of BMI reported in these studies are consistent with the view that the association between BMI and wage is likely nonlinear, differ across gender, and that a specification based on clinical thresholds might not optimally capture important turning points in its true relationship.

[8] In this context, a socially acceptable BMI is assumed to be determined by the median regional BMI adjusted for gender and broad age groups.

3.2 Sample definition

As in Fahr (2006), Brunello and d'Hombres (2007) and Atella et al. (2008) our sample is restricted to waves and countries that provide valid data on respondents' weight and height, that is, for Austria, Denmark, Finland, Ireland, Italy, Greece, Portugal and Spain in the years 1998 to 2001, leading to a raw sample of approximately 280,000 observations.[9]

In the ECHP, respondent's BMI is calculated from self-reported measures of height and weight. It is well-known that self reported height and weight are measured with errors.[10] Following Atella et al. (2008), we drop respondents with a reported BMI below 15 (147 observations) and over 50 (201 observations). As pointed out by Sanz-de-Galdeano (2005) and Brunello and d'Hombres (2007), the absence of true height and weight data for the countries under study prevents us from applying further corrective methods such as the one proposed by Cawley (2004). We use the longitudinal nature of our data, however, to remove from our sample individuals reporting either clearly inconsistent height or highly suspicious weight using variations in reported height and weight across waves. This is done by comparing period t weight (or height) for respondent i to her average height (or weight) reported in all other periods.[11] We drop all observations with a difference in height larger than 5 centimeters or a difference in weight of 12 kilograms or more, compared to other period average. Interestingly, the rates of inconsistent self-reporting do not seem to differ significantly across gender but varies greatly across countries.[12]

We restrict our sample to all employees (not in agriculture) working at least 15 hours per week. To prevent estimates from being driven by a limited number of outlying observations, we also drop respondents with hourly wage either less than 1.5 euros (251 observations) or over 50 euros (72 observations). For comparability with Atella et al. (2008), we keep all respondents between 25 and 64 years of age. The resulting sample includes 43,300 male and 33,501 female respondents with non-missing data on wage and all relevant explanatory variables including age, indicator variables for being married, the highest level of completed education, reporting being in poor or bad health, being a smoker, working part-time and four occupation group dummies.[13] The dependent vari-

[9] Unlike earlier studies, we also exclude Belgium due to the abnormally large number of missing data on respondents' main sector of activity and occupation in waves 5 and 6.

[10] See Danubio et al. (2008) for a recent comparison study between self-reported and measured height and weight among young Italian adults.

[11] A similar procedure is adopted by Fahr (2006).

[12] For instance, this procedure leads us to reduce our female sample by just 0.74% in Finland but as much as 11.7% in Spain. All numbers are available from the authors upon request.

[13] Our occupational group dummy variables were constructed by grouping nine occupational categories available in the ECHP User database into four groups which we label *Professional*, *Clerks*, *Craft* and *Elementary*, which broadly reflect decreasing skill requirements.

able of our wage models is the natural logarithm of hourly wage expressed in constant 1996 PPP euros.[14]

3.3 Descriptive statistics

Summary statistics of our sample are reported separately for male and female respondents living in Northern European countries (Table 1) and Southern European countries (Table 2). According to clinical thresholds, the average European man is overweight with a mean BMI just under 26. European women report on average a healthier BMI just over 23 in the south and just over 24 in the north. The distribution of the population by BMI categories reveals that more than half of all male respondents in our sample are overweight or obese. This observation broadly holds across regions and countries. Obesity rates vary significantly across countries ranging between just under 7% in Italy to over 12% in Spain for males and between just over 3% in Italy and just over 10% in Finland for females.[15]

The incidence of underweight among males is extremely low in all countries. Our pooled sample of Southern European (Northern) countries, only includes 79 (29) underweight male respondents. This implies that the wage effect of underweight males in each separate country would be identified on just a few cases. While the incidence of underweight is higher among females at about just over 2.5% in Northern Europe (or 356 observations) and just over 4% (or 798 observations) in Southern Europe, the number of underweight respondents in each separate country remains small.[16] As a result, we limit our analysis to the estimated wage effects of body size for the pooled samples of Southern –Greece, Italy, Portugal and Spain– and Northern European countries –Austria, Denmark, Finland and Ireland– separately. As in Brunello and d'Hombres (2007), the implicit assumption behind this pooling is that these Southern and Northern European countries share some unobserved regional cultural traits. However, since we dropped Belgium, our pooled sample of Northern European countries is not strictly comparable to their so called *beer belt* countries.[17]

Figure 1 compactly presents the distribution of the population and the unconditional average hourly wages by BMI levels, separately by gender, for the pooled samples of Northern and Southern European countries. The figure confirms that a large share of male

[14] We have constructed hourly wage following Arulampalam et al. (2007), that is, as gross monthly earnings from main job including overtime divided by 4.5 times weekly hours in main job including overtime.

[15] Note that the incidence of obesity in our sample is largely consistent with the numbers reported by Atella et al. (2008) using the same age sample restriction than this study.

[16] With the exception of Italy and Spain due to their larger samples and higher incidence of underweight respondents.

[17] Single country results are available from the authors upon request.

respondents in our sample is concentrated in the 24–26 BMI range and a large share of females is found in the 21–23 BMI range in both regions. Interestingly, the (unconditional) BMI–wage profiles of females peak at a relatively low BMI level (around 22) and only decline modestly thereafter regardless the region considered. Heavier females in Southern Europe, however, appear to experience a somewhat more important wage penalty. In contrast, the wage profiles of males peak at higher BMI scores (around 26) without significantly decreasing thereafter. Furthermore, it is worthwhile to note that males with low BMI – not necessarily in the unhealthy range – appear to earn significantly less than any other male respondents – more so in Northern Europe.

Overall, these descriptive plots suggest that important turning points of the unconditional wage function likely differ across gender and are not usually occurring at points consistent with commonly used clinical thresholds. In particular, female wages usually peak at a much lower BMI score than males and, unlike the latter, appear to consistently monotonically decrease as BMI increases thereafter.

4 Estimation Results

4.1 Parametric Results

We first replicate earlier work and estimate equation (2) parametrically by assuming that BMI enters the wage equations (A) linearly, (B) as a quadratic function and (C) as piecewise constant in BMI categories. Parametric estimates from these regressions provide convenient benchmarks to contrast our results with earlier ECHP studies and to compare the expected wage function obtained semi-parametrically. Our baseline model includes demographic and human capital controls which are thought to be potentially BMI-determined. The latter includes a quadratic function of age, indicator variables for educational attainment (one for secondary and one for tertiary education), being married, being in bad health,[18] being a smoker and a set of time and country dummies. Our second specification adds job related characteristics to the baseline specification including indicator variables for part-time work, whether a respondent works in the private sector and a set of four occupational group dummy variables.

Coefficient estimates are reported in Table 3 separately for females and males. The first two columns report the estimated wage effects from our two model specifications on

[18]Conventionally defined as when respondents self-report being either in poor or very poor health.

the pooled sample of Northern European countries followed by the estimated wage effects in Southern Europe. Our discussion, primarily focus on the pooled sample estimates from the less parsimonious model specification (Model 2).

Estimates from the linear model do not reveal any significant association between wage and BMI among males. In contrast, we find a significant wage penalty among females, but an effect relatively small in magnitude. In particular, a 10% increase in the BMI of female respondents is associated with a modest decrease in wages of about 0.48% in Northern Europe and about 0.93% in the South. The insignificant linear wage effect of BMI for males, however, appears to mask a more complex association. As pointed out by Gregory and Ruhm (2009), if individual BMI is negatively associated with both being obese and underweight –as suggested by our unconditional wage profiles– , linear estimates may misleadingly suggest the absence of a significant association. Once we model the wage effect of BMI with a quadratic form, we find a statistically significant inverted U-shaped association for males. The peak in the relationship is found at a BMI of about 27 or 28 in both Northern and Southern Europe. We do not find a statistically significant quadratic wage effect for females. Male estimates from the piecewise constant model with BMI categories corroborate the existence of an inverted U-shaped association for males by indicating the existence of a wage penalty for being underweight or obese and a premium for being overweight. These estimates, however, are only statistically significant in Northern Europe for underweight and overweight respondents and never significant in Southern Europe.[19]

Taken together, we interpret these results as evidence that the association between BMI and wage might be inverted U-shaped for males with a peak at a BMI level in the overweight range. Clinical thresholds defining BMI categories do not seem to capture accurately critical turning points in this association; possibly more so, in Southern Europe. These results corroborate recent evidence found in Germany (Cawley et al., 2005) and Denmark (Greve, 2008) and are consistent with previous ECHP-based estimates by Fahr (2006) and Atella et al. (2008). In contrast, estimates reveal that overweight or obese females earn significantly less than those in the clinically recommended BMI category. These estimates, however, are only significant in Southern Europe indicating that overweight and obese female earn about 2.7% and 5% less than their 'healthy ' counterparts. This regional difference is consistent with Brunello and d'Hombres (2007) finding of a stronger association in so called *olive belt* countries.[20] These results provide support for

[19]Estimates based on single country samples corroborate pooled sample estimates in signs and magnitude but, not unexpectedly given the small sample sizes, are usually insignificant for both males and females. These results are available upon request.

[20]As pointed out by Brunello and d'Hombres (2007), this regional difference might just be the result of the smaller sample size in Northern Europe. Our more parsimonious specification suggests the existence of a

the existence of a monotonically decreasing wage effects in BMI for females as implied by the linear model.

4.2 Semi-parametric model estimates of the BMI-wage relationship

As argued in the Introduction, parametric regression results may mask the complexity of the functional relationship between wage and BMI. There is interest in considering an unconstrained specification to check whether sufficient flexibility is achieved with a quadratic or a classic piecewise constant parametric model.

Our non-parametric estimates of the effect of BMI on log-wage are presented graphically, separately for males and females in Northern (Figures 2 and 3) and Southern Europe (Figures 4 and 5). Each figure presents the BMI-wage profiles implied by our two model specifications (the parametric components are estimated as explained in Section 2.1). For each level of BMI on the x-axis, the plots show the expected log-wage as given by equation (2), that is, $\mathrm{E}(\log(y)|\bar{X}, \mathrm{BMI}) = \bar{X}\hat{\beta} + \hat{f}(\mathrm{BMI})$, where $\bar{X}$ is the vector of means of other covariates in the sample considered and $\hat{\beta}$ and $\hat{f}$ are the model estimates. We overlay estimates from the semi-parametric model over predictions implied by the parametric estimates presented in the previous section. Point-wise 90 percent confidence bootstrap variability bands for the predicted BMI-wage profile from the semi-parametric model are represented by the vertical bars around the predictions.[21] The red bars at the bottom of each graph are kernel density function estimates of the distribution of BMI in the sample.

Our semi-parametric results corroborate our earlier conjecture that the association between BMI and wage reveals an inverted U-shaped for Northern European males (see Figures 2). However, while the quadratic results suggested a single peak at a BMI of about 28, the semi-parametric estimates rather suggest that there is a plateau with maximum wage in the range 24–31 and a penalty above or beyond these. This suggests that there is a wage penalty for people beyond a 'normal' body size. The linear model is clearly misspecified. The quadratic model underestimates the wage penalty beyond the 'normal'

statistically significant wage penalty of about 2.2% and 3.9% for overweight and obese females in Northern Europe. Alternatively, it might also reflect a true North-South differences in norms making clinical BMI thresholds less relevant for Northern European female respondents.

[21] We implemented the repeated half-sample bootstrap algorithm of Saigo et al. (2001). To take into account the stratification of the survey we resample within stratum identified in the data (for Ireland, Spain, Portugal and Finland) or within NUTS-1 level region if detailed stratum identifier are not provided in the public-use dataset (all other countries). The resampling unit is the wave 1 household, so that all dependence of responses for same household respondents and for repeated responses over time is properly taken into account. All estimates reported are based on 500 replications.

range. The piecewise constant model fails to capture the variations within the 'healthy' (20–24) and 'obese' (above 30) ranges.

Figure 4 reveals quite a different expected wage profile for Southern European males. There is no plateau, but the profile is not more quadratic. Surprisingly, we also observe an inflection point at BMI around 32 suggesting the existence of a large obesity premium for Southern European males (which is not at all apparent in the obesity dummy in the piecewise constant specification). Extremely few observations are observed with a BMI above 32 however. Interestingly, single country figures (not reported here but available upon request from the authors) indicate that this surprising wage increase in Southern Europe is not confined to a single outlying country but is observed consistently in Italy, Portugal and Spain.

Figures 3 and 5 globally corroborate our earlier parametric estimates indicating a general decline in expected wage with BMI for females. But note that the non-parametric estimates show the existence of a peak at a BMI of about 21 (for Northern European women) or 22 (for Southern European women). There is therefore also a penalty for underweight among females, yet a much smaller one than the penalty for overweight or obesity.

In sum, while the overall relationship seem to be relatively well approximated by a quadratic model for men and a linear model for women, the non-parametric estimation reveals fine details that are missed by all other models. Formal tests based on equation (3) of the null hypothesis of equality of the non-parametric curves with any of the parametric models considered, all strongly support rejection.[22]

4.3 Modeling body size without the body mass index

We finally consider the predicted wage effects from body size when the latter is captured by a smooth function of height and weight estimated non-parametrically, rather than through the body mass index. Our motivation is to test whether the parametric association between weight and height in effect implied by the BMI functional form –weight in kilograms over height squared– adequately relates the combined effects of height and weight on wages. To achieve this, we have estimated directly the bivariate function g of equation (1) using the same partial linear model as in the previous sub-section. The key difference is that the non-parametric component is now an unspecified, smooth, bivariate

[22]Test statistics are not reported here to save space (all p-values are well below 0.001) but are available from the authors upon request.

function of height and weight instead of a univariate function of BMI.

Figures 6 to 9 illustrate the differences between the expected wage profiles implied by this model with the one relying on the BMI functional form. Each element in these figures illustrate the relationship between height and log-wage for a fixed level of BMI in $\{18, 20, 22, 24, 26, 28, 30, 32, 34\}$. After fixing the BMI level, the height–wage relationship is constant in the univariate BMI-based model, while it may vary in the flexible, bivariate model if height has an independent effect on wage after fixing the BMI level. Figures therefore illustrate the presence of this independent effect of height by plotting the difference in expected wage obtained from the two models at different levels of height for each selected BMI:

$$\Delta(h, \text{BMI}, \bar{X}) = \text{E}^g(\log(y)|\bar{X}, h, w = \text{BMI} \times h^2) - \text{E}^f(\log(y)|\bar{X}, \text{BMI}) \qquad (4)$$

where

$$\text{E}^g(y|\bar{X}, h, w = \text{BMI} \times h^2) = \bar{X}\hat{\beta}^g + \hat{g}(h, w = \text{BMI} \times h^2)$$

and

$$\text{E}^f(y|\bar{X}, \text{BMI}) = \bar{X}\hat{\beta}^f + \hat{f}(\text{BMI}).$$

Whenever $\Delta(h, \text{BMI}, \bar{X})$ is positive, this is indicative of an additional wage premium for people of height h. Observing any significant deviation of $\Delta(h, \text{BMI}, \bar{X})$ from zero is therefore indicative that the univariate model relating body size to wage is too restrictive and does not adequately describe the relationship between height, weight and wage.[23]

Except for the Southern European women, there appears to be a clear positive and significant association between height and weight, at most levels of BMI (see Figures 6 to 9). In particular, men of below average height are exposed to a statistically significant wage penalty in both the Northern and Southern European samples. This adverse height effect is stronger for relatively low BMI levels. For instance, our estimates suggest that the expected wage of a short, 1.65m tall, European man (either from the North or the South), with a BMI in the 22–26 range, is between 3% to 6.5% lower than the expected wage implied by the less flexible BMI model which does not consider height separately. It is also worth noting that this estimated wage penalty for below average height is decreasing with BMI.

In contrast, Northern European women of above average height enjoy a significant wage premium. For instance, the expected wage of a woman who is 1.80 m tall with a

[23] Vertical bars are point-wise 90 percent confidence bootstrap variability bands. See *infra* for details on the bootstrap resampling algorithm.

BMI of 22 is approximately 6% higher than the expected wage implied by our less flexible model. This wage premium rises to 10% for a clinically overweight woman with a BMI of 26 and 12% for clinically obese women with a BMI of 30.

5 Discussion

All ECHP studies which investigate the association between body size and wage in Europe (Fahr, 2006; Brunello and d'Hombres, 2007; Atella et al., 2008; Villar and Quintana-Domeque, 2009) rely on the BMI to measure body size. In this regard, the semi-parametric estimates reported in this study shed further light on the nature of this association in Europe and strengthen our current understanding of the relationship between height, weight and wage in general.

Both our parametric and semi-parametric results suggest that for males the association between BMI and wage is, broadly speaking, an inverted U-shaped peaking in overweight territory, but a shape that cannot be adequately reduced to a quadratic relationship. Our results also suggest that, for males, being too thin might be more detrimental than being obese. These results are in line with Atella et al. (2008) pooled sample OLS estimates for Europe at large (Table 3, page 311), and Cawley (2004) OLS estimates for the US.

In contrast, female wages seem to peak in healthy BMI territory (around 22) and to decrease monotonically thereafter. This finding is consistent with the large number of US (parametric) studies either reporting a significant linear negative association with BMI (Cawley, 2004) or a significant wage penalty for obesity (Baum II and Ford, 2004; Han et al., 2009). Likewise, our female results are consistent with Atella et al. (2008) pooled OLS estimates for Europe at large (Table 2, page 310).

Interestingly, our semi-parametric estimates fully corroborate Gregory and Ruhm (2009) recent semi-parametric analysis for the US. In particular, Gregory and Ruhm (2009) also find that "women's wage peak at BMI of 23 or lower." They interpret this finding as possible evidence for the existence of a wage premium for physical attractiveness rather than a wage penalty reflecting the adverse (health) effects of unhealthy weight since respondents start experiencing a significant wage penalty at BMI levels well below conventional thresholds defining unhealthy weight. In this context, BMI becomes a proxy for societal views on physical attractiveness which in turn is rewarded in the market. Early contributions by Loh (1993) and Hamermesh and Biddle (1994) show the existence of a beauty premium in the US labour market. The wage penalty incurred by obese females may mirror the penalty for deviating from a socially acceptable weight. Our estimates could therefore suggest that the socially acceptable bmi of European females is approximately 22 - deviations from this ideal physical trait is associated with lower earnings, possibly of larger magnitude in Southern Europe.

Similarly, our analysis shows that males wage peak in the mostly overweight BMI range 24–31 in Northern Europe. This result is again consistent with the existence of a wage penalty triggered by deviations from socially acceptable weight which, for the case of males, would be determined by the average (or median) BMI of the working population. In this context, it is therefore not surprising that being too thin is associated with a significant wage penalty since the European males included in our sample are on average slightly overweight. This is consistent with Fahr (2006) who posits that, in some European countries, social norms "set the relevant standard to evaluate men's physical appearance."[24] While we find a comparable pattern in Southern Europe, the sudden wage premium enjoy by men with a BMI over 32 is rather puzzling. It is, however, again consistent with the view that the association between BMI and wage is not necessarily driven by the adverse health effects of abnormal weight.

Finally, our fully flexible model reveals that shorter males suffer an additional significant height penalty independent of BMI and that Northern European females of above average height enjoy a wage premium. The existence of a height-wage premium is now a well documented empirical regularity. Recent height studies have consistently documented that taller workers earn significantly more than their shorter counterparts in Australia (Kortt and Leigh, 2010), Germany (Heineck, 2005), the UK (Case et al., 2009) and Europe at large (Cinnirella and Winter, 2009). The existence of differences in cognitive skills between shorter and taller workers is one possible pathway explaining this height-wage premium (Case and Paxson, 2008). In a recent European study, Cinnirella and Winter (2009) argue that, without being exclusive, a large part of this premium could also be due to employer discrimination. While we do not formally explore these issues, we believe that overall, results from our most flexible specification are also consistent with the existence of a premium for physical attractiveness.

In sum, our study corroborates the view that the shape of the BMI association differs across gender and suggests that the BMI functional form may be too restrictive to adequately capture the complexities of the association between height, weight and wages for men. As Gregory and Ruhm (2009), we posit that this association could be driven by physical attractiveness rather than unhealthy weight. In this context, gender differences stem from differences in judgment regarding desirable body types. A *good height* might be more important than an healthy weight for males while an healthy BMI is more desir-

[24]Fahr (2006) defines social norm as the gender, age group and region specific median BMI.

able for females. This conjecture is consistent with Rooth (2010) who finds that obese job applicants in Sweden experience lower call back rates and that this differential treatment is mostly driven by obesity for women and attractiveness for men. However, one needs to keep in mind that, even if significant, the estimated wage effects of body size reported in this study are overall fairly small.

Unlike Gregory and Ruhm (2009), our data do not allow us to control for endogeneity of BMI in the semi-parametric setting. This makes it difficult to give a strictly causal interpretation to the estimated wage effect of body weight discussed in this study. Reverse causality and the possibility that body weight could be correlated with unobserved factors also affecting wages are the two main sources of endogeneity bias identified in the obesity literature. Reverse causality is usually controlled for by instrumenting contemporaneous BMI with a sufficiently distant BMI measure (Gortmaker et al., 1993; Averett and Korenman, 1996; Cawley, 2004; Gregory and Ruhm, 2009). As in other ECHP studies (Brunello and d'Hombres, 2007; Atella et al., 2008), we are not able to control for this potential source of bias since our exploitable longitudinal sample only covers three years of data. However, most of these studies find that the estimated wage effect of body weight using a lagged measure of BMI is virtually identical to that using current BMI score. The obvious response to the second source of bias is again to use instrumental variable estimation techniques. While implementing such strategy is empirically straightforward, identifying strong instruments turns out to be challenging. A vast majority of studies that adopted instrumental variable estimation (Cawley, 2004; Cawley et al., 2005; Brunello and d'Hombres, 2007; Norton and Han, 2008; Gregory and Ruhm, 2009) have used the BMI of genetically related family members following Cawley (2004). These studies usually find that controlling for potential endogeneity does not affect their results substantially.[25] In addition, the reliability of IV estimates using the body weight of a genetically related family member as instrument on ECHP data has been forcefully questioned by Atella et al. (2008). The latter is suspected to yield significant bias on ECHP data due to severe non-random sample selection (see Atella et al., 2008, for further discussion) from imposed sample restrictions. Given this concern and in the absence of any convincing alternative instruments in our data, we deliberately do not address the potential endogeneity of weight (or BMI) in this study. Caution should therefore be exercised to give a fully causal interpretation to the estimates of wage effects of body size presented in this paper.

[25]See Kortt and Leigh (2010) for a comprehensive survey of previous IV studies.

References

Arulampalam, W., Booth, A. L. and Bryan, M. L. (2007), 'Is there a glass ceiling over Europe? Exploring the gender pay gap across the wage distribution', *Industrial and Labor Relations Review* **60**(2), 163–186.

Atella, V., Pace, N. and Vuri, D. (2008), 'Are employers discriminating with respect to weight? European evidence using quantile regression', *Economics and Human Biology* **6**, 305–329.

Averett, S. A. and Korenman, S. (1996), 'The economic reality of the beauty myth', *Journal of Human Resources* **31 (2)**, 304–330.

Baum II, C. L. and Ford, W. F. (2004), 'The wage effects of obesity: A longitudinal study', *Health Economics* **13**, 885–899.

Brunello, G. and d'Hombres, B. (2007), 'Does body weight affect wages? Evidence from Europe', *Economics and Human Biology* **5**(1), 1–19.

Burkhauser, R. and Cawley, J. (2008), 'Beyond BMI: The value of more accurate measures of fatness and obesity in social science research', *Journal of Health Economics* **27**, 519–529.

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

Case, A., Paxson, C. and Islam, M. (2009), 'Making sense of the labor market height premium: Evidence from the British Household Panel Survey', *Economics Letters* **102**, 174–176.

Cawley, J. (2004), 'The impact of obesity on wages', *Journal of Human Resources* **39**(2), 451–474.

Cawley, J., Grabka, M. M. and Lillard, D. R. (2005), 'A comparison of the relationship between obesity and earnings in the U.S. and Germany', *Schmollers Jahrbuch* **125**(1), 119–129.

Cinnirella, F. and Winter, J. (2009), Size matters! Body height and labor market discrimination: A cross-european analysis, Technical report, CESifo Working Paper 2733.

Conley, D. and Glauber, R. (2005), Gender, body mass and economic status, Technical Report 11343, National Bureau of Economic Research, Cambridge MA.

Danubio, M. E., Miranda, G., Vinciguerra, M. G., Vecchi, E. and Rufo, F. (2008), 'Comparison of self-reported and measured height and weight: Implications for obesity research among young adults', *Economics and Human Biology* **6**(1), 181–190.

EUROSTAT (2003), *DOC.PAN 168/2003-12: ECHP UDB manual, Waves 1 to 8*, Eurostat, European Commission, Luxembourg.

Fahr, R. (2006), The wage effects of social norms: Evidence of deviations from peers' body-mass in Europe, Discussion Paper No. 2323, Institute for the Study of Labor (IZA), Bonn.

Gortmaker, S., Must, A., Perrin, J., Sobol, A. and Dietz, W. (1993), 'Social and economic consequences of overweight in adolescence and young adulthood', *New England Journal of Medicine* **329**, 1008–1012.

Gregory, C. A. and Ruhm, C. J. (2009), Where does the wage penalty bite?, NBER Working Paper 14984, National Bureau of Economic Research, Cambridge MA.

Greve, J. (2008), 'Obesity and labor market outcomes in Denmark', *Economics and Human Biology* **6**, 350–362.

Hamermesh, D. S. and Biddle, J. E. (1994), 'Beauty and the labor market', *American Economic Review* **84**(5), 1174–1194.

Han, E., Norton, E. C. and Stearns, S. C. (2009), 'Weight and wage: Fat versus lean paychecks', *Health Economics* **18**, 535–548.

Heineck, G. (2005), 'Up in the skies? The relationship between body height and earnings in Germany', *Labour* **19**(3), 469–489.

Kan, K. and Lee, M. (2009), Lose weight for money only if over-weight: Marginal integration for semi-linear models, HEDG Working Paper 09/19, University of York.

Kortt, M. and Leigh, A. (2010), 'Does size matter in Australia?', *Economic Record* **86**(272), 71–83.

Lehmann, P. and Wirtz, C. (2003), *The EC Household Panel Newsletter (01/02)*, Methods and Nomenclatures, Theme 3: Population and social conditions, Eurostat, European Commission, Luxembourg.

Loh, E. S. (1993), 'The economic effects of physical appearance', *Social Science Quarterly* **74**(2), 420–438.

Lundborg, P., Nystedt, P. and Rooth, D.-O. (2010), No country for fat men? Obesity, earnings, skills and health among 450,000 swedish men, IZA Working Paper 4775, IZA.

Mocan, N. H. and Tekin, E. (2009), Obesity, self-esteem and wages, NBER Working Paper 15101, National Bureau of Economic Research, Cambridge MA.

Morris, S. (2006), 'Body mass index and occupational attainment', *Health Economics* **25**, 347–364.

Norton, E. and Han, E. (2008), 'Genetic information, obesity, and labor market outcomes', *Health Economics* **17**, 1089–1104.

Pagan, J. A. and Davila, A. (1997), 'Obesity, occupational attainment and earnings', *Social Science Quarterly* **78**(3), 756–770.

Peracchi, F. (2002), 'The European Community Household Panel: A review', *Empirical Economics* **27**, 63–90.

Register, C. A. and Williams, D. R. (1990), 'Wage effects of obesity among young workers', *Social Science Quarterly* **71**(1), 130–141.

Robinson, P. M. (1988), 'Root-n-consistent semiparametric regression', *Econometrica* **56**(4), 931–954.

Rooth, D.-O. (2010), 'Obesity, attractiveness, and differential treatment in hiring: A field experiment', *Journal of Human Resources* **44(3)**, 710–735.

Ruppert, D., Wand, M. P. and Carroll, R. J. (2003), *Semiparametric Regression*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, New York.

Saigo, H., Shao, J. and Sitter, R. R. (2001), 'A repeated half-sample bootstrap and balanced repeated replications for randomly imputed data', *Survey Methodology* **27**(2), 189–196.

Sanz-de-Galdeano, A. (2005), The obesity epidemic in Europe, Discussion Paper 1814, Institute for the Study of Labor (IZA), Bonn.

Shimokawa, S. (2008), 'The labour market impact of body weight in China: A semiparametric analysis', *Applied Economics* **40**, 949–968.

StataCorp (2007), *Stata Statistical Software: Release 10*, StataCorp LP, College Station, TX.

Steckel, R. H. (1995), 'Height and the standard of living', *Journal of Economic Literature* **33**, 1903–1940.

Villar, J. G. and Quintana-Domeque, C. (2009), 'Income and body mass index in Europe', *Economics and Human Biology* **7**, 73–83.

Wada, R. and Tekin, E. (2007), Body composition and wages, NBER Working Paper 13595, National Bureau of Economic Research, Cambridge MA.

Wand, M. P. (2003), 'Smoothing and mixed models', *Computational Statistics* **18**, 223–249.

Yatchew, A. (1997), 'An elementary estimator of the partial linear model', *Economics Letters* **57**, 135–143.

Yatchew, A. (2003), *Semiparametric regression for the applied econometrician*, Themes in modern econometrics, Cambridge University Press, Cambridge, UK.

Yatchew, A. and No, J. A. (2001), 'Household gasoline demand in Canada', *Econometrica* **69**(6), 1697–1709.

6 Tables and Figures

Table 1: Sample averages by gender, ECHP, Northern European Countries

	All		**Austria**		**Denmark**		**Finland**		**Ireland**	
	M	F	M	F	M	F	M	F	M	F
Earnings										
Hourly wage	12.72	10.38	10.99	8.70	15.53	13.22	11.38	9.04	13.67	11.02
Body Size										
BMI	25.75	24.10	25.72	23.45	25.48	23.98	25.92	24.63	25.91	24.13
Underweight	0.002	0.025	0.001	0.045	0.003	0.027	0.001	0.014	0.002	0.017
Overweight	0.443	0.258	0.431	0.231	0.411	0.247	0.442	0.283	0.503	0.264
Obese	0.099	0.082	0.099	0.059	0.092	0.079	0.117	0.106	0.085	0.074
Height (in cms.)	178.3	165.5	177.9	166.1	180.0	167.0	178.1	164.8	176.8	164.0
Weight (in kgs.)	81.89	66.02	81.43	64.68	82.56	66.91	82.39	66.91	81.03	64.77
Demographic and employment variables										
Age	41.41	41.20	40.54	39.18	42.15	42.02	41.78	43.06	41.28	39.06
Primary Level of Education	0.168	0.175	0.101	0.196	0.124	0.114	0.163	0.168	0.337	0.254
Secondary Level of Education	0.565	0.487	0.810	0.676	0.532	0.514	0.441	0.357	0.400	0.439
Tertiary Level of Education	0.267	0.338	0.089	0.127	0.344	0.372	0.396	0.475	0.263	0.307
Married	0.673	0.664	0.677	0.628	0.641	0.663	0.669	0.698	0.716	0.648
Poor or Bad Health	0.013	0.018	0.014	0.016	0.011	0.019	0.018	0.024	0.004	0.005
Smoker	0.374	0.294	0.431	0.324	0.392	0.350	0.333	0.229	0.317	0.297
Private	0.711	0.522	0.716	0.647	0.731	0.416	0.707	0.453	0.682	0.650
Part-Time Work	0.019	0.171	0.013	0.288	0.015	0.134	0.028	0.066	0.022	0.274
Occupation										
Professionals	0.440	0.451	0.334	0.294	0.514	0.522	0.524	0.541	0.392	0.384
Clerks	0.139	0.417	0.184	0.523	0.114	0.381	0.102	0.350	0.151	0.459
Craft	0.351	0.055	0.412	0.060	0.293	0.042	0.329	0.044	0.361	0.088
Elementary	0.070	0.077	0.069	0.123	0.078	0.055	0.045	0.065	0.096	0.069
Sample size	15620	14158	4652	3328	3857	3676	4175	4760	2936	2394

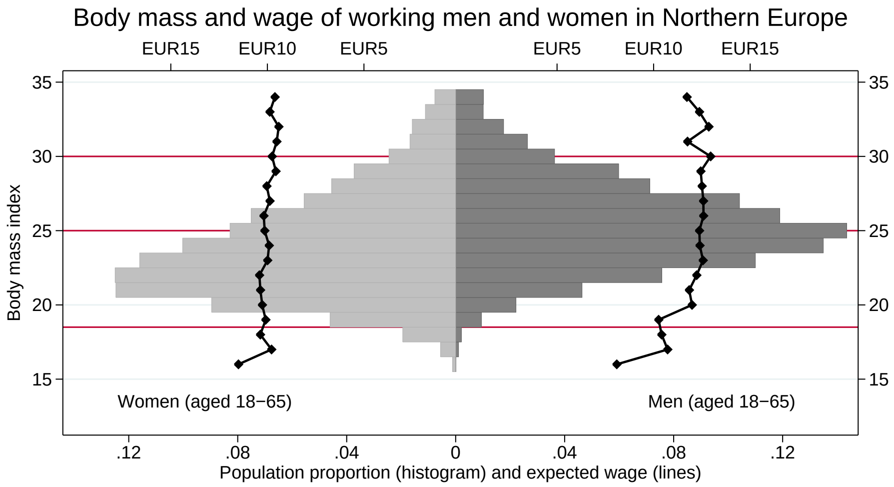
Note: Own calculation based on the ECHP data. M=Males, F=Females. Hourly wages figures are expressed in real 1996 prices and PPP.

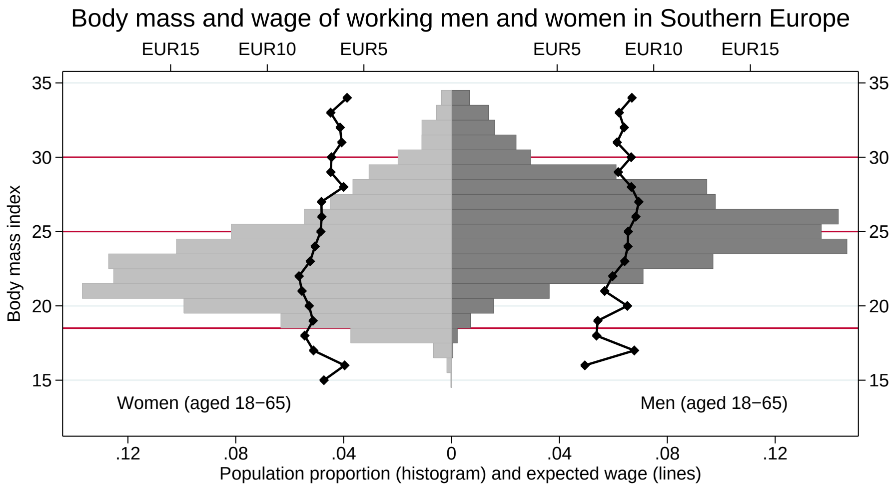
Table 2: Sample averages by gender, ECHP, Southern European Countries

	All		**Italy**		**Greece**		**Portugal**		**Spain**	
	M	F	M	F	M	F	M	F	M	F
Earnings										
Hourly wage	8.33	7.53	9.90	9.31	7.65	6.78	5.63	5.51	9.53	8.28
Body Size										
BMI	25.76	23.38	25.33	22.73	26.14	23.60	25.73	24.17	26.08	23.06
Underweight	0.003	0.041	0.003	0.059	0.001	0.028	0.002	0.025	0.004	0.048
Overweight	0.466	0.214	0.419	0.177	0.545	0.246	0.471	0.262	0.468	0.182
Obese	0.091	0.050	0.072	0.033	0.085	0.044	0.086	0.072	0.123	0.046
Height (in cms.)	172.9	162.4	173.5	162.8	175.7	164.8	170.2	160.9	173.2	162.4
Weight (in kgs.)	77.00	61.59	76.23	60.15	80.62	64.05	74.46	62.51	78.23	60.75
Demographic and employment variables										
Age	40.29	38.83	40.82	39.48	40.95	38.38	39.65	39.22	39.91	37.75
Primary Level of Education	0.523	0.393	0.440	0.284	0.347	0.249	0.771	0.625	0.480	0.323
Secondary Level of Education	0.285	0.333	0.433	0.557	0.377	0.389	0.132	0.169	0.207	0.217
Tertiary Level of Education	0.192	0.274	0.127	0.159	0.277	0.362	0.097	0.206	0.312	0.460
Married	0.745	0.680	0.751	0.690	0.740	0.701	0.746	0.713	0.738	0.611
Poor or Bad Health	0.028	0.036	0.028	0.033	0.007	0.010	0.049	0.062	0.019	0.022
Smoker	0.488	0.306	0.431	0.297	0.659	0.493	0.442	0.153	0.498	0.402
Private	0.722	0.611	0.672	0.542	0.617	0.562	0.781	0.654	0.784	0.676
Part-Time Work	0.013	0.088	0.013	0.091	0.020	0.060	0.008	0.061	0.015	0.138
Occupation										
Professionals	0.235	0.327	0.228	0.334	0.274	0.357	0.171	0.271	0.284	0.372
Clerks	0.241	0.392	0.300	0.452	0.271	0.411	0.211	0.336	0.183	0.375
Craft	0.417	0.112	0.377	0.120	0.386	0.093	0.466	0.138	0.434	0.079
Elementary	0.107	0.169	0.095	0.094	0.069	0.140	0.152	0.255	0.100	0.174
Sample size	27680	19343	8711	6000	4390	2821	7230	5954	7349	4568

Note: Own calculation based on the ECHP data. M=Males, F=Females. Hourly wages figures are expressed in real 1996 prices and PPP.

Figure 1: Body mass and wage of working men and women in Northern Europe (top) and Southern Europe (bottom)





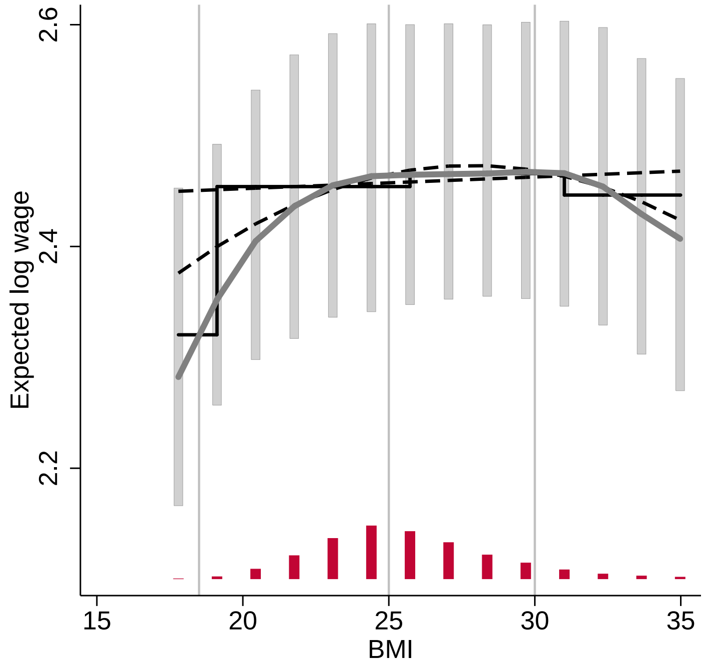
Note: BMI distribution histograms in the range 15–35 are shown horizontally and labelled on the bottom axis (in light grey at left for women, in dark grey at right for men). Mean wages at each BMI level is marked by a diamond and labelled on the top axis.

Table 3: Coefficients on BMI parameters (Northern and Southern European countries)

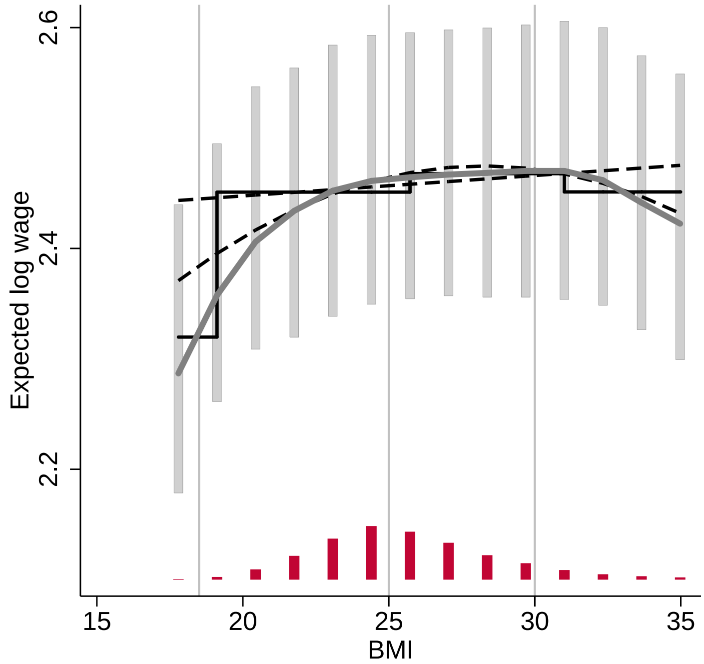
	North		South	
	(1)	(2)	(1)	(2)
Men, 18–65				
Linear specification				
BMI	$0{\cdot}001$	$0{\cdot}002$	$0{\cdot}001$	$0{\cdot}001$
Quadratic specification				
BMI	$0{\cdot}054*$	$0{\cdot}054*$	$0{\cdot}040*$	$0{\cdot}035*$
BMI squared	$-0{\cdot}001*$	$-0{\cdot}001*$	$-0{\cdot}001*$	$-0{\cdot}001*$
Estimated peak BMI	$27{\cdot}8$	$28{\cdot}3$	$27{\cdot}4$	$27{\cdot}5$
Piecewise constant specification				
BMI<18.5 (underweight)	$-0{\cdot}134\dagger$	$-0{\cdot}131\dagger$	$-0{\cdot}080$	$-0{\cdot}072$
25≤BMI<30 (overweight)	$0{\cdot}012$	$0{\cdot}017\dagger$	$0{\cdot}009$	$0{\cdot}009$
BMI≥30 (obese)	$-0{\cdot}008$	$0{\cdot}000$	$-0{\cdot}015$	$-0{\cdot}012$
Women, 18–65				
Linear specification				
BMI	$-0{\cdot}004*$	$-0{\cdot}002\dagger$	$-0{\cdot}004*$	$-0{\cdot}004*$
Quadratic specification				
BMI	$-0{\cdot}016$	$-0{\cdot}016$	$0{\cdot}006$	$0{\cdot}003$
BMI squared	$0{\cdot}000$	$0{\cdot}000$	$-0{\cdot}000$	$-0{\cdot}000$
Estimated peak BMI	> 35	$30{\cdot}8$	< 15	< 15
Piecewise constant specification				
BMI<18.5 (underweight)	$0{\cdot}011$	$0{\cdot}011$	$-0{\cdot}009$	$0{\cdot}008$
25≤BMI<30 (overweight)	$-0{\cdot}022\dagger$	$-0{\cdot}015$	$-0{\cdot}038*$	$-0{\cdot}027*$
BMI≥30 (obese)	$-0{\cdot}039*$	$-0{\cdot}016$	$-0{\cdot}059*$	$-0{\cdot}050*$

Notes: Model specification (1) includes a quadratic function of age, indicator variables for educational attainment, marital status, bad health, being a smoker and a set of time and country dummies. Model specification (2) is as (1) with additional controls for occupation, sector and part-time employment. $*$ and † indicate significance at 1 and 5 percent levels respectively based on cluster robust standard error estimates

Figure 2: BMI and Expected Wages of Northern European Males



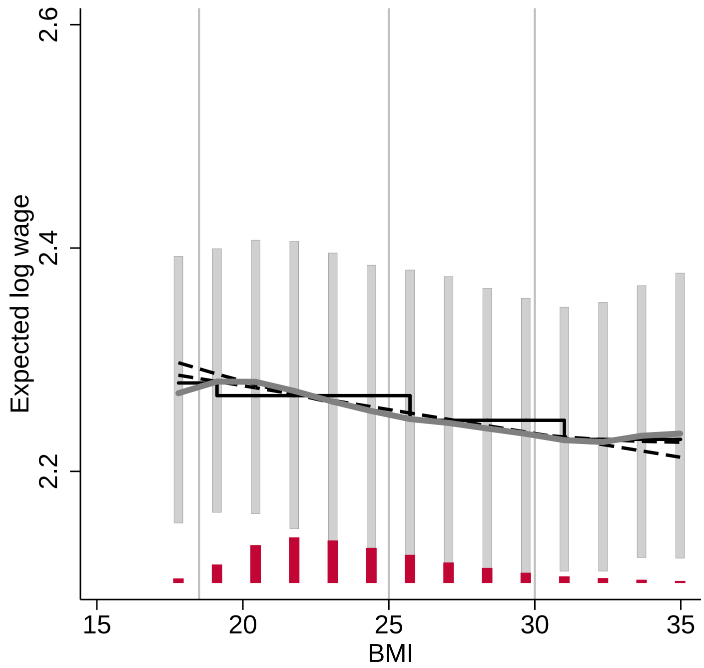
Northern Europe, Men, Model specification (1)



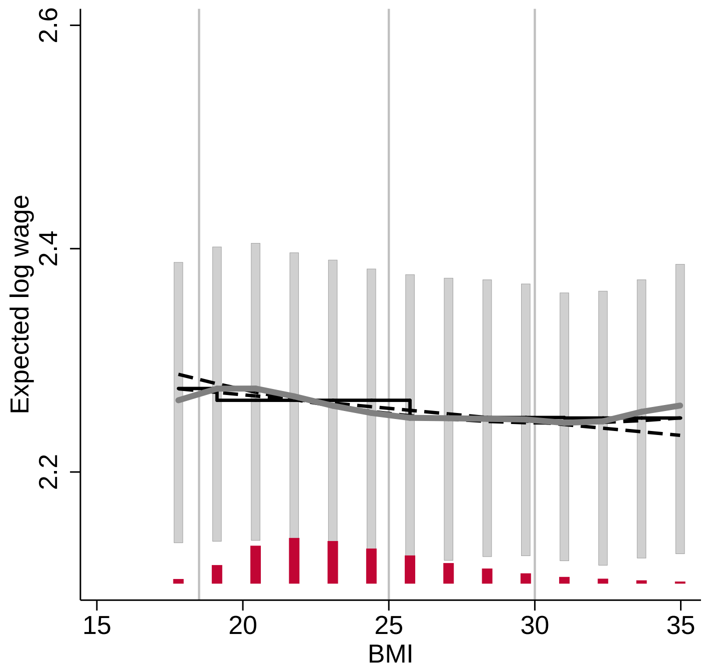
Northern Europe, Men, Model specification (2)

Note: Grey lines show semi-parametrically estimated wage-BMI profiles (with point-wise bootstrap variability bands). Black lines are the corresponding parametric predictions from a piece-wise constant, a linear and a quadratic model. Predictions are computed with all covariates (except BMI) set at their sample means. Density estimates of the distribution of BMI in the sample is reported at the bottom of each plot.

Figure 3: BMI and Expected Wages of Northern European Females



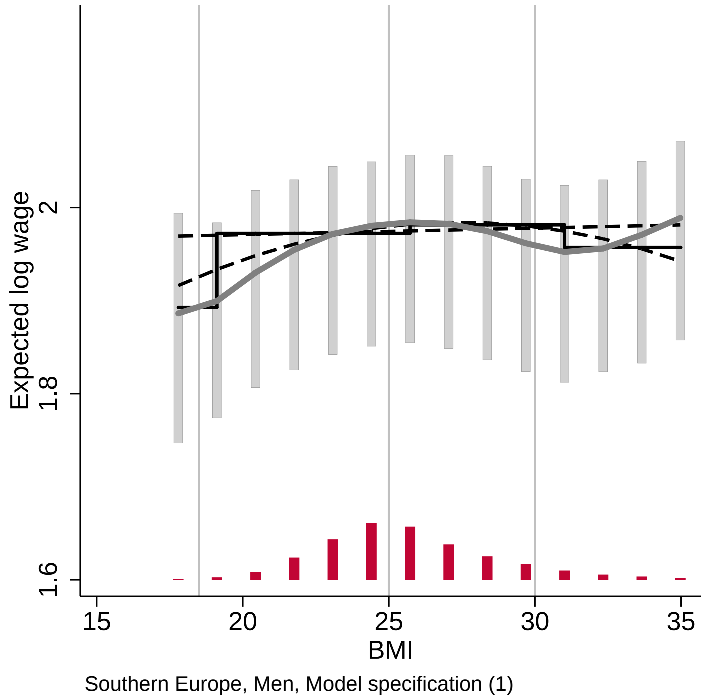
Northern Europe, Women, Model specification (1)

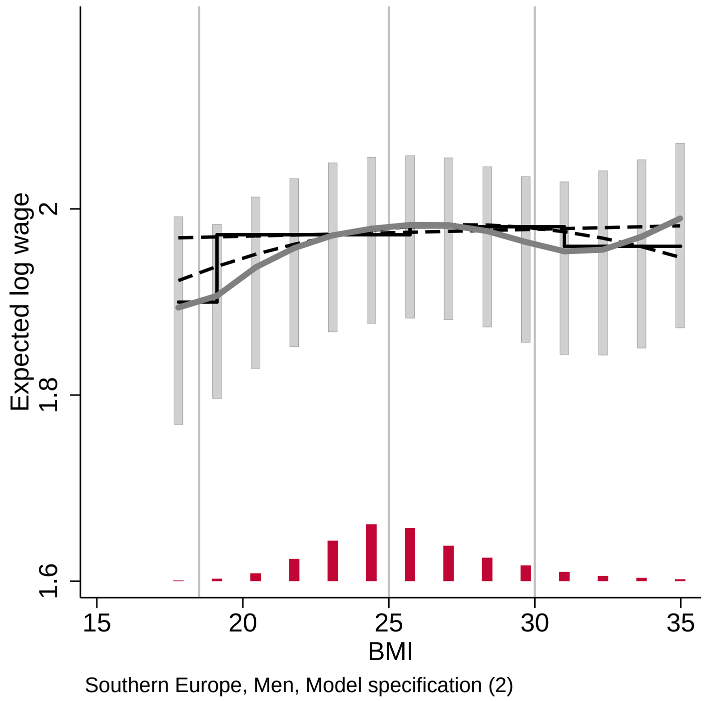


Northern Europe, Women, Model specification (2)

Note: Grey lines show semi-parametrically estimated wage-BMI profiles (with point-wise bootstrap variability bands). Black lines are the corresponding parametric predictions from a piece-wise constant, a linear and a quadratic model. Predictions are computed with all covariates (except BMI) set at their sample means. Density estimates of the distribution of BMI in the sample is reported at the bottom of each plot.

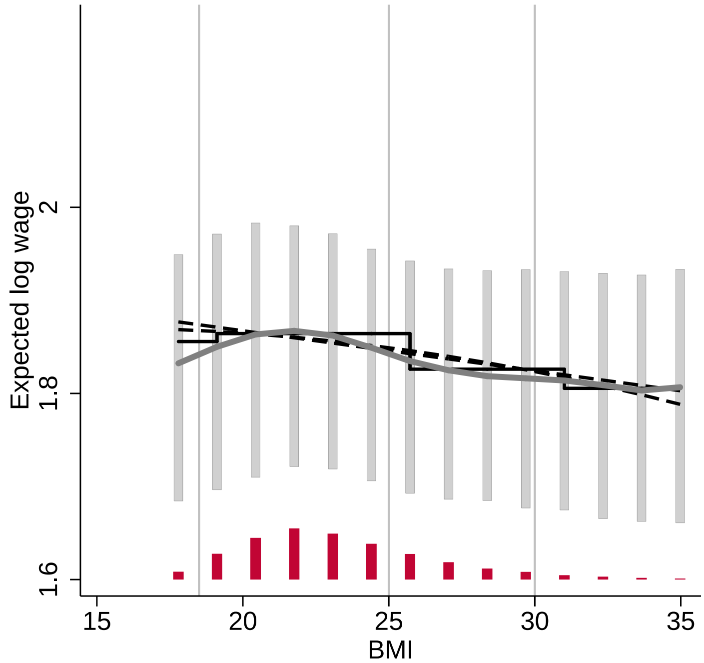
Figure 4: BMI and Expected Wages of Southern European Males



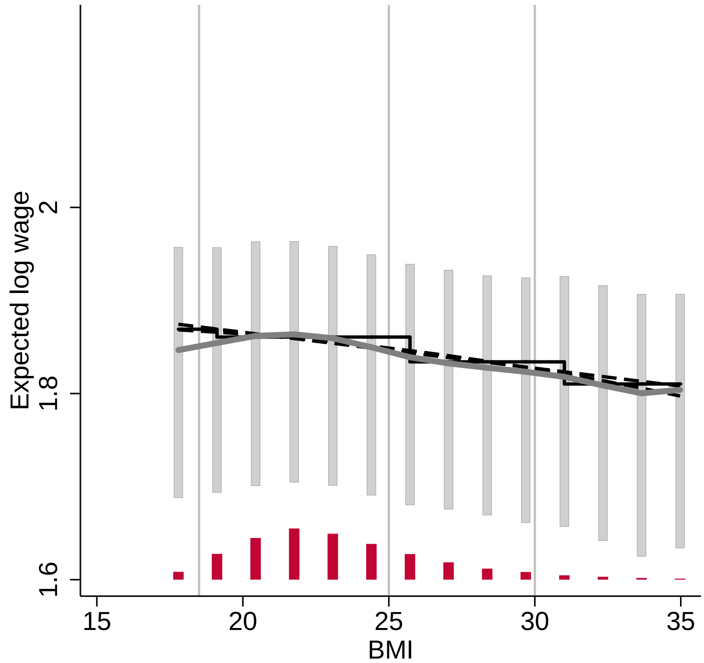


Note: Grey lines show semi-parametrically estimated wage-BMI profiles (with point-wise bootstrap variability bands). Black lines are the corresponding parametric predictions from a piece-wise constant, a linear and a quadratic model. Predictions are computed with all covariates (except BMI) set at their sample means. Density estimates of the distribution of BMI in the sample is reported at the bottom of each plot.

Figure 5: BMI and Expected Wages of Southern European Females



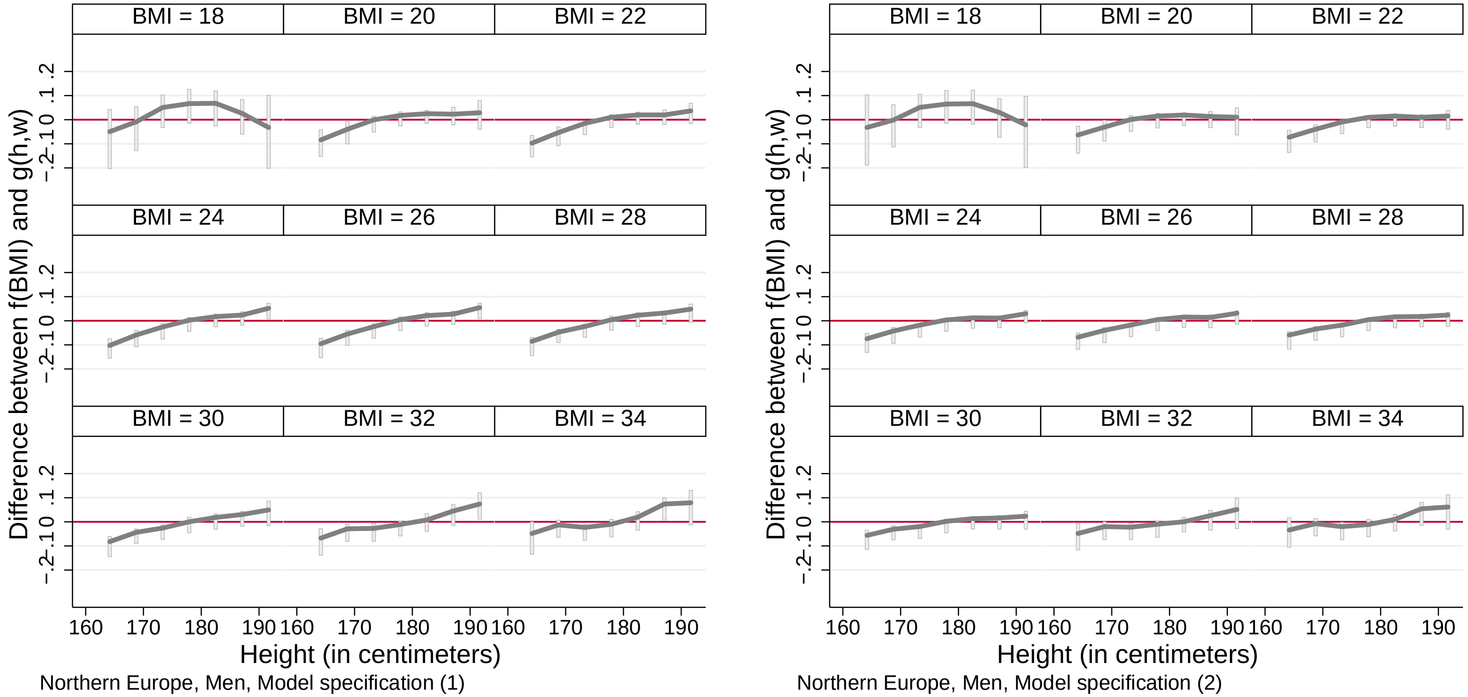
Southern Europe, Women, Model specification (1)



Southern Europe, Women, Model specification (2)

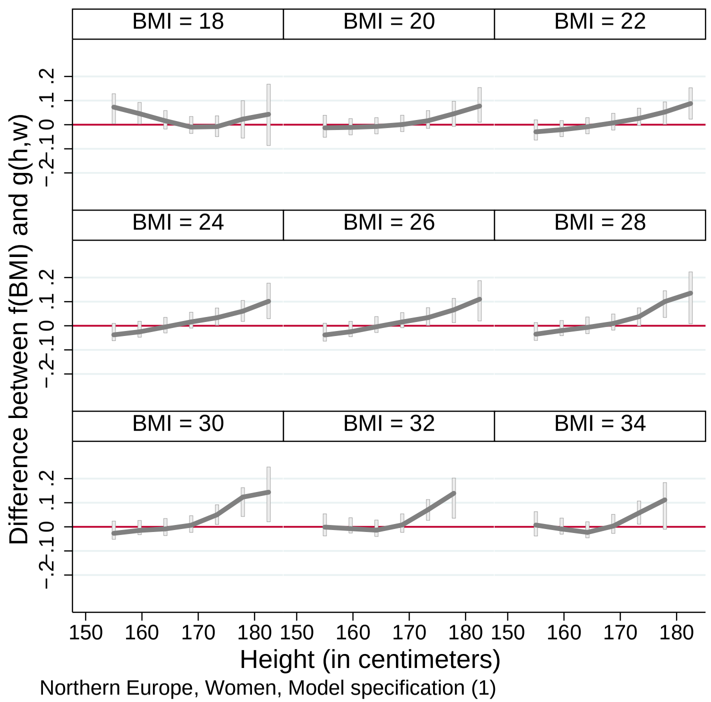
Note: Grey lines show semi-parametrically estimated wage-BMI profiles (with point-wise bootstrap variability bands). Black lines are the corresponding parametric predictions from a piece-wise constant, a linear and a quadratic model. Predictions are computed with all covariates (except BMI) set at their sample means. Density estimates of the distribution of BMI in the sample is reported at the bottom of each plot.

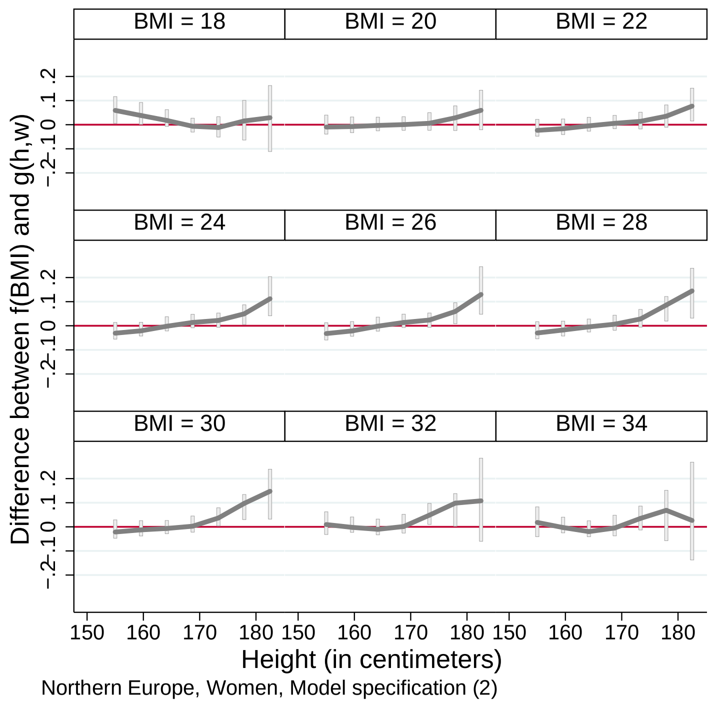
Figure 6: BMI, Height and Weight: Difference in Expected Wages of Northern European Males



Note: Grey lines show the difference between log wage predictions from the unconstrained, bivariate model with height and weight entered separately and log wage predictions from the univariate model based on the BMI index. Differences are reported for nine different BMI levels and, within each, at different heights (x axis). Any deviation of the grey line from the horizontal axis at zero indicates that height has an additional effect on wages. Predictions are computed with all covariates (except height and weight) set at their sample means. Vertical bars are point-wise bootstrap variability bands of the prediction differences. Estimates with variability bands larger than 0.45 are not reported.

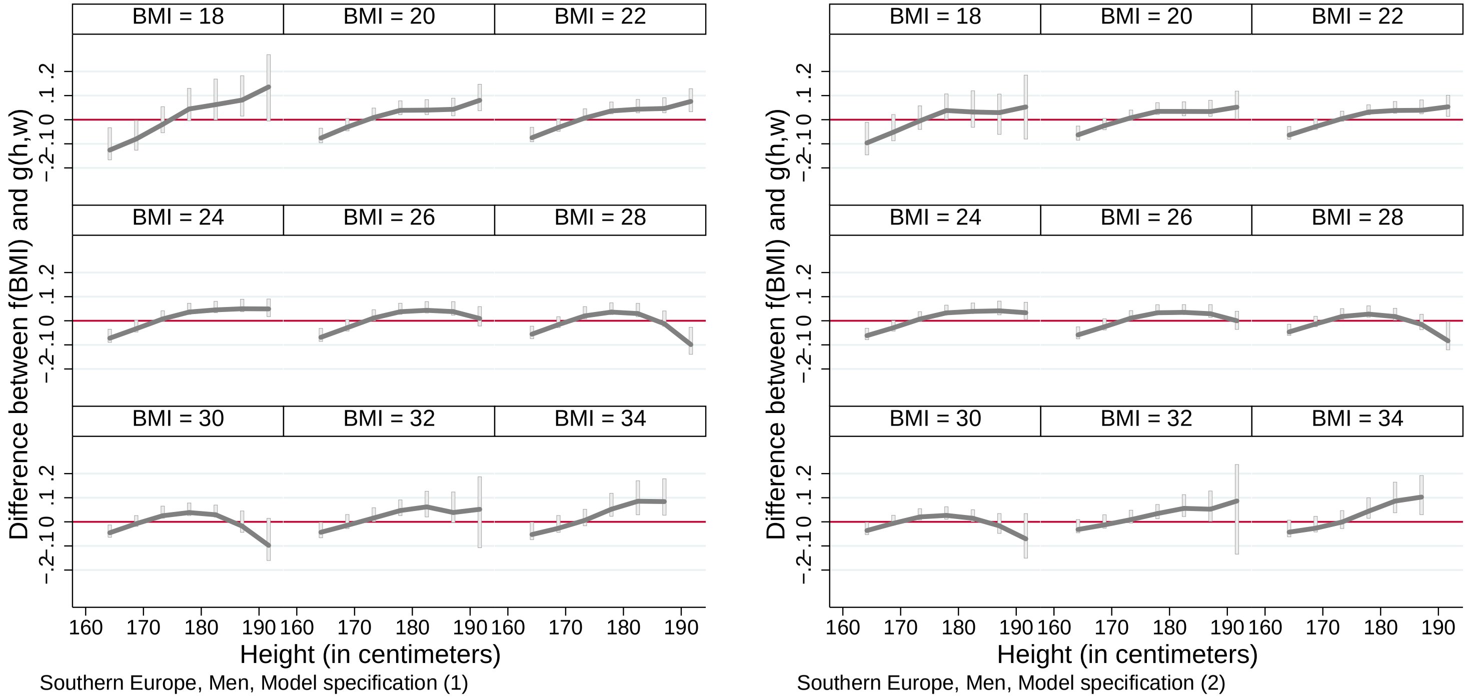
Figure 7: BMI, Height and Weight: Difference in Expected Wages of Northern European Females





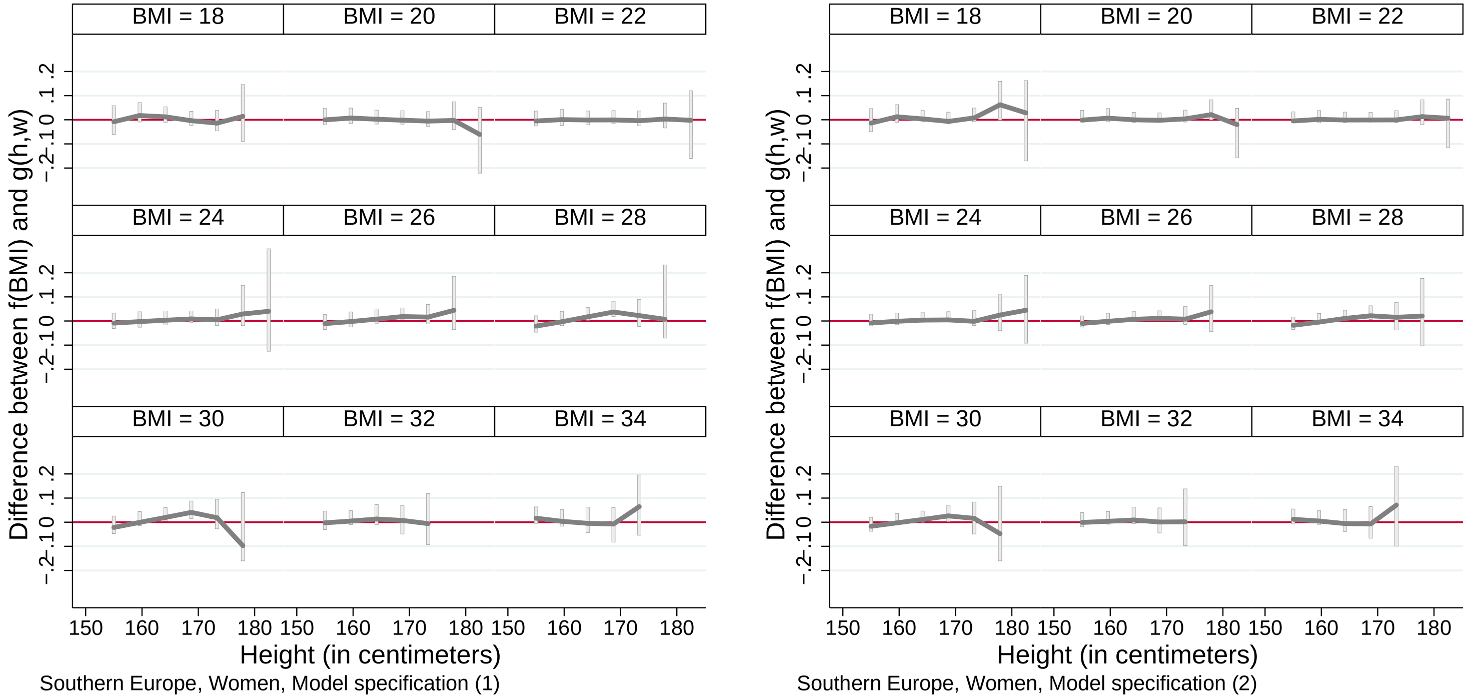
Note: Grey lines show the difference between log wage predictions from the unconstrained, bivariate model with height and weight entered separately and log wage predictions from the univariate model based on the BMI index. Differences are reported for nine different BMI levels and, within each, at different heights (x axis). Any deviation of the grey line from the horizontal axis at zero indicates that height has an additional effect on wages. Predictions are computed with all covariates (except height and weight) set at their sample means. Vertical bars are point-wise bootstrap variability bands of the prediction differences. Estimates with variability bands larger than 0.45 are not reported.

Figure 8: BMI, Height and Weight: Difference in Expected Wages of Southern European Males



Note: Grey lines show the difference between log wage predictions from the unconstrained, bivariate model with height and weight entered separately and log wage predictions from the univariate model based on the BMI index. Differences are reported for nine different BMI levels and, within each, at different heights (x axis). Any deviation of the grey line from the horizontal axis at zero indicates that height has an additional effect on wages. Predictions are computed with all covariates (except height and weight) set at their sample means. Vertical bars are point-wise bootstrap variability bands of the prediction differences. Estimates with variability bands larger than 0.45 are not reported.

Figure 9: BMI, Height and Weight: Difference in Expected Wages of Southern European Females



Note: Grey lines show the difference between log wage predictions from the unconstrained, bivariate model with height and weight entered separately and log wage predictions from the univariate model based on the BMI index. Differences are reported for nine different BMI levels and, within each, at different heights (x axis). Any deviation of the grey line from the horizontal axis at zero indicates that height has an additional effect on wages. Predictions are computed with all covariates (except height and weight) set at their sample means. Vertical bars are point-wise bootstrap variability bands of the prediction differences. Estimates with variability bands larger than 0.45 are not reported.

CEPS
INSTEAD

B.P. 48
L-4501 Differdange
Tél.: +352 58.58.55-801
www.ceps.lu