Journal of Applied Psychology
2011, Vol. 96, No. 1, 95–112

0021-9010/10/$12.00 DOI: 10.1037/a0020860

When It Comes to Pay, Do the Thin Win? The Effect of Weight on Pay for Men and Women

Timothy A. Judge
University of Florida

Daniel M. Cable
London Business School

Cultivation theory suggests that society holds very different body standards for men versus women, and research indicates that the consequences of defying these social norms may not be linear. To test these notions in the employment context, we examined the relationship between weight and income and the degree to which the relationship varies by gender. For women, we theorized a negative weight–income relationship that is steepest at the thin end of the distribution. For men, we predicted a positive weight–income relationship until obesity, where it becomes negative. To test these hypotheses, we utilized 2 longitudinal studies, 1 German and 1 American. In Study 1, weight was measured over 2 time periods, and earnings were averaged over the subsequent 5 years. Study 2 was a multilevel study in which weight and earnings were within-individual variables observed over time, and gender was a between-individual variable. Results from the 2 studies generally support the hypotheses, even when examining within-individual changes in weight over time.

Keywords: weight, gender, sex, earnings, pay

> *Nemo liber est qui corpori servit* (No one is free who is a slave to his body).
>
> —Lucius Annaeus Seneca

> The ideal image of a woman is almost impossible for anybody to achieve.
>
> —Peter J. Brown (anthropologist at Emory University)

The standard of attractiveness portrayed on television and in magazines is slimmer for women than for men (Silverstein, Perdue, Peterson, & Kelly, 1986), slimmer than in the past (Silverstein et al., 1986), thinner than the actual female population (e.g., Fouts & Burggraf, 1999), and often thinner than the criteria for anorexia (Wiseman, Gray, Moismann, & Ahrens, 1992). In the fashion industry, models have become so frail and emaciated that shows have started banning models below certain weight ratios (e.g., Cable News Network [CNN], 2008; Wilson, 2006). On the basis of these pervasive social norms surrounding weight, it is perhaps no surprise that some women have internalized these values: About 90% of cases of anorexia and bulimia are women (Fouts & Burggraf, 1999).

What may be more surprising is the degree to which employers also seem to have internalized the notion that employees' weight matters. Roehling's (1999) comprehensive review suggests that obese individuals are rated as being less desirable as subordinates, coworkers, and bosses, and they are viewed as less conscientiousness, less agreeable, less emotionally stable, and less extraverted than their "normal-weight" counterparts. Even though these stereotypes are inaccurate (Roehling, Roehling, & Odland, 2008), it appears that in the United States, obese employees are viewed by their employers as lazy and lacking in self-discipline (Puhl & Brownell, 2003). Roehling's (1999) review also revealed that overweight women are consistently judged more harshly in the workplace than overweight men, and Griffin (2007) reported that 60% of overweight women and 40% of overweight men describe themselves as having been discriminated against in the course of employment.

One of the most important employment outcomes that may be affected by weight is income, because discrimination against obese employees at any stage of the employment cycle—including hiring, training, or promotions—should ultimately be reflected in employees' income. In addition to affecting the type of life one can live, income can be viewed as a culmination of employees' social and human capital investments (Becker, 1993), and it is a symbol of what society values. As such, income often influences people's identities because it serves as mark of success relative to others and also is an index of career success (Judge, Cable, Boudreau, & Bretz, 1995). Though economists have studied the wage penalties applied to obesity (Baum & Ford, 2004; Cawley, 2004; Conley & Glauber, 2007), we are not aware of any study in the organizational behavior literature that has examined the weight–earnings relationship or gender differences therein. Beyond the obvious relevance of the topic to organizational behavior, there are three additional limitations in past research that we address with the present investigation.

First, existing theory on the weight–income relationship has focused on the effects of being obese but not being thin, and existing empirical research has assumed a linear relationship between weight and income. Past research has not examined whether the weight–income relationship is actually nonlinear because of employers expecting and rewarding specific weight ideals. In fact,

This article was published Online First September 20, 2010.

Timothy A. Judge, Department of Management, Warrington College of Business, University of Florida; Daniel M. Cable, Department of Organisational Behaviour, London Business School, London, England.

Correspondence concerning this article should be addressed to Timothy A. Judge, who is now at the University of Notre Dame, 360 Mendoza College of Business, Notre Dame, IN 46556. E-mail: tjudge@nd.edu

we could not locate any research that developed theory or empirically examined the differences between rewarding thin employees versus penalizing obese employees or whether the mechanism is different for women versus men. Given the extreme emphasis that American society places on women being thin, the literature's focus on obesity may be obscuring important theoretical distinctions and empirical trends. Thus, we develop theory regarding the rewards and penalties that employers apply to men and women on the basis of body weight, which we test using curvilinear analyses.

Second, although existing weight–income research has suggested that the labor market punishes obese people in terms of income, past research has not used a multilevel analysis of longitudinal data to explore within-person relationships between weight and income. A multilevel perspective, although having many theoretical and analytical advantages (Kozlowski & Klein, 2000), also limits interpretational confounding (confounding within- vs. between-individual effects) and lessens the possibility that background variables function as a common cause of weight and career success. For example, someone born into a lower socioeconomic class could possess both a tendency to be overweight and a tendency to hold lower career aspirations. To rule out these alternative explanations for the weight–income relationship, we use a multilevel analysis to examine whether a given person who gains weight is punished in terms of income and how changes in within-person weight affect income differently for thin versus obese individuals. In other words, any "starting" differences in weight are controlled in our analyses because our analyses focus on within-person change and because we model an intercept in the analyses.

Finally, most research on the weight–income relationship has been conducted in the United States. However, the United States is one of the most weight-conscious societies in the world while simultaneously being one of the most obese (e.g., Critser, 2003), suggesting that weight–income research based on U.S. inhabitants may not generalize to other countries. We begin to address this issue by examining whether the trend of results from the United States extends to Germany, a country with some key differences in social and employment expectations.

Theory and Hypotheses

Stereotyping theory predicts that obese individuals will be penalized in the labor market. As described by Greenberg, Eastin, Hofschire, Lachlan, and Brownell (2003, p. 1342), negative stereotypes are attached to obese individuals, who are often thought to be "undisciplined, dishonest, sloppy, ugly, socially unattractive, sexually unskilled, and less likely to do productive work, among other attributes." An open aversion toward fat so thoroughly permeates Western society that even obese individuals hold negative attitudes toward other obese people (Finkelstein, DeMuth, & Sweeney, 2007). These pervasive negative stereotypes also appear to be held by employers, accounting for the widespread obesity discrimination that has been documented, for both men and women, across every part of the employment cycle in the American workplace in both experimental and field studies (Larkin & Pines, 1979; Pingitore, Dugoni, Tindale, & Spring, 1994; Roehling, 1999).

Although past obesity research has relied on stereotyping theory as the rationale for why obese employees experience discrimination in the workplace, the logic of stereotyping is silent about the thin-to-average range of weight, and it also is silent about differences between men and women. In the next section, we use cultivation theory as a starting point to develop theory on the weight–income relationship for nonobese individuals, focusing on the logic for curvilinear relationships and differences between men and women.

Cultivation Theory

According to cultivation theory, the media (especially television) is the most powerful storyteller in Western culture, continually repeating the myths, ideologies, and patterns of relationships that legitimize the social order (Brown, 2002). The key cultivation hypothesis is that the images depicted in the media, over time, act like the pull of gravity toward an imagined center. This pull results in a shared set of social conceptions and expectations such that media portrayals become ideal representations of reality (see Gerbner, Gross, & Morgan, 2002).

In the American media, standards of attractiveness are substantially slimmer for women than for men, and more recent standards for women are slimmer than they were in the past (Silverstein et al., 1986). This media ideal is quite pervasive in society, with female cartoon characters, movie/television actresses, Playboy centerfolds, and Miss America Pageant winners all having become increasingly thinner over the decades (Grabe, Shibley, & Ward, 2008; Silverstein et al., 1986). On the basis of the media's standards, people generally perceive average female weight as overweight, and they perceive very thin women as average in weight. Although media presentations of women's bodies are so skewed as to be out of reach to most, consistent exposure even leads women themselves to adopt this reality, resulting in decreased satisfaction with their bodies (e.g., Brown, 2002; Cash, Winstead, & Janda, 1986).

From a cultivation theory perspective, very thin women are idealized and more socially valued compared with their normal-weight and obese peers. Consistent with this prediction, empirical research shows that very thin women are considered more attractive, better mate choices, and more positive people compared with those of average weight (e.g., Wade & DiMaria, 2003). In fact, women who fail to live up to society's unrealistic weight expectations are often viewed as "rebels" who do not support the classic American Protestant values of self-discipline, thrift, and hard work (Crandall, 1994).

Integrating stereotyping theory with the logic and evidence that very thin women are more socially valued compared with their normal-weight peers, we predict a curvilinear relationship between weight and income. Specifically, we hypothesize that—for women—the strongest relationship between weight and income occurs in the very thin to average weight range. Conceptually, this prediction reflects the premise that when employers encounter very thin female job applicants and employees, who by definition are rare and therefore stand out, they celebrate these individuals with higher pay. However, as women reach average weight, they have already "fallen from grace" according to media images and social expectations (Crandall, 1994), and the relationship between weight and income should level off but still remain negative. Thus, we hypothesize the following:

> *Hypothesis 1:* For women, weight will have a negative linear effect on earnings (Hypothesis 1a), but the quadratic term will be positive (Hypothesis 1b), such that the effect of increasing weight on women's earnings will be particularly negative at low (below-average) levels of weight.

A very different set of social standards exists regarding men's weight. The media's ideal male body is chunky and muscular (Hargreaves & Tiggemann, 2004), characterized by wide shoulders and well-developed chest and arms (Pope, Phillips, & Olivardia, 2000). Like the Barbie doll (which symbolizes an unrealistic female body ideal), today's G. I. Joe figure is just as unattainable, with a bicep almost as big as the waist and bigger than that of the greatest body-builders of all time (Olivardia, 2002). Consistent with cultivation theory, the media images of this male ideal are internalized by many. Compared with average weight men, underweight men are more dissatisfied with their build, feel less handsome with less sex appeal, and feel lonelier (Cohane & Pope, 2001).

Although being skinny is not an ideal for most men, it also clearly is not a social ideal for men to become obese in Western society (e.g., Hebl & Turchin, 2005). As noted by Mishkind, Rodin, Silberstein, and Striegel-Moore (1986), mesomorphic male physiques (i.e., of well-proportioned build) are considered more attractive than ectomorphic (thin) or endomorphic (obese) physiques. Studies have demonstrated that people assign positive personality traits to drawings or photographs of mesomorphic men and mostly negative traits to nonmesomorphic men. For example, Kirkpatrick and Sanders (1978, pp. 91–92) showed that the traits ascribed to mesomorphic men were very positive (i.e., best friend, has lots of friends, polite, happy, helps others, brave, healthy, smart, and neat). By contrast, endormorphic men were characterized by one set of negative traits (i.e., sloppy, dirty, worries, lies, tired, stupid, lonely, and lazy), whereas ectomorphic men were described with a different set of negative traits (i.e., nervous, sneaky, afraid, sad, weak, and sick).

Thus, the research literature reveals deeply entrenched cultural preferences toward well-proportioned men and aversions to skinny and obese men. Conceptually, men who fail to resemble the masculine body ideal fail to live up to gender-role norms and may thus experience negative consequences of violating such norms (Mishkind et al., 1986). To the extent that employers also have internalized these social norms, men should be rewarded for weight gain until the point of obesity, at which point they should be penalized. Although past research on male–female differences in the weight–income relationship has assumed a linear relationship, has been limited to the United States, and has not used multilevel analyses of longitudinal data, existing research generally supports the notion that overweight women are punished more than overweight men (Baum & Ford, 2004; Cawley, 2004; Conley & Glauber, 2007). On the basis of this theorizing, we hypothesize the following:

> *Hypothesis 2:* For men, weight will have a positive linear effect on earnings (Hypothesis 2a), but the quadratic term will be negative (Hypothesis 2b), such that men's earnings will increase with weight up to a point, at which point (at above-average levels of weight) the relationship becomes less positive or negative.

Generalizing From the United States

America appears to be fascinated with people's weight, and there is strong evidence of pervasive "antifat" sentiment in the United States (Crandall, 1994). In contrast to this social fixation with weight, constant access to cheap fast food is part of the American way of supersized life (Pollan, 2006). Americans are eating out more than ever before (Briefel & Johnson, 2004), and eating out is associated with higher fat, calories, cholesterol, and larger portions (Duffey, Gordon-Larsen, Jacobs, Williams, & Popkin, 2007). Perhaps as a result of diet, U.S. obesity percentages began to spike in the late 1980s across all levels of income (Chang & Lauderdale, 2005), and now over 30% of the U.S. population is obese—twice the percentage in 1990 (The Economist, 2004). In fact, it could be said that becoming obese is a predictable response to the American environment (Critser, 2003, p. 3).

Given the idiosyncratic dynamics in American society regarding weight, we offer a rare glimpse at the weight–income relationship in Germany. Although the United States and Germany are both considered prototypical examples of Western industrialized democratic societies, there are some key differences in the sociostructural conditions between the countries (Staudinger, Fleeson, & Baltes, 1999). First, the income Gini coefficient (Gini, 1921), which is used as an indicator of income inequality, is about 30% higher in the United States than in Germany (Staudinger et al., 1999). Because income inequity is more culturally acceptable in the United States, income differences between men and women on the basis of weight also may be more acceptable in the United States than in Germany.

Next, the German social welfare state differs from the United States. For example, the public expenditure on health in the United States is 6.6% of the gross domestic product, whereas in Germany it is 8.2% (Staudinger et al., 1999). Given the well-publicized health effects of obesity (e.g., Rundle, 2003), Americans may be less willing than Germans to subsidize the health costs of obesity, resulting in more negative attitudes toward obesity in the United States. Finally, German employees are better legally protected in employment decisions regarding wages and benefits than their U.S. counterparts. This implies that employers have less opportunity to discriminate against overweight individuals, in particular women who gain weight over the course of employment. Thus, although both the United States and Germany are Western industrialized democratic societies, several structural differences imply that weight may be less predictive of income in Germany. By exploring the hypotheses in both countries, we hope to begin to address the cross-cultural generalizability of our results.

Study 1

Overview

The purpose of Study 1 was to use a German sample to test the differential effect of weight on income by gender. Below, we describe the control variables that we used in Study 1, and then we proceed to describe the methodology and results of the study.

Control variables. In estimating the differential effect of weight on earnings for men and women, we controlled for a number of variables that have been shown to affect weight, earnings, or both variables. The most obvious control variable is height. This is not only importantly conceptually—whether a particular weight would be considered overweight fundamentally depends on height, and height is correlated with both gender and earnings (Judge & Cable, 2004)—it is empirically relevant as well.[1] Because weight tends to increase with age (Flegal, Carroll, Ogden, & Johnson, 2002), and because age is correlated with earnings for both men and women (Barnum, Liden, & Ditomaso, 1995), we also controlled for age. Next, we controlled for marital status because a "relatively consistent finding is that married individuals achieve higher levels of objective success than unmarried individuals" (Judge et al., 1995, p. 487) and because marriage may reward men more than women (Pfeffer & Ross, 1982). Having children may have long-term effects on women's weight (Linné, Barkeling, & Rössner, 2002), and having school-age children disproportionately penalizes women in the labor market (Tilly & Albelda, 1994); thus, we also controlled for the presence of school-age children in the household. Likewise, having children is associated taking time off from work, and therefore we controlled for maternal/paternal leave. Next, because health problems can stem from weight and can lead to absenteeism, lower job performance, and lower pay, and to separate weight effects from other unhealthy behaviors, we controlled for overall health as well as drinking and smoking behavior (Sturm, 2002).

Hours worked is a consistent predictor of earnings (Judge et al., 1995) and yet is one in which men and women tend to differ (Feldman, 2002). Accordingly, we also controlled for hours worked. Human capital theory (Becker, 1993) posits that employers reward workers' accumulated stocks of skills and knowledge. Common indicators of human capital are education and seniority (Ballou & Podgursky, 2002), and we used educational attainment and organizational tenure as control variables. Likewise, because people with greater ability gravitate to higher level and better paying jobs (Wilk, Desmarais, & Sackett, 1995), at the job level we controlled for necessary training and intrinsic job characteristics. We also controlled for self-esteem because it is related to job performance (Sekiguchi, Burton, & Sablynski, 2008) and to other behaviors that may undermine pay, such as deviance (Ferris, Brown, Lian, & Keeping, 2009). We also controlled for industry because research has shown income levels to depend on industry (Judge et al., 1995). Because it is "well documented that foreign workers earn less than natives in the German labor market" (Constant & Massey, 2005, p. 489), we controlled for whether an individual was a German native. Finally, because gender segregation has been found to be lower in the civil service, and because this segregation may explain male–female earnings differentials (Lewis, 1996), we controlled for whether the individual was employed in a civil service position.

Method

Participants and procedure. Participants in Study 1 were individuals enrolled in the German Socio-Economic Panel Study (GSOEP), a panel study originally initiated in 1984 by *Deutsches Institut für Wirtschaftsforschung* (DIW Berlin). GSOEP participants were initially chosen at random from a representative cross-section of the adult population living in private households in Germany. Originally, only individuals in the Federal Republic of Germany (West Germany) were included, but after reunification in 1990, individuals in the former German Democratic Republic (East Germany) were added.

Nearly all of the GSOEP data were collected via personal interviews (or, starting in 1998, both traditional and computer-assisted interviews), in which interviewers followed a protocol in asking participants questions, listened to their responses, and recorded their answers. To preserve the causal logic of our study, we limited our study to the GSOEP variables that we collected concurrent with and after the weight variables were first assessed in 2002 (and 2004).

Because sample attrition has occurred over time, and because participants were added in five different waves after the 1984 initial survey (1990, 1995, 1998, 2000, and 2002), response rates are difficult to quantify. In general, however, sample attrition has been relatively low. Among the original participants, the attrition rate has averaged 3.0% per year. Across all waves of data, since the time frame of the current study (2002), 83.8% of the individuals in the sample in 2002 remained in the sample in 2006.

Analyses were limited to individuals working 20 hr or more per week and who had complete data over the 5-year (2002–2006) time interval. All told, 11,253 individuals were included in the analyses. Of those individuals, 57.6% were men, 92% were natives of Germany, and 63% were married. The average participant was 44.23 years of age, had 10.51 years of organizational tenure, worked 37.02 hr per week, earned a monthly salary of €2,407 ($3,228), and had a net household income of €2,931 ($3,930). The average man was 178.86 cm tall (5 ft, 10 in.) and weighed 82.67 kilograms (182.26 lbs). The average woman was 166.33 cm tall (5 ft, 5 in.) and weighed 66.29 kilograms (146.14 lbs).

Focal variable measures.

Wages/salary. Participant wages/salary was measured by asking them to report their "wages or salary as employee, including wages for training (*Vorruhestand*) and wages for sick time (*Lohnfortzahlung*), Gross amount per month EURO." Because this was measured each year, 2002–2006, wages/salary was computed to

[1] Body mass index (BMI)—devised by Belgian scientist Lambert Adolphe Jacques Quételet (1796–1874)—is a frequently used measure of obesity. Although BMI continues to be widely used in health and medical research, it has significant measurement limitations. First, BMI measurements are inferior to other measurements of obesity (e.g., Akpinar, Bashan, Bozdemir, & Saatci, 2007; Mascie-Taylor & Goto, 2007). The measurement problems can be traced to unintended biases of the measure: BMI measures appear to be artifactually affected by age, exercise, physical conditioning, race, gender, and height, among other factors (see Prentice & Jebb, 2001). Second, because BMI is an algebraic combination of height and weight $\left(\mathrm{BMI} = \frac{\mathrm{Weight}\,[\mathrm{lbs}] \times 703}{\mathrm{Height}\,[\mathrm{in.}]^2}\right)$ it obscures or confounds these constituent elements. Thus, as is the case with difference scores and other indices (Edwards, 1994), the measure implicitly contains numerous constraints that are left untested in the composite index. Finally, and related, many algebraic measures have conceptual ambiguities in which the psychological meaning implicit in some of the operations is unclear. In this case, the necessity or meaning of multiplying the numerator (weight) or squaring the denominator (height) is unclear. A less constrained measurement approach is to "unpack" the elements and to estimate the separate impact of height and weight on the criteria.

the average score over the 5 years. The reliability of this scale was $\alpha = .96$ ($\alpha = .95$ for men; $\alpha = .97$ for women).[2]

Weight. Participant weight was measured in 2002 and 2004 when interviewers asked participants, "How many kilograms do you currently weigh?" The responses for the two years were then averaged. The reliability of this scale was $\alpha = .96$ ($\alpha = .96$ for both men and women). The quadratic term was computed by squaring the variable (Cohen, Cohen, West, & Aiken, 2003).

Gender. Gender was measured with a GSOEP generated variable that was scored 1 = *male*, 2 = *female*.

Control variable measures.

Height. Participant height was measured in 2002 and 2004 when interviewers asked individuals, "How tall are you?" Individuals reported their height in centimeters for both years, and these two measurements ($\alpha = .98$) were averaged.

Age. Participant age was assessed by subtracting the participant's reported birth year from 2008.

German native. Whether the participant was German native was assessed with a 2002 question that asked individuals to report in which country they were born. Responses were coded such that a score of 1 was assigned if they were born in Germany, and a score of 0 was assigned otherwise.

Technical, university, and graduate degree. Participant education was measured by their response on the most recent (2006) survey indicating their level of educational attainment. The responses were coded to reflect the German education system, such that one dummy variable (1 = *yes*, 0 = *no*) was created indicating whether the participant held a technical degree (akin to a community college degree). Another dummy variable was coded to represent a university degree (equivalent to a U.S. undergraduate degree). A third dummy variable was created to represent whether the individual held a graduate degree.

Smoking and alcohol use. Smoking behavior was assessed by averaging, 1999–2006, responses to the question, "Do you smoke?" Responses were coded 1 = *yes* and 0 = *no*, and then they were averaged over the years. Alcohol use was measured by whether individuals regularly reported consuming, during the only year such questions were asked (2006), the following alcoholic beverages: (a) beer, (b) wine/champagne, (c) spirits (schnapps, brandy, etc.), and (d) mixed drinks. Thus, scores on this variable represent whether the responded reported being a regular user of any of these alcoholic beverages.

Married. Whether the participant was married was measured with GSOEP variables that coded whether individuals were married in each survey year from 2002 to 2006. The responses were scored to reflect 1 if the individual was married and 0 if they were single, widowed, divorced, or separated. An average score was computed across the 5 years ($\alpha = .94$).

Children less than 16 years of age. Whether the participant had any children less than 16 years of age in the household was measured with a question asking participants, "Do children who were born in [year] or later live in your household?" in which the year given was 16 years prior to the survey year. Participants' responses were averaged over the 5 years ranging from 2002 to 2006 ($\alpha = .95$).

Maternity/paternity leave. Whether individuals took maternity or paternity leave was assessed through a series of questions, asked in the 1999–2006 interviews, "Are you currently on maternity leave or on paid leave as a new parent?" Responses were coded 1 = *yes* and 0 = *no*, and they were averaged over the years so that scores on the variable represent the proportion of time the individual reported being on maternity/paternity leave.

Hours worked. Hours worked were measured by variables in the GSOEP database that assessed agreed upon work time per week, 2002–2006. These measurements ($\alpha = .93$) were averaged over the five time periods.

Organization tenure. Participants' tenure with their organization was assessed with GSOEP variables that reflected the "tenure length of time with (current) firm" in years for 2002–2006. These five measurements were averaged ($\alpha = .96$) to compute an average tenure variable.

Civil service job. Whether the participant held a civil service job was measured with a question each year, 2002–2006, asking participants, "Does the company in which you are employed belong to the public sector?" The responses to this question were scored 1 = *yes* and 0 = *no*, and then they were averaged over the five time periods ($\alpha = .59$).

Intrinsic job characteristics. Respondents' perceptions of intrinsic job characteristics were measured by their responses to seven questions (four in 1995 and three in 2001) that assessed the degree to which their job afforded autonomy and challenge and enhanced development. Example questions include the following: "Do you determine the way your work is done?" and "Do you often learn something new on the job, something which is relevant for your

[2] As is nearly always the case, the measure of pay was not normally distributed. Both the skew (*SK*) and kurtosis (*KT*) statistics were significant for both men ($SK_M = 2.16$, $p < .01$; $KT_M = 12.71$, $p < .01$) and women ($SK_W = 2.42$, $p < .01$; $KT_W = 25.67$, $p < .01$), revealing that for both genders, the distribution has a positive skew (observations clustered toward left-hand of distribution, with a long right-hand tail) and is platykurtotic (flatter than normal). We performed two nonlinear transformations of the pay variable—taking the natural log and square root of pay—and both transformations substantially reduced the skew and kurtosis of the distribution for both genders. Similarly, ordinalizing pay (breaking pay into 10 ordinal categories) also reduced skew and kurtosis for men and women. When we used these transformed variables in the analyses, the significance of both the linear and quadratic weight terms were unchanged for men and women (the size of the coefficients became considerably stronger for men and slightly weaker, but still significant, for women). This suggests that nonnormality, or differences in nonnormality, does not explain the differential weight effects for men and women. We did not use the transformed variables in the analyses for two reasons. First, any data transformations change the meaning of the distribution of the variable, in this case, replacing a linear relationship with a nonlinear one (Russell & Dean, 2000). As Gullikson (2006, "Interpreting OLS Regression and Transformations") noted with respect to the most common transformation, the natural log: "Taking the natural log of the independent variable, for example, implies a 'diminishing returns' relationship." In this study, such an assumption would seemingly play mischief with our hypothesized functional forms, which are already nonlinear (differing, of course, by gender). Second, Busemeyer and Jones (1983) and Russell and Dean (2000, p. 168) cautioned against the conventional use of such transformations on statistical grounds. Russell and Dean noted that log transformations of positively skewed dependent variables (which is exactly the case here) "greatly enhance" the probability of committing a Type II error (concluding that an effect is nonsignificant when, in fact, it is significant). Despite our decision not to use transformed measures of pay, the analyses of such measures do suggest that the differential weight effects by gender are not due to distributional differences in pay between men and women.

career?" Responses were coded on a 3-point scale (1 = *applies*, 2 = *partially applies*, and 3 = *does not apply*), which was subsequently recoded so that high scores reflected high intrinsic characteristics. The reliability of this seven-item scale was α = .70.

Job training requirements. Job training requirements were assessed by averaging—from 1986 to1999, on an interviewer-coded 7-point scale ranging from 1 (*no training*) to 7 (*university or college*)—the required training for the job held by the participant (α = .84).

Perceived health. Perceived health was measured by participants' responses to two questions, asked over several time periods, about their current health status. One question asked, "How would you describe your health at present?" to which participants responded, using a 5-point scale ranging from 1 (*very good*) to 5 (*very poor*), from 1996 to 2001. The other item asked participants to rate their "current self-rated health status," using a 5-point scale ranging from 1 (*very good*) to 5 (*bad*), from 1994 to 2001. These responses were reverse-scored so that high scores indicated good health, and then they were averaged across the two items over the time periods (α = .92).

Self-esteem. Self-esteem was assessed by participants' responses to 17 statements asked from 1994 to 1999. Sample statements include, "Most plans I make are successful" and "In comparison to others, I haven't achieved what I deserve." Participants responded using a 4-point scale ranging from 1 (*applies completely*) to 4 (*does not apply*). Responses were coded to make high scores reflect high self-esteem and were averaged across the items (α = .76).

Analyses. To test the linear and nonlinear effects of weight on income, we estimated a regression equation using income as the dependent variable, the weight linear and quadratic terms as independent variables, and the aforementioned control variables. Because a pooled (common) regression equation rests on the untested assumption that the effects of the independent variables are the same for two subgroups (Dougherty, 2006), consistent with other research on gender and earnings (e.g., Hirsch & Leppel, 1982), we estimated separate regression equations for men and women. We report both unstandardized ($\hat{B}$) and standardized ($\hat{\beta}$) regression coefficients.

To test the equality of the coefficients on weight for men and women, we used Kennedy's (2003) subset test, derived from Chow's (1960) widely used test. Specifically, whereas Chow's omnibus test assesses the equivalence of an entire regression equation (intercept and slopes of all independent variables) across two groups, Kennedy's procedure allows tests of single coefficient estimates across two groups or subsets of coefficients. Because the present study concerned the differential effects of weight for men and women, only the weight coefficients (the weight linear and quadratic terms) were tested using Kennedy's procedure. To verify these results, we also performed the analysis using the general linear model procedure in SPSS, which tests polynomial (in this case, quadratic) effects and interactions (including, in this case, an interaction with a quadratic effect).

Results

Table 1 provides the descriptive statistics of and intercorrelations among the study variables. The regression results are provided in Table 2. As the table shows, many of the control variables predicted wages for both men and women. Of the control variables, height, age, education (all three degrees), marital status, hours worked, organizational tenure, intrinsic job characteristics, job training requirements, and self-esteem all positively predicted earnings for both men and women, though the relative strength of the effects varied (e.g., marital status mattered more for men, whereas hours worked mattered more for women). Perceived health negatively predicted wages for both men and women. Three of the control variables predicted wages for men but not women. Smoking and alcohol use significantly negatively predicted wages for men but not women, and having children less than 16 years of age in the household significantly positively predicted wages for men but not for women. Two variables—taking maternity/paternity leave and working in a civil service job—negatively predicted wages for men but positively predicted wages for women.

Turning to the weight variables, the linear term negatively and significantly predicted wages for women ($\hat{B} = -21.29$, $\hat{\beta} = -.24$, $p < .01$), supporting Hypothesis 1a. Supporting Hypothesis 2a, the linear term positively and significantly predicted wages for men ($\hat{B} = 31.47$, $\hat{\beta} = .24$, $p < .01$). Supporting Hypothesis 1b, the quadratic term positively and significantly predicted wages for women ($\hat{B} = .11$, $\hat{\beta} = .18$, $p < .05$). For men, the opposite was true—the quadratic negatively and significantly predicted wages for men ($\hat{B} = -.17$, $\hat{\beta} = -.24$, $p < .01$), supporting Hypothesis 2b. To test whether the associations differed by gender, we first conducted the previously described Chow test. The test statistic was highly significant ($F = 42.81$, $p < .001$), indicating that the effect of weight on earnings differed significantly by gender. Second, estimating a multivariate regression using the general linear model procedure, results indicated that the Gender × Weight effect was significant ($F = 36.39$, $p < .001$), as was the Gender × Weight^2 effect ($F = 14.22$, $p < .001$). In this omnibus analysis, the incremental variance explained for the interactions was $R^2 = 9.4\%$ ($p < .01$) for the Gender × Weight interaction and $R^2 = 6.9\%$ ($p < .01$) for the Gender × Weight^2 interaction. The incremental variance explained by the weight and weight^2 "main effect" terms was $\Delta R^2 = 1.24\%$ (*ns*) and $\Delta R^2 = 0.83\%$ (*ns*), respectively. For the main effect for gender, that is, effect of gender on earnings, $\Delta R^2 = 0.90\%$ ($p < .01$).[3]

To depict the nature of these effects, we plotted regression results for men and women. The predicted values for men and women are provided in Figure 1. In plotting the predicted values, we standardized weight such that the weight is expressed on the x-axis as a deviation from each group's average (e.g., –15 repre-

[3] For linear regression models, as noted by Hagquist and Stenbeck (1998), "The utility of R^2 as a measure of goodness of fit is highly controversial among methodologists" (p. 234). There are defenders of the utility of R^2 (Lewis-Beck & Skalaban, 1990) and critics (Achen, 1982, 1990). Although one might reasonably argue over the merits of the coefficient of determination for linear models, for nonlinear models, the situation is much less sanguine. As noted by Cameron and Windmeijer (1997), the coefficient of determination is not well-suited for nonlinear models, and with such models, estimates of R^2 tend to be inefficient. For these reasons, in the economics, political science, and other literatures, for nonlinear models, "R^2 measures are very rarely used" (Cameron & Windmeijer, 1997, p. 330). Nevertheless, we report them here for informational purposes, with the aforementioned interpretational caveats.

Table 1
Means, Standard Deviations, and Intercorrelations Among Variables (Study 1)

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1. Gender (1 = *male*, 2 = *female*)	1.42	0.49	1.00																					
2. Height (cm)	173.53	9.14	−.68	1.00																				
3. Age (years)	44.18	12.28	−.06	−.09	1.00																			
4. German native†	0.92	0.26	.03	.10	.02	1.00																		
5. Technical degree†	0.14	0.19	−.02	.05	−.01	.04	1.00																	
6. University degree†	0.38	0.29	.03	.03	.03	.10	−.10	1.00																
7. Graduate degree†	0.19	0.39	.01	.02	.18	.07	.15	.61	1.00															
8. Smoking†	0.46	0.50	−.08	.06	−.13	−.07	−.04	−.12	−.16	1.00														
9. Alcohol use†	0.39	0.49	−.27	.21	−.01	.07	.06	.10	.10	.08	1.00													
10. Married†	0.63	0.43	−.10	−.01	.45	−.05	−.03	−.04	.06	−.14	−.07	1.00												
11. Children <16 years of age†	0.35	0.44	−.09	.09	−.17	−.07	.01	−.02	−.02	.02	.01	.25	1.00											
12. Pregnancy leave†	0.06	0.23	.25	−.15	−.15	.00	.02	.05	.01	−.02	−.08	.00	.22	1.00										
13. Hours worked per week	37.03	5.16	−.36	.25	−.08	−.01	.02	−.05	−.04	.06	.09	−.05	−.05	−.17	1.00									
14. Organization tenure	10.49	8.83	−.09	−.01	.56	.02	.00	−.01	.07	−.10	.03	.27	−.11	−.09	−.03	1.00								
15. Civil service job†	0.13	0.22	.11	−.05	.14	.10	−.01	.11	.13	−.09	−.09	.09	−.04	−.01	−.09	.20	1.00							
16. Job characteristics	2.20	0.33	.00	.11	−.21	.18	.06	.16	.14	−.04	.01	−.11	−.01	.01	.04	−.11	.07	1.00						
17. Training requirements	3.39	0.83	−.06	.04	.22	.12	.05	.13	.22	−.05	.05	.11	−.02	−.03	.03	.18	.06	−.23	1.00					
18. Perceived health	2.64	0.47	.05	−.05	.11	.08	.01	−.03	−.02	−.02	−.03	.03	−.09	−.07	−.03	.07	.05	.28	−.24	1.00				
19. Self-esteem	2.90	0.61	.01	.08	−.12	.26	.04	.09	.07	−.04	.03	−.05	.00	.03	.00	−.06	.04	.18	.05	−.04	1.00			
20. Weight (kg)	75.84	15.25	−.53	.59	.21	.04	.03	−.04	−.01	.00	.14	.17	.07	−.15	.19	.15	−.04	−.04	.08	.03	−.02	1.00		
21. Weight^2 (kg^2)	5,984.44	2,493.85	−.49	.56	.19	.04	.03	−.04	−.01	.00	.13	.16	.06	−.14	.18	.14	−.03	−.03	.07	.04	−.02	.98	1.00	
22. Wages/salary per month (€)	2,400.63	1,562.91	−.28	.24	.41	.05	.09	.26	.36	−.11	.07	.28	.05	−.10	.19	.33	.08	.12	.21	.00	.05	.25	.23	1.00

Note. N = 11,340. For $|r| \geq .03$, $p < .01$. Pregnancy leave = maternity/paternity leave; Job characteristics = intrinsic job characteristics; Training requirements = job training requirements.
† Denotes dummy variables coded as 1 = *yes*, 0 = *no*.

Table 2
Regressions Predicting Effect of Weight on Earnings for Men and Women (Study 1)

Variable	Men			Women		
	B	SE_B	β	B	SE_B	β
Constant	−9,857.09**	594.58		−5,256.58**	432.12	
Height	22.98**	2.79	.093**	17.93**	2.26	.102**
Age	51.23**	2.05	.364**	31.21**	1.52	.344**
German native	−51.48	66.59	−.008	−107.06	56.15	−.023
Technical degree	528.71**	88.94	.060**	265.54**	74.41	.042**
University degree	706.57**	80.14	.115**	712.52**	54.77	.189**
Graduate degree	862.64**	59.97	.195**	368.52**	41.88	.131**
Smoking	−70.95*	33.43	−.021*	30.65	26.64	.014
Alcohol use	−170.43**	33.24	−.049**	−9.13	30.59	−.003
Married	385.82**	50.27	.094**	85.17*	33.62	.033*
Children <16 years of age	471.26**	41.73	.122**	10.08	35.20	.004
Maternity/paternity leave	−384.35	207.46	−.017	117.58**	43.07	.035**
Hours worked	42.04**	4.45	.090**	46.73**	2.27	.253**
Organization tenure	22.58**	2.22	.121**	21.53**	1.90	.158**
Civil service job	−232.84*	93.08	−.028*	426.41**	61.94	.091**
Intrinsic job characteristics	1,090.42**	59.98	.204**	428.44**	47.89	.126**
Job training requirements	127.05**	21.01	.064**	100.92**	18.65	.071**
Perceived health	−290.07**	38.35	−.079**	−58.42*	29.73	−.025*
Self-esteem	78.88**	27.72	.028**	102.36**	22.58	.055**
Weight	31.47**	9.16	.244**	−21.29**	7.06	−.235**
Weight^2	−0.17**	0.05	−.236**	0.11**	0.05	.178*
R			.661**			.615**
R^2			.437**			.379**

Note. Though not reported, all regressions include eight industry indicator variables (farm/mining, construction, manufacturing, transportation/utilities, retail/service, health care, government/education, and banking/finance). B = unstandardized coefficient (in dollars); SE_B = standard error of B; β = standardized coefficient.
* $p < .05$. ** $p < .01$.

sents 15 kg below the average weight for each gender [for women, the average was 51 kg; for men, the average was 67 kg]). As the figure shows, there is a substantial average wage differential—such that regardless of weight, men earn more than women. However, the figure also shows that that the slope of weight is negative for women and positive for men. As the results in Table 2 demonstrate, there are curvilinearities in the results, such that the positive effect of weight on earnings for men becomes negative at above-average levels of weight (so that increasing levels of weight positively impact earnings at low levels of weight but negatively impact earnings at high levels of weight). For women, the curvilinearity is different—although increasing levels of weight continue to be negatively associated with earnings at high levels of weight, the slope is less steep (so that increasing levels of weight is more damaging to women's earnings at below-average weight than above-average weight).

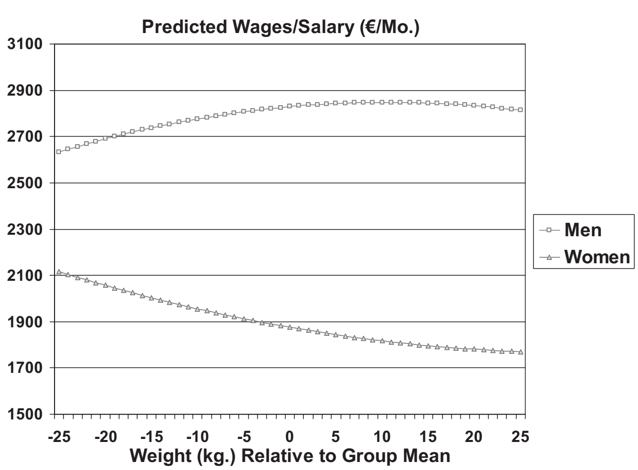

Figure 1. Effect of weight on earnings for men and women (Study 1).

Effect size estimates. Finally, to illustrate the practical nature of the results more fully, we estimated the predicted effects of weight on earnings for men and women across five weight ranges. These predicted values were obtained by taking the predicted monthly earnings produced by the regression results in Table 2 and by calculating predicted values for men and for women at ±1 and ±2 *SD*s. To render the results comparable across studies, we annualized the monthly earnings and translated the currency (from Euros to U.S. dollars) using current exchange rates. The results for Study 1 reveal that, for men, both heavier-than-average and thinner-than-average men earn less—men who are 2 *SD*s below average in weight earn \$4,057 less, and men who are 2 *SD*s above average earn \$146 less, than the average man. However, these statistics also show that the penalty for being very thin (−2 *SD*s) is nearly 28 times that for being very heavy (2 *SD*s). The results also show that very thin (−2 *SD*s in weight) women earn the most—\$3,981 more per year than their average weight counterparts. Conversely, whereas relatively heavy (2 *SD*s) women earn the least—\$1,848 per year less than the average weight women—a woman loses more than twice as much from going to very thin (−2 *SD*s) to average than going from average to relatively heavy (2 *SD*s).

Study 2

Overview

A limitation of Study 1 is that weight was measured only twice, making within-individual analyses impossible. Thus, in Study 1, generalizations are limited to those between individuals. A unique feature of Study 2 was that weight was measured 15 times over a 25-year time span, opening up investigation of within-individual changes in weight and earnings. Though, of course, no field study meets the criteria for strong causal inference, the within-individual design of Study 2 tests whether between-individual Gender × Weight effects extend to multiple levels of analysis (i.e., whether an individual difference variable—gender—predicts within-individual variation between weight and earnings/income). Whereas Study 1 was based on a large sample of German participants, Study 2 is based on American participants. As in Study 1, we first describe the control variables used, and then we describe the methodology and results of Study 2.

Control variables. For the most part, we controlled for the same or similar variables in Study 2 as in Study 1. Specifically, as in Study 1, we controlled for height, age, gender, marital status, children less than 16 years of age in the household, maternity leave, perceived health problems, smoking and drinking behavior, educational attainment, hours worked, self-esteem, tenure, job complexity, industry, and whether the individual worked in the civil service (public sector). There were two differences in the control variables between Study 1 and Study 2. First, in Study 2, we controlled for race rather than nationality. Second, in Study 2, although we did not control for job training, we were able to control for spouse's wages/salary and for childhood socioeconomic status, the latter of which may be related to individuals' weight, income, or both (Chang & Lauderdale, 2005; Conley & Glauber, 2007; Pollan, 2006).

Method

Participants and procedure. Participants in Study 2 were individuals enrolled in the National Longitudinal Surveys of Youth 1979 (NLSY79), a panel study funded by the U.S. Department of Labor and the Bureau of Labor Statistics. The initial participant sampling in the NLSY79 began when the National Opinion Research Center (NORC) at the University of Chicago created a list of U.S. households for the first interview. Interviewers then went to a random sample of these households and performed a brief interview designed to collect basic information on every household resident. Drawing a stratified random sample that oversampled minority group members and the economically disadvantaged, NORC researchers asked interviewed individuals to participate in the first interview. Of the 14,574 individuals designated for participation, a total of 12,686 individuals agreed, representing a response rate of 87%.

NLSY79 participants were interviewed annually through 1994 and were interviewed biannually thereafter. Although most (77.9%) NLSY79 data were collected via face-to-face interviews, where NORC interviewers interviewed participants in their homes, over time more interviews have been conducted over the phone. Since 2000, most (61.8%) of the interviews have been conducted over the phone. From 1979 to 1994, participants were paid $10 upon completion of the interview. In 1996, this stipend was increased to $20 and then to $40 beginning with the 2002 interview.

Given the 25-year time span of the NLSY79, attrition has been quite low. The response rate to each interview has averaged more than 90%. Of the original sample of 12,686 individuals, 7,661 (60.4%) remained in the study as of 2004. However, voluntary attrition rates were considerably lower—of the 5,025 original participants no longer in the study as of 2004, 2,722 of these individuals were dropped because of budget cutbacks in 1985 (n = 1,079) and 1991 (n = 1,643). Of the remaining 2,303 individuals who were in the study in 1979 but not in 2004, the reasons for dropping out were as follows: death (n = 399), inability to locate (n = 452), difficult cases (n = 116), refusal to participate further (n = 1,134), and other (n = 202). Thus, counting the last four categories as voluntary nonparticipation, the effective response rate from 1979 to 2004 is 85%. Participants reside in all 50 states, the District of Columbia, U.S. territories, and countries abroad. Participants ranged in age (as of 2008) from 43 to 51 years (M = 46.90). The sample was composed of 50.5% men. White individuals composed 68.9% of the sample. Of the sample, 42% had attended some college, 19% had earned a bachelor's degree (or equivalent), and 6% had earned a graduate degree (or equivalent).

Between-individual (Level 2) measures.

Gender. In the Time 1 interview, interviewers recorded the gender of the participant as 1 = *male*, 2 = *female*.

Height. Height was measured during the 1981, 1982, and 1985 surveys, when the interviewer asked the participant, "How tall are you?" Their responses were recorded in feet and inches. We averaged the responses over these three time periods (α = .92).

Age. Participants' age was measured with an item, collected in the Time 1 interview, that asked individuals to report their birth date, from which their age was verified. We then added 29 to this variable to make it reflect participants' current age as of 2008.

Race (White). In the 1979 interview, interviewers recorded participants' race as 1 = *White*, 2 = *Black*, 3 = *other*. We recoded this to 1 = *White*, 0 = *otherwise*.

Childhood socioeconomic status. Childhood socioeconomic status was measured with a composite of five variables reflecting the social and economic environment the child experienced: (a) father's years of education, (b) mother's years of education, (c) whether the individual's household was below U.S. government poverty guidelines at the initiation of the study, (d) father's occupational prestige (using Duncan's, 1961, index of occupational prestige), and (e) mother's occupational prestige (again using Duncan's, 1961, measure). These five items were standardized and then were averaged. The reliability of this five-item composite scale was α = .72.

Associate's, undergraduate, and graduate degrees. Participants' educational attainment was measured by participants' response to the question, asked during 2004: "What is the highest grade or year of regular school that you have completed and got credit for?" We then used that variable to form dummy variables representing whether the participant had obtained the equivalent of an associate's degree (1 = *yes*, 0 = *no*), an undergraduate degree (1 = *yes*, 0 = *no*), or a graduate degree (1 = *yes*, 0 = *no*).

Smoking and excessive drinking behavior. Smoking behavior was assessed by taking a composite for four variables (α = .95): (a) whether the individual reported current smoking behavior, averaged over the four time periods the question was asked; (b) whether the individual reported ever smoking, averaged over the four time periods the question was asked; (c) the quantity of cigarettes consumed, averaged over the five time periods the question was asked; and (d) whether the individual reported smoking at least 100 cigarettes, averaged across the three time periods the question was asked. Excessive drinking was measured (α = .77) by averaging a question asked in eight waves: "How often have you had six or more drinks on one occasion during the last 30 days?" Responses were coded using a 7-point scale (0 = *never*, 1 = *once*, 2 = *two or three times*, 3 = *four or five times*, 4 = *six or seven times*, 5 = *eight or nine times*, and 6 = *10 or more times*).

Married. Each survey year participants were asked to report their current marital status, which was coded by the interviewer as 1 = *never married*, 2 = *married*, and 3 = *other*. We recoded these responses to 1 = *married*, 0 = *otherwise*, and then we averaged the responses so that the variable reflects the proportion of time the participant was married (M = 0.42, SD = 0.27).

Children less than 16 years of age in household. Whether participants had children less than 16 years of age in the household was computed by, for each interview year, recording whether the participants' children averaged less than 16 years of age (1 = *yes*, 0 = *no*) and then averaging that variable over the years (α = .90) so that the variable reflected the proportion of time the participants' children averaged less than 16 years of age (M = 0.70, SD = 0.46).

Spouse's wages/salary. Each year, participants were asked: "How much did your husband/wife receive from wages, salary, commissions, or tips from all jobs, before deductions for taxes or anything else?" The responses to this question were averaged over the 15 years (α = .88). For those participants without a spouse, this variable was coded 0.

Maternity leave. Female participants (paternity leave was not assessed in this study) were asked, over the past year, "Are there any periods of a full week or more during which you took any (paid) leave from (work/your business) with (name of employer) because of a pregnancy or the birth of a child?" Participants' responses were coded 1 = *yes* and 0 = *no* and then were averaged over the 12 years the question was asked (asked annually from 1988 to 1993 and then biannually from 1994 to 2004).

Hours worked. Each year participants responded to the interviewer question "How many hours per week do/did you usually work?" at their current jobs. Participants' responses were then averaged over the interviewer years (α = .76) to form a variable representing the average hours worked per week (M = 36.5, SD = 9.7) over the course of the study.

Job tenure. Job tenure was assessed by taking the average amount of "total tenure (in weeks) with employer as of interview date," as reported during each of the interviews (α = .93; M = 135.1 weeks, SD = 118.2 weeks).

Public sector job. Each study year, participants were asked to report whether they worked for the government. For those who responded affirmatively, their responses were scored 1 = *federal*, 2 = *state*, and 3 = *local*. We recoded this variable to represent 1 = *government employee*, 0 = *otherwise* (M = 0.05, SD = 0.13), and then we averaged these responses over the time periods (α = .66).

Job complexity. For each year, job complexity was coded on the basis of participants' job titles, using the Dictionary of Occupational Titles (U.S. Department of Labor, 1991) classification. Specifically, participants' occupational codes were translated into the job's complexity in dealing with people, data, and things. For each occupation, we summed these three ratings so that scores ranged from 0 (jobs of low complexity, such as janitor) to 19 (jobs of high complexity, such as surgeon). These job complexity scores were then averaged over the 15 years (α = .91).

Perceived health (problems). Each year, participants were asked "Are you limited in the kind of work you do on a job for pay because of your health?" Responses were coded with a scale in which 1 = *yes*, 0 = *no*. The scale was formed by averaging the responses to this question over each time period (α = .79).

Self-esteem. Self-esteem was assessed with a 20-item scale that consisted of four questions from 1980 ("I take a positive attitude toward myself"), four questions from 1987 ("On the whole, I am satisfied with myself"), and 12 questions from 1992 ("I often feel helpless in dealing with the problems of life"). Individuals responded to the items using a 4-point scale ranging from 1 (*strongly agree*) to 4 (*strongly disagree*); responses were reverse-scored so that high scores reflected high self-esteem. Responses to the 20 items were then averaged (α = .80).

Within-individual (Level 1) measures.

Time. Because the average individual gains weight over time (Williams & Wood, 2006), we controlled for temporal effects by using time period as a Level 1 variable. Consistent with Bryk and Raudenbush (1987), we created the time variable representing the time of each relevant time period, expressed in terms of years since the onset of the study. Specifically, since weight was first measured in 1981 and the study began in 1979, for that time period, Time = 2. For the most recent time period, 2004, Time = 25 (i.e., 25 years after 1979).

Weight. Participants' weight was assessed in 15 surveys from 1981–2004 with their responses to the interviewer question "How much do you weigh?" Their responses (in pounds) were recorded by the interviewer (averaged across all time periods, M_{M} = 182.20, M_{F} = 147.12). As in Study 1, the quadratic term was computed by squaring the variable (Cohen et al., 2003). The reliability of this variable, in which each year was treated as a scale item, was α = .99 (α = .99 for both men and women).

Wages/salary. Each year, participants were asked "During ____, how much did you receive from wages, salary, commissions, or tips from all (other) jobs, before deductions for taxes or anything else?" Participants' responses were used as the measure of wages/salary received each year, matched to the years in which weight was reported. To separate real from inflationary wage growth over time, we converted wages/salary from the earlier time periods to 2008 dollars using the Bureau of Labor Statistics CPI Inflation Calculator (http://www.bls.gov/bls/inflation.htm). The reliability of this variable, in which each year was treated as a scale item, was α = .91 (α = .92 for men; α = .90 for women).

Analyses. Data were analyzed using hierarchical linear modeling (HLM, Version 6.0; Raudenbush, Bryk, Cheong, & Congdon, 2004). HLM is well suited to analyze multilevel data, permitting investigation of within-individual change (Bryk & Raudenbush, 1987) and multilevel interactions. In this case, this involved ascertaining the degree to which an individual difference or between-individual (or Level 2) variable, gender, predicted the relationship between two within-individual or Level 1 variables, weight and pay (wages/salary). The primary advantage of multilevel modeling in this study is the ability to study within-individual change in weight. As we explain shortly, all individuals enter the study's time frame with their own weight as a baseline, and we then examine, on a within-individual basis, how their weight changes over time (and how the pecuniary implications of these within-individual changes vary by gender).

Because weight and pay were measured 15 times for each individual over the course of the study, there are 15 observations per person. Though HLM involves simultaneous estimation (Raudenbush et al., 2004), for purposes of explanation, if one assumes that there are 1,000 individuals, multilevel modeling involves estimating a regression of earnings on weight for each of the 1,000 individuals, and then these 1,000 regression weights (one for each person) are themselves predicted by the between-individual variables (including gender). Moderation is observed when a between-individual variable predicts the regression weights corresponding to the within-individual relationships. For example, a moderating effect of gender is observed if gender (Level 2) predicts the (Level 1) weight–earnings regression weights.

In constructing the HLM Level 1 data set, we matched the other Level 1 variables to the weight variables. Specifically, because weight was assessed in 15/21 of the time periods from 1979 to 2004 (all possible years except 1979, 1980, 1983, 1984, 1987, and 1991, including all of the biannual time periods after 1993), we constructed the Level 1 data set by matching the other Level 1 variables to these time periods.

The within-individual independent variables (weight, weight^2, and time) were individual-mean-centered prior to analysis. What this means is that within-individual scores on these variables are expressed as deviations from each individual's mean. Such centering provides both conceptual and methodological advantages. Conceptually, because Study 2 focused on changes in weight and earnings over time (and the role that gender plays in these relationships), centering explicitly models weight and earnings as deviations from an individual's mean. This provides for more appropriate weight comparisons because average weight varies significantly between men and women (i.e., changes in weight are relative to the group mean for men and women). Methodologically, centering partials out all between-individual (individual difference) variables that affect the predictor (in this case, weight; Hofmann, Griffin, & Gavin, 2000). Thus, no individual difference variables (such as childhood exercise or diet) that affect an individual's starting weight or average weight can affect the weight results (unless these variables are related to changes in weight beyond the initial or average baseline).

Results

Table 3 contains the descriptive statistics and correlations of the study variables (though we caution that, because the correlation matrix is based on a between-individual design, it is not truly informative as to within-individual [Level 1] or multilevel [cross-level] relationships). HLM results predicting wages are provided in Table 4. As is shown in the table, most of the Level 2 control variables significantly predicted wages in the expected direction. Specifically, height, race, childhood socioeconomic status, education (all three degrees), children less than 16 years of age in the household, spouse's wages/salary, hours worked, job tenure, job complexity, and self-esteem significantly and positively predicted the intercept—meaning that tall individuals; those who were White and who were raised in high status households; those who had earned associate's, undergraduate, and graduate degrees; those with children less than 16 years of age in the household; those whose spouses earned more; those who worked more hours per week; those with higher levels of job tenure; those who worked in complex jobs; and those with higher self-esteem earned more on average. Gender, smoking, being married, taking maternity leave, having a public sector job, and having health problems negatively predicted the intercept—meaning that women, those who smoke, those who are married and took maternity leave, and those working in government jobs earned less on average.[4]

As shown by the intercept coefficient on Time (B_{10} = 2,594.82, $p < .01$), time significantly predicted wages, meaning that, in real terms, individuals' pay increased over time. As predicted, gender moderated the relationship between weight and wages (i.e., gender predicted the Level 1 relationship between weight and wages). The coefficient was negative ($B_{21} = -819.97$, $p < .01$), indicating that the weight–earnings association was more positive for men (coded as 1) than women (coded as 2). Indeed, as shown in Figure 2, for men, changes in weight are associated with increases in earnings, whereas for women, changes in weight are associated with decreases in earnings.[5]

Table 4 also shows that the slope of the quadratic term was significant and negative ($B_{30} = -1.92$, $p < .01$), indicating diminishing returns as weight increases. Also, as shown in the table, gender positively and significantly predicted the quadratic term ($B_{31} = 1.65$, $p < .01$), meaning that the diminishing returns were stronger for women than for men. Figure 2 also clarifies how the curvilinear relationship between weight and earnings varies by gender. For men, there are slightly diminishing returns to increasing weight. For women, there are diminishing penalties to gaining weight that are stronger, in absolute magnitude, to the diminishing returns (curvilinearities) for men. Overall, these results confirm

[4] As noted by Bliese and Ployhart (2002), "It is very unlikely that longitudinal data will be independent; consequently, one typically assumes that data collected from individuals over time will display significant nonindependence" (p. 379). The intraclass correlation coefficient (ICC) can be used to estimate the degree of nonindependence. When estimating a null model in Study 2, the corresponding $ICC = .283$ suggests a moderate degree of interdependency of pay over time.

[5] As with Study 1, pay was not normally distributed in Study 2. Accordingly, we explored whether nonlinear transformations to the pay variable changed the functional form of the relationship. In no case did any of the three transformations (natural log, square root, ordinalizing into 10 categories) change the direction or significance of the gender effects in Table 4. In each case, gender negatively predicted the linear weight–earnings coefficient (the weight–earnings linear relationship is more positive for men than women) and positively predicted the weight^2–earnings coefficient (the weight–earnings quadratic relationship is stronger for women than men). Moreover, a visual inspect of the graphs revealed the same functional form: For men, weight was positively related to earnings, with diminishing returns to higher levels of weight; for women, weight was negatively related to earnings, with diminishing penalties at higher levels of weight. Though, like Study 1, the results are not significantly different using these transformations, for the same reasons as Study 1, we utilize the untransformed measures in our primary analyses.

Table 3
Means, Standard Deviations, and Intercorrelations Among Variables (Study 2)

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1. Gender (1 = *male*, 2 = *female*)	1.49	0.50	1.00																						
2. Height	61.85	3.95	−.62	1.00																					
3. Age	46.83	2.30	.01	−.01	1.00																				
4. White (1 = *White*, 0 = *other*)	0.69	0.46	.01	.04	.04	1.00																			
5. Childhood SES	0.00	0.69	−.02	.12	.07	.25	1.00																		
6. Associate's degree	0.43	0.50	.08	.03	.04	.03	.36	1.00																	
7. College degree	0.20	0.40	.04	.07	.02	.09	.36	.58	1.00																
8. Graduate degree	0.06	0.24	.01	.03	.01	.06	.23	.30	.51	1.00															
9. Smoking	0.03	0.93	−.03	.05	−.04	−.05	−.03	−.16	−.17	−.11	1.00														
10. Excessive drinking	0.86	1.00	−.34	.25	−.07	.09	.01	−.14	−.10	−.07	.27	1.00													
11. Married	0.42	0.27	.10	−.03	.23	.28	.07	.01	.01	.00	−.10	−.16	1.00												
12. Children <16 years of age in HH	0.72	0.45	.16	−.06	−.03	−.02	−.08	−.05	−.04	−.03	.10	−.06	.39	1.00											
13. Spouse's wages/salary	6,127.40	9,809.40	.28	−.10	.05	.17	.22	.19	.20	.14	.00	−.13	.47	.26	1.00										
14. Maternity leave	0.14	0.47	.30	−.17	−.07	.02	.06	.13	.11	.06	−.03	−.11	.09	.18	.24	1.00									
15. Hours worked	36.45	9.53	−.28	.20	.12	−.03	−.04	−.10	−.09	−.03	.12	.14	.04	.07	−.10	−.07	1.00								
16. Job tenure	136.55	117.72	−.05	.09	.11	.02	.07	.08	.06	.02	.00	−.01	.17	.11	.16	.19	.19	1.00							
17. Public sector job	0.09	0.29	.08	−.04	−.02	−.03	.05	.16	.17	.17	−.03	−.06	.06	.08	.13	.08	−.01	.15	1.00						
18. Job complexity	7.03	2.62	.12	.01	.09	.13	.31	.38	.37	.29	−.10	−.10	.12	.02	.22	.14	.01	.16	.12	1.00					
19. Health problems	0.04	0.08	.08	−.06	.02	−.05	−.06	−.06	−.07	−.04	.14	−.03	−.06	.03	−.02	−.03	−.04	−.07	−.02	−.09	1.00				
20. Self-esteem	3.32	0.36	−.03	.09	.05	.07	.27	.28	.25	.15	−.06	−.04	.13	.01	.18	.07	.04	.14	.09	.27	−.13	1.00			
21. Weight†	162.12	34.56	−.48	.55	.02	−.09	−.02	−.03	−.04	−.03	.04	.20	−.06	−.02	−.15	−.15	.22	.13	.01	−.07	.02	.02	1.00		
22. Weight^{2}†	27,811.38	13,008.70	−.41	.49	.02	−.07	−.02	−.02	−.03	−.02	.04	.17	−.06	−.02	−.13	−.13	.20	.12	.01	−.07	.03	.02	.97	1.00	
23. Earnings†	24,693.93	35,693.90	−.19	.20	.04	.09	.20	.20	.23	.18	−.04	.06	.07	.06	.12	.04	.20	.27	.03	.23	−.10	.19	.13	.12	1.00

Note. For $|r| \geq .03$, $p < .01$. SES = socioeconomic status; HH = household.
† Because these variables were Level 1 (within-individual variables), for purposes of the correlation matrix, they were averaged over time and, thus, do not properly represent multilevel relationships.

Table 4
Multilevel Regression Equation Predicting Total Earnings (Study 2)

Variable	B	SE_B	t
Intercept, B_0			
Intercept, B_{00}	−77,813.79	17,384.34	−4.476**
Gender (1 = *male*, 2 = *female*), B_{01}	−12,330.16	1,556.94	−7.919**
Height, B_{02}	436.45	157.82	2.766**
Age, B_{03}	353.79	192.00	1.843
White (1 = *White*, 0 = *other*), B_{04}	3,202.28	726.36	4.409**
Childhood socioeconomic status, B_{05}	2,324.97	565.69	4.110**
Associate degree, B_{06}	4,738.30	923.80	5.129**
College degree, B_{07}	6,067.59	1,449.82	4.185**
Graduate degree, B_{08}	12,914.25	3,325.61	3.883**
Smoking, B_{09}	−1,853.23	354.52	−5.227**
Excessive drinking, B_{010}	452.04	495.55	0.912
Married, B_{011}	−6,642.31	3,246.37	−2.046*
Children <16 years of age in household, B_{012}	6,232.83	1,031.24	6.044**
Spouse's wages/salary, B_{013}	0.33	0.16	2.032*
Maternity leave, B_{014}	−1,325.68	568.16	−2.333**
Hours worked, B_{015}	628.97	61.77	10.182**
Job tenure, B_{016}	34.34	3.63	9.468**
Public sector job, B_{017}	−6,908.87	1,260.19	−5.482**
Job complexity, B_{018}	1,941.92	251.14	7.732**
Health problems, B_{019}	−11,104.19	2,986.59	−3.718**
Self-esteem, B_{020}	7,440.60	1,102.85	6.747**
Time, B_1			
Intercept, B_{10}	2,594.82	67.45	38.473**
Weight, B_2			
Intercept, B_{20}	1,051.57	129.76	8.104**
Gender (1 = *male*, 2 = *female*), B_{21}	−819.97	72.43	−11.320**
Weight^2, B_3			
Intercept, B_{30}	−1.92	0.25	−7.810**
Gender (1 = *male*, 2 = *female*), B_{31}	1.65	0.16	10.526**

Note. Though not reported, hierarchical linear modeling (HLM) includes eight industry indicator variables (farm/mining, construction, manufacturing, transportation/utilities, professional–technical, retail/service, government/education, and banking/finance). B = HLM coefficient; SE_B = standard error of B.
* $p < .05$. ** $p < .01$.

Hypothesis 1 and Hypothesis 2 in that, for men, increases in weight have positive linear effects of pay but at diminished returns at above-average levels of weight. For women, increases in weight have negative linear effects on pay, but the negative effects are stronger at below-average than at above-average weight levels.[6]

Prospective analyses. To determine whether the weight effects persisted over time, and to bolster the causal logic underlying the hypothesized relationships among the within-individual variables, we estimated the same model as that displayed in Table 4, except that we forward lagged the earnings variable (so that weight and weight^2 were measured at Time t, and earnings were measured at Time $t + 1$ [one year after the weight variable was measured]). The results from this estimation—for parsimony, only the pertinent Level 1 slope results are shown—are shown in Table 5. As the table shows, these prospective results are similar to those provided in Table 4. First, gender negatively predicts the weight–earnings linear relationship (B_{21} = –293.41, $p < .01$), meaning that weight is more positively related to earnings for men than for women. Second, gender positively predicts the weight–earnings quadratic relationship (B_{31} = .25, $p < .01$), meaning that the curvilinear relationship is stronger for women than for men (i.e., the penalties to increased weight diminish for women more strongly than the rewards for increased weight diminish for men). Though these results utilize a subset of the available data (i.e., the Time 15 earnings variable is necessarily lost in the data set), the fact that the results are similar increases confidence in the support for the hypotheses.

A reviewer on a previous version of the article argued that such an analysis, although helpful, only considers half of the causal logic. Namely, weight might be affected by prior earnings. To address this concern, we undertook an instrumental variable analysis (Heckman, 1997, 2008; Kennedy, 2003), wherein the weight variables were specified as endogenous to prior income. This deals explicitly with the possibility that weight may be affected by earlier income. In conducting this instrumental variables analysis, we used two-stage least squares regression (Davidson & MacKinnon, 1993; Hsiao,

[6] The incremental variance explained values—calculated as recommended by Bliese and Ployhart (2002)—were as follows: weight slope intercept (B_{20}), ΔR^2 = 2.69%; weight^2 slope intercept (B_{30}), ΔR^2 = 2.56%; gender as predictor of weight–earnings slope (B_{21}), ΔR^2 = 4.14%; gender as predictor of weight^2–earnings slope (B_{31}), ΔR^2 = 3.78%. (For consistency, we use the notation ΔR^2 here; the experienced HLM user will know that incremental variance estimates are derived from reductions in error variance by adding the variable of interest [see Hofmann, 1997].)

1997). Specifically, as before, weight and weight^2 at Time t are used to predict earnings at Time $t + 1$, but weight and weight^2 at Time t also are considered endogenous to earnings at Time $t - 1$. The results of this two-stage least squares analysis revealed that accounting for the endogeneity of weight to prior earnings had little effect on the relationship of weight or weight^2 to prospective earnings or on the degree to which gender predicted these Level 1 relationships. Specifically, the intercepts for weight and weight^2 changed from $B_{20} = 419.891$ ($p < .01$) and $B_{30} = -0.403$ ($p < .01$)—as in Table 5—to $B_{20} = 419.276$ ($p < .01$) and $B_{30} = -0.401$ ($p < .01$) when the instrumental variables were controlled. The gender coefficients for weight and weight^2 changed from $B_{21} = -303.100$ ($p < .01$) and $B_{31} = 0.277$ ($p < .01$)—as in Table 5—to $B_{21} = -302.834$ ($p < .01$) and $B_{31} = 0.276$ ($p < .01$) when the instrumental variables were controlled. The two-stage least squares regression analysis did suggest that prior earnings, as one of the instrumental variables, had a significant effect on weight^2. Specifically, the prior earnings coefficients on weight and weight^2 were $B = 0.016$ ($t = 1.72$, *ns*) and $B = -0.026$ ($t = -2.28$, $p < .05$). However, it appears that this endogeneity of weight (i.e., that weight—in its quadratic form—is affected, to some degree, by income) has little effect on the substantive results, even for the analyses in which income is measured prospectively (i.e., a year after weight).

Effect size estimates. The right-hand portion of Table 6 provides the effect size estimates for men and women for Study 2. As the table shows, a decrease in weight of 2 *SD*s results in a predicted earnings decrement of $17,535 for men and a predicted earnings increment of $22,283 for women. Conversely, an increase in weight of 2 *SD*s results in a predicted earnings increment of $14,889 for men and a predicted earnings decrement of $18,902 for women. Thus, as in Study 1, weight variability is more important to predicted earnings at the low end of the weight distribution than at the high end of the distribution. Though the linear effects are stronger in Study 2 than in Study 1, the nonlinear effects are somewhat weaker (though still significant). To some degree, those differences may reflect cultural (United States vs. Germany) differences or levels of analysis (between-individual vs. within-individual) differences.[7]

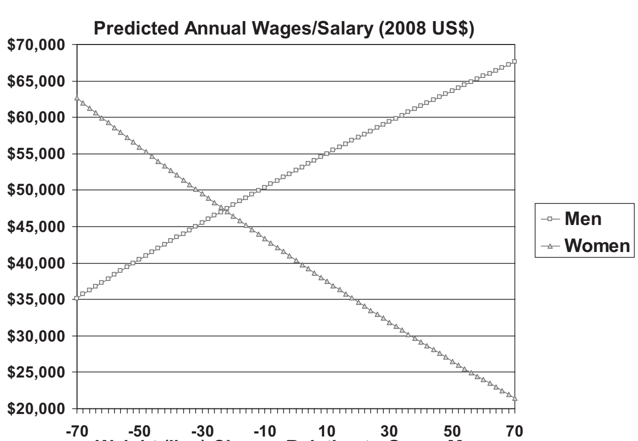

Figure 2. Multilevel effect of gender on the weight–earnings relationship (Study 2).

Table 5
Multilevel Regression Using Prospective Total Earnings (Study 2)

Variable	B	SE_B	t
Time, B_1			
Intercept, B_{10}	2,735.52	50.85	53.795**
Weight, B_2			
Intercept, B_{20}	419.89	56.48	7.435**
Gender (1 = *male*, 2 = *female*), B_{21}	−303.10	31.65	−9.576**
Weight^2, B_3			
Intercept, B_{30}	−0.40	0.07	−5.438**
Gender (1 = *male*, 2 = *female*), B_{31}	0.28	0.05	5.939**

Note. Though the control variables in this regression are the same as the regression in Table 4, for parsimony, only the Level 1 slope results are presented. Weight and Weight^2 are measured at Time t. Total earnings is measured at Time $t + 1$. B = hierarchical linear modeling coefficient; SE_B = standard error of B.
** $p < .01$.

General Discussion

Because of the society in which we live, physical appearance plays an important role in workplace interactions and outcomes (Möbius & Rosenblat, 2006). Body weight is one of the most obvious physical characteristics, and obesity has become a major social and health issue. Past research shows consistent evidence of discrimination against obese employees in every stage of the employment cycle, including career counseling, selection, placement, compensation, promotion, discipline, training, and discharge (Roehling, 1999). Some firms are even charging obese employees higher premiums for their insurance and creating incentive programs for obese employees who lose weight (Wysocki, 2004).

Although the management literature's past focus on obesity has been instructive, it has assumed a linear relationship between body weight and employment outcomes, and it has not theorized why women might experience different conditions than men on the basis of their weight. The goal of the present article was to develop theory for why the relationship between weight and employment outcomes should be curvilinear, and why the specific form of the curvilinear relationship should be different for men and women. We then tested our theorizing using pay level, which is an important employment outcome because it reflects the value placed on employees' full set of human and social capital.

Perhaps the most startling finding of this investigation is that men and women experience opposite incentives regarding weight in the very thin to average weight range. Whereas women are punished for any weight gain, very thin women receive the most severe punishment for their first few pounds of weight gain. This

[7] Although the patterns of results in Study 1 and Study 2 are quite consistent, the effect size magnitudes are not. As shown in Table 6, the thin penalties for men, and the thin rewards to women (relative to their heavier counterparts), are much stronger in Study 2 than in Study 1. Given the many differences across the studies (methodology, time span, levels of analysis, measurement, and culture), isolating a single explanation for the differences is difficult. However, one clear explanation is the stronger curvilinearities in Study 1. If only linear terms were used in both studies, the differences in predicted values across genders between the studies would be much smaller.

Table 6
Effect Size Estimates of Weight on Earnings for Men and Women

Weight deviation (from mean)	Predicted change in salary (relative to mean)			
	Study 1		Study 2	
	Men	Women	Men	Women
−2 *SD*s	−$4,056.67	$3,980.88	−$17,535.00	$22,283.45
−1 *SD*	−$1,503.00	$1,650.81	−$8,436.75	$10,719.10
0 *SD*s (mean)	0	0	0	0
1 *SD*	$452.32	−$1,276.64	$7,775.25	−$9,873.85
2 *SD*s	−$146.04	−$1,847.49	$14,889.00	−$18,902.45

Note. Table entries are deviations from average predicted (net of influence of all other independent variables) salary, annualized and in U.S. dollars for both studies. Study 1 weight deviations are between-individual; Study 2 weight deviations are within-individual.

finding is consistent with research showing that the media's consistent depiction of an unrealistically thin female ideal leads people to see this ideal as normative, expected, and central to female attractiveness (Brown, 2002). Indeed, both our German and American results show that once women reach an average weight, subsequent weight gains are actually penalized to a lesser extent, presumably because the social preferences for a feminine body have already been violated. For American women, gaining 25 lbs produces an average predicted decrease in salary of approximately –$15,572 at below average weight and –$13,847 at above average weight; the per pound penalty at 25 lbs below the group mean is 12% harsher than at 25 lbs above the group mean. This means that, all else equal, a woman who is average weight earns $389,300 less across a 25-year career than a woman who is 25 lbs below average weight. Thus, our results suggest that both German and American societies reward women who conform to the improbably thin female standard perpetuated by the media and mete out the stiffest punishments for the initial "rebellion" from this standard.

Very thin men, conversely, are punished relative to their average weight peers, and men are rewarded for gaining weight until the point of obesity. For American men, gaining 25 lbs produces a predicted increase in wages of roughly $8,437 per year at below average weights and a predicted increase of approximately $7,775 per year at above average weights. Thus, for men, there are only slightly diminishing returns to increasing weight. All else equal, a man who is 25 lbs below average weight is predicted to earn $210,925 less across a 25-year career than a man who is of average weight. Thus, notwithstanding any health benefits from being thin, both German and American societies appear to value heavier men, and men who fail to resemble gender-role norms experience the consequences of rebelling against these norms.

Theoretical Implications

Past research has relied primarily on stereotype theory to account for the effects of weight on social and employment outcomes (Crandall, 1994; Finkelstein et al., 2007; Larkin & Pines, 1979; Pingitore et al., 1994; Roehling, 1999; Wade & DiMaria, 2003). However, stereotype theory is mute about the very thin to average range of weight, and it also does not differentiate between men and women. For example, stereotype theory predicts that an average weight man or woman who gains 30 lbs would be labeled with the negative stereotypes attached to obese individuals and subsequently penalized, whereas a very thin man or woman who gains 30 lbs (and becomes average weight) should not be labeled with negative stereotypes or penalized.

In the present study, we used the logic of cultivation theory to extend stereotyping theory. By considering how the social expectations propagated by the media lead society to accept body size portrayals as ideal representations of reality (Gerbner et al., 2002), our logic anticipates both curvilinearity and differences between men and women. Specifically, our logic suggests an inflection point occurs when men and women "rebel" against the media's pervasive but unrealistic weight standards. As such, our theory has direct implications for past weight research using stereotyping theory and making assumptions about linearity. It will be interesting to see whether our logic and predictions hold for other employment outcomes besides pay (e.g., hiring, promotions, career/job attitudes). Moreover, given that weight discrimination also has been documented in education and health care (Puhl & Brownell, 2001), it will be interesting for future research to examine whether the effects are curvilinear and different for men and women outside the employment setting.

It also would be worthwhile for future research to consider the genesis of society's differential and curvilinear standards regarding men's and women's weight. For example, although not directly testable, sexual strategies theory (Buss & Schmitt, 1993) addresses basic human preferences for certain body types by considering how such preferences become intertwined in the narrative of human evolution. According to the theory, men and women evolved distinct mate selection strategies based on what they needed from mates to maximize reproduction. Human evolution should have produced preferences for men who could provide scarce resources such as food and protect their mates from aggressors, such that larger men offered credible signals about access to food and dominance against adversaries. Human evolution likewise should have produced preferences for women that provided cues about reproductive value and the two most obvious cues—according to Buss and Schmitt (1993)—are youthful appearance and health. Thus, to the extent that certain physical traits are in high social demand, these traits' historical relationship with mate selection potentially could be carried over into contemporary norms and social decisions, including employment decisions.

Practical Implications

Only by making weight–income trends public can employers become aware of how much people's weight affects employment decisions. Although it is possible that truly obese employees may create additional employment costs (e.g., medical costs, workspace accommodations), weight may have little relationship with true performance in most jobs. In particular, it would seem quite unlikely that weight affects job performance in the very thin to average weight range. As such, it is troubling that average weight women and thin men are penalized in the employment context, whereas very thin women and men of average or even above average weights are rewarded. Thus, it may be possible and competitively advantageous for employers to try and recognize—and then reduce—the role that weight plays in their employment decisions.

It is interesting to think about pinpointing employers' "idealized" weights on the basis of our results, although it is important to note that our models are based on the nonextreme weight distributions observed in our samples. In Study 1, the "ideal weight" (i.e., local income maximum) for men was 207 lbs, meaning that men's income reached a maximum at this weight, with income declining for both higher weights and lower weights. However, for Study 2, the local income maximum was an unlikely 540 lbs for men, suggesting that our model cannot and should not be applied to either extremely low (emaciated) or extremely high (morbidly obese) weight levels.

Limitations and Strengths

Like all investigations, ours has limitations that should be noted. Perhaps most important, our results do not prove that employer discrimination is the theoretical mechanism linking weight and income. Although we examined within-person effects across time and included a number of important control variables to help rule out alternative explanations (e.g., health issues), experimental designs are needed to confirm discrimination effects (Hebl & Turchin, 2005; Pingitore et al., 1994). For example, it is possible that employee performance is the causal mechanism linking weight and income, although at first brush it is difficult to understand why women's performance would decrease most as they moved from being very thin to average weight, whereas men's performance would increase most with these same weight gains. However, perhaps the weight–income trends that we observed are due to performance in the sense that employees are more able to influence others and get things accomplished when they conform to the media's ideal body form. In this sense, employees who conform to societal body expectations may perform better, and employers may simply be rewarding good performance in a nondiscriminatory manner (i.e., disparate impact, not disparate treatment).

Next, given the media norms surrounding men's bodies, it would have been useful to have examined not only body weight in our study but also muscularity. Another limitation is that our German sample only reported their weights twice, prohibiting within-person analyses. Thus, the German results are open to criticisms of reverse causation (e.g., more successful people become heavier because of richer meals) or omitted variable bias (e.g., personality traits serve as a common cause of weight and income), though, by themselves, these omitted variables would not explain the differential associations we observed by gender. Although we addressed causal limitations in Study 1 with the longitudinal, multilevel analysis in Study 2, it would be useful to conduct this type of careful analysis in countries outside the United States. It also should be noted that we controlled for perceived health problems, and it would be useful for future research to examine objective measures of health (such as blood pressure and diabetes).

These drawbacks are offset by several important assets of the studies. Perhaps most important, we developed theory regarding weight, which is a visible human characteristic with considerable social ramifications, and income, which is an index of social value. Specifically, we introduced a logic for curvilinearity in the weight–income relationship, and we developed theory for why the form of this curvilinearity should be different for men versus women. Thus, our investigation makes an important conceptual contribution to past weight theory and research, which has assumed a linear relationship and has not focused on the thin side of the body weight range, despite the social emphasis on women being thin. Next, our two-sample investigation improves our knowledge about how the hypotheses hold up in other cultures, whereas our within-person analysis allowed us to focus on changes in weight across time. In addition to revealing how weight change affects income differently for thin versus obese individuals, this analysis helps rule out alternative explanations of the results and increases our confidence that individual traits were not serving as a common cause of weight and career success.

Conclusion

Although weight has received some attention by researchers, we provide theory and evidence that the "weight double standard" may be more complex than past research has suggested. Specifically, the double standard depends not only on gender but where on their respective weight distributions men and women find themselves. Across two large studies, using two different methodologies, the present study suggests that the smallest income gap between genders occurs at thin weights (where men are penalized and women are rewarded) and that the slopes for both genders are steepest at the thin end of the weight distribution (more steeply negative for women and more steeply positive for men). Although to date most conceptual and empirical weight research across disciplines has focused on obesity, our results suggest that to understand the essence of the weight double standard, the thin end of the weight distribution may prove at least as incisive.

References

Achen, C. H. (1982). *Interpreting and using regression.* Beverly Hills, CA: Sage.

Achen, C. H. (1990). What does "explained variance" explain? Reply. *Political Analysis, 2,* 173–184.

Akpinar, E., Bashan, I., Bozdemir, N., & Saatci, E. (2007). Which is the best anthropometric technique to identify obesity: Body mass index, waist circumference or waist–hip ratio? *Collegium Antropologicum, 31,* 387–393.

Ballou, D., & Podgursky, M. (2002). Returns to seniority among public school teachers. *Journal of Human Resources, 37,* 892–912.

Barnum, P., Liden, R. C., & Ditomaso, N. (1995). Double jeopardy for women and minorities: Pay differences with age. *Academy of Management Journal, 38,* 863–880.

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

Becker, G. S. (1993). *Human capital: A theoretical and empirical analysis, with special reference to education* (3rd ed.). Chicago, IL: University of Chicago Press.

Bliese, P. D., & Ployhart, R. E. (2002). Growth modeling using random coefficient models: Model building, testing, and illustrations. *Organizational Research Methods, 5,* 362–387.

Briefel, R. R., & Johnson, C. L. (2004). Secular trends in dietary intake in the United States. *Annual Review of Nutrition, 24,* 401–431.

Brown, J. D. (2002). Mass media influences on sexuality. *Journal of Sex Research, 39,* 42–45.

Bryk, A. S., & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin, 101,* 147–158.

Busemeyer, J. R., & Jones, L. E. (1983). Analysis of multiplicative combination rules when the causal variables are measured with error. *Psychological Bulletin, 93,* 549–562.

Buss, D. M., & Schmitt, D. P. (1993). Sexual strategies theory: An evolutionary perspective on human mating. *Psychological Review, 100,* 204–232.

Cable News Network (CNN). (2008, September 11). *Models still too skinny?* Retrieved from http://edition.cnn.com/2006/WORLD/europe/09/13/spain.models/index.html

Cameron, C. A., & Windmeijer, F. A. G. (1997). An *R*-squared measure of goodness of fit for some common nonlinear regression models. *Journal of Econometrics, 77,* 329–342.

Cash, T. F., Winstead, B. A., & Janda, L. H. (1986). The great American shape-up. *Psychology Today, 20,* 30–37.

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

Chang, V. W., & Lauderdale, D. S. (2005). Income disparities in body mass index and obesity in the United States, 1971–2002. *Archives of Internal Medicine, 165,* 2122–2128.

Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica, 28,* 591–605.

Cohane, G. H., & Pope, H. G., Jr. (2001). Body image in boys: A review of the literature. *International Journal of Eating Disorders, 29,* 373–379.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Hillsdale, NJ: Erlbaum.

Conley, D., & Glauber, R. (2007). Gender, body mass, and socioeconomic status: Evidence from the PSID. *Advances in Health Economics and Health Services Research, 17,* 253–254.

Constant, A., & Massey, D. S. (2005). Labor market segmentation and the earnings of German guestworkers. *Population Research and Policy Review, 24,* 489–512.

Crandall, C. S. (1994). Prejudice against fat people: Ideology and self-interest. *Journal of Personality and Social Psychology, 66,* 882–894.

Critser, G. (2003). *Fat land: How Americans became the fattest people in the world.* Boston, MA: Houghton Mifflin.

Davidson, R., & MacKinnon, J. G. (1993). *Estimation and inference in econometrics.* New York, NY: Oxford University Press.

Dougherty, C. (2006). *Introduction to econometrics.* Oxford, England: Oxford University Press.

Duffey, K. J., Gordon-Larsen, P., Jacobs, D. R., Williams, O. D., & Popkin, B. M. (2007). Differential associations of fast food and restaurant food consumption with 3-y change in body mass index. *American Journal of Clinical Nutrition, 85,* 201–208.

Duncan, O. D. (1961). A socioeconomic index for all occupations. In A. J. Reiss Jr., O. D. Duncan, P. K. Hatt, & C. C. North (Eds.), *Occupations and social status* (pp. 109–138). New York, NY: Free Press.

Edwards, J. R. (1994). The study of congruence in organizational behavior research: Critique and a proposed alternative. *Organizational Behavior and Human Decision Processes, 58,* 51–100.

Feldman, D. C. (2002). Managers' propensity to work longer hours: A multilevel analysis. *Human Resource Management Review, 12,* 339–357.

Ferris, D. L., Brown, D. J., Lian, H., & Keeping, L. M. (2009). When does self-esteem relate to deviant behavior? The role of contingencies of self-worth. *Journal of Applied Psychology, 94,* 1345–1353.

Finkelstein, L. M., DeMuth, R. L. F., & Sweeney, D. L. (2007). Bias against overweight job applicants: More explorations of when and why. *Human Resource Management, 46,* 203–222.

Flegal, K. M., Carroll, M. D., Ogden, C. L., & Johnson, C. L. (2002). Prevalence and trends in obesity among U.S. adults, 1999–2000. *Journal of the American Medical Association, 288,* 1723–1727.

Fouts, G., & Burggraf, K. (1999). Television situation comedies: Female body images and verbal reinforcement. *Sex Roles, 40,* 473–481.

Gerbner, G., Gross, L., & Morgan, M. (2002). Growing up with television: Cultivation processes. In J. Bryant & D. Zillman (Eds.), *Media effects: Advances in theory and research* (2nd ed., pp. 43–67). Mahwah, NJ: Erlbaum.

Gini, C. (1921). Measurement of inequality of incomes. *Economic Journal, 31,* 124–126.

Grabe, S. H., Shibley, J., & Ward, L. M. (2008). The role of the media in body image concerns among women: A meta-analysis of experimental and correlational studies. *Psychological Bulletin, 134,* 460–476.

Greenberg, B. S., Eastin, M., Hofschire, L., Lachlan, K., & Brownell, K. (2003). Portrayals of overweight and obese individuals on commercial television. *American Journal of Public Health, 93,* 1342–1348.

Griffin, A. W. (2007). Women and weight-based employment discrimination. *Cardozo Journal of Law and Gender, 13,* 631–662.

Gullikson, A. (2006). *Interpreting OLS regression and transformations.* Retrieved from http://www.uoregon.edu/~aarong/teaching/V3212_Outline/node8.html

Hagquist, C., & Stenbeck, M. (1998). Goodness of fit in regression analysis—R^2 and G^2 reconsidered. *Quality & Quantity, 32,* 229–245.

Hargreaves, D. A., & Tiggemann, M. (2004). Idealized media images and adolescent body image: Comparing boys and girls. *Body Image, 1,* 351–361.

Hebl, M. R., & Turchin, J. M. (2005). The stigma of obesity: What about men? *Basic and Applied Social Psychology, 27,* 267–275.

Heckman, J. (1997). Instrumental variables: A study of implicit behavioral assumptions used in making program evaluations. *Journal of Human Resources, 32,* 441–462.

Heckman, J. (2008). Econometric causality. *International Statistical Review, 76,* 1–27.

Hirsch, B. T., & Leppel, K. (1982). Sex discrimination in faculty salaries: Evidence from a historically women university. *American Economic Review, 72,* 829–835.

Hofmann, D. A. (1997). An overview of the logic and rationale of hierarchical linear models. *Journal of Management, 23,* 723–744.

Hofmann, D. A., Griffin, M. A., & Gavin, M. B. (2000). The application of hierarchical linear modeling to organizational research. In K. Klein & S. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations* (pp. 467–511). San Francisco, CA: Jossey-Bass.

Hsiao, C. (1997). Statistical properties of the two-stage least squares estimator under cointegration. *Review of Economic Studies, 64,* 385–398.

Judge, T. A., & Cable, D. M. (2004). The effect of physical height on workplace success and income. *Journal of Applied Psychology, 89,* 428–441.

Judge, T. A., Cable, D. M., Boudreau, J. W., & Bretz, R. D. (1995). An empirical investigation of the predictors of executive career success. *Personnel Psychology, 48,* 485–519.

Kennedy, P. (2003). *A guide to econometrics* (5th ed.). Cambridge, MA: MIT Press.

Kirkpatrick, S. W., & Sanders, D. M. (1978). Body image stereotypes: A developmental comparison. *Journal of Genetic Psychology, 132,* 87–95.

Kozlowski, S. W. J., & Klein, K. J. (2000). A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions* (pp. 3–90). San Francisco, CA: Jossey-Bass.

Larkin, J. C., & Pines, H. A. (1979). No fat persons need apply: Experimental studies of the overweight stereotype and hiring preference. *Sociology of Work and Occupations, 6,* 312–327.

Lewis, G. B. (1996). Gender integration of occupations in the federal civil service: Extent and effects on male–female earnings. *Industrial and Labor Relations Review, 49,* 472–483.

Lewis-Beck, M. S., & Skalaban, A. (1990). When to use *R*-squared. *The Political Methodologist, 3,* 9–11.

Linné, Y., Barkeling, B., & Rössner, S. (2002). Long-term weight development after pregnancy. *Obesity Reviews, 3,* 75–83.

Mascie-Taylor, C. G. N., & Goto, R. (2007). Human variation and body mass index: A review of the universality of BMI cut-offs, gender and urban-rural differences, and secular changes. *Journal of Physiological Anthropology, 26,* 109–112.

Mishkind, M. E., Rodin, J., Silberstein, L. R., & Striegel-Moore, R. H. (1986). The embodiment of masculinity: Cultural, psychological, and behavioural dimensions. *American Behavioral Scientist, 29,* 545–562.

Möbius, M., & Rosenblat, T. (2006). Why beauty matters. *American Economic Review, 96,* 222–235.

Olivardia, R. (2002). Body image and muscularity. In T. F. Cash & T. Pruzinsky (Eds.), *Body image: A handbook of theory, research, and clinical practice* (pp. 210–218). New York, NY: Guilford Press.

Pfeffer, J., & Ross, J. (1982). The effects of marriage and a working wife on occupational and wage attainment. *Administrative Science Quarterly, 27,* 66–80.

Pingitore, R., Dugoni, B. L., Tindale, R. S., & Spring, B. (1994). Bias against overweight job applicants in a simulated employment interview. *Journal of Applied Psychology, 79,* 909–917.

Pollan, M. (2006). *The omnivore's dilemma: A natural history of four meals.* New York, NY: Penguin Press.

Pope, H. G., Phillips, K. A., & Olivardia, R. (2000). *The Adonis complex: The secret crisis of male body obsession.* New York, NY: Free Press.

Prentice, A. M., & Jebb, S. A. (2001). Beyond body mass index. *Obesity Reviews, 2,* 141–147.

Puhl, R., & Brownell, K. D. (2001). Bias, discrimination, and obesity. *Obesity Research, 9,* 788–805.

Puhl, R., & Brownell, K. D. (2003). Psychosocial origins of obesity stigma: Toward changing a powerful and pervasive bias. *Obesity Reviews, 4,* 213–227.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., & Congdon, R. T., Jr. (2004). *HLM 6: Hierarchical linear and nonlinear modeling.* Chicago, IL: Scientific Software International.

Roehling, M. V. (1999). Weight-based discrimination in employment: Psychological and legal aspects. *Personnel Psychology, 52,* 969–1016.

Roehling, M. V., Roehling, P. V., & Odland, L. M. (2008). Investigating the validity of stereotypes about overweight employees: The relationship between body weight and normal personality traits. *Group & Organization Management, 33,* 392–424.

Rundle, R. A. (2003, March 12). Obesity tops smoking for medical costs. *Wall Street Journal,* p. B7.

Russell, C. J., & Dean, M. A. (2000). To log or not to log: Bootstrap as an alternative to the parametric estimation of moderation effects in the presence of skewed dependent variables. *Organizational Research Methods, 3,* 166–185.

Sekiguchi, T., Burton, J. P., & Sablynski, C. J. (2008). The role of job embeddedness on employee performance: The interactive effects with leader–member exchange and organization-based self-esteem. *Personnel Psychology, 61,* 761–792.

Silverstein, B., Perdue, L., Peterson, B., & Kelly, E. (1986). The role of the mass media in promoting a thin standard of bodily attractiveness for women. *Sex Roles, 14,* 519–532.

Staudinger, U. M., Fleeson, W., & Baltes, P. B. (1999). Predictors of subjective physical health and global well-being during midlife: Similarities and differences between the U.S. and Germany. *Journal of Personality and Social Psychology, 76,* 305–319.

Sturm, R. (2002). The effects of obesity, smoking, and drinking on medical problems and costs. *Health Affairs, 21,* 245–253.

The Economist. (2004, March 27). Obesity: A staple in time. *The Economist,* p. 50.

Tilly, C., & Albelda, R. (1994). Family structure and family earnings: The determinants of earnings differences among family types. *Industrial Relations, 33,* 151–167.

U.S. Department of Labor. (1991). *Dictionary of occupational titles, revised* (4th ed.). Retrieved from ftp://ftp.xwalkcenter.org/download/

Wade, T. J., & DiMaria, C. (2003). Weight halo effects: Individual differences in perceived life success as a function of women's race and weight. *Sex Roles, 48,* 461–465.

Wilk, S. L., Desmarais, L. B., & Sackett, P. R. (1995). Gravitation to jobs commensurate with ability: Longitudinal and cross-sectional tests. *Journal of Applied Psychology, 80,* 79–85.

Williams, P. T., & Wood, P. D. (2006). The effects of changing exercise levels on weight and age-related weight gain. *International Journal of Obesity, 30,* 543–551.

Wilson, E. (2006, September 21). When is thin too thin? *The New York Times*. Retrieved from http://www.nytimes.com/2006/09/21/fashion/21MODELS.html

Wiseman, C., Gray, J., Moismann, J., & Ahrens, A. (1992). Cultural expectations of thinness in women: An update. *International Journal of Eating Disorders, 11,* 85–89.

Wysocki, B., Jr. (2004, October 12). Companies get tough with smokers, obese to trim costs. *Wall Street Journal,* p. B1.

Received April 15, 2009
Revision received May 27, 2010
Accepted July 2, 2010 ■