SHAPE MATTERS: EVIDENCE FROM MACHINE LEARNING ON BODY SHAPE-INCOME RELATIONSHIP

A PREPRINT

Suyong Song†
Department of Economics
University of Iowa
Iowa City, IA 52242
suyong-song@uiowa.edu

Stephen Baek
Department of Industrial and Systems Engineering
University of Iowa
Iowa City, IA 52242
stephen-baek@uiowa.edu

December 19, 2021

ABSTRACT

We study the association between physical appearance and family income using a novel data which has 3-dimensional body scans to mitigate the issue of reporting errors and measurement errors observed in most previous studies. We apply machine learning to obtain intrinsic features consisting of human body and take into account a possible issue of endogenous body shapes. The estimation results show that there is a significant relationship between physical appearance and family income and the associations are different across the gender. This supports the hypothesis on the physical attractiveness premium and its heterogeneity across the gender.

Keywords Physical attractiveness premium · Geometric data · Graphical autoencoder

1 Introduction

This paper studies the relationship between physical appearance and family income using unique three-dimensional (3D) whole-body scan data. Recent development in machine learning is adapted to extract intrinsic body features from the scanned data. Our approach underscores the importance of reporting errors and measurement errors on conventional measurements of body shape such as height and body mass index (BMI).

In the literature on the association between physical attractiveness and the labor market outcomes, sparse measurements such as facial attractiveness, height and BMI are mainly considered as measurements of the physical appearance. For instance, Hamermesh and Biddle (1994) studied the impact of facial attractiveness on wages and showed that there is significant beauty premium. Persico et al. (2004) and Case and Paxson (2008) analyzed the effects of height on wages. They found apparent height premium in the labor market outcomes. Cawley (2004) estimated the effects of BMI on wages and reported that weight lowers the wages of white females. In most previous studies, physical appearance was measured by imperfect proxies from subjective opinion based on surveys. This concerns a possibility of attenuation bias from reporting errors on the physical appearance in the estimation of the relation between physical appearance and labor market outcomes. In addition, measurements such as height, weight, and BMI are too sparse to characterize detailed body shapes. As a result, the issue of the measurement errors on the body shapes would make it difficult to correctly estimate the true relation so that many studies ended up with mixed results.

*The authors would like to express their appreciation to Petri Böckerman, Gordon B. Dahl, David Frisvold, Dan Hamermesh, M. Ali Khan, Dan Silverman, and Harald Uhlig for numerous helpful suggestions. This paper has also benefited from the reactions of seminar audiences at Microsoft Research in Redmond, Federal Reserve Bank at Kansas City, Georgia Institute of Technology, New York University, as well as from participants at Conference on New Frontiers in Econometrics at Stamford, Joint Statistical Meetings at Vancouver, Midwest Econometrics Group at University of Wisconsin at Madison, Annual Conference of the Southern Economic Association at Washington D.C., and Econometrics Mini-conference at University of Iowa.

†Corresponding author
We use a novel dataset, the Civilian American European Surface Anthropometry Resource (CAESAR) dataset. The dataset contains detailed demographics of subjects and 40 anthropometric measurements such as height and weight, obtained using tape measures and calipers. It also contains the height and weight reported by subjects. This allows us to calculate the reporting errors in height and weight for each subject and investigate their properties and impacts on the estimation results. We found that the reporting error in males’ height is correlated with their characteristics such as family income, age, race and birth region, while the reporting error in females’ height is correlated with their age. We also found that the reporting error in males’ weight is dependent with their true weight, while the reporting error in females’ weight is associated with their true weight and hours of exercises.

We further investigate properties of reporting errors using a nonparametric conditional mean and nonlinear quantile functions. The nonparametric estimation of the conditional expectations of the reporting errors in height given the true height shows that the reporting error for female’s height is nonclassical in the sense that the reporting error and the true height are dependent. The quantile regression provides that the conditional median of the reporting error is independent of the true height. Thus it would be more plausible to impose a restriction on the conditional quantile of the reporting error of height than the conditional mean (see Bolleller (1998), Hu and Schennach (2008), and Song (2015)). On the other hand, the nonparametric conditional mean and nonlinear quantile regressions show that there are substantial nonclassical errors in both genders’ reported weight.

The estimation results for the association between height (or BMI) and the family income confirm that the reporting errors have substantial impacts on the estimated coefficients. Furthermore, such conventional measurements on body shape are too sparse to describe whole body structure. So the analyses with the sparse measurements are very sensitive to the variable selection, which implies that regressions with the measured height and BMI might suffer from the issue of the measurement errors on the body shape. A handful of papers address the issue by proposing statistical methods such as bias-correction methods or instrumental-variables approaches which deliver consistent estimators at the expense of strong assumptions.

The dataset encloses digital 3D whole-body scans of subjects, which is a very unique feature. The scanned data on human body shapes would mitigate the issue of possible measurement errors due to the sparse measurements. Since the observed variable on body shapes in the dataset is three-dimensional, nevertheless, it is not straightforward to incorporate the data into the model of the family income. Indeed, there are 45,534 covariates for each individual’s body scan. To this end, we adopt methods based on machine learning to identify important features from 3D body scan data. Autoencoders are a certain type of artificial neural networks that possesses an hour-glass shaped network architecture. They are useful in extracting the intrinsic information from the high dimensional input and in finding the most effective way of compressing such information into the lower dimensional encoding. As shown in this paper, the graphical autoencoder can effectively extract the body features and is not sensitive to random noises. To the best of our knowledge, we are not aware of any studies published in Economics which use three-dimensional graphical data with such high dimensionality.

There have been increasing attentions to geometric data such as human body shapes, social networks, firm networks, product reviews, geographical models, etc, in economic studies. In this paper, we introduce new methodology built on deep neural networks and show how it can be utilized to analyze the economic model when the available data has a geometric structure. When one attempts to incorporate geometric data in statistical analyses, there is no trivial grid-like representation for the data. As a result, encoding the features and characteristics of each data point into a numerical form is neither straightforward nor consistent. Most existing studies simplify the geometric features with some sparse characteristics. For instance, in the human body data, many of the relevant studies quantify the geometric characteristics of a human body shape with some sparse measurements, such as height and weight. However, such methods do not always capture detailed geometric variations and often lead to an incorrect statistical conclusion due to the measurement errors. As a better alternative, we propose a graphical autoencoder that can interface with the three-dimensional graphical data. The graphical autoencoder permits incorporation of geometric manifold data into economic analyses. As we will discuss, direct incorporation of the graphical data can reduce measurement errors because graphical data in general provides more comprehensive information on geometric data compared to discrete geometric measurements.

From the proposed method using the graphical autoencoder, we successfully identify intrinsic features of the body shape from 3D body scan data. Interestingly, intrinsic features of the body type are significantly important to explain the family income. Using the graphical autoencoder, we identify two intrinsic features forming male’s body type and three intrinsic features for female’s body type. In contrast to the conventional principle component analysis, the graphical autoencoder renders us to interpret the extracted features. For both genders, the first feature captures how tall a person is

---

3 A 3D whole-body scan for each subject contains 15, 178 number of vertices/nodes and each vertex consists of geometric $(x, y, z)$ coordinate. Thus there are $15, 178 \times 3 = 45, 534$ inputs which is high dimensional.
Figure 1: Summary of the Estimation Results for Family Income Equation. Estimated coefficients and bootstrapped 90% confidence bands are reported. The left panel presents results from the conventional body measures and the right panel reports results from the deep-learned body parameters through the graphical autoencoder.

As acknowledged in the literature, body types could be endogenous in that these can be also driven by unobserved factors of income such as nutrition, personality, ability, and family background. In order to identify the causal impact of body types on family income, we correct for possible endogeneity issues of body types. We utilize proxy variables approach and control functions approach. In particular, our identification strategy is to use variations in shoe size, jacket size for males (blouse size for females), and pants size as legitimate instrumental variables for stature in the control functions approach. By testing the null of exogenous stature, we find that female’s stature is endogenous but not male’s.

We summarize the main findings in Figure 1. In the estimation results from the deep-learned body parameters (right panel), we find that for males stature has a positive impact on family income and is statistically significant at 5% significance level, while obesity is insignificant. We estimate one centimeter increase in stature (converted in height) is associated with approximately $998 increase in the family income for a male who earns $70,000 of median family income. For females, the coefficient of obesity is negative and statistically significant at 1% significance level. On the other hand, coefficients of other features such as height and hip-to-waist ratio are statistically insignificant. One unit decrease in obesity (converted in BMI) is associated with approximately $934 increase in the family income for a female who earns $70,000 family income. The results imply there still exist physical attractiveness premium and its heterogeneity across the gender in the relationship between body shapes and income, even after controlling for unobserved confounding factors. Education is statistically significant for both genders but experience is significant only for the female samples. However, in the estimation results from the conventional body measures such as height and BMI (left panel), the magnitude of the estimated coefficients are much smaller than those from the deep-learned body parameters, which supports the possibility of attenuation bias due to measurement errors. This could suggest that height and BMI have limited powers to describe body shapes so any statistical analysis based on those simple measurements would lead to wrong economic inferences. Our findings also highlight the importance of correctly measuring body shapes to provide adequate public policies for the healthcare.

The rest of the article is organized as follows. Section 2 presents the model of interest. Section 3 introduces and summarizes the CAESAR dataset. Section 4 discusses the impact of reporting errors in height and weight. Section 5 discusses estimation results for the impact of the physical appearance on family income. Section 6 concludes. Empirical estimation results are contained in Appendix.

2 Model

We consider the association between family income and body shapes as follows:

\[ \text{Family Income}_i = \alpha X_i + \beta \text{Body Shapes}_i + \epsilon_i, \quad i = 1, \ldots, N, \]

where Family Income\(_i\) is log family income. Body Shapes\(_i\) is a measure of body types, \(X_i\) is a set of covariates, and \(\epsilon_i\) is unobserved causes of family income for individual \(i\). We are particularly interested in the parameter \(\beta\), but we also discuss the relationship between family income and other individual characteristics through the vector of parameters \(\alpha\). We consider family income instead of individual income, since family income is available in the data. So we identify the combined effects of body shapes on income through the labor market and marriage market. In fact, as documented
We use a unique dataset, called the Civilian American European Surface Anthropometry Resource (CAESAR) dataset. In Chiappori et al. (2012), various studies found assortative matching on income, wages, education, and anthropometric characteristics such as weight or height in marriage market. As a result, total effects of body shapes in the labor market and marriage market are identifiable from family income so that it is worthwhile to investigate the impact of physical attractiveness on family income.

A large body of literature has analyzed the presence of earnings differentials based on physical appearance. A strand of literature has focused on facial attractiveness. Hamermesh and Biddle (1994) analyzed the effect of physical appearance on earnings using interviewers’ ratings of respondents’ physical appearance. They found evidence of a positive impact of looks on earnings. Mobius and Rosenblat (2006) examined the sources of the beauty premium and decomposed the beauty premium that arises during the wage negotiation process between employer and employee in an experimental labor market. They identified transmission channels through which physical attractiveness raises an employer’s estimate of a worker’s ability. Scholz and Sicinski (2015) studied the impact of facial attractiveness on the lifetime earnings. They found there exists the beauty premium even after controlling for other factors which enhance productivity in the labor market earnings.

Other threads of literature have analyzed the effects of height on labor market outcomes. Persico et al. (2004) found that an additional inch of height is associated with an increase in wages, which is a consistent finding in the literature in addition to racial and gender bias. They showed that how tall a person is as a teenager is the source of the height wage premium. This implies that there are positive effects of social factors associated with the accumulation of productive skills and attributes on the development of human capital and the distribution of economic outcomes. Case and Paxson (2008) also found there are substantial returns to height in the labor market. However, they showed that the height premium is the result of positive correlation between height and cognitive ability. Lundborg et al. (2014) found that the positive height-earnings association is explained by both cognitive and noncognitive skills observed in tall people. Deaton and Arora (2009) reported that taller people evaluate their lives more favorably and the findings are explained by the positive association between height and both family income and education. Böckerman and Vainiomäki (2013) used twin data to control for unobserved ability and found a significant height premium in wage for women but not for men. Lindqvist (2012) studied the relationship between height and leadership and confirmed that tall men are significantly more likely to attain managerial positions.

Cawley (2004) considered the effects of obesity on wages. He found that weight lowers the wages of white females and noted that one possible reason for the result is that obesity has adverse impact on the self-esteem of white females. Rooth (2012) used a field experimental approach to find differential treatment against obese applicants in terms of the number of callbacks for a job interview in the hiring process in the Swedish labor market. He found the callback rate to interview was lower for both obese male and female applicants than for nonobese applicants.

Mathematically, human body shapes can be viewed as arbitrary 2-manifolds \( M \) embedded in the Euclidean 3-space \( \mathbb{R}^3 \). In statistical analyses as in equation (1), quantifying geometric characteristics of different manifold shapes and encoding them into a numerical form is not straightforward. Thus, these continuous manifolds are approximated by proxies in a tensor form. Due to this reason, many of the relevant works in the literature on the physical appearance quantify the geometric characteristics of a human body shape with some sparse measurements, such as height, weight, or BMI. As we will see in the later sections, however, such kind of quantification methods do not always capture detailed geometric variations and often lead to an erroneous explanation of statistical data. For instance, with height and BMI alone, one can hardly distinguish muscular individuals from individuals with round body shapes. The situation does not improve even if some new variables, such as chest circumference, are added, since these variables still are not quite enough to codify all the subtle variations in body shapes. Moreover, oftentimes, such additional variables merely add redundancy, without adding any substantial statistical description of data, as the commonly-used anthropometric parameters are highly correlated to each other. In addition, it is also noteworthy that the manual selection of measurement variables can also introduce one’s bias into the model. In this paper, we compare several common ways of quantifying manifold structured data with a newly-proposed graphical autoencoder method.

### 3 Data

We use a unique data, called the Civilian American European Surface Anthropometry Resource (CAESAR) dataset. It was collected from a survey of the civilian populations of three countries representing the North Atlantic Treaty Organization (NATO) countries: the U.S., The Netherlands, and Italy. The survey was primarily conducted by the U.S. Air Force and the sample from the U.S. was used for our study. The survey of the U.S. sample was conducted from 1998 to 2000 and carried out in 12 different locations which were selected to obtain subjects approximately in proportion to the population of each of 4 regions of the U.S. Census.\(^4\)

---

\(^4\)The U.S. data is referred to as the North American sample since one site in Ottawa, Canada was added to the sample.
The dataset contains 2,383 individuals whose ages vary from 18 to 65 with a diverse demographical population. The dataset contains detailed demographics of subjects, anthropometric measurements done with a tape measure and caliper, and digital 3D whole-body scans of subjects. In contrast to other traditional surveys, the data contains both reported and measured height and weight. This feature makes it possible to calculate reporting errors in the survey data and analyze their relations to the correctly measured height/weight as well as individual characteristics. In addition, the existence of 3D whole-body scan data makes the CAESAR data serve as a good proxy to physical appearance such that potential issue of measurement errors can be mitigated.

Some of the total 2,383 subjects in the database have missing demographic and anthropometric information; these have been deleted in our study. In addition, there are also subjects who elected not to disclose and/or were not aware of their income, race, education, etc. These individuals have also been removed in this study. In the analysis, we divide the sample by gender to take into account the differential treatment across genders.

Tables 1-2 provide summary statistics of the variables in the database for males and females, respectively. The data has a single question about family income (grouped into ten classes). Average family income is $76,085 for males and $65,998 for females. The differences in average family income across genders would be due to the fact that the male sample includes more married people than the female sample. Median family income is slightly lower than the mean family income, which amounts to $70,000 for males and $52,500 for females. For males, on average, reported height is 179.82 centimeters and measured height is 178.26 centimeters, which shows a tendency of over-reporting. The gap is larger when median reported height (180.34 centimeters) and measured height (177.85 centimeters) are compared. We observe a similar pattern in the female sample: reported height is 164.96 centimeters and measured height is 164.22 centimeters on average; median reported height is 165.1 centimeter and median measured height is 164 centimeters.

The males’ average reported weight is 86.03 kilograms and the average of the measured weight is 86.76 kilograms. The median of two measurements are the same. For females, reported weight is 67.88 kilograms and measured weight is 68.81 kilograms on average. Median reported weight is 63.49 kilograms and median measured weight is 64.85 kilograms. In both subsamples, the standard errors of the weight are large, which are approximately 17 kilograms. BMI has been commonly used as a screening tool for determining whether a person is overweight or obese. BMI is calculated as weight in kilograms divided by height in meters squared. We refer reported BMI (measured BMI) to the one based on reported height and weight (measured height and weight). In the tables, height, weight and BMI are those measured by professional tailors at the survey sites. For both genders, reported BMI is slightly larger than measured BMI on average.

In addition to the bio-metric measurements, the data contains other variables for individual characteristics and socio-economic backgrounds. Education grouped into nine categories is 16.29 years for males and 15.75 years for females on average. Experience is calculated as potential experience = age – education – 6 and its mean is 17.54 years for males and 18.62 years for females. Fitness is defined as exercise hours per week. Its mean and median are 4.24 hours and 2.5 hours, respectively, for males. For females, its mean and median are 3.74 hours and 2.5 hours, respectively.

The data also include the number of children. Marital status is classified as three groups: single, married, divorced/widowed. Occupation consists of white collar, management, blue collar, and service. Race has four groups including White, Hispanic, Black, and Asian. Birth region is grouped into five groups including Midwest, Northeast, South, West, and Foreign. The majority in the dataset are white collar married White males and females born in Midwest. As we will discuss later, the data also contains 40 body measures which includes height and weight. The list of the body measures are provided in Table 3.

4 Reporting Errors in Height and Weight

Several studies in the literature use survey data so that they assume there are no reporting errors in height and weight or reporting errors are classical in that they are not correlated with true measures. Since our data contains both reported and measured height and weight, we can further investigate the properties of the reporting errors. We consider measured height and weight as the true height and weight since they are measured by professional tailors. The reporting errors are calculated as Reporting Error_\text{H} = \text{Reported Height} - \text{Height} and Reporting Error_\text{W} = \text{Reported Weight} - \text{Weight}, respectively.

According to Centers for Disease Control and Prevention (CDC), the standard weight status categories associated with BMI ranges for adults are as follows: below 18.5 (underweight), 18.5-24.9 (normal or healthy weight), 25.0-29.9 (overweight), 30.0 and above (obese).

Exceptionally, Persico et al. [2004] and Case and Paxson [2008] use measured height from the British National Child Development Survey, even though they also use self-reported height from the British Cohort Study and the National Longitudinal Survey of Youth, respectively. Lundborg et al. [2014] use measured height from the Swedish National Service Administration.
The following equation estimates which personal background explains reporting error in height and weight:

\[
\text{Reporting Error}^H = \pi X_i + \mu \text{Height}_i + \varepsilon_i, \tag{2}
\]

\[
\text{Reporting Error}^W = \pi X_i + \mu \text{Weight}_i + \varepsilon_i, \tag{3}
\]

where \(X_i\) is a set of covariates including family income, age, age squared, occupation, education, marital status, fitness, race, and birth region. \(\text{Height}_i\) is the true height in millimeters, and \(\text{Weight}_i\) is the true weight in kilograms. We found dependence between reporting errors and some covariates. Table 4 reports the estimation results. The standard errors are estimated by bootstrapping and are reported inside the parentheses. In the equation (2), the coefficient of the true height is not statistically significant for both genders. We observe different results across the gender. For males, family income is negatively correlated with the reporting error in height at 1% significance level. Age squared is positively correlated with the reporting error. Hispanic males are more likely to under-report their height compared to White males. Males who were born in Northeast are more likely to over-report their height relative to those from Midwest. On the other hand, the coefficient of family income is not statistically significant for females. Older females are more likely to under-report their height. The estimation results are summarized in Figure 2.

In the equation (3), the true weight is negatively correlated with the reporting error in weight (at 1% significance level) for both genders: heavier people have a tendency to under-report their weight. For females, it is interesting to find that the coefficient of fitness is statistically significant at 5% significance level and it is negatively correlated with reporting-error in weight. Thus females who spend more time on exercise have a tendency to under-report their weight. However, we find little evidence that other personal background are correlated with reporting error in weight. The estimation results are summarized in Figure 3.

Figures 4 and 5 plot the estimation of the conditional expectations of the reporting errors in height/weight given the true measures, namely, \(E[\text{Reporting Error}^H | \text{Height}]\) and \(E[\text{Reporting Error}^W | \text{Weight}]\) with their 95% confidence bands, respectively. We use nonparametric kernel estimation where the kernel function is an Epanechnikov kernel and the bandwidth is chosen by the Silverman’s rule-of-thumb. The confidence bands are estimated by a nonparametric bootstrap method. The solid line represents zero reporting error. The nonparametric plots for the height show that the reported height is larger than the true height at almost all height level in both genders showing over-reporting patterns. It displays no significant relation between the reporting error and the true height for males. For females, we observe more reporting error at low height level than at the average height level, which indicates that the reporting error for females’ height is nonclassical in the sense that the reporting error and the true height are dependent. This result is a finding which is not captured by the linear mean regression in Table 4 where the reporting error in height is not related to the true height for both genders.

The plots for the reported weight show more substantial nonclassical errors. Males who are at the low weight level (below approximately 75 kilogram) have a tendency to over-report their weight, but males at the weight level above approximately 75 kilogram under-report their weight. Similarly, females who are at the low weight level (below approximately 50 kilogram) have a tendency to over-report their weight, but females at the weight level above
approximately 50 kilogram under-report their weight. The both plots display an apparent dependence between the reporting error and the true weight. This confirms the significant negative relation between the reporting error and the true weight shown in Table 4.

Conditional quantile function is a useful tool to estimate heterogeneity in a conditional distribution. It also measures what proportion of reporting errors are positive or negative. Figures 6-7 present the estimation of the conditional quantiles of the reporting errors in height and weight conditional on the true measures, namely, $Q_\tau[\text{Reporting Error}_H \mid \text{Height}]$ and $Q_\tau[\text{Reporting Error}_W \mid \text{Weight}]$ for $\tau \in (0, 1)$ with their 95% confidence bands, respectively. We estimate the conditional quantiles using the nonlinear polynomial regression. The figures display the median and the 10%, 25%, 75%, and 90% quantiles across genders. In Figure 6, the results show that there is heterogeneity in the conditional distribution of the reporting error in height for both genders. It is shown that over-reporting of height is more pronounced for males than females. We notice that more than 75% of the sample of the males over-report their height. Interestingly, the median regression lines in both genders are approximately parallel with the horizontal line, which implies that the conditional median of the reporting error is independent of the true height. [Bollinger (1998)] also found that median regressions for earnings will be more robust to the reporting error than mean regressions. Thus, it would be more natural to impose a restriction on the conditional quantile of the reporting error of height than the conditional mean, as in [Hu and Schennach (2009)].
Figure 5: Conditional Mean of Reporting Errors in Weight Conditional on True Weight.

Figure 6: Conditional Quantile of Reporting Errors in Height Conditional on True Height.

Figure 7 also displays apparent heterogeneity in the conditional distribution for both genders. It shows that when heavier people than the average are concerned, under-reporting of weight is more pronounced for females than males. We notice that within this group, almost 75% of the sample of the females under-report their weight. The median regressions in both genders are dependent on the level of the measured weight. This indicates that there are substantial nonclassical errors in the reported weight so that a restriction on the conditional quantile may not be valid.

5 Estimation of the Association between Physical Appearance and Labor Market Outcomes

In this section, we estimate the association between the physical appearance and family income using various methods.

5.1 Height, Weight and Reporting Errors

Most papers in the literature estimate the relationship in the equation (1) by replacing body shapes with their observed proxies such as height or weight. However, these measurements are hardly accurate to fully describe body shapes. Furthermore, only including either height or weight without controlling for the other as in the literature could suffer from an omitted variable problem. For instance, consider there are two people who have the same height but different weight. Then comparing height only will fail to identify the difference in their body shapes. Thus, we consider the
following two regression equations:

\[
\text{Family Income}_i = \alpha X_i + \beta_1 \text{Height}_i + \epsilon_i, \\
\text{Family Income}_i = \alpha X_i + \beta_1 \text{Height}_i + \beta_2 \text{Weight}_i + \epsilon_i,
\]

where \( X_i \) is a set of controls including experience, experience\(^2 \), race, occupation, education, marital status, and number of children. We can test the importance of controlling for weight by comparing the estimated coefficients from two equations. In addition, as mentioned before, the data contains measurements on height and weight both reported by subjects and measured by on-site measurers. Therefore, by comparing estimates of measured one with reported counterpart, we can see how much the reporting errors affect the estimation results. Table 5 reports estimation results from reported height and weight. Table 6 provides estimation results from measured height and weight.

The hypothesis that the coefficient on height is zero is tested across gender. Results for both genders are presented in each table. In equation (4) of Table 5, reported weight is not included. The column for males shows that education is statistically significant in the income equation. The coefficient of the reported height is positive and statistically significant at 10\% significance level. The column for females is somewhat different from that for males: the coefficient of experience, experience\(^2 \), and education are statistically significant. In addition, the coefficient on the reported height is positive and statistically significant at 5\% significance level. In equation (5), we add the reported weight to the set of regressors. The column for males shows that the coefficient of the reported height becomes statistically insignificant, and the coefficient of the reported weight is also insignificant. However, in the column for females the coefficient of the reported height is still positive and statistically significant, but the coefficient on the reported weight is insignificant.

In Table 6, we instead use the measured height and weight to estimate the income equation. Interestingly, the coefficients on the height for males in both equations are positive and statistically significant at 1\% significance level. Their magnitudes are larger than those from Table 5. When the measured weight is added, its coefficient is still insignificant for males. For females, the coefficients on the measured height are statistically significant in both equations and their magnitudes are larger than those from Table 5. When the measured weight is added, its coefficient becomes significant at 10\% significance level, which shows a negative association between family income and weight. Thus, we confirm there are apparent reporting errors in height and weight. Particularly, the impacts of the reporting errors on the estimation results are more severe in males than females. These reporting errors bring attenuation bias to the estimates. Furthermore, the estimation results from two equations (4)-(5) are different. It shows that using height only as a proxy to body shapes might be too simple to describe delicate figures of the physical appearance.

As in [Cawley (2004)], we consider BMI as the primary regressor in the regression equations (4)-(5) to estimate the impact of obesity on income. Again, we add weight or height as additional regressor to take into account a possible omitted variable problem. So we consider the income equations as following:

\[
\text{Family Income}_i = \alpha X_i + \beta_1 \text{BMI}_i + \epsilon_i, \\
\text{Family Income}_i = \alpha X_i + \beta_1 \text{BMI}_i + \beta_2 \text{Weight}_i + \epsilon_i, \\
\text{Family Income}_i = \alpha X_i + \beta_1 \text{BMI}_i + \beta_2 \text{Height}_i + \epsilon_i,
\]

where BMI\(_i\) is the body mass index. We first estimate the equations using the reported variables and summarize the estimation results in Table 7. From the columns for males in the table, the coefficient of the reported BMI in equation
Figure 8: Comparison of Reported and Measured Body Measures. Estimated coefficients and bootstrapped 90% confidence bands are reported. We provide results from reported body measures (Reported) and measured body measures (Measured). Note that the unit for height is converted into centimeter (cm).

(6) is statistically insignificant. Adding the reported weight or height does not change the result for the reported BMI as in equations (7)-(8). Instead, the estimated coefficient of the reported height or weight is significant. For females, the coefficients of the reported BMI are insignificant in equations (6) and (8). However, in equation (7), the reported BMI has a negative impact on family income and the relation is statistically significant at 1% significance level. It also shows that the coefficient of the reported height or weight is positive and significant at 5% significance level.

We next estimate equations (6)-(8) using the measured BMI, height and weight. For all equations in Table 8, the estimation results are somewhat different from Table 7. In equation (6) for males, the coefficient of BMI is still insignificant. When weight is included as in equation (7), its coefficient for males is positive and statistically significant at 1% significance level. The coefficient of BMI becomes negative and statistically significant at 5% significance level. When height is included as in equation (8), its coefficient for males is positive and statistically significant at 1% significance level. However, the coefficient of BMI is statistically insignificant. For females, the results are different from those in Table 7. The coefficient of BMI is always negative and statistically significant. The coefficient of height or weight is positive and also statistically significant. The results for equation (8) are highlighted in Figure 8. It is shown that the magnitudes of the coefficients for height are larger when measured height are used. For females, the impact of the measured BMI on family income is significant in contrast with insignificant impact of the reported BMI. Different signs on the effect of BMI across genders are also observed: positive effect of male BMI as opposed to negative effect of female BMI. Thus, the analysis confirms that the reporting errors in body measures have significant impacts on the estimated coefficients.

Interestingly, we also observe that the estimation results have significantly changed as different set of measures of body types were included in the equations. One possible explanation for this is that even the measured height and BMI might not be perfect proxies to the body types, although they are less prone to reporting errors. In fact, height, weight and BMI are simple measures of body types so that they might miss useful information on the true body types (e.g., see Wada and Tekin (2010) for BMI).

In order to further investigate the role of the measurement errors on the body types, we run the following regression equation:

$$\text{Family Income}_i = \alpha X_i + \beta \text{Body}_i + \epsilon_i,$$

where Body is a set of body measurements which include 40 number of measurements on various parts of body. Since these are more sophisticated than simple measurements of height and BMI, it is less likely that the measurement errors on body type is prevalent.

Table 9 presents the estimation results. Except height and weight, for brevity, we only report measures of body parts which are statistically significant. Coefficients on age, race, occupation, education, and marital status are very similar.

Several studies propose statistical methods to reduce the measurement errors in body-shape measurements. Among others, Courtemanche et al. (2015) propose a rank-based correction method for using validation data to correct the measurement errors in obesity. Murillo et al. (2019) reduce bias in obesity by applying regression calibration, simulation extrapolation, and multiple imputation approaches.

A full list of the measurements is provided in Table 3.
Figure 9: Various Measures and Family Income. Estimated coefficients and bootstrapped 90% confidence bands are reported for male (left) and female (right). Note that units for all measurements except weight are converted into centimeter (cm).

to those in Table 8 for both genders. Interestingly, we found eight statistically-significant body measurements for males and four for females (see Figure 9). For instance, in the sample of males, Acromial Height (Sitting), Chest Circumference, and Waist Height (Preferred) have positive association with the family income, while Arm Length (Shoulder-to-Elbow), Buttock (Knee Length), Elbow Height (Sitting), Subscapular Skinfold, Waist Circumference (Preferred) are negatively correlated with the family income. However, the coefficients on Face Length, Hand Length, Neck Base Circumference are all negative. The most distinctive result is that the coefficients on height and weight for both genders are statistically insignificant in the regression. This implies that there are useful information on body types which are embedded into various body measures. The body shapes or types cannot be fully captured by simple measures such as height or weight.

Moreover, it is possible that there are interactions between different body measurements since they have close relationships to construct a body shape. So we consider the original regressors \( X_i \) and Body \( i \) in equation (9)) and interaction terms of Body \( i \) as a set of regressors. This gives 797 number of covariates, which makes the OLS regression inconsistent. In order to mitigate the issue of high-dimensional data, we use the following Lasso regression which is valid under a sparsity assumption:

\[
\min_{\psi} \left( \frac{1}{2N} \sum_{i=1}^{N} (\text{Family Income}_i - \psi Z_i)^2 + \lambda \sum_{j=1}^{p} |\psi_j| \right)
\]

where \( Z \) is a vector of covariates including the interaction terms with size \( p = 797 \) and \( \lambda \in (0, 1) \) is a regularization parameter. We construct the lasso fit using 10-fold cross-validation. Figure 10 plots mean-squared-error (MSE) over the sequence of the regularization parameter \( \lambda \) for each gender. For males, the minimum MSE is 0.231 at \( \lambda = 0.024 \) and the minimum MSE plus one standard error is 0.241 at \( \lambda = 0.047 \). For females, the minimum MSE is 0.216 at \( \lambda = 0.017 \) and the minimum MSE plus one standard error is 0.228 at \( \lambda = 0.058 \). The regression results show that many interaction terms are statistically significant. When equation (9) is re-estimated with these interaction terms (so-called "post-Lasso"), it obtains higher adjusted \( R^2 \) squared (0.393 for males and 0.470 for females) than those in Table 9. Thus it is highly likely that these body measures are interrelated. However, constructing stylized body types based on these relevant body measures is not straightforward and a nonstandard problem.

5.2 Physical Appearance and Graphical Autoencoder

Characterization of geometric quantity such as physical appearance of human body shape using a sparse set of canonical features (e.g., height and weight) often causes unwanted bias and misinterpretation of data. For simple shapes like rectangles, canonical measures such as width and height already provide a complete description of the shape. Hence, shape variation among rectangles could easily be described using the two canonical parameters without much issues. However, this seldom applies to more sophisticated shape variations, if at all. Instead, the canonical shape descriptors, often hand-selected, might cause nonignorable error in capturing genuine statistical distribution by overlooking some

\(^9\)We note that OLS is consistent under some regularity conditions only if the number of observations is larger than the number of covariates.

\(^{10}\)To save the space, we omit the results here.
important geometric features or measuring highly-correlated variables redundantly, which can be thought of as a measurement error of some sort.

Unfortunately, however, extracting a complete and unbiased set of shape descriptors is not a trivial task. Furthermore, the task is highly problem-specific such that, for example, the shape descriptors for car shapes would not be appropriate for describing human body shapes. To this end, we propose a novel data-driven framework for extracting complete, unbiased shape descriptors from a set of geometric data in this paper. The proposed framework utilizes an autoencoder neural network (Bourlard and Kamp, 1988) defined on a graphical model. In this section, we present an overview of the approach and demonstrate that the shape descriptors obtained through the new approach can actually provide a better description of data.

5.2.1 Graphical Autoencoder

Mathematically, human body shapes can be represented as curved surfaces, or more formally manifolds $M^{(i)}$ embedded in $\mathbb{R}^3$, where $i$ is an index identifying each individual. A manifold is a topological space that locally looks like Euclidean space near each point. The statistical models that this paper concerns are, in a generic form, a regression of an economic variable $Y$ with respect to a manifold-structured regressor $M$ and other covariates $X$:

$$Y = \phi(M, X; \theta) + \epsilon. \quad (11)$$

where $\phi$ is a known function up to unknown parameter $\theta$ and $\epsilon$ is an error term. Here, a problem arises regarding the manifold regressor $M$ as the regressor $M$ is an abstract, geometric object and not a usual vector variable as in other typical economic and statistical models. In other words, there is no statistical model that naturally accepts the manifold regressor $M$, unless $M$ is somehow converted into a vector form.

Due to the above bottleneck, one may consider measuring a few geometric dimensions, such as lengths and girths, and use those measurements to encode body shapes. However, as being shown in the paper, such simplistic measurements are not an accurate characterization of complex geometric objects such as human body shapes. Instead, data driven parameterizations such as in [Wang 2005; Baek and Lee 2012; Pishchulin et al. 2017] provide more comprehensive and reliable codification of body shapes, but many of these works assume that the human body shape distribution is linear, leading to inaccurate encoding of body shapes [Freifeld and Black 2012; Baek 2013].

In this paper, we employ a data-driven, nonlinear parameterization of body shapes achieved via a graphical autoencoder. An autoencoder is a certain type of artificial neural network that possesses a hour-glass shaped network architecture. An autoencoder can be thought of as two multilayer perceptron (MLP) models cascaded sequentially, where the first MLP codifies a high-dimensional input into a lower dimensional embedding (encoder) and the second MLP reconstructs the original input back from the encoded embedding (decoder). Because of the dimensional bottleneck created in the middle, the neural network is promoted to search for the most effective way of compressing the high dimensional input into the lower dimensional embedding.

The concept of a graphical autoencoder we propose here is an extension of such notion of autoencoders to manifold-structured data. As in many geometric data analysis applications, we discretize a manifold $M$ to a triangular mesh, achieving piece-wise linear approximation of the original surface. A triangular mesh is a graph $\mathcal{G} = \{V, E, F\}$ where $V$...
is a set of vertices/nodes, $E$ are edges interconnecting the vertices, and $F$ are triangular facets. We equip the meshes $G_1,\ldots,N$ with a semantic correspondence structure such that the graph elements of the same index correspond to the same body part location across all meshes $G_1,\ldots,N$. This process is commonly called “registration” or “correspondence matching” in computer graphics and geometry processing literature and can be achieved via methods such as Zuffi and Black (2015); Wei et al. (2016); Sun et al. (2017, 2018).

In this setting, the graphical autoencoder is defined as follows:

$$p = (f_1 \circ f_2 \circ \cdots \circ f_m)(V \in V), \quad \text{(encoder)}$$

$$V = (g_1 \circ g_2 \circ \cdots \circ g_m)(p). \quad \text{(decoder)}$$

Here, each of the layers $f_1 \cdots f_m$ and $g_1 \cdots g_m$ are modeled as a simple perceptron:

$$f_i(h) \text{ or } g_i(h) = \sigma \left( \sum_j W_i^T h + b_i \right),$$

where $W_i$ are neural weights and $b_i$ are bias. $\sigma$ is the activation function where we empirically decide to be rectified linear unit (ReLU) activation for $f_1 \cdots f_{m-1}$ and $g_1 \cdots g_{m-1}$. We set linear activation for the terminal layers $f_m$ and $g_m$ (i.e. no rectification).

Finally, the graphical autoencoder is trained to minimize the mean square error between the original mesh and the reconstructed mesh:

$$\min_{\theta_f, \theta_g} \|V - g(p)\|$$

where $p = f(V)$ by definition, $\theta_f$ and $\theta_g$ are the model parameters of $f$ and $g$ respectively, and $V$ is the list of vertex coordinates of a graphical model.

Figure 11 illustrates a schematic overview of the graphical autoencoder. As shown in the figure, the vertices of a topology-normalized graphical model act as input neurons in the autoencoder model. Then the input neurons are connected to the hidden neurons in the next layer which then are connected in chain through the “bottleneck” layer. The bottleneck layer has a significantly small number of neurons compare to the input neurons and, hence, the dimensionality compression occurs there. The latter half of the autoencoder is symmetric to the first half and finally reconstructs the bottleneck encoding into the original graphical model. The training process of the graphical autoencoder attempts to minimize the discrepancy between the reconstructed model and the original input by tuning the neural weights of the hidden layers.

5.2.2 Graphical Autoencoder on CAESAR Dataset

The CAESAR scan dataset includes 15,178 number of vertices as well as $(x, y, z)$ coordinate. This gives us 45,534 inputs for each individual. In order to extract body shape parameters that encode the geometric characteristics of a person’s appearance, we designed a graphical autoencoder consisting of seven hidden layers. Each of the hidden layers are comprised of 256-64-16-$d$-16-64-256 neurons respectively, where $d$ is the intrinsic data dimension, or the dimensionality of the embedding. The RMSprop optimizer was used for the training. The dataset was randomly split to a training group used for training and a validation group that were set aside during the training. The ratio between the number of data samples in such groups were 80:20 respectively. The training continued until 5,000 epochs.
with the batch size of 200 samples. As a criterion to evaluate the performance of the graphical autoencoder, we used the reconstruction error measured in mean-squared-error (MSE). As described above, the graphical autoencoder first embeds graphical data into a lower dimensional embedding through the encoder part of the network, which then is reconstructed back into a graphical model through the decoder part. We compared how the reconstructed output is different from the original input to the network.

The first experiment was conducted to test the ability of the graphical autoencoder in embedding the geometric information underlying in data. To achieve this, we applied the aforementioned graphical encoder to the entire CAESAR dataset, with varying embedding dimension \( d \) from 1 to 20 as reported in Figure 12. The embedding accuracy was below \( 3 \times 10^{-4} \) for most cases. Particularly, when the dimension \( d \) was 3, it showed the lowest MSE, in both training and validation losses, which provides a justification for estimating \( d = 3 \) as the intrinsic dimension.

For the meaning of the embedded parameters of the third dimension, the first component, \( P_1 \), discerned to be related to height of a person and \( P_2 \) to the body volume (obesity/leanness). Interestingly, as \( P_3 \) increases the body shape became more feminine, (namely, more prominent chest and hip-to-waist ratio) and, conversely, as it decreases the body shape became more masculine with less prominent chest and curves.

Based on such observation, we further conducted another similar experiment for training the graphical autoencoder with separate genders. Among 2,383 subjects in the CAESAR dataset, there were 1,122 males and 1,261 females. The two groups had been separated to two experiment sessions in which they were further separated to training and validation groups with the same 80:20 ratio.

The new experiment with separate genders demonstrated a similar trend to the first experiment in terms of how the intrinsic dimension affects the reconstruction error, as visualized in Figure 13. However, interestingly, this time, the reasonable intrinsic dimension \( d \) was observed to be 2 for male subjects. We interpret this result that, since now the two genders are separated, the role of \( P_3 \) (feminine/masculine) is now less significant than before and, thus, the gain of accuracy by including the third dimension becomes negligible for males. We also note that, however, such interpretation was not true with the female population, since the accuracy was in fact higher when \( P_3 \) was included. Our explanation to such is that, for the female body shapes, there is a greater variation in body curves compared to male population, and therefore, the third component has a greater significance for females. We, therefore, select \( d = 2 \) for males and \( d = 3 \) for females. Lastly, we also note that the convergence was slower when the two genders were separated and measurable gain of accuracy could be observed even after 1,000 epochs, which was not the case when the two genders were combined in the training. This could be because the number of training samples in the training dataset is significantly smaller (about a half) than the previous case, rendering a drop of the representative power of the data.

Figure 14 illustrates the body shape spanned by the two parameters obtained from the graphical autoencoder for each gender. 3D body shape models are arranged in accordance with their body shape parameters with increments of \(-3\sigma, -1.5\sigma, 1.5\sigma, \) and \(3\sigma\) with respect to the mean in each direction where \( \sigma \) is the standard deviation of each parameter. Body shapes for male (left) and female (right) display similar patterns over changes in the two parameters. Overall, the first parameter \( P_1 \) affects how tall a person is. That is, a smaller value in \( P_1 \) indicates the person is not tall compared to the other population and vice versa. \( P_2 \) is how lean a person is. That is, a large value in \( P_2 \) results in an obese person, while a small value in \( P_2 \) results in a more slim and fit person.

![Figure 12: Result of Training Graphical Autoencoder with the Entire CAESAR Dataset. The abscissa is the number of epochs for the training and the ordinate is the model loss in terms of MSE. The left shows the loss on training dataset (training loss) while the right shows the loss on validation dataset (validation loss). The accuracy did not show any significant improvement after 1,000 epochs for all cases and thus removed from the figure for the sake of better visualization.](image-url)
In order to better understand these parameters, we consider a linear fit of BMI, height, or weight on each parameter. Figure 15 displays the relation between body shape parameters and the conventional body measurements for male. $P_1$ is positively correlated with BMI, height, and weight. Among these body measurements, height is the most highly correlated with $P_1$ (approximately $R^2 = 0.70$). $P_2$ is negatively correlated with height, but is positively correlated with BMI and weight. BMI has the highest correlation with $P_2$ (approximately $R^2 = 0.69$). Figure 16 displays the relation between body shape parameters and the conventional body measurements for female. The patterns are close to those for male in Figure 15. As discussed before, the female sample produces an additional feature, $P_3$. We visualize the third parameter for female in Figure 17. As shown in the figure, $P_3$ captures the ratio of hip to waist for females, which is unique to female dataset. For simplicity, thus, we will interpret $P_1$, $P_2$, and $P_3$ as features associated with a person’s stature, obesity, and hip-to-waist ratio, respectively. It is worthwhile to note that the extracted features $P_1$, $P_2$, and $P_3$ perform better in explaining nonlinear variations in body shapes than simple measures such as height, BMI, and hip-to-waist ratio. As shown in Figure 14, the body shape spanned by $P_1$ and $P_2$ displays dynamic and nonlinear patterns as the parameters vary. In addition, we consider a linear prediction of female $P_3$ using various body measures and report the results in Table 10 (also in Figure 18). It shows that there are many body parts which are highly associated with $P_3$. This confirms that $P_3$ captures complexity in female body shapes and reflects not only hip-to-waist ratio but also variations in other body parts. Nevertheless, we call $P_3$ hip-to-waist ratio for convenience.

---

We only report statistically-significant variables.
5.2.3 Extracted Body Types and Family Income

We now use the measurements of body type which are extracted by graphical autoencoder in the previous section. We estimate the equation (1) with the extracted body types in place of a set of body measurements for Bodyi as following:

\[
\text{Family Income}_i = \alpha X_i + P_{1i} + \epsilon_i, \quad (15)
\]
\[
\text{Family Income}_i = \alpha X_i + P_{2i} + \epsilon_i, \quad (16)
\]
\[
\begin{cases}
\text{Family Income}_i = \alpha X_i + \beta_1 P_{1i} + \beta_2 P_{2i} + \epsilon_i & \text{if male}, \\
\text{Family Income}_i = \alpha X_i + \beta_1 P_{1i} + \beta_2 P_{2i} + \beta_3 P_{3i} + \epsilon_i & \text{if female}, 
\end{cases} \quad (17)
\]

where \( P_{1i}, P_{2i} \) and \( P_{3i} \) are body types for each individual \( i \). Table 12 reports estimation results across the gender with the same set of controls. In equation (17), we add all intrinsic features of the body shape to the income equation. For males, only the coefficient of the \( P_{1i} \) measurement is statistically significant and \( P_{2i} \) does not explain the family income. One standard deviation increase in males’ \( P_{1i} \) is associated with 0.052 increase in log family income. For females, on the other hand, only the coefficient of the \( P_{2i} \) measurement is statistically significant, and \( P_{1i} \) and \( P_{3i} \) are not correlated with the family income. When these insignificant variables are dropped as in equations (15) and (16), the regression equations get higher adjusted R squared. The results show that one standard deviation decrease in females’ \( P_{2i} \) is associated with 0.056 increase in log family income.

For comparison, we replace the extracted body types with height, BMI, and hip-to-waist ratio and re-estimate the above equations. The estimated results are reported in Table 11. In both genders, height has positive impact on log family income and is statistically significant. The estimated coefficients of BMI and hip-to-waist ratio are insignificant. In particular, BMI was significant at 10% significance level in Table 8 when height and BMI are included. However, BMI is not significant anymore when hip-to-waist ratio is added. We observe no gender differential in the impact of body types. The results confirm that the estimation of the income equation with conventional body measurements is susceptible to variable selection and provides different conclusions than those with our proposed method.

5.3 Endogenous Body Types

5.3.1 Proxy Variables Approach

If unobserved determinants of family income such as individual personality and childhood nutrition are correlated with the physical appearance, the estimates in the previous section are inconsistent. To recover the causal relationship, we

---

12 Hip-to-waist ratio is calculated as \( \text{Hip Circumference, Maximum} \times 100. \)
We choose these variables as relevant proxies since fitness and car size could be related to individual preference to body types, and birth state and survey site could reflect local nutrition environments. To control for the endogeneity, thus, we assume the conditional independence of body types and unobserved determinants of family income, conditional on the observed proxies.

We estimate equation (17) by controlling for various subset of the proxy variables. As a comparison, we use measured Height, BMI, and hip-to-waist Ratio in place of $P_1$, $P_2$, and $P_3$, respectively. The estimation results are reported in Table 13. The estimated coefficients of height are similar to those in columns for equation (8) for both genders in Table 8. In particular, height still has positive effects on family income in both genders. On the other hand, the estimated coefficient of BMI for female becomes significant; BMI has negative effects in female subsample.

Table 14 reports estimation results for the equation (17) with $P_1$, $P_2$, and $P_3$. It is worth noting that fitness and car size are not statistically significant but birth state (Northeast) and survey site (Dayton, OH for male; Marlton, NJ for female) are statistically significant. When all proxy variables are included, the estimated coefficient of $P_1$ is 0.056 for males. It is statistically significant at 1% significance level. Thus taller males have a tendency to have higher family income. But we do not find statistically meaningful relationship between the males’ obesity and the family income. For females, the $P_2$ measurement is negatively associated with the family income and its coefficient is statistically significant at 1% significance level. Thus we find that females’ obesity matters for the family income in a negative sense but their stature and hip-to-waist ratio are not associated with the family income. The results are qualitatively similar to those from Table 12.

5.3.2 Control Functions Approach

Even though some unobserved determinants of family income have been controlled for, there would be other possible unobserved factors such as individual ability. As reported in Persico et al. (2004), Case and Paxson (2008), and

---

13Fitness is measured as hours of exercise per week. Car size is classified as two groups: Sedan (compact, economy, intermediate, full size, luxury, sports car) and Non-sedan (SUV, minivan, station wagon, truck, van). Birth state are classified as five groups: Foreign, West, Midwest, South, and Northeast. Survey site includes LA (CA), Detroit (MI), Ames (IA), Dayton (OH), Greensboro (NC), Marlton (NJ), Ottawa (Ontario, CAN), Minneapolis (MN), Houston (TX), Portland (OR), San Francisco (CA), and Atlanta (GA).
Figure 16: Relation between Body Shape Parameters and the Conventional Body Measurements for Female. The straight line displays the linear fit. The R squared is reported in the parentheses.

Lundborg et al. (2014), stature would be highly correlated with individual cognitive and noncognitive abilities. Thus we conjecture $P_1$ or stature would be possibly endogenous regressor. To address this issue, we look for a set of valid instrumental variables (IV). Our identification strategy is to assume that unobserved determinants of Shoe size, Pants size, and Jacket Size for male (or Blouse Size for female) are uncorrelated with individual cognitive and noncognitive abilities. Given the condition, we adopt these unobserved size determinants or variations as valid IVs and incorporate the instrumental variables or control functions (CF) approach into the economic model based on the graph convolution method.

Figure 19 shows a graphical depiction of the causal diagram (see e.g. [Pearl (2000) and Chalak and White (2011)]). Complete circles denote observed variables and dashed circles denote unobserved determinants. Arrows denote direct causal relations. A line without arrows denotes dependence between two variables. The primary parameter of interest is the impact of body type on family income. But body type (stature) is endogenous due to the dependence of body type determinants and cognitive and noncognitive abilities. Shoe size, jacket (blouse) size, and pants size are determined by body type as well as unobserved determinants such as personal size preference. We assume that such unobserved size determinants are uncorrelated with ability. Given the condition, they can serve as legitimate IVs to identify the causal relation. Indeed, it is plausible to assume that these variables are legitimate IVs. First, they are unlikely to be correlated with the unobserved determinants of family income such as cognitive and noncognitive abilities. For instance, it is natural to assume that highly able people do not necessarily wear bigger shoes or pants given their foot length or waist circumference. In fact, it is unlikely that these IVs directly cause the family income. Second, as shown in the estimation results below, they are also strongly correlated with the body type.

Since each size determinants or variations are unobserved, we estimate them from the projection of the observed size on the most relevant body part. In practice, we consider the projection of the reported shoe size, pants size and jacket (or blouse size) on the measured foot length, waist circumference, and chest circumference, respectively. The residuals from each projection are used as the estimated size determinants or variations (analogous to the ‘residuals as instruments’ in [Hausman and Taylor (1983)]).

We consider the following first-step reduced form equation:

$$P_{1i} = \delta X_i + \gamma_1 \text{Shoe Size Determinants}_i + \gamma_2 \text{Jacket Size/Blouse Size Determinants}_i + \gamma_3 \text{Pants Size Determinants}_i + \nu_i,$$

(18)
Figure 17: The Third Body Shape Parameter $P_3$ for Females. The third parameter tends to capture the hip-to-waist ratio of the body shape among the female subsample.

Figure 18: Various Measures and $P_3$ for Females. Estimated coefficients and bootstrapped 90% confidence bands are reported for female. Note that units for all measurements except cup size are converted into centimeter (cm).

where Shoe Size Determinants$_i$ is the estimated individual $i$’s variation or determinants in shoe size and Jacket Size/Blouse Size Determinants$_i$ is the estimated variation in jacket size for male (or blouse size for female), and Pants Size Determinants$_i$ is variation in pants size, and where $X_i$ are a set of exogenous regressors and $\nu_i$ is idiosyncratic shocks to the $P_1i$. By construction, $\nu_i$ is the component which generates the endogeneity. From the reduced-from equations (18), we estimate the control function $\hat{\nu}_i$. In the second-step, we then estimate the income equation by adding the control function as following:

\[
\begin{align*}
\text{Family Income}_i &= \alpha X_i + \beta_1 P_{1i} + \beta_2 P_{2i} + \pi \hat{\nu}_i + \epsilon_i & \text{if male}, \\
\text{Family Income}_i &= \alpha X_i + \beta_1 P_{1i} + \beta_2 P_{2i} + \beta_3 P_{3i} + \pi \hat{\nu}_i + \epsilon_i & \text{if female}.
\end{align*}
\]

(19)

Since $\hat{\nu}$ corrects for the sources of the endogeneity, we can consistently estimate the parameters associated with the physical appearances. Other nice feature of the control functions approach is that we can test whether the physical appearances are endogenous by checking if $\pi = 0$.

Table 16 reports estimation results for the equations (18)-(19). In the columns for the equation (18), all IVs are statistically significant and positively correlated with stature in both genders. Experience is also positively associated
Figure 19: Causal Diagram for the Impact of Body Type on Income. Shoe/Pants/Jacket size determinants are independent of individual ability so that they can serve as valid instrumental variables.

with individual’s stature and the relation is nonlinear. People who were born in Foreign countries or Northeast are likely to be shorter than those born in Midwest. For males, education is negatively correlated with stature. Asian people are less likely to be taller than White people. For females, Hispanic and Asian people are less likely to be taller than White people.

In the columns for the equation (19), the estimated coefficient of $P_1$ is 0.097 for male, which is larger than that from Table [14]. It is statistically significant at 5% significance level. Thus taller males have a tendency to have higher family income. Interestingly, we do not find statistically significant causal relationship between males’ obesity and the family income. We estimate that one standard deviation increase in $P_1$ measurement is associated with $0.097 \times 70,000 = $6,790 increase in the family income for a male who earns $70,000 of median family income. This is equivalent to $\frac{0.097}{1.645} \times 70,000 = $998.5 increase of family income per centimeter. Note that for males one standard deviation in $P_1$ is equivalent to 6.8 centimeter in height and one standard deviation in $P_2$ is equivalent to 4.07 $kg/m^2$ in BMI.

The estimation results for the covariates resemble those in previous tables. As shown in the literature on the returns to education, education has a positive impact on family income. Its estimated coefficient is 0.046 and statistically significant at 1% significance level. Males born in Northeast have tendency to have higher family income than those born in Midwest. Interestingly, the estimated coefficient of $\hat{\nu}_1$ is statistically insignificant. Thus, we find no strong evidence that $P_1$ is a endogenous regressor.

For females, the estimated coefficient of $P_1$ is negative and the estimated coefficient of $P_3$ is positive, but they are not statistically significant. The $P_2$ measurement is negatively associated with the family income. Its coefficient is $-0.069$ and statistically significant at 1% significance level. Thus we find that a female’s obesity negatively matters for her family income, but we do not find a strong causal impacts of her stature and hip-to-waist ratio on the family income. One standard deviation decrease in $P_2$ measurement is associated with $0.069 \times 70,000 = $4,830 increase in the family income for a female who earns $70,000 family income. This can be interpreted to $\frac{0.069}{3.17} \times 70,000 = $934.2 increase per one unit of BMI. Note that for females one standard deviation in $P_1$ is equivalent to 6.8 centimeter in height and one standard deviation in $P_2$ is equivalent to 5.17 $kg/m^2$ in BMI.

For females, experience is important to have higher family income. As commonly reported in the literature on the wage equation, the experience displays a quadratic functional form. Education has a positive impact on the family income and its coefficient is statistically significant at 1% significance level, which is similar to the finding in the male case. Similarly, females born in Northeast have tendency to have higher family income than those born in Midwest. The estimated coefficient of $\hat{\nu}$ is statistically significant. Thus, we find substantial evidence that stature is endogenous in the female’s income equation.

As a comparison, we apply the control functions approach to the income equations where height, BMI, and hip-to-waist ratio are used in place of the extracted body features. The corresponding reduced-form and structural equations are the same as the equations (18)-(19). Table 15 reports estimation results. For male, we estimate that one centimeter increase
Figure 20: Comparison of Conventional Body Measures and Deep-learned Body Parameters (Male). Estimated coefficients and bootstrapped 90% confidence bands are reported. We provide results from conventional body measures (Conventional) and deep-learned body parameters (Deep) with/without control function (CF) approach. Note that the unit for height is converted into centimeter (cm).

in height is associated with $0.01 \times 70,000 = $700 increase in the family income for a male who earns $70,000 of median family income. The estimated effect is smaller than that when the extracted features from the deep learning are used. For females, pants size is not associated with height in the first step regression. The estimated coefficient of $\hat{\nu}$ is significant in female sample, which implies evidence of endogenous female height. For females, interestingly, coefficients of height, BMI, and hip-to-waist ratio are all insignificant. Thus we do not find strong evidence of body shape effects when height, BMI, and hip-to-waist ratio are used. As a result, we observe that the estimation results with the conventional measurements are volatile across different regression models – OLS, proxy variable approach, and control functions approach. On the other hands, those with the deep-learned body parameters are very stable across different models and interestingly captures gender differential in the impact of body types on income.

Finally, we summarize the estimated coefficients in Figure 20 for males and in Figure 21 for females. We compare the results from the conventional body measures to those from the deep-learned body parameters with/without controlling for the endogeneity. They show that the estimated effects from the measured height, BMI, and hip-to-waist ratio are substantially different than those from the deep learned parameters. One can possibly interpret such difference as a limitation of conventional body measures on describing appearances. In fact, it is widely known in literature (CDC document, accessed 2019) that BMI is a surrogate measure of body fatness. Neither does it distinguish fat, muscle, or bone mass, nor does it describe distribution of fat among people. As such, there is a chance where the difference in family income is falsely correlated to stature while the true underlying statistics suggests otherwise. To illustrate this problem, consider a tall and visually obese male and a short and muscular male with the same body mass. In this case, since there is no difference in BMI, the difference in family income must be explained by stature, which may lead to an inaccurate conclusion. This, however, was not the case for the deep-learned parameters.

6 Conclusion

This paper studies the relationship between the physical appearance and family income. We show there are significant reporting errors in the reported height and weight, and show that these discrete measurements are too sparse to provide complete description of the body shape. In fact, these reporting errors are shown to be correlated with individual
backgrounds. We also find that the regression of family income on the self-reported measurements suffers from the issue of reporting errors and delivers biased estimates compared to the regression on the true measurements. The findings shed light on the importance of measuring body types instead of simply relying on subjects’ self-reports for public policies.

We introduce a new methodology built on graphical autoencoder in deep machine learning. From the three dimensional whole-body scan data, we identify two intrinsic features consisting of human body shapes for males and three intrinsic features for females. These body features are presumably less likely to suffer from measurement errors on the physical appearances. We also take into account a possible issue of endogenous body shapes by utilizing proxy variables and control functions approaches. The empirical results document positive impact of stature and negative impact of obesity on family income for males. On the other hand, results for females show that obesity is the only significant feature and it negatively affects family income. The findings support the hypotheses on the physical attractiveness premium and the differential treatment across the gender in the labor market outcomes. Finally, we believe that the proposed method can be applied to many interesting research questions in Economics which deal with geometric data such as graphical, image, spatial, and social networks data.

References

Baek, S. and Lee, K. (2012). Parametric human body shape modeling framework for human-centered product design. Computer-Aided Design, 44:56–67.

Baek, S.-Y. (2013). Nonlinear Analysis of the Space of Human Body Shapes and Its Application to Parametric Human Modeling System. PhD thesis, Seoul National University.

Böckerman, P. and Vainiomäki, J. (2013). Stature and life-time labor market outcomes: Accounting for unobserved differences. Labour Economics, 24:86–96.

Bollinger, C. (1998). Measurement error in the current population survey: A nonparametric look. Journal of Labor Economics, 16:576–594.
Bourlard, H. and Kamp, Y. (1988). Auto-association by multilayer perceptrons and singular value decomposition. Biological Cybernetics, 59:291–294.

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

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

Chalak, K. and White, H. (2011). An extended class of instrumental variables for the estimation of causal effects. Canadian Journal of Economics, 44:1–51.

Chiappori, P.-A., Oreffice, S., and Quintana-Domeque, C. (2012). Fatter attraction: Anthropometric and socioeconomic matching on the marriage market. Journal of Political Economy, 120:659–695.

Courtemanche, C., Pinkston, J., and Stewart, J. (2015). Adjusting body mass for measurement error with invalid validation data. Economics and Human Biology, 19:275–293.

Deaton, A. and Arora, R. (2009). Life at the top: The benefits of height. Economics and Human Biology, 7:133–136.

Freifeld, O. and Black, M. J. (2012). Lie bodies: A manifold representation of 3D human shape. In European Conference on Computer Vision (ECCV), Part I, LNCS 7572, pages 1–14. Springer-Verlag.

Hamermesh, D. and Biddle, J. (1994). Beauty and the labor market. American Economic Review, 84:1174–1194.

Hausman, J. and Taylor, W. (1983). Identification in linear simultaneous equations models with covariance restrictions: An instrumental variables interpretation. Econometrica, 51:1527–1550.

Hu, Y. and Schennach, S. (2008). Instrumental variable treatment of nonclassical measurement error models. Econometrica, 76:195–216.

Lindqvist, E. (2012). Height and leadership. Review of Economics and Statistics, 94:1191–1196.

Lundborg, P., Nystedt, P., and Rooth, D.-O. (2014). Height and earnings: The role of cognitive and noncognitive skills. Journal of Human Resources, 49:141–166.

Mobius, M. and Rosenblat, T. (2006). Why beauty matters. American Economic Review, 96:222–235.

Murillo, A., Affuso, O., Peterson, C., Li, P., Wiener, H., Tekwe, C., and Allison, D. (2019). Illustration of measurement error models for reducing bias in nutrition and obesity research using 2-d body composition data. Obesity, 27:489–495.

Pearl, J. (2000). Causality: Models, Reasoning, and Inference. Cambridge University Press, New York, NY.

Persico, N., Postlewaite, A., and Silverman, D. (2004). The effect of adolescent experience on labor market outcomes: The case of height. Journal of Political Economy, 112:1019–1053.

Pishchulin, L., Wuhrer, S., Helten, T., Theobalt, C., and Schiele, B. (2017). Building statistical shape spaces for 3D human modeling. Pattern Recognition, 67:276–286.

Rooth, D.-O. (2012). Obesity, attractiveness, and differential treatment in hiring. Journal of Human Resources, 44:710–735.

Scholz, J. K. and Sicinski, K. (2015). Facial attractiveness and lifetime earnings: Evidence from a cohort study. Review of Economics and Statistics, 97:14–28.

Song, S. (2015). Semiparametric estimation of models with conditional moment restrictions in the presence of nonclassical measurement errors. Journal of Econometrics, 185:95–109.

Sun, Z., He, Y., Gritsenko, A., Lendasse, A., and Baek, S. (2017). Deep Spectral Descriptors: Learning the point-wise correspondence metric via Siamese deep neural networks. ArXiv e-prints.

Sun, Z., Rooke, E., He, Y., Lu, J., and Baek, S. (2018). Zernet: Convolutional neural networks on arbitrary surfaces via zernike local tangent space estimation.

U.S. Department of Health and Human Services Centers for Disease Control (Accessed January 25, 2019). Body mass index: Considerations for practitioners. https://www.cdc.gov/obesity/downloads/bmiforpractitioners.pdf.

Wada, R. and Tekin, E. (2010). Body composition and wages. Economics and Human Biology, 8:242–254.

Wang, C. C. (2005). Parameterization and parametric design of mannequins. Computer-Aided Design, 37(1):83–98.

Wei, L., Huang, Q., Ceylan, D., Vouga, E., and Li, H. (2016). Dense human body correspondences using convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 1544–1553.

Zuffi, S. and Black, M. J. (2015). The stitched puppet: A graphical model of 3d human shape and pose. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 3537–3546.
## A Appendix — Empirical Results

| Variable                   | Mean  | Median | S.D.  | Min  | Max   |
|----------------------------|-------|--------|-------|------|-------|
| Family Income ($)          | 76,085| 70,000 | 41,470| 7,500| 150,000|
| Reported Height (mm)       | 1,798.2| 1,803.4| 82.5  | 1,498.6| 2,108.2|
| Reported Weight (kg)       | 86.0  | 83.9   | 17.3  | 48.5 | 188.2 |
| Reported BMI (kg/m\(^2\)) | 26.5  | 25.8   | 4.6   | 14.0 | 59.5  |
| Height (mm)                | 1,782.6| 1,778.5| 78.1  | 1,497.0| 2,084.0|
| Weight (kg)                | 86.8  | 83.9   | 17.5  | 45.8 | 181.4 |
| BMI (kg/m\(^2\))          | 27.2  | 26.4   | 4.8   | 17.4 | 55.1  |
| Experience (years)         | 17.5  | 17.0   | 10.2  | 0    | 47.0  |
| Education (years)          | 16.3  | 16.0   | 2.5   | 12.0 | 24.0  |
| # of Children              | 1.3   | 1.0    | 1.4   | 0    | 7.0   |
| Fitness (hours)            | 4.2   | 2.5    | 3.0   | 0.5  | 10.0  |

| Variable                  | # of Samples | Variable          | # of Samples |
|----------------------------|--------------|-------------------|--------------|
| Marital Status (Single)   | 240          | Race (White)      | 644          |
| Marital Status (Married)  | 473          | Race (Hispanic)   | 18           |
| Marital Status (Div./Wid.)| 61           | Race (Black)      | 68           |
| Occupation (White Collar) | 461          | Race (Asian)      | 44           |
| Occupation (Management)   | 144          |                   |              |
| Occupation (Blue Collar)  | 101          |                   |              |
| Occupation (Service)      | 68           |                   |              |
| Birth Region (Foreign)    | 159          |                   |              |
| Birth Region (Midwest)    | 275          |                   |              |
| Birth Region (Northeast)  | 106          |                   |              |
| Birth Region (South)      | 106          |                   |              |
| Birth Region (West)       | 128          |                   |              |
| # of Total Observations   | 774          |                   |              |

**Table 1:** Summary Statistics (Male)
| Variable               | Mean   | Median  | S.D.   | Min   | Max   |
|------------------------|--------|---------|--------|-------|-------|
| Family Income ( $)     | 65,998 | 52,500  | 38,853 | 7,500 | 150,000 |
| Reported Height (mm)   | 1,649.6| 1651.0  | 76.1   | 1,320.8| 1,930.4 |
| Reported Weight (kg)   | 67.9   | 63.5    | 16.9   | 37.2  | 172.3 |
| Reported BMI (kg/m²)   | 24.9   | 23.3    | 5.9    | 12.9  | 57.8 |
| Height (mm)            | 1,642.2| 1,640.0 | 71.3   | 1,382.0| 1,879.0 |
| Weight (kg)            | 68.8   | 64.9    | 17.3   | 39.2  | 156.5 |
| BMI (kg/m²)            | 25.5   | 23.8    | 6.1    | 15.2  | 57.1 |
| Experience (years)     | 18.6   | 19.0    | 10.8   | 0     | 50.0 |
| Education (years)      | 15.8   | 16.0    | 2.1    | 12.0  | 24.0 |
| # of Children          | 1.0    | 0       | 1.2    | 0     | 6.0 |
| Fitness (hours)        | 3.7    | 2.5     | 2.7    | 0.5   | 10.0 |

| Variable               | # of Samples | Variable            | # of Samples |
|------------------------|--------------|---------------------|--------------|
| Marital Status (Single)| 248          | Race (White)        | 644          |
| Marital Status (Married)| 407        | Race (Hispanic)     | 11           |
| Marital Status (Div./Wid.)| 134         | Race (Black)        | 88           |
| Occupation (White Collar)| 607         | Race (Asian)        | 46           |
| Occupation (Management)| 52           |                      |              |
| Occupation (Blue Collar)| 49          |                      |              |
| Occupation (Service)   | 81           |                      |              |
| Birth Region (Foreign) | 105          |                      |              |
| Birth Region (Midwest) | 318          |                      |              |
| Birth Region (Northeast)| 103         |                      |              |
| Birth Region (South)   | 122          |                      |              |
| Birth Region (West)    | 141          |                      |              |

| # of Total Observations| 789          |

**Table 2: Summary Statistics (Female)**
### Table 3: List of Various Body Measures

| Variable (mm)                                                                 | Variable (mm)                  |
|------------------------------------------------------------------------------|--------------------------------|
| Acromial Height, Sitting                                                     | Head Length                    |
| Ankle Circumference                                                          | Hip Breadth, Sitting           |
| Arm Length (Spine to Wrist)                                                  | Hip Circumference, Maximum     |
| Arm Length (Shoulder to Wrist)                                               | Hip Circumference Max Height   |
| Arm Length (Shoulder to Elbow)                                               | Knee Height                    |
| Armscye Circumference (Scye Circumference Over Acromion)                    | Neck Base Circumference        |
| Bizygomatic Breadth                                                          | Shoulder Breadth               |
| Chest Circumference                                                          | Sitting Height                 |
| Bust/Chest Circumference Under Bust                                          | Height                         |
| Buttock-Knee Length                                                          | Subscapular Skinfold           |
| Chest Girth at Scye (Chest Circumference at Scye)                            | Thigh Circumference            |
| Crotch Height                                                                | Thigh Circumference Max Sitting|
| Elbow Height, Sitting                                                        | Thumb Tip Reach                |
| Eye Height, Sitting                                                          | Triceps Skinfold               |
| Face Length                                                                  | Total Crotch Length (Crotch Length) |
| Foot Length                                                                  | Vertical Trunk Circumference   |
| Hand Circumference                                                           | Waist Circumference, Preferred |
| Hand Length                                                                  | Waist Front Length             |
| Head Breadth                                                                 | Waist Height, Preferred        |
| Head Circumference                                                           | Weight (kg)                    |


| Variable              | Male (Eq. (2)) | Female (Eq. (2)) | Male (Eq. (3)) | Female (Eq. (3)) |
|-----------------------|----------------|-----------------|----------------|-----------------|
| Intercept             | 114.481***     | 41.826          | 8.574**        | 4.056*          |
| (36.056)              | (38.067)       | (3.438)         | (2.329)        |                 |
| Height (mm)           | -0.006         | 0.004           | -0.056***      | -0.040***       |
| (0.011)               | (0.019)        | (0.018)         | (0.015)        |                 |
| Weight (kg)           |                |                 |                |                 |
| Family Income         | -5.847***      | -0.105          | -0.477         | -0.112          |
| (2.738)               | (1.810)        | (0.365)         | (0.194)        |                 |
| Age                   | -1.067         | -1.801*         | 0.074          | 0.018           |
| (0.728)               | (0.920)        | (0.111)         | (0.061)        |                 |
| Age²                  | 0.014*         | 0.021**         | -7.6e-4        | -1.0e-6         |
| (0.008)               | (0.011)        | (0.001)         | (7.1e-4)       |                 |
| Occupation            | -1.401         | -2.376          | -0.566         | -0.172          |
| (Management)          | (4.335)        | (6.439)         | (1.105)        | (0.592)         |
| Occupation            | -0.398         | -4.486          | -0.181         | 0.211           |
| (Blue Collar)         | (4.924)        | (6.960)         | (1.115)        | (0.573)         |
| Occupation            | 3.736          | 2.051           | -1.444         | -0.698          |
| (Service)             | (4.938)        | (5.863)         | (1.241)        | (0.549)         |
| Education             | -0.612         | -0.393          | -0.039         | -0.065          |
|                      | (0.378)        | (0.471)         | (0.050)        | (0.043)         |
| Marital Status        | 5.146          | -1.538          | 0.185          | -0.285          |
| (Married)             | (6.544)        | (3.527)         | (0.511)        | (0.535)         |
| Marital Status        | -2.960         | -1.026          | -0.003         | -0.761          |
| (Div./Wid.)           | (5.293)        | (3.790)         | (0.483)        | (0.596)         |
| Fitness               | 0.146          | 0.682           | 0.020          | -0.075**        |
|                      | (0.311)        | (0.423)         | (0.054)        | (0.031)         |
| Race                  | -9.340*        | 6.936           | -0.036         | 0.045           |
| (Hispanic)            | (5.558)        | (6.252)         | (0.975)        | (0.948)         |
| Race                  | -2.383         | 3.050           | -0.066         | -0.610          |
| (Black)               | (5.788)        | (7.302)         | (0.974)        | (0.723)         |
| Race                  | -2.506         | 11.208          | -1.143         | -0.670          |
| (Asian)               | (6.554)        | (7.640)         | (0.916)        | (0.636)         |
| Birth Region          | 1.597          | 1.277           | 0.119          | -0.008          |
| (Foreign)             | (2.469)        | (2.792)         | (0.376)        | (0.279)         |
| Birth Region          | 6.668*         | 0.546           | 0.019          | 0.344           |
| (Northeast)           | (3.728)        | (2.612)         | (0.559)        | (0.234)         |
| Birth Region          | 4.893          | -1.405          | -0.425         | -0.151          |
| (South)               | (4.043)        | (2.807)         | (0.504)        | (0.388)         |
| Birth Region          | 1.562          | -0.617          | -0.438         | -0.009          |
| (West)                | (2.754)        | (2.841)         | (0.628)        | (0.239)         |

| $R^2$                | 0.011          | 0.009           | 0.034          | 0.070           |
| $F$-statistic vs. constant model | 1.47 | 1.42 | 2.50 | 4.33 |
| p-value              | 0.094          | 0.114           | 0.001          | 6.4e-09         |
| N                    | 778            | 793             | 776            | 792             |

Table 4: The Association between Reporting Error in Height/Weight and Personal Background
| Variable                | Income (Eq. (4)) | Income (Eq. (5)) |
|-------------------------|------------------|------------------|
|                         | Male             | Female           | Male             | Female           |
| Intercept               | 9.108***         | 8.747***         | 9.259***         | 8.666***         |
|                         | (0.517)          | (0.510)          | (0.529)          | (0.521)          |
| Reported Height (mm)    | 4.0e-4*          | 5.3e-4**         | 2.3e-4           | 6.5e-4***        |
|                         | (2.2e-4)         | (2.1e-4)         | (2.7e-4)         | (2.2e-4)         |
| Reported Weight (kg)    |                  |                  | 0.002            | -0.002           |
|                         |                  |                  | (0.001)          | (0.001)          |
| Experience              | 0.005            | 0.017***         | 0.004            | 0.018***         |
|                         | (0.006)          | (0.006)          | (0.006)          | (0.006)          |
| Experience²             | 1.0e-4           | -4.1e-4***       | 1.0e-4           | -4.3e-4***       |
|                         | (1.5e-4)         | (1.4e-4)         | (1.5e-4)         | (1.3e-4)         |
| Occupation (Management) | 0.301            | 0.329            | 0.304            | 0.327            |
|                         | (0.253)          | (0.227)          | (0.256)          | (0.234)          |
| Occupation (Blue Collar)| -0.160           | -0.122           | -0.161           | -0.121           |
|                         | (0.275)          | (0.251)          | (0.273)          | (0.247)          |
| Occupation (Service)    | -0.015           | -0.007           | -0.012           | 4.0e-4           |
|                         | (0.244)          | (0.198)          | (0.251)          | (0.204)          |
| Education               | 0.055***         | 0.052***         | 0.056***         | 0.050***         |
|                         | (0.008)          | (0.009)          | (0.008)          | (0.009)          |
| Marital Status (Married)| 0.418            | 0.688            | 0.415            | 0.681            |
|                         | (0.387)          | (0.601)          | (0.385)          | (0.604)          |
| Marital Status (Div./Wid.)| 0.005          | 0.095            | 0.001            | 0.098            |
|                         | (0.269)          | (0.466)          | (0.267)          | (0.439)          |
| Race (Hispanic)         | -0.102           | -0.014           | -0.102           | -0.011           |
|                         | (0.149)          | (0.105)          | (0.153)          | (0.098)          |
| Race (Black)            | -0.174           | -0.141           | -0.178           | -0.126           |
|                         | (0.126)          | (0.112)          | (0.132)          | (0.106)          |
| Race (Asian)            | -0.173           | -0.028           | -0.168           | -0.038           |
|                         | (0.140)          | (0.107)          | (0.142)          | (0.102)          |
| # of Children           | -0.016           | -0.004           | -0.017           | -0.004           |
|                         | (0.017)          | (0.017)          | (0.017)          | (0.015)          |
| $R^2$                   | 0.410            | 0.407            | 0.334            | 0.410            |
| $F$-statistic vs.       |                  |                  |                  |                  |
| constant model          | 31.3             | 43.8             | 29.2             | 40.7             |
| p-value                 | 1.9e-62          | 6.1e-84          | 5.4e-62          | 3.7e-83          |
| N                       | 790              | 801              | 788              | 799              |

Table 5: The Association between Reported Height/Weight and Family Income
| Variable          | Income (Eq. (4)) | Income (Eq. (5)) |
|-------------------|------------------|------------------|
|                   | Male             | Female           | Male             | Female           |
| Intercept         | 8.608***         | 8.573***         | 8.775***         | 8.466***         |
|                   | (0.509)          | (0.535)          | (0.505)          | (0.539)          |
| Height (mm)       | 6.8e-4***        | 6.4e-4***        | 5.0e-4**         | 7.9e-4***        |
|                   | (2.2e-4)         | (2.4e-4)         | (2.5e-4)         | (2.5e-4)         |
| Weight (kg)       | 0.002            | -0.002*          | 0.001            | 0.001            |
|                   | (0.001)          | (0.001)          | (0.001)          | (0.001)          |
| Experience        | 0.005            | 0.016***         | 0.004            | 0.018***         |
|                   | (0.006)          | (0.006)          | (0.006)          | (0.006)          |
| Experience^2      | 1.1e-4           | -3.9e-4***       | 1.2e-4           | -4.1e-4***       |
|                   | (1.5e-4)         | (1.3e-4)         | (1.5e-4)         | (1.3e-4)         |
| Occupation (Management) | 0.302       | 0.328           | 0.305            | 0.325            |
|                   | (0.254)          | (0.231)          | (0.252)          | (0.232)          |
| Occupation (Blue Collar) | -0.152         | -0.123          | -0.155           | -0.120           |
|                   | (0.274)          | (0.252)          | (0.274)          | (0.248)          |
| Occupation (Service) | -0.014        | -0.006          | -0.015           | -0.010           |
|                   | (0.248)          | (0.204)          | (0.248)          | (0.201)          |
| Education         | 0.055***         | 0.051***         | 0.056***         | 0.050***         |
|                   | (0.008)          | (0.009)          | (0.008)          | (0.008)          |
| Marital Status (Married) | 0.418     | 0.688           | 0.413            | 0.682            |
|                   | (0.397)          | (0.598)          | (0.385)          | (0.591)          |
| Marital Status (Div./Wid.) | 0.003     | 0.097           | -0.002           | 0.093            |
|                   | (0.274)          | (0.476)          | (0.270)          | (0.481)          |
| Race (Hispanic)   | -0.090           | -0.004          | -0.089           | 0.003            |
|                   | (0.145)          | (0.105)          | (0.146)          | (0.101)          |
| Race (Black)      | -0.170           | -0.138          | -0.174           | -0.113           |
|                   | (0.127)          | (0.052)          | (0.124)          | (0.101)          |
| Race (Asian)      | -0.150           | -0.013          | -0.145           | -0.022           |
|                   | (0.142)          | (0.112)          | (0.144)          | (0.102)          |
| # of Children     | -0.017           | -0.005          | -0.017           | -0.006           |
|                   | (0.018)          | (0.015)          | (0.017)          | (0.016)          |
| $R^2$             | 0.338            | 0.412           | 0.339            | 0.414            |
| $F$-statistic vs. constant model | 32.0   | 44.2           | 29.9            | 41.4             |
| p-value           | 1.0e-63          | 1.0e-84         | 2.1e-63          | 1.4e-84          |
| N                 | 791              | 802            | 791              | 802              |

**Table 6:** The Association between Height/Weight and Family Income
| Variable                   | Income (Eq. (6))                      | Income (Eq. (7))                      | Income (Eq. (8))                      |
|----------------------------|---------------------------------------|---------------------------------------|---------------------------------------|
|                            | Male                                  | Female                               | Male                                  | Female                               | Male                                  | Female                               |
| Intercept                  | 9.685***                              | 9.747***                              | 9.668***                              | 9.746***                              | 8.965***                              | 8.885***                              |
|                            | (0.332)                               | (0.393)                               | (0.336)                               | (0.389)                               | (0.523)                               | (0.544)                               |
| Reported BMI               | 0.005                                 | -0.005                                | -0.007                                | -0.019***                             | 0.005                                 | -0.004                                |
|                            | (0.004)                               | (0.003)                               | (0.009)                               | (0.007)                               | (0.005)                               | (0.003)                               |
| Reported Height (mm)       |                                       |                                       |                                       |                                       | 3.9e-4*                               | 5.2e-4**                              |
|                            |                                       |                                       |                                       |                                       | (2.1e-4)                              | (2.3e-4)                              |
| Reported Weight (kg)       | 0.004                                 | 0.019***                              | 0.004                                 | 0.018***                              | 0.005                                 | 0.018***                              |
|                            | (0.006)                               | (0.006)                               | (0.007)                               | (0.006)                               | (0.006)                               | (0.005)                               |
| Experience                 | 9.3e-5                                | -4.5e-4***                            | 1.0e-4                                | -4.3e-4***                            | 1.0e-5                                | -4.3e-4***                            |
|                            | (1.5e-4)                              | (1.3e-4)                              | (1.5e-4)                              | (1.3e-4)                              | (1.5e-4)                              | (1.2e-4)                              |
| Occupation (Management)    | 0.300                                 | 0.326                                 | 0.305                                 | 0.328                                 | 0.304                                 | 0.327                                 |
|                            | (0.258)                               | (0.230)                               | (0.255)                               | (0.233)                               | (0.255)                               | (0.232)                               |
| Occupation (Blue Collar)   | -0.170                                | -0.134                                | -0.161                                | -0.121                                | -0.160                                | -0.121                                |
|                            | (0.279)                               | (0.252)                               | (0.281)                               | (0.246)                               | (0.282)                               | (0.245)                               |
| Occupation (Service)       | -0.010                                | -0.001                                | -0.012                                | -7.6e-4                               | -0.012                                | 1.5e-4                                |
|                            | (0.246)                               | (0.203)                               | (0.253)                               | (0.206)                               | (0.253)                               | (0.202)                               |
| Education                  | 0.055***                              | 0.051***                              | 0.056***                              | 0.050***                              | 0.056***                              | 0.050***                              |
|                            | (0.008)                               | (0.009)                               | (0.008)                               | (0.009)                               | (0.008)                               | (0.009)                               |
| Marital Status (Married)   | 0.418                                 | 0.676                                 | 0.415                                 | 0.681                                 | 0.415                                 | 0.681                                 |
|                            | (0.390)                               | (0.581)                               | (0.392)                               | (0.595)                               | (0.379)                               | (0.593)                               |
| Marital Status (Div./Wid.) | 0.002                                 | 0.093                                 | 0.002                                 | 0.100                                 | 0.001                                 | 0.098                                 |
|                            | (0.263)                               | (0.472)                               | (0.262)                               | (0.456)                               | (0.267)                               | (0.462)                               |
| Race (Hispanic)            | -0.126                                | -0.034                                | -0.103                                | -0.013                                | -0.102                                | -0.011                                |
|                            | (0.161)                               | (0.096)                               | (0.152)                               | (0.100)                               | (0.144)                               | (0.101)                               |
| Race (Black)               | -0.184                                | -0.135                                | -0.179                                | -0.129                                | -0.178                                | -0.126                                |
|                            | (0.126)                               | (0.115)                               | (0.130)                               | (0.105)                               | (0.128)                               | (0.101)                               |
| Race (Asian)               | -0.201                                | -0.074                                | -0.171                                | -0.049                                | -0.166                                | -0.039                                |
|                            | (0.147)                               | (0.103)                               | (0.149)                               | (0.096)                               | (0.149)                               | (0.098)                               |
| # of Children              | -0.015                                | -0.005                                | -0.016                                | -0.003                                | -0.016                                | -0.004                                |
|                            | (0.017)                               | (0.015)                               | (0.017)                               | (0.016)                               | (0.017)                               | (0.015)                               |
| $R^2$                      | 0.332                                 | 0.407                                 | 0.334                                 | 0.410                                 | 0.334                                 | 0.410                                 |
| $F$-statistic vs. constant model | 31.1                                | 43.1                                 | 29.2                                 | 40.6                                 | 29.2                                 | 40.7                                 |
| p-value                    | 4.5e-62                               | 7.7e-83                               | 5.7e-62                               | 4.9e-83                               | 5.4e-62                               | 3.7e-83                               |
| N                          | 788                                   | 799                                   | 788                                   | 799                                   | 788                                   | 799                                   |

Table 7: The Association between Reported BMI and Family Income
### Table 8: The Association between BMI and Family Income

| Variable          | Male (Eq. (6)) | Female (Eq. (6)) | Male (Eq. (7)) | Female (Eq. (7)) | Male (Eq. (8)) | Female (Eq. (8)) |
|-------------------|----------------|------------------|----------------|------------------|----------------|------------------|
| Intercept         | 9.678***       | 9.751***         | 9.678***       | 9.768***         | 8.499***       | 8.737***         |
|                   | (0.345)        | (0.382)          | (0.342)        | (0.384)          | (0.511)        | (0.537)          |
| BMI               | 0.006          | -0.005*          | -0.018**       | -0.024***        | 0.005          | -0.005*          |
|                   | (0.004)        | (0.003)          | (0.009)        | (0.007)          | (0.004)        | (0.003)          |
| Height (mm)       | -0.007***      | -0.007***        | -0.007***      | -0.007***        | 6.6e-4***      | 6.3e-4***        |
|                   | (0.002)        | (0.003)          | (0.002)        | (0.003)          | (2.2e-4)       | (2.4e-4)         |
| Weight (kg)       | 0.007***       | 0.007***         | 0.007***       | 0.007***         | 0.007***       | 0.007***         |
|                   | (0.002)        | (0.003)          | (0.002)        | (0.003)          | (0.003)        | (0.003)          |
| Experience        | 0.004          | 0.019***         | 0.004          | 0.018***         | 0.004          | 0.018***         |
|                   | (0.006)        | (0.006)          | (0.006)        | (0.006)          | (0.006)        | (0.006)          |
| Experience²       | 9.6e-4         | -4.4e-4***       | 1.2e-4         | -4.0e-4***       | 1.2e-4         | -4.1e-4***       |
|                   | (1.5e-4)       | (1.3e-4)         | (1.5e-4)       | (1.3e-4)         | (1.5e-4)       | (1.3e-4)         |
| Occupation (Mgmt) | 0.299          | 0.325            | 0.306          | 0.326            | 0.305          | 0.325            |
|                   | (0.259)        | (0.237)          | (0.255)        | (0.231)          | (0.251)        | (0.234)          |
| Occupation (Coll) | -0.172         | -0.134           | -0.154         | -0.123           | -0.154         | -0.120           |
|                   | (0.280)        | (0.251)          | (0.274)        | (0.246)          | (0.267)        | (0.254)          |
| Occupation (Serv) | -0.014         | -0.012           | -0.015         | 0.012            | -0.015         | -0.010           |
|                   | (0.253)        | (0.196)          | (0.244)        | (0.202)          | (0.252)        | (0.209)          |
| Education         | 0.055***       | 0.052***         | 0.056***       | 0.051***         | 0.056***       | 0.050***         |
|                   | (0.008)        | (0.009)          | (0.008)        | (0.009)          | (0.008)        | (0.009)          |
| Marital Status (M) | 0.416          | 0.677            | 0.414          | 0.681            | 0.413          | 0.681            |
|                   | (0.394)        | (0.587)          | (0.386)        | (0.584)          | (0.389)        | (0.593)          |
| Marital Status (D) | 3.5e-4         | 0.088            | -2.2e-4        | 0.094            | -0.002         | 0.093            |
|                   | (0.273)        | (0.478)          | (0.264)        | (0.475)          | (0.268)        | (0.468)          |
| Race (Hispanic)   | -0.125         | -0.029           | -0.088         | 0.004            | -0.090         | 0.004            |
|                   | (0.161)        | (0.101)          | (0.145)        | (0.103)          | (0.141)        | (0.098)          |
| Race (Black)      | -0.185         | -0.127           | -0.175         | -0.115           | -0.174         | -0.113           |
|                   | (0.133)        | (0.105)          | (0.125)        | (0.097)          | (0.125)        | (0.100)          |
| Race (Asian)      | -0.201         | -0.072           | -0.147         | -0.033           | -0.144         | -0.024           |
|                   | (0.147)        | (0.095)          | (0.140)        | (0.101)          | (0.145)        | (0.100)          |
| # of Children     | -0.015         | -0.007           | -0.017         | -0.005           | -0.017         | -0.006           |
|                   | (0.018)        | (0.015)          | (0.017)        | (0.016)          | (0.017)        | (0.016)          |

*R²: 0.332  0.410  0.339  0.414  0.338  0.414

F-statistic vs. constant model: 31.2  43.8  29.9  41.4  29.9  41.4

p-value: 2.1e-62  5.2e-84  1.7e-63  1.5e-84  2.3e-63  1.4e-84

N: 791  802  791  802  791  802
| Variable                              | Male     | Female    |
|--------------------------------------|----------|-----------|
| Intercept                            | 7.173*** | 9.138***  |
|                                      | (1.428)  | (1.339)   |
| Height (mm)                          | 0.001    | 0.002     |
|                                      | (0.002)  | (0.002)   |
| Weight (kg)                          | -0.003   | 0.007     |
|                                      | (0.007)  | (0.007)   |
| Acromial Height (Sitting, mm)        | 0.006*** |           |
|                                      | (0.002)  |           |
| Arm Length (Shoulder-to-Elbow, mm)   | -0.005** |           |
|                                      | (0.002)  |           |
| Chest Circumference                  | 0.002*   |           |
|                                      | (9.5e-4) |           |
| Buttock (Knee Length, mm)            | -0.003** |           |
|                                      | (0.002)  |           |
| Elbow Height (Sitting, mm)           | -0.006***|           |
|                                      | (0.002)  |           |
| Subscapular Skinfold (mm)            | -0.006*  |           |
|                                      | (0.003)  |           |
| Waist Circumference (Preferred, mm)  | -0.001*  |           |
|                                      | (5.0e-4) |           |
| Waist Height (Preferred, mm)         | 0.003**  |           |
|                                      | (0.001)  |           |
| Face Length (mm)                     | -0.006** |           |
|                                      | (0.003)  |           |
| Hand Length (mm)                     | -0.007** |           |
|                                      | (0.003)  |           |
| Neck Base Circumference (mm)         | -0.002*  |           |
|                                      | (0.001)  |           |
| Shoulder Breadth (mm)                | 0.002*   |           |
|                                      | (0.001)  |           |

Table 9: The Association between Various Body Measures and Family Income
| Variable          | Male     | Female    |
|-------------------|----------|-----------|
| Experience        | 0.005    | 0.018***  |
|                   | (0.007)  | (0.006)   |
| Experience\(^2\)  | 1.5e-4   | -3.3e-4** |
|                   | (1.6e-4) | (1.4e-4)  |
| Occupation (Mgmt) | 0.300    | 0.300     |
|                   | (0.249)  | (0.206)   |
| Occupation (Bl Coll) | -0.145  | -0.084    |
|                   | (0.270)  | (0.223)   |
| Occupation (Srv)  | -0.024   | -0.017    |
|                   | (0.240)  | (0.183)   |
| Education         | 0.051*** | 0.052***  |
|                   | (0.009)  | (0.009)   |
| Marital Status (M) | 0.414   | 0.682     |
|                   | (0.387)  | (0.593)   |
| Marital Status (D/W) | 0.009  | 0.091     |
|                   | (0.264)  | (0.475)   |
| Race (Hispanic)   | -0.070   | 0.088     |
|                   | (0.151)  | (0.119)   |
| Race (Bl)         | -0.196   | -0.088    |
|                   | (0.141)  | (0.110)   |
| Race (As)         | -0.092   | 0.010     |
|                   | (0.150)  | (0.127)   |
| # of Children     | -0.019   | -0.013    |
|                   | (0.018)  | (0.017)   |
| \(\bar{R}^2\)    | 0.350    | 0.417     |
| \(F\)-statistic vs. constant model | 9.32 | 11.9 |
| \(p\)-value       | 1.7e-51  | 2.3e-67   |
| \(N\)             | 788      | 792       |

**Table 9:** The Association between Various Body Measures and Family Income (Continued)
| Variable                                      | $P_3$     |
|-----------------------------------------------|-----------|
| Bust/Chest Circumference Under Bust           | -0.003*   |
|                                               | (0.002)   |
| Chest Circumference                           | 0.003**   |
|                                               | (0.001)   |
| Crotch Height                                 | -0.005*   |
|                                               | (0.003)   |
| Cup Size                                      | -0.066*   |
|                                               | (0.036)   |
| Eye Height, Sitting                           | -0.006*   |
|                                               | (0.003)   |
| Face Length                                   | 0.019***  |
|                                               | (0.005)   |
| Hip Breadth, Sitting                          | 0.004*    |
|                                               | (0.002)   |
| Hip Circumference Max Height                  | 0.003***  |
|                                               | (0.001)   |
| Subscapular Skinfold                          | -0.010*   |
|                                               | (0.006)   |
| Triceps Skinfold                              | -0.014**  |
|                                               | (0.006)   |
| Waist Front Length                            | -0.003**  |
|                                               | (0.001)   |

$R^2$ = 0.109

$F$-statistic vs. constant model = 4.49

p-value = 1.2e-18

N = 1205

Table 10: The Association between Female $P_3$ and Various Body Measures
| Variable            | Income (Eq. (15)) | Income (Eq. (16)) | Income (Eq. (17)) |
|---------------------|------------------|------------------|------------------|
|                     | Male             | Female           | Male             | Female           | Male             | Female           |
| Intercept           | 8.608*** (0.492) | 8.573*** (0.534) | 9.680*** (0.347) | 9.751*** (0.386) | 8.499*** (0.513) | 8.490*** (0.625) |
| Height (mm)         | 6.8e-4*** (2.1e-4) | 6.4e-4*** (2.4e-4) | 6.6e-4*** (2.2e-4) | 6.2e-4*** (2.4e-4) | 8.499*** (0.513) | 8.490*** (0.625) |
| BMI                 |                  |                  |                  |                  |                  |                  |
| Male                | 0.006 (0.004)    | -0.005* (0.003)  | 0.005 (0.004)    | -0.004 (0.003)   |                  |                  |
| Female              |                  |                  |                  |                  |                  |                  |
| Hip-to-waist Ratio  |                  |                  |                  |                  | 0.002 (0.002)    |                  |
| Experience          | 0.005 (0.006)    | 0.016*** (0.006) | 0.004 (0.006)    | 0.019*** (0.006) | 0.004 (0.006)    | 0.018*** (0.006) |
| Male                |                  |                  |                  |                  |                  |                  |
| Female              |                  |                  |                  |                  |                  |                  |
| Experience^2        | 1.2e-4 (1.6e-4)  | -3.9e-4*** (1.3e-4) | 9.6e-5 (1.5e-4) | -4.4e-4*** (1.3e-4) | 1.2e-4 (1.5e-4) | -4.0e-4*** (1.3e-4) |
| Male                |                  |                  |                  |                  |                  |                  |
| Female              |                  |                  |                  |                  |                  |                  |
| Occupation (Management) | 0.302 (0.242) | 0.328 (0.234) | 0.299 (0.257) | 0.325 (0.234) | 0.305 (0.257) | 0.320 (0.229) |
| Male                |                  |                  |                  |                  |                  |                  |
| Female              |                  |                  |                  |                  |                  |                  |
| Occupation (Blue Collar) | -0.152 (0.269) | -0.123 (0.251) | -0.172 (0.274) | -0.134 (0.263) | -0.154 (0.269) | -0.121 (0.249) |
| Male                |                  |                  |                  |                  |                  |                  |
| Female              |                  |                  |                  |                  |                  |                  |
| Occupation (Service) | -0.014 (0.247) | -0.006 (0.210) | -0.014 (0.240) | -0.012 (0.210) | -0.015 (0.255) | -0.009 (0.200) |
| Male                |                  |                  |                  |                  |                  |                  |
| Female              |                  |                  |                  |                  |                  |                  |
| Education           | 0.055*** (0.008) | 0.051*** (0.009) | 0.055*** (0.008) | 0.052*** (0.009) | 0.056*** (0.009) | 0.050*** (0.009) |
| Male                |                  |                  |                  |                  |                  |                  |
| Female              |                  |                  |                  |                  |                  |                  |
| Marital Status (Married) | 0.418 (0.395) | 0.688 (0.593) | 0.416 (0.388) | 0.677 (0.593) | 0.413 (0.390) | 0.677 (0.595) |
| Male                |                  |                  |                  |                  |                  |                  |
| Female              |                  |                  |                  |                  |                  |                  |
| Marital Status (Div./Wid.) | 0.003 (0.272) | 0.097 (0.475) | 3.5e-4 (0.272) | 0.088 (0.464) | -0.002 (0.269) | 0.089 (0.472) |
| Male                |                  |                  |                  |                  |                  |                  |
| Female              |                  |                  |                  |                  |                  |                  |
| Race (Hispanic)     | -0.090 (0.137)   | -0.004 (0.106)   | -0.125 (0.165)   | -0.029 (0.100)   | -0.090 (0.143)   | 0.005 (0.102)   |
| Male                |                  |                  |                  |                  |                  |                  |
| Female              |                  |                  |                  |                  |                  |                  |
| Race (Black)        | -0.170 (0.124)   | -0.138 (0.117)   | -0.185 (0.130)   | -0.127 (0.108)   | -0.174 (0.122)   | -0.116 (0.103)  |
| Male                |                  |                  |                  |                  |                  |                  |
| Female              |                  |                  |                  |                  |                  |                  |
| Race (Asian)        | -0.150 (0.142)   | -0.013 (0.113)   | -0.201 (0.156)   | -0.072 (0.099)   | -0.144 (0.144)   | -0.020 (0.105)  |
| Male                |                  |                  |                  |                  |                  |                  |
| Female              |                  |                  |                  |                  |                  |                  |
| # of Children       | -0.017 (0.017)   | -0.005 (0.016)   | -0.015 (0.017)   | -0.007 (0.015)   | -0.017 (0.017)   | -0.005 (0.016)  |
| Male                |                  |                  |                  |                  |                  |                  |
| Female              |                  |                  |                  |                  |                  |                  |
| $R^2$               | 0.338            | 0.412            | 0.332            | 0.410            | 0.338            | 0.413            |
| $F$-statistic vs. constant model | 32.0 | 44.2 | 31.2 | 43.8 | 29.9 | 38.5 |
| p-value             | 1.0e-63          | 1.0e-84          | 2.1e-62          | 5.2e-84          | 2.3e-63          | 2.1e-83          |
| N                   | 791              | 802              | 791              | 802              | 791              | 799              |

Table 11: The Association between BMI/Height/Hip-to-waist-Ratio and Family Income
| Variable     | Income (Eq. (15)) | Income (Eq. (16)) | Income (Eq. (17)) |
|--------------|-------------------|-------------------|-------------------|
|              | Male              | Female            | Male              | Female            | Male              | Female            |
| Intercept    | 9.823***          | 9.629***          | 9.841***          | 9.620***          | 9.823***          | 9.638***          |
|              | (0.309)           | (0.392)           | (0.307)           | (0.392)           | (0.317)           | (0.382)           |
|              | 0.052***          | 0.033*            | 2.0e-4            | -0.056***         | 0.052***          | 0.024             |
|              | (0.020)           | (0.018)           | (0.002)           | (0.017)           | (0.019)           | (0.020)           |
|              | 0.004             | 0.016***          | 0.005             | 0.020***          | 0.004             | 0.019***          |
|              | (0.007)           | (0.006)           | (0.007)           | (0.006)           | (0.006)           | (0.006)           |
|              | 1.1e-4            | -4.0e-4***        | 9.5e-5            | -4.6e-4***        | 1.1e-4            | -4.4e-4***        |
|              | (1.6e-4)          | (1.3e-4)          | (1.6e-4)          | (1.3e-4)          | (1.5e-4)          | (1.3e-4)          |
|              | 0.303             | 0.328             | 0.296             | 0.325             | 0.303             | 0.325             |
| Occupation   | (0.247)           | (0.226)           | (0.248)           | (0.237)           | (0.245)           | (0.221)           |
| (Management) | -0.156            | -0.128            | -0.170            | -0.130            | -0.155            | -0.126            |
|              | (0.274)           | (0.251)           | (0.273)           | (0.257)           | (0.281)           | (0.247)           |
| Occupation   | -0.010            | -0.006            | -0.013            | -0.007            | -0.010            | -0.004            |
| (Blue Collar)| (0.245)           | (0.200)           | (0.244)           | (0.201)           | (0.237)           | (0.199)           |
| Occupation   | 0.055***          | 0.052***          | 0.054***          | 0.050***          | 0.055***          | 0.049***          |
| (Service)    | (0.008)           | (0.009)           | (0.008)           | (0.009)           | (0.008)           | (0.009)           |
| Education    | 0.415             | 0.687             | 0.422             | 0.681             | 0.415             | 0.685             |
| Marital Status| 2.6e-4            | 0.095             | 0.005             | 0.095             | 3.5e-4            | 0.097             |
| (Married)    | (0.265)           | (0.459)           | (0.278)           | (0.482)           | (0.256)           | (0.456)           |
| Race         | -0.014            | -0.008            | -0.126            | -0.016            | -0.104            | -0.001            |
| (Hispanic)   | (0.135)           | (0.101)           | (0.157)           | (0.103)           | (0.137)           | (0.098)           |
| Race         | -0.158            | -0.142            | -0.180            | -0.126            | -0.158            | -0.121            |
| (Black)      | (0.119)           | (0.112)           | (0.126)           | (0.108)           | (0.121)           | (0.101)           |
| Race         | -0.157            | -0.028            | -0.209            | -0.065            | -0.156            | -0.041            |
| (Asian)      | (0.135)           | (0.109)           | (0.149)           | (0.101)           | (0.139)           | (0.101)           |
| # of Children| -0.017            | -0.004            | -0.015            | -0.008            | -0.017            | -0.007            |
|              | (0.017)           | (0.016)           | (0.017)           | (0.015)           | (0.017)           | (0.016)           |
| R²           | 0.337             | 0.410             | 0.330             | 0.415             | 0.336             | 0.415             |
| F-statistic vs. constant model | 31.9 | 43.8 | 31.0 | 44.7 | 29.5 | 38.9 |
| p-value      | 1.6e-63           | 5.4e-84           | 6.9e-62           | 1.9e-85           | 9.4e-63           | 3.0e-84           |
| N            | 791               | 802               | 791               | 802               | 791               | 802               |

Table 12: The Association between Body-type Parameters and Family Income
| Variable                | Male (Eq. (17)) | Female (Eq. (17)) | Male (Eq. (17)) | Female (Eq. (17)) | Male (Eq. (17)) | Female (Eq. (17)) |
|-------------------------|----------------|------------------|----------------|------------------|----------------|------------------|
| Intercept               | 8.482***       | 8.626***         | 8.510***       | 8.626***         | 8.510***       | 8.626***         |
|                         | (0.511)        | (0.585)          | (0.513)        | (0.571)          | (0.513)        | (0.571)          |
| Height (mm)             | 6.5e-4***      | 5.8e-4**         | 7.2e-4***      | 5.8e-4**         | 7.2e-4***      | 6.1e-4**         |
|                         | (2.1e-4)       | (2.4e-4)         | (2.0e-4)       | (2.3e-4)         | (2.0e-4)       | (2.4e-4)         |
| BMI                     | 0.006          | -0.005*          | 0.005          | -0.007**         | 0.003          | -0.007**         |
|                         | (0.004)        | (0.003)          | (0.004)        | (0.003)          | (0.004)        | (0.003)          |
| Hip-to-waist Ratio      | 0.001          | 0.001            | 0.001          | 0.001            | 0.001          | 0.001            |
|                         | (0.002)        | (0.002)          | (0.002)        | (0.002)          | (0.002)        | (0.002)          |
| Experience              | 0.005          | 0.018***         | 0.005          | 0.021***         | 0.008          | 0.023***         |
|                         | (0.007)        | (0.006)          | (0.006)        | (0.006)          | (0.006)        | (0.006)          |
| Experience²             | 9.2e-5         | -4.2e-4***       | 9.8e-5         | -4.7e-4***       | 7.6e-5         | -4.8e-4***       |
|                         | (1.6e-4)       | (1.3e-4)         | (1.5e-4)       | (1.3e-4)         | (1.5e-4)       | (1.3e-4)         |
| Occupation (Management) | 0.298          | 0.313            | 0.313          | 0.305            | 0.236          | 0.243            |
|                         | (0.245)        | (0.225)          | (0.250)        | (0.214)          | (0.198)        | (0.172)          |
| Occupation (Blue Collar)| -0.147         | -0.118           | -0.141         | -0.097           | -0.119         | -0.089           |
|                         | (0.270)        | (0.243)          | (0.272)        | (0.22)           | (0.216)        | (0.191)          |
| Occupation (Service)    | -0.030         | -0.012           | -0.021         | 3.9e-4           | 0.008          | 0.028            |
|                         | (0.234)        | (0.194)          | (0.246)        | (0.183)          | (0.199)        | (0.154)          |
| Education               | 0.054***       | 0.051***         | 0.050***       | 0.049***         | 0.054***       | 0.049***         |
|                         | (0.008)        | (0.009)          | (0.008)        | (0.008)          | (0.008)        | (0.009)          |
| Marital Status (Married)| 0.416          | 0.681            | 0.386          | 0.652            | 0.388          | 0.634            |
|                         | (0.379)        | (0.596)          | (0.362)        | (0.571)          | (0.363)        | (0.564)          |
| Marital Status (Div./Wid.)| -0.003     | 0.097            | 0.004          | 0.068            | -0.014         | 0.041            |
|                         | (0.277)        | (0.465)          | (0.250)        | (0.432)          | (0.256)        | (0.441)          |
| Race (Hispanic)         | -0.089         | 0.003            | -0.110         | 0.021            | -0.135         | -0.020           |
|                         | (0.140)        | (0.109)          | (0.142)        | (0.110)          | (0.140)        | (0.129)          |
| Race (Black)            | -0.153         | -0.129           | -0.172         | -0.119           | -0.174         | -0.175           |
|                         | (0.125)        | (0.108)          | (0.124)        | (0.112)          | (0.129)        | (0.135)          |
| Race (Asian)            | -0.139         | -0.017           | -0.091         | 0.040            | -0.076         | -0.057           |
|                         | (0.138)        | (0.111)          | (0.139)        | (0.118)          | (0.140)        | (0.122)          |
| # of Children           | -0.010         | -0.007           | -0.006         | -0.016           | -0.004         | -0.010           |
|                         | (0.017)        | (0.016)          | (0.016)        | (0.016)          | (0.015)        | (0.017)          |
| Fitness                 | 0.005          | -0.004           | 0.004          | -0.001           | 0.003          | -1.3e-4          |
|                         | (0.006)        | (0.006)          | (0.006)        | (0.006)          | (0.005)        | (0.006)          |
| Car Size (Sedan)        | -0.012         | -0.014           | 0.005          | -0.029           | 0.005          | -0.029           |
|                         | (0.033)        | (0.037)          | (0.034)        | (0.037)          | (0.034)        | (0.037)          |
| Birth State (Foreign)   | 0.042          |                  | 0.042          |                  |                |                  |
|                         | (0.075)        |                  | (0.075)        |                  |                |                  |
| Birth State (Northeast) | 0.126*         | 0.193***         |                |                  |                |                  |
|                         | (0.064)        | (0.057)          |                |                  |                |                  |
| Birth State (South)     | 0.024          |                  | 0.024          |                  |                |                  |
|                         | (0.067)        |                  | (0.067)        |                  |                |                  |
| Birth State (west)      | -0.009         | -0.005           | -0.009         | -0.005           |                |                  |
|                         | (0.066)        | (0.067)          | (0.066)        | (0.067)          |                |                  |
| Survey Site (Detroit, MI)| -0.070        | 0.015            | -0.070         | 0.015            |                |                  |
|                         | (0.175)        | (0.176)          | (0.175)        | (0.176)          |                |                  |

**Table 13:** The Association between BMI/Height/Hip-to-waist-Ratio and Family Income - Proxy Variable Approach
| Survey Site                          | Income (Eq. (17)) Male | Income (Eq. (17)) Female | Income (Eq. (17)) Male | Income (Eq. (17)) Female |
|-------------------------------------|------------------------|--------------------------|------------------------|--------------------------|
| (Ames, IA)                          | -0.250                 | (0.178)                  | -0.258                 | (0.181)                  |
| (Dayton, OH)                        | -0.351**               | (0.178)                  | -0.186                 | (0.180)                  |
| (Greensboro, NC)                    | -0.217                 | (0.177)                  | -0.157                 | (0.179)                  |
| (Marlton, NJ)                       | -0.224                 | (0.168)                  | -0.305                 | (0.186)                  |
| (Ontario, CAN)                      | -0.198                 | (0.172)                  | -0.066                 | (0.180)                  |
| (Minneapolis, MN)                   | -0.224                 | (0.162)                  | -0.163                 | (0.187)                  |
| (Houston, TX)                       | -0.223                 | (0.165)                  | -0.169                 | (0.192)                  |
| (Portland, OR)                      | -0.170                 | (0.168)                  | -0.244                 | (0.190)                  |
| (San Francisco, CA)                 | -0.189                 | (0.164)                  | 0.010                  | (0.174)                  |
| (Atlanta, GA)                       | -0.130                 | (0.172)                  | -0.113                 | (0.173)                  |
| \( \bar{R}^2 \)                     | 0.336                  | 0.414                    | 0.335                  | 0.413                    | 0.349                  | 0.427                  |
| \( F \)-statistic vs. constant model| 27.5                   | 35.9                     | 24.8                   | 32.3                     | 13.9                   | 18.6                   |
| p-value                             | 7.2e-62                | 3.6e-82                  | 9.9e-59                | 1.7e-77                  | 1.2e-54                | 1.5e-73                |
| N                                   | 786                    | 792                      | 759                    | 757                      | 750                    | 754                    |

Table 13: The Association between BMI/Height and Family Income - Proxy Variable Approach (Continued)
| Variable            | Income (Eq. (17)) | Income (Eq. (17)) | Income (Eq. (17)) | Income (Eq. (17)) |
|---------------------|-------------------|-------------------|-------------------|-------------------|
|                     | Male              | Female            | Male              | Female            |
| Intercept           | 9.819***          | 9.638***          | 9.937***          | 9.669***          |
|                     | (0.315)           | (0.403)           | (0.303)           | (0.367)           |
|                     | 9.937***          | 9.669***          | 9.933***          | 9.784***          |
|                     | (0.326)           | (0.390)           | (0.326)           | (0.390)           |
| $P_1$               | 0.054***          | 0.020             | 0.054***          | -0.020            |
|                     | (0.019)           | (0.020)           | (0.019)           | (0.019)           |
| $P_2$               | 0.002             | -0.059***         | -0.007            | -0.070***         |
|                     | (0.018)           | (0.019)           | (0.018)           | (0.018)           |
| $P_3$               | 0.005             | 0.004             | 0.004             | 0.003             |
|                     | (0.020)           | (0.020)           | (0.020)           | (0.018)           |
| Experience          | 0.005             | 0.020***          | 0.005             | 0.023***          |
|                     | (0.006)           | (0.006)           | (0.006)           | (0.006)           |
|                     | 0.008             | 0.024***          | 0.008             | 0.024***          |
|                     | (0.006)           | (0.006)           | (0.006)           | (0.006)           |
| Experience$^2$      | 8.4e-5            | -4.5e-4***        | 9.0e-5            | -5.0e-4***        |
|                     | (1.5e-4)          | (1.3e-4)          | (1.5e-4)          | (1.3e-4)          |
|                     | -5.0e-4***        | (1.5e-4)          | -5.0e-4***        | (1.4e-4)          |
|                     | (1.4e-4)          | (1.4e-4)          | (1.4e-4)          | (1.4e-4)          |
| Occupation          | 0.297             | 0.318             | 0.311             | 0.309             |
| (Management)        | (0.247)           | (0.233)           | (0.255)           | (0.208)           |
|                     | 0.234             | 0.249             | 0.184             | 0.191             |
|                     | (0.199)           | (0.199)           | (0.199)           | (0.198)           |
|                     | 0.234             | 0.249             | 0.184             | 0.191             |
|                     | (0.199)           | (0.199)           | (0.199)           | (0.198)           |
| Education           | 0.053***          | 0.050***          | 0.048***          | 0.048***          |
|                     | (0.008)           | (0.009)           | (0.008)           | (0.008)           |
|                     | 0.053***          | 0.048***          | 0.053***          | 0.048***          |
|                     | (0.008)           | (0.008)           | (0.008)           | (0.008)           |
| Marital Status      | 0.417             | 0.689             | 0.387             | 0.660             |
| (Married)           | (0.392)           | (0.603)           | (0.356)           | (0.574)           |
|                     | 0.388             | 0.388             | 0.388             | 0.642             |
|                     | (0.365)           | (0.365)           | (0.365)           | (0.570)           |
| Marital Status      | -6.3e-4           | 0.106             | 0.007             | 0.077             |
| (Div./Wid.)         | (0.276)           | (0.463)           | (0.255)           | (0.452)           |
|                     | -0.013            | -0.013            | -0.013            | 0.048             |
|                     | (0.448)           | (0.448)           | (0.448)           | (0.448)           |
| Race                | -0.103            | 6.2e-4            | -0.126            | 0.020             |
| (Hispanic)          | (0.132)           | (0.108)           | (0.130)           | (0.106)           |
|                     | -0.151            | -0.151            | -0.151            | -0.151            |
|                     | (0.144)           | (0.144)           | (0.144)           | (0.144)           |
| Race                | -0.134            | -0.139            | -0.158            | -0.126            |
| (Black)             | (0.118)           | (0.114)           | (0.119)           | (0.109)           |
|                     | -0.163            | -0.163            | -0.163            | -0.163            |
|                     | (0.144)           | (0.144)           | (0.144)           | (0.144)           |
| Race                | -0.151            | -0.036            | -0.106            | 0.024             |
| (Asian)             | (0.142)           | (0.108)           | (0.136)           | (0.116)           |
|                     | -0.087            | -0.087            | -0.087            | -0.087            |
|                     | (0.137)           | (0.137)           | (0.137)           | (0.129)           |
| # of Children       | -0.011            | -0.008            | -0.006            | -0.018            |
|                     | (0.016)           | (0.016)           | (0.015)           | (0.016)           |
|                     | -0.005            | -0.005            | -0.005            | -0.010            |
|                     | (0.015)           | (0.015)           | (0.015)           | (0.017)           |
| Fitness             | 0.004             | -0.004            | 0.004             | -0.002            |
|                     | (0.006)           | (0.006)           | (0.006)           | (0.006)           |
|                     | 0.003             | -8.3e-4           | 0.003             | -8.3e-4           |
|                     | (0.006)           | (0.006)           | (0.006)           | (0.006)           |
| Car Size            | -0.017            | -0.014            | 2.2e-4            | -0.029            |
| (Sedan)             | (0.060)           | (0.036)           | (0.035)           | (0.035)           |
|                     | 0.037             | -0.012            | 0.037             | -0.012            |
|                     | (0.073)           | (0.091)           | (0.073)           | (0.091)           |
| Birth State         | 0.125**           | 0.182***          | 0.125**           | 0.182***          |
| (Foreign)           | (0.063)           | (0.055)           | (0.063)           | (0.055)           |
| Birth State         | -0.005            | -0.006            | 0.032             | 0.052             |
| (Northeast)         | (0.069)           | (0.061)           | (0.069)           | (0.061)           |
| Birth State         | -0.005            | -0.006            | 0.032             | 0.052             |
| (South)             | (0.064)           | (0.067)           | (0.064)           | (0.067)           |
| Birth State         | -0.070            | 0.003             | 0.070             | 0.003             |
| (west)              | (0.172)           | (0.164)           | (0.172)           | (0.164)           |

Table 14: The Association between Body-type Parameters and Family Income - Proxy Variable Approach
| Variable                     | Income (Eq. (17)) Male | Income (Eq. (17)) Female | Income (Eq. (17)) Male | Income (Eq. (17)) Female |
|------------------------------|------------------------|--------------------------|------------------------|--------------------------|
| Survey Site (Ames, IA)       | -0.240                 | (0.172)                  | -0.256                 | (0.176)                  |
| Survey Site (Dayton, OH)     | -0.352**               | (0.173)                  | -0.198                 | (0.179)                  |
| Survey Site (Greensboro, NC) | -0.227                 | (0.168)                  | -0.168                 | (0.172)                  |
| Survey Site (Marlton, NJ)    | -0.215                 | (0.168)                  | -0.294*                | (0.175)                  |
| Survey Site (Ontario, CAN)   | -0.193                 | (0.165)                  | -0.068                 | (0.182)                  |
| Survey Site (Minneapolis, MN)| -0.229                 | (0.163)                  | -0.169                 | (0.172)                  |
| Survey Site (Houston, TX)    | -0.225                 | (0.164)                  | -0.158                 | (0.171)                  |
| Survey Site (Portland, OR)   | -0.174                 | (0.171)                  | -0.249                 | (0.184)                  |
| Survey Site (San Francisco, CA)| -0.192              | (0.164)                  | 0.002                  | (0.171)                  |
| Survey Site (Atlanta, GA)    | -0.134                 | (0.168)                  | -0.106                 | (0.169)                  |

| Variable                     | Income (Eq. (17)) Male | Income (Eq. (17)) Female | Income (Eq. (17)) Male | Income (Eq. (17)) Female |
|------------------------------|------------------------|--------------------------|------------------------|--------------------------|
| Survey Site (Ontario, CAN)   | 0.334                  | 0.416                    | 0.332                  | 0.418                    |
| Survey Site (Minneapolis, MN)| -0.229                 | (0.163)                  | -0.169                 | (0.172)                  |
| Survey Site (Houston, TX)    | -0.225                 | (0.164)                  | -0.158                 | (0.171)                  |
| Survey Site (Portland, OR)   | -0.174                 | (0.171)                  | -0.249                 | (0.184)                  |
| Survey Site (San Francisco, CA)| -0.192              | (0.164)                  | 0.002                  | (0.171)                  |
| Survey Site (Atlanta, GA)    | -0.134                 | (0.168)                  | -0.106                 | (0.169)                  |

| Variable                     | Income (Eq. (17)) Male | Income (Eq. (17)) Female | Income (Eq. (17)) Male | Income (Eq. (17)) Female |
|------------------------------|------------------------|--------------------------|------------------------|--------------------------|
| Survey Site (Ontario, CAN)   | 0.334                  | 0.416                    | 0.332                  | 0.418                    |
| Survey Site (Minneapolis, MN)| -0.229                 | (0.163)                  | -0.169                 | (0.172)                  |
| Survey Site (Houston, TX)    | -0.225                 | (0.164)                  | -0.158                 | (0.171)                  |
| Survey Site (Portland, OR)   | -0.174                 | (0.171)                  | -0.249                 | (0.184)                  |
| Survey Site (San Francisco, CA)| -0.192              | (0.164)                  | 0.002                  | (0.171)                  |
| Survey Site (Atlanta, GA)    | -0.134                 | (0.168)                  | -0.106                 | (0.169)                  |

| \( R^2 \)                     |                        |                          |                        |                          |
|------------------------------|------------------------|--------------------------|------------------------|--------------------------|
| \( F \)-statistic vs.        |                        |                          |                        |                          |
| constant model               |                        |                          |                        |                          |
| \( p \)-value                | 2.5e-61                 | 3.3e-83                   | 5.3e-58                 | 4.3e-79                   |
| \( N \)                      | 786                    | 795                      | 759                    | 760                      |

**Table 14:** The Association between Body-type Parameters and Family Income - Proxy Variable Approach (Continued)
| Variable              | Height (1<sup>st</sup>-step Eq. (18)) | Income (2<sup>nd</sup>-step Eq. (19)) |
|-----------------------|----------------------------------------|----------------------------------------|
|                       | Male                                   | Female                                 | Male                                   | Female                                 |
| Intercept             | 1832.318*** (32.212)                   | 1634.629*** (27.136)                   | 7.785*** (1.104)                       | 10.674*** (1.294)                      |
| Height (mm)           | 0.001* (5.9e-4)                        | -6.4e-4 (7.5e-4)                       |                                       |                                       |
| BMI                   | 0.004 (0.004)                           | -0.005 (0.003)                         |                                       |                                       |
| Hip-to-waist Ratio    |                                       |                                       | 0.002                                 | (0.002)                                |
| $\hat{\nu}$          | -6.0e-4 (6.5e-4)                       | 0.001* (8.0e-4)                        |                                       |                                       |
| Shoe Size             | 13.502*** (4.958)                      | 19.847*** (4.187)                      |                                       |                                       |
| Jacket Size (Blouse Size) | 10.132*** (1.309)                  | 7.425*** (1.874)                       |                                       |                                       |
| Pants Size            | 8.372*** (2.384)                       | 1.010 (1.416)                         |                                       |                                       |
| Experience            | 1.463 (0.958)                          | 1.729* (0.895)                        | 0.005                                 | 0.020*** (0.006)                      |
| Experience$^2$        | -0.053** (0.023)                       | -0.066*** (0.020)                     | 5.7e-5 (1.6e-4)                       | -5.0e-4*** (1.4e-4)                   |
| Occupation (Management) | -4.203 (14.216)                   | 5.024 (17.568)                        | 0.315                                 | 0.293                                 |
| Occupation (Blue Collar) | -20.229 (15.689)                  | -23.463 (20.184)                       | -0.129                                | -0.157                                |
| Occupation (Service)  | 2.690 (16.849)                         | 0.906 (15.647)                        | 0.021                                 | -0.024                                |
| Education             | -2.367* (1.353)                        | 1.674 (1.115)                         | 0.047*** (0.008)                      | 0.047*** (0.009)                      |
| Marital Status (Married) | 2.222 (8.797)                      | -6.362 (8.130)                        | 0.372                                 | 0.616                                 |
| Marital Status (Div./Wid.) | 2.327 (11.445)                   | -3.087 (8.850)                        | -0.028                                | 0.038                                 |
| Race (Hispanic)       | -48.276 (38.699)                       | -42.993* (23.101)                     | -0.097                                | -0.071                                |
| Race (Black)          | -32.369 (30.115)                       | -14.088 (26.216)                      | -0.117                                | -0.182                                |
| Race (Asian)          | -67.490*** (26.640)                   | -40.836* (23.785)                     | 0.006                                 | -0.073                                |
| Fitness               | 0.178 (0.987)                          | -0.271 (0.836)                        | 0.005                                 | 8.0e-5                                 |
| Car Size (Sedan)      | -5.187 (5.922)                         | -1.996 (5.696)                        | 0.001                                 | -0.019                                 |
| Birth State (Foreign) | -10.789 (7.837)                       | -24.459*** (7.717)                    | 0.061                                 | -0.009                                 |
| Birth State (Northeast)| -9.752 (7.639)                      | -21.534*** (8.096)                    | 0.138** (0.058)                       | 0.115** (0.057)                       |

Table 15: The Association between BMI/Height/Hip-to-waist-Ratio and Family Income - Control Function Approach
| Variable      | Height (1st-step Eq. 18) | Income (2nd-step Eq. 19) |
|--------------|--------------------------|--------------------------|
|              | Male    | Female | Male    | Female |
| Birth State (South) | 7.755   | -8.738 | -0.033  | 0.023  |
|               | (9.776) | (8.156) | (0.065) | (0.049) |
| Birth State (West)   | 12.373  | 1.445  | 0.010   | -0.021 |
|               | (7.976) | (6.638) | (0.055) | (0.056) |
| \(R^2\)     | 0.228   | 0.203  | 0.327   | 0.404  |
| \(F\)-statistic vs. constant model | 10.7    | 10.1   | 17.0    | 24.1   |
| p-value      | 3.39e-29| 9.4e-28| 5.8e-47 | 1.1e-68|
| N            | 660     | 716    | 660     | 716    |

Table 15: The Association between BMI/Height and Family Income - Control Function Approach (Continued)
| Variable       | Male          | Female        | Male          | Female        |
|----------------|---------------|---------------|---------------|---------------|
| P₁ (1st-step Eq. (18)) | 0.546 (0.410) | -0.181 (0.347) | 9.945*** (0.336) | 9.735*** (0.365) |
| P₂ | 0.097** (0.049) | -0.074 (0.056) |
| P₃ | 6.5e-4 (0.019) | -0.069*** (0.018) |
| \(\hat{\nu}\) | -0.058 (0.055) | 0.101* (0.058) |
| Shoe Size | 0.159** (0.062) | 0.193*** (0.053) |
| Jacket Size | 0.118*** (0.016) | 0.099** (0.025) |
| Pants Size | 0.123*** (0.030) | 0.040** (0.020) |
| Experience | 0.030** (0.010) | 0.033*** (0.012) | 0.005 (0.007) | 0.023*** (0.006) |
| Experience\(^2\) | -6.5e-4** (3.0e-4) | -0.001*** (2.8e-4) | 5.2e-5 (1.6e-4) | -5.7e-4*** (1.5e-4) |
| Occupation | -0.016 (0.153) | 0.062 (0.212) | 0.312 (0.250) | 0.299 (0.232) |
| Occupation (Management) | -0.190 (0.164) | -0.273 (0.232) | -0.132 (0.275) | -0.160 (0.250) |
| Occupation (Blue Collar) | -0.015 (0.170) | -0.003 (0.194) | 0.025 (0.241) | -0.021 (0.196) |
| Occupation (Service) | -0.032** (0.016) | 0.025 (0.016) | 0.046*** (0.008) | 0.047*** (0.009) |
| Education | -0.066 (0.126) | -0.100 (0.117) | 0.369 (0.340) | 0.617 (0.422) |
| Marital Status | 0.109 (0.158) | -0.037 (0.124) | -0.033 (0.244) | 0.042 (0.553) |
| Marital Status (Div./Wid.) | -0.288 (0.525) | -0.825** (0.366) | -0.123 (0.125) | -0.098 (0.129) |
| Race | -0.651 (0.410) | -0.202 (0.366) | -0.091 (0.136) | -0.187 (0.147) |
| Race (Hispanic) | -0.795* (0.421) | -0.567* (0.339) | 0.003 (0.149) | -0.097 (0.121) |
| Race (Black) | -0.004 (0.012) | -0.010 (0.011) | 0.005 (0.006) | -0.002 (0.006) |
| Fitness | -0.007 (0.073) | -0.021 (0.073) | -0.007 (0.035) | -0.015 (0.039) |
| Car Size | -0.200* (0.104) | -0.322*** (0.100) | 0.064 (0.050) | -0.016 (0.060) |
| Birth State | -0.204* (0.110) | -0.312*** (0.108) | 0.144*** (0.055) | 0.107* (0.057) |

Table 16: The Association between Body-type Parameters and Family Income - Control Function Approach
| Variable                  | $P_1$ (1st-step Eq. (18)) | Income (2nd-step Eq. (19)) |
|--------------------------|----------------------------|----------------------------|
|                          | Male          | Female      | Male          | Female      |
| Birth State (South)      | -0.035        | -0.078      | -0.023        | 0.034       |
|                          | (0.112)       | (0.109)     | (0.065)       | (0.049)     |
| Birth State (West)       | 0.109         | -0.020      | 0.012         | -0.021      |
|                          | (0.098)       | (0.096)     | (0.056)       | (0.055)     |
| $R^2$                    | 0.217         | 0.203       | 0.325         | 0.408       |
| $F$-statistic vs. constant model | 10.1       | 10.2        | 16.9          | 24.6        |
| p-value                  | 1.9e-27       | 8.1e-28     | 1.4e-46       | 6.9e-70     |
| N                        | 660           | 718         | 660           | 718         |

Table 16: The Association between Body-type Parameters and Family Income - Control Function Approach (Continued)