DISENTANGLING THE CONTEMPORANEOUS AND DYNAMIC EFFECTS OF
HUMAN AND HEALTH CAPITAL ON WAGES OVER THE LIFE CYCLE

Donna B. Gilleskie
Euna Han
Edward C. Norton

Disentangling the Contemporaneous and Dynamic Effects of Human and Health Capital on
Wages over the Life Cycle
Donna B. Gilleskie, Euna Han, and Edward C. Norton
NBER Working Paper No. 22430
July 2016
JEL No. I12,J12,J13

ABSTRACT

In this study we quantify the life-cycle effects of human and health capital on the wage
distribution of females, with a focus on health measured by body mass. We use NLSY79 data on
women followed annually up to twenty years during the time of their lives when average annual
weight gain is greatest. We allow body mass to explain variation in wages contemporaneously
conditional on observed measures of human capital and productivity histories (namely, education,
employment experience, marital status, and family size) and dynamically over the life cycle
through its impact on the endogenous histories of behaviors that determine wages. We find
significant differences in the contemporaneous effect and the dynamic effect of body mass on
wages, both across females of different races and over the distribution of wages.

Donna B. Gilleskie
Department of Economics
University of North Carolina, Chapel Hill
CB #3305, 6B Gardner Hall
Chapel Hill, NC 27599-3305
donna_gilleskie@unc.edu

Euna Han
Associate Professor
College of Pharmacy
Yonsei University
162-1 Songdo-Dong
Yeosu-Gu
Incheon, South Korea
eunahan@yonsei.ac.kr

Edward C. Norton
Department of Health Management and Policy
Department of Economics
University of Michigan
School of Public Health
1415 Washington Heights, M3108 SPHII
Ann Arbor, MI 48109-2029
and NBER
ecnorton@umich.edu

1   Introduction

It is common for economists to evaluate the contemporaneous effects of policy and behavioral variables on human capital outcomes such as education, employment, and wages. With increasing interest lately, researchers are studying the effect of early life (and in utero) exposures on these later life outcomes (Heckman, et al., 2013). However, what is lacking in the literature is a dynamic analysis of life-cycle impacts of evolving variables on time-varying outcomes. For example, economists may examine the impact of early life health (e.g., birth weight, childhood illness, mother's smoking behavior while pregnant) on later life outcomes (e.g., adult weight, human capital accumulation, mortality) but, generally, are unable to demonstrate the mechanisms through which variation in early life health affects later life outcomes. This omission often stems from lack of data on individuals at multiple times in their lives over a long period. In addition to the fact that the outcome of interest may vary over the life cycle, the explanatory variable itself (e.g., health) may evolve over time as a dynamic process. Hence, as initial health impacts subsequent health, time-varying health may impact the outcome of interest over time. Additionally, causality may run in both directions: for example, health impacts employment and employment impacts subsequent health.[1] Thus, it is important to model the endogeneity of both time-varying variables and to capture the dynamic impacts of each on each other.

In this study we quantify the life-cycle effects of human and health capital on wages of females, with a focus on health measured by body mass. We use data on women followed annually up to twenty years during the time of their lives when average annual weight gain is greatest. We allow body mass to explain variation in wages contemporaneously conditional on observed measures of human capital and productivity histories (namely, education, employment experience, marital status, and family size) and dynamically over the life cycle through its impact on the endogenous histories of behaviors that determine wages.

It has been shown that the body mass of females is negatively correlated with wages (Cawley, 2004). The evidence in much of the literature reflects the contemporaneous relationship between current body mass and current wages. Even in the cases where longitudinal

---

[1]Employment may impact health directly through injury or work-related stress or indirectly through income and consumption of health inputs.

observations on individuals have been available, researchers focus on the impact of current weight status on current earnings, and ignore the evolutionary process of weight gain.[2] While theory suggests avenues through which the history of an individual's body mass may impact wages in the present (e.g., through educational attainment and work experience), the causal relationship between past health (body mass) and accumulation of human capital is rarely included in published analyses of weight on wages.[3] In our effort to understand how wages vary with a determinant of productivity such as body mass that evolves over time, we argue that it is important to consider the effects of this characteristic on other decisions made jointly with employment over the life cycle. Using a long panel of observations on young women as they approach middle age, we capture the influence of an evolving body mass on life-cycle decisions pertaining to schooling, employment, marital status, and family size. The history of these decisions, as well as one's body mass history, define cumulative measures of human capital and other productivity measures that explain variation in observed wages. In turn, these behaviors may also impact the evolution of body mass over the life cycle.

Even after accounting for the effect of an evolving body mass on the productive determinants of wages that accumulate over the life cycle (e.g., educational attainment, work experience, marital status, and family size), variations in contemporaneous body mass may still explain observed variation in contemporaneous wages. A significant contemporaneous effect may reflect wage penalties associated with lower productivity (i.e., through reduced health) conditional on other observed measures of productivity. That is, the variables that traditionally summarize one's human capital may not fully capture productivity. Alternatively, a contemporaneous effect of body mass may reflect employer preference for employees with a particular outward appearance due to concerns about high expected medical care expenditure affecting firm health insurance costs or consumer distaste for obesity. Finally, employers may simply discriminate (either statistically or based on own tastes) against women with higher body mass. In this paper, we establish the existence (or not) of a contemporaneous wage penalty of body mass while also measuring the dynamic impact of body

---

[2]Only a few papers, to our knowledge, consider the effects of current and past body mass on current earnings (Chen, 2012; Pinkston, 2015). Chen includes body mass from a decade prior and Pinkston includes body mass from one year ago. Pinkston's use of lagged wages to explain current wages and twice lagged wages as instruments allows for indirect life-cycle body mass effects.

[3]Han, et al. (2011) find an indirect wage penalty of body weight in one's late teens through its effect on education and occupation outcomes measured in the late thirties.

mass on wages; we do not differentiate between individual productivity, employer preference, or employer discrimination as explanations for the contemporaneous effect. Admittedly, they each may partially explain the current market value (i.e., wage), or marginal productivity, of an individual's time.

Our contribution is the careful, quantifiable distinction between the wage effects of body mass characterized by life-cycle human capital differences (i.e., the indirect or dynamic effect) and those reflecting additional differences in current valuations of market work (i.e., the direct or contemporaneous effect). To disentangle these life-cycle and contemporaneous effects we use a multiple equation, dynamic empirical model to capture the annual employment decisions and wage distribution of women aged 18 to 45 while jointly explaining education, marital, and family size histories, and the evolution of body mass, over time.

Three potential sources of bias compromise measurement of the role of health capital (i.e., body mass) in explaining wages: selection, endogeneity, and measurement error.[4] First, observed wages of employed individuals reflect accepted wage offers, yet we want to measure the labor market's valuation of particular individual wage determinants unconditional on the observed employment decision. Individuals who chose to work in the formal labor market are a self-selected group: only those who receive wage offers above an individual reservation wage are observed to work. Estimation using the self-selected sample will result in biased coefficients on explanatory variables for wages if non-random selection into employment is not modeled jointly with the observed wages.[5] Thus, we model non-, part-time, and full-time employment over the life cycle jointly with wages conditional on employment, allowing the observed employment outcome and wages to be correlated through common unobservables. Second, the human capital and other productivity characteristics that determine pay for work reflect life-cycle decisions of an individual. These wage determinants include educational attainment, work experience, marital status, number of children in household, and

---

[4]These same biases impact evaluation of human capital and other endogenous wage determinants, which we also measure.

[5]In addition to the effects of body mass on the wage distribution, the probability of receiving an employment offer may be impacted by one's weight, thus influencing accumulation of work experience over the life cycle. Much of the literature on weight and wages ignores this employment selection.

health (measured by body mass in this paper). Endogeneity suggests that unobserved characteristics that explain these (time-varying) decisions may also explain differences in wages.[6] Thus, we use a flexible, latent factor approach to model the permanent and time-varying unobserved heterogeneity (UH) that is correlated with wages and their endogenous determinants. Third, the measurement error (i.e., reporting, subjectivity, and rounding errors) that accompanies self-reported individual-level characteristics (e.g., weight and height) can lead to inconsistent estimates.[7] Our latent factor method that accounts for individual-level unobservables addresses potential attenuation bias.

We extend the literature in several dimensions. First, we analyze the effects of body mass on wages using yearly observations on a cohort of women followed for twenty consecutive years (National Longitudinal Survey of Youth (NLSY) 1979 cohort). Our research sample includes annual observations from the ages of 18-26 in 1983 through the ages of 37-45 in 2002. These twenty years of data on the same individuals provide a quite comprehensive picture of life changes from young adulthood to middle age. This point of reference allows for a better understanding of both body mass and wage dynamics and provides a perspective that differs from (repeated) cross-sectional analyses. We are able to quantify the effects of (and determinants of) body mass as an individual ages, and broaden the interpretation from one of aggregate time effects to individual behaviors. Second, we jointly model, along with the evolution of body mass and wages, several individual behaviors that may be affected by body mass and that also explain observed wage variation. Hence, we are able to decompose the correlation between body mass and wages reported in the literature into life-cycle effects captured through human capital variables and contemporaneous (direct) influences. Third, we apply dynamic estimation techniques that go beyond traditional methods to better capture impacts of interest over time and across the impacted distribution and to eliminate

---

[6]For example, a permanent individual trait that may affect both body mass and wages (conditional on body mass) is self-confidence, which is often unobserved in survey data. Although used infrequently in this literature, first-differenced methods and panel data can address endogeneity that may bias the measured weight/wage relationship. Receiving even less attention has been the role of time-varying unobservables that affect both body mass and employment or wages. For example, an unobserved temporary negative health event (e.g., a broken leg) may lead to weight gain as well as a reduction in observed wages if the researcher cannot fully account for accumulation (or depreciation) of human capital (e.g., work absences).

[7]There is evidence in the literature that women understate their weight (Lakdawalla and Philipson, 2002; Chen, 2012). Some authors in this literature have re-scaled self-reported height and weight following Cawley (2004); others have used different measures of adiposity (Burkhauser and Cawley, 2008).

bias in both the observed and unobserved dimensions. More specifically, we make few distributional assumptions about the jointly-correlated unobservables; rather, we discretize the distributions and estimate the event-specific impacts, mass points, and weights of these unobserved factors using a flexible, discrete factor random effects method. We also seek to understand the impact of body mass on the full support of wages, not just the first moment. That is, we allow body mass (and other variables) to influence wages differently at different points of support of the wage distribution by modeling the density of wages without imposing a distributional assumption. Similarly, we allow determinants of body mass to have different effects at different points of the support of the body mass distribution. We estimate the conditional densities for wages and body mass jointly with endogenous behaviors that determine both wages and body mass, allowing permanent and time-varying observed and unobserved heterogeneity to affect all outcomes.

Using our research sample consisting of twenty years of annual observations on the same individuals, we have replicated, with similar specifications and estimation techniques, the results in the literature. The biased findings suggest that current body mass, measured by the conventional body mass index (BMI), has a statistically significant negative effect on the current wages of white women. We find a smaller, but statistically significant effect on the wages of black women in the most sparse specifications. When we take advantage of the longitudinal nature of the data, and estimate a model with fixed individual effects, the upwardly-biased significant effect of BMI on wages of white women falls, but is still significant. The effect for black women in a model that allows for permanent individual unobserved heterogeneity actually reverses signs. We move beyond the standard approaches in the literature to jointly estimate 17 equations that capture the dynamic life-cycle behavior described above and that allow for selection into employment, endogeneity of several determinants of wages, and measurement error. Here, calculation of the marginal effect of BMI on wages is not so straightforward because of its contemporaneous and life-cycle effects captured by the dynamic model. We quantify those marginal effects through simulation using the estimated model. We find a different effect of BMI at different levels of wages, which is captured by the conditional density estimation procedure. More specifically, the impact of a contemporaneous improvement in body mass on wages of white women is positive and increases over the

support of the distribution of wages. That is, the largest impacts occur at the highest deciles of wages. For black women, the contemporaneous weight-related wage penalty is smaller and is actually negative at lower wages, yet approaches that of white women at higher wages. When we account for the impact of better health (i.e., a lower body mass) over the life cycle on other determinants of wages, we find that the total effect on wages is slightly smaller than that implied by the contemporaneous effect. The life-cycle impact of body mass (which includes the contemporaneous effect) is significantly different from zero for all women over the support of wages.

2   Review of the Relevant Literature

The economics literature examining wage determinants and wage disparities is large. Rather than discuss this literature, we focus on research that explores the impact of health on wages, with health specifically measured by body mass. We also briefly discuss the relationship between body mass and other determinants of employment and wages (Cawley and Ruhm, 2012).

2.1   Measured Associations between Body Mass and Wages

An association between obesity and labor market outcomes, particularly wages, has been documented in the economic literature. The majority of these studies estimate a reduced-form static model of the relationship between body mass and wages. A few exceptions explore the underlying mechanisms behind the relationship (Harper, 2000; Baum and Ford, 2004; Bhattacharya and Bundorf, 2004; Carpenter, 2006; Han, Norton, and Stearns, 2009; Han, Norton, and Powell, 2009). Harper (2000) shows a positive effect of being physically attractive on the probability of working in managerial or professional specialty and clerical occupations for women. However, physical attractiveness is not found to be associated with sorting into customer-oriented occupations and also no wage differences are found between non-attractive and attractive women in customer-oriented jobs in the same study. Baum and Ford (2004) examine four potential pathways linking obesity to labor market outcomes: less productivity due to health problems from obesity (measured by health limitations), less

investment in human capital by obese workers (measured by work experience), employer distaste for obese employees due to high health care cost (measured by employer provision of health insurance), and consumer distaste for obese workers (measured by classification of occupation as customer-related). Bhattacharya and Bundorf (2004) report no statistically significant wage differential between obese and non-obese individuals by employer provided health insurance coverage. Carpenter (2006) shows that the employment rate increased four percent for obese women and two percent for obese men compared to their respective non-obese counterparts after a 1993 court case (Cook vs. Rhode Island) in which a federal appeals court ruled that obesity is covered under the Rehabilitation Act of 1973 and the Americans with Disabilities Act. Han, Norton, and Stearns (2009) report larger negative relationships between adult contemporaneous body mass and wages in occupations requiring interpersonal skills. Han, Norton, and Powell (2009) examine the direct effect of body mass on wages and the indirect effects operating through education and occupation choice. They find that body weight in the upper tail of the distribution significantly affects educational and occupational outcomes of men and women, which in turn have significant effects on wages. The 11 percent adult wage penalty of obesity during a woman's late teen years can be partly attributed to the effect of body mass on human capital accumulation. While almost all empirical work on this subject estimates ordinary least squares models with or without attempts to address endogeneity of body mass, Johar and Katayama (2012) use quantile regression to assess the role of body mass at different wage levels. They find that the body mass wage penalty is larger at higher wages for white women and at median wages for white males.

Regardless of the estimation techniques in the previous economic studies, most studies report gender differences in the association between obesity and wages. Women consistently show a statistically significant negative association of BMI or obesity with hourly wages (Averett and Korenman 1996; Cawley 2004; Baum and Ford 2004; Conley and Glauber 2005; Han, Norton, and Stearns 2009). BMI is also associated with women's total household income by affecting their spouses' earnings and occupational prestige (Conley and Glauber, 2005). Some studies also report racial differences in the association between body mass and labor market outcomes for women. For example, Cawley (2004) reports that only white women's hourly wages are causally associated with their body mass.

2.2 Endogenous Body Mass

Only a few of the studies attempt to address the potential endogeneity of body mass, typically measured by BMI or an obesity indicator (Averett and Korenman, 1996; Pagan and Davila, 1997; Behrman and Rosenzweig, 2001; Cawley, 2004; Baum and Ford, 2004; Conley and Glauber, 2005; Norton and Han, 2008; Johar and Katayama, 2012; Sabia and Rees, 2012; and Pinkston, 2015). Instrumental variable estimation techniques require a variable that is correlated with body mass in the current period but uncorrelated with wages in the current period. Instrumental variables (IVs) used in this literature include a lag of respondents' own body weight or BMI (Averett and Korenman, 1996; Conley and Glauber, 2005), or siblings' or children's BMI (Cawley, 2000; Cawley, 2004; Johar and Karayama, 2012). The lag of current BMI is not a valid instrument if there is any serial or inter-temporal correlation in the wage residuals. Likewise, children's body weight is not a valid instrument if unobserved heterogeneity in the wage residual is associated with both children's body weight and the mother's employment behavior. Also, siblings' BMI is not a valid instrument if siblings share unobserved endowment factors that influence earnings. Norton and Han (2008) use genetic variation among young adults to identify their model (using multiple instruments). They find no statistically significant effect of adolescents' BMI on wages in young adulthood. There are potential limitations of using own genetic information as IVs for own behavioral outcomes given that genes that predict body weight outcomes are related to brain chemistry and, thus, are likely to affect labor market outcomes such as wages via various channels. Pinkston (2015) estimates the impacts of current and lagged BMI on wages using a differenced GMM estimator which relies on additional lags of body mass as identifying instruments.

2.3 Body Mass and Other Outcomes

The growing interest among economists in the dynamic effects of body mass on many socio-economic outcomes (e.g, employment, schooling, marriage) emphasizes the importance of understanding the effects of body mass on the accumulated determinants of wages (i.e., human capital and productivity) over the life cycle. Some studies assess the association of obesity with labor market outcomes at the extensive margin by measuring its effect on the

probability of employment or the different effect by different job sectors. Cawley (2000) finds no statistically significant effect of obesity on the amount of paid work and limitations on types of paid work. Morris (2007) estimates a negative relationship between obesity and the probability of employment for British people for both genders. Paraponaris and colleagues (2005) report that a one standard deviation increase of BMI from the mean at age 20 raises the percentage of time spent unemployed during the working years and lowers the probability of employment after a period of unemployment for both genders. Pagan and Davila's (1997) study reports that both obese men and women are less likely to sort into managerial, professional and technical occupations among fourteen Census occupation categories.

There are also a few papers estimating the effect of obesity on educational achievement. Some studies report no effect of obesity on grade point average among adolescents (Crosnoe and Muller, 2004), scores on Peabody Individual Achievement Tests among pre-teen children (Kaestner and Grossman, 2009), and grade progression and drop out status (Kaestner, Grossman, and Yarnoff, 2009) among adolescents. In contrast, Sabia (2007) reports that white female adolescents have a grade point average penalty for being obese, whereas non-white female and male adolescents do not.[8]

The role of body mass on marriage behavior has received much attention from economists. Using a model of the marriage market, Chiappori, et al. (2012) estimate the trade-offs (in terms of income and education) that men and women are willing to accept to compensate for body mass penalties. Malcolm and Kaya (2014) argue that an observed negative correlation between BMI and marriage is not the only consequence of a lower desirability for individuals with higher BMI, which reduces their competitiveness in the marriage market. They demonstrate that individuals with higher BMI who marry do so at a younger age. Averett, et al. (2013) show that the direction of causality also flows in the reverse direction; that is, marriage impacts the health of individuals. Lin et al.(2014) show that sex ratios are correlated with body mass.

---

[8]Sabia's work provides another example of attempts to account for endogneous BMI. He uses parent-reported self-classification of own obesity as an IV for adolescent BMI, which is not likely to be a valid IV if the parents' self-classification reflects the level of their self-esteem or time preference which may also influence their children's educational achievement.

2.4 Measurement of Body Mass

The body mass index is a simple means for classifying sedentary individuals by weight: underweight, ideal weight, overweight, and obese.[9] The original proponents of this index stressed its advantageous use in making population comparisons. They warned against using it for individual health diagnosis. BMI, as a function of weight and height, does not allow one to distinguish between fat mass and fat-free mass (such as muscle and bone). It may overestimate adiposity in those with more lean body mass (e.g., athletes), while underestimating adiposity on those with less lean body mass (e.g., the elderly).

Different methods of assessing overweight and obesity may allow researchers to capture adiposity (or body weight) as it relates to productivity versus appearance. For example, fat mass can be measured by percentage of body fat (using skinfold, underwater weighing, or fat-free mass techniques) or by measures that account for mass and volume location (waist and arm circumference and the body volume index). Johansson, et al. (2009) and Wada and Tekin (2010) show that the measure of adiposity affects analysis of the obesity wage penalty. Burkhauser and Cawley (2008) find that obesity measured by BMI is only weakly correlated with obesity defined by other measures of fatness. Unfortunately, however, many data sets that contain information on employment and wages contain only measures of height and weight that can be used to construct the body mass index. Another concern is that the data on height and weight are often self-reported. These self-reports are likely to suffer from reporting error, subjectivity error, and rounding error. Because it is necessary for us to rely on the traditional BMI index to measure adiposity, we use an estimation technique that addresses potential measurement errors.

3   Description of the Research Sample

The National Longitudinal Survey of Youth (NLSY) provides data on a 1979 cohort of 12,686 individuals aged 14 to 22 followed annually through 1994. These individuals continue to be followed every two years after 1996 (with data available currently through 2012). Our

---

[9]Specifially, BMI equals weight in kilograms divided by height in meters squared. A BMI less than or equal to 18.5 indicates underweight; between 18.5 and 25, normal weight; greater than or equal to 25 and less than 30, overweight; and 30 or greater, obese.

research sample includes those who have not attrited by 1983 and who have responses for important variables in our analysis. Because the literature on the relationship between body mass and wages has detected a wage effect that differs between black and white females, we restrict our analysis in this paper to women of these races. We drop sample-eligible individuals who report an Asian race or Hispanic ethnicity. From 1983, we follow 3,213 females through 2002 or until they attrit from the NLSY survey. Table 1 displays the research sample size by year and the percent of individuals who attrit each year.[10] The research sample contains 51,884 person-year observations.

Table 1
Empirical Distribution of Research Sample

Year	Sample Size	Attriters	Attrition Rate
1983	3,213	-	-
1984	3,213	67	2.09
1985	3,146	81	2.57
1986	2,065	101	3.30
1987	2,964	97	3.27
1988	2,867	52	1.81
1989	2,815	57	2.02
1990	2,758	46	1.67
1991	2,712	60	2.21
1992	2,652	35	1.32
1993	2,617	39	1.49
1994	2,578	134	5.20
1995	2,444	78	3.19
1996	2,366	102	4.31
1997	2,264	70	3.09
1998	2,194	58	2.64
1999	2,136	114	5.34
2000	2,022	42	2.08
2001	1,980	102	5.15
2002	1,878	-	-

Person-year observations: 51,884

---

[10]We restrict the initial 1983 sample to include individuals who are observed for at least two consecutive periods. Hence, attrition does not occur between 1983 and 1984. In fact, the 1983 data serve as initial conditions for the subsequent period of observation. Additionally, we have no need to model attrition at the end of 2002 since this is the last year of data that we use.

The NLSY data provide a unique understanding of individual body mass dynamics over time because the survey follows the same individuals for such a long time (e.g., over 20 years in some cases) relative to most data sets.[11] Since the evolution of body mass as individuals age is important in this study, we depict the trends in body mass over time in Figure 1. Average body mass increases by age for women of both races, with the mean increasing from 21.5 and 22.7 at age 20 to 25.9 and 29.8 at age 40, for white and black females respectively. More importantly, the value of the 75th percentile of the BMI distribution is increasing and the right tail of the distribution is larger as individuals age. .[12] Across all ages, one-year body mass deviations (not pictured) average 0.21 and 0.32 BMI units for white and black females respectively. For an average height female (i.e., 5' 4"), this amounts to an increase of 1.2 to 1.9 pounds per year. Despite the fact that much of the interquartile range of BMI deviations from one year to another lies above zero at each age, the median BMI deviation is zero.

Since we intend to explore how wages vary with body mass, we emphasize that body mass increases with age, on average, in the 18 to 45 year age range. It is also the case that the strongest predictors of wages — education and work experience — also increase with age. In general, while additional education may be obtained at any age, it is often completed in the early twenties. Work experience, however, accumulates over the life cycle. If individuals gain weight as they age, but are also gaining work experience as they age, a depiction of wages by age will almost surely reveal that those with larger body mass also have higher wages. Hence, it is important to control for human capital variables (i.e., education and work experience) that increase with age, as well as other cumulative variables that influence productivity (i.e., marital status and tenure, number of children in the household, and body mass). Because these important wage determinants may be influenced by body mass and,

---

[11]Weight and height are observed in the NLSY79 data in most years with the exception of 1983, 1984, 1991, and the non-survey years of 1995, 1997, 1999, and 2001. We linearly interpolate or extrapolate values of weight in years that we do not observe weight. We assume that height is the same across all years an individual is observed. Using these height and weight values, we construct the body mass index for each year an individual is in our research sample.

[12]The horizontal lines in Figure 1 indicate upper thresholds for underweight (18.5), normal (25), and overweight (30) with a BMI greater than 30 indicating obese. The dark shaded regions indicate the interquartile range. The light shaded regions extend up to 1.5 times the interquartile range. Points on either side of the shaded areas represent remaining outliers.
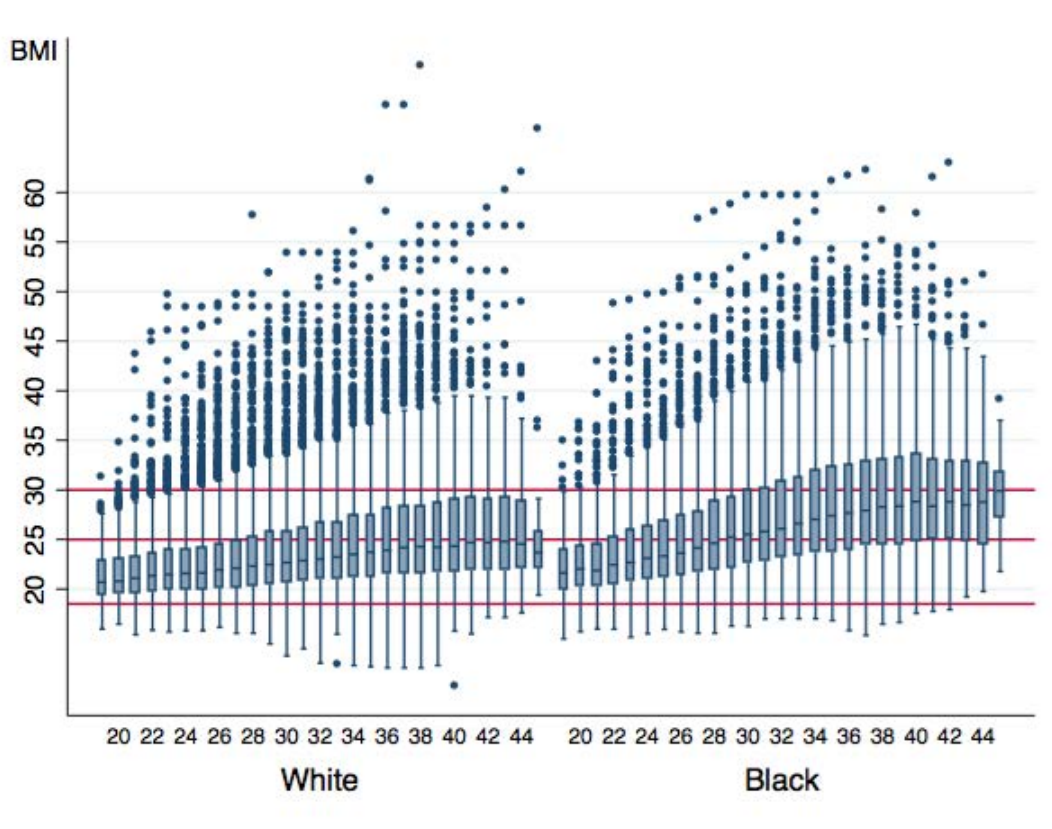
Figure 1: Distribution of Body Mass as Females Age, by Race

inversely, influence body mass, we model their dynamic evolutions jointly while allowing for common observed and unobserved heterogeneity over the life cycle.

Educational attainment and work experience are determined by schooling and work behaviors at each age. Figure 2 displays these probabilities by age and by predominant body mass status.[13] We model the probability of part-time employment or non-employment relative to full-time employment, and depict full-time employment probabilities by race and weight status in this figure.
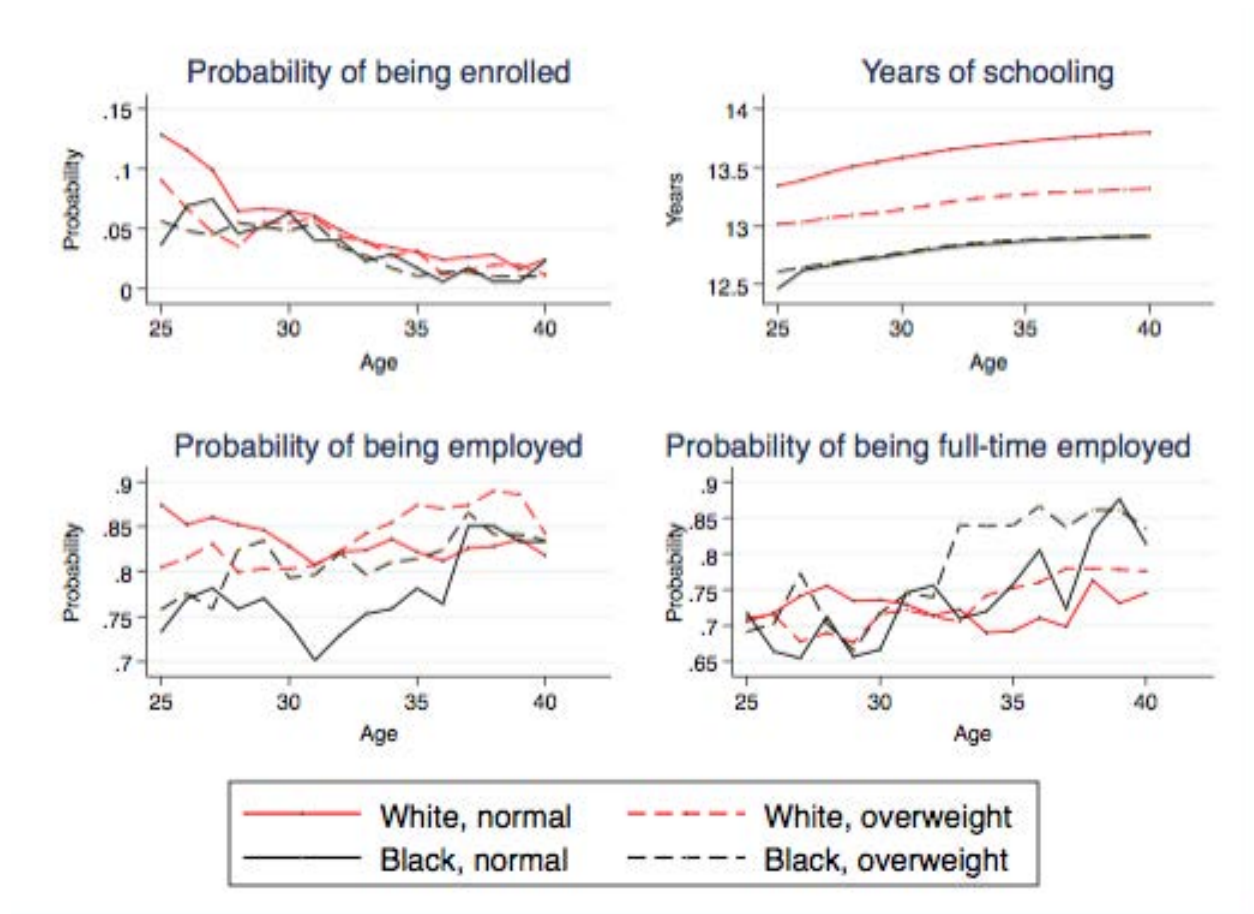
Figure 2: Schooling and Employment Statistics by Age, Race, and Body Mass

Marital status and fertility may change as one ages. These annual behaviors define accumulated stocks (namely, marital tenure and number of children in the household) that

---

[13]These graphs are based on those individuals in the research sample who are observed every year from age 25 to age 40. The normal and overweight distinction reflects the body mass status of an individual in more than half of those 16 years.

may shift preferences for schooling and employment. They may also impact productivity at work conditional on employment, which may be reflected in wages. Rather than focus only on births as a mechanism by which household size may change, we broaden the process to include increases through birth, adoption, and blended families and reductions through death, marriage dissolution, removal, and matriculation. In addition to the effects children in the household may have on employment behavior and productivity, children may impact body mass of the parents. Most obviously, childbearing imposes physical changes in a woman's body mass that may not be temporary. Additionally, children impose both time and financial requirements that may alter caloric intake and expenditure of their caregivers. Figure 3 describes marriage and child accumulation probabilities for females by predominant body mass status.
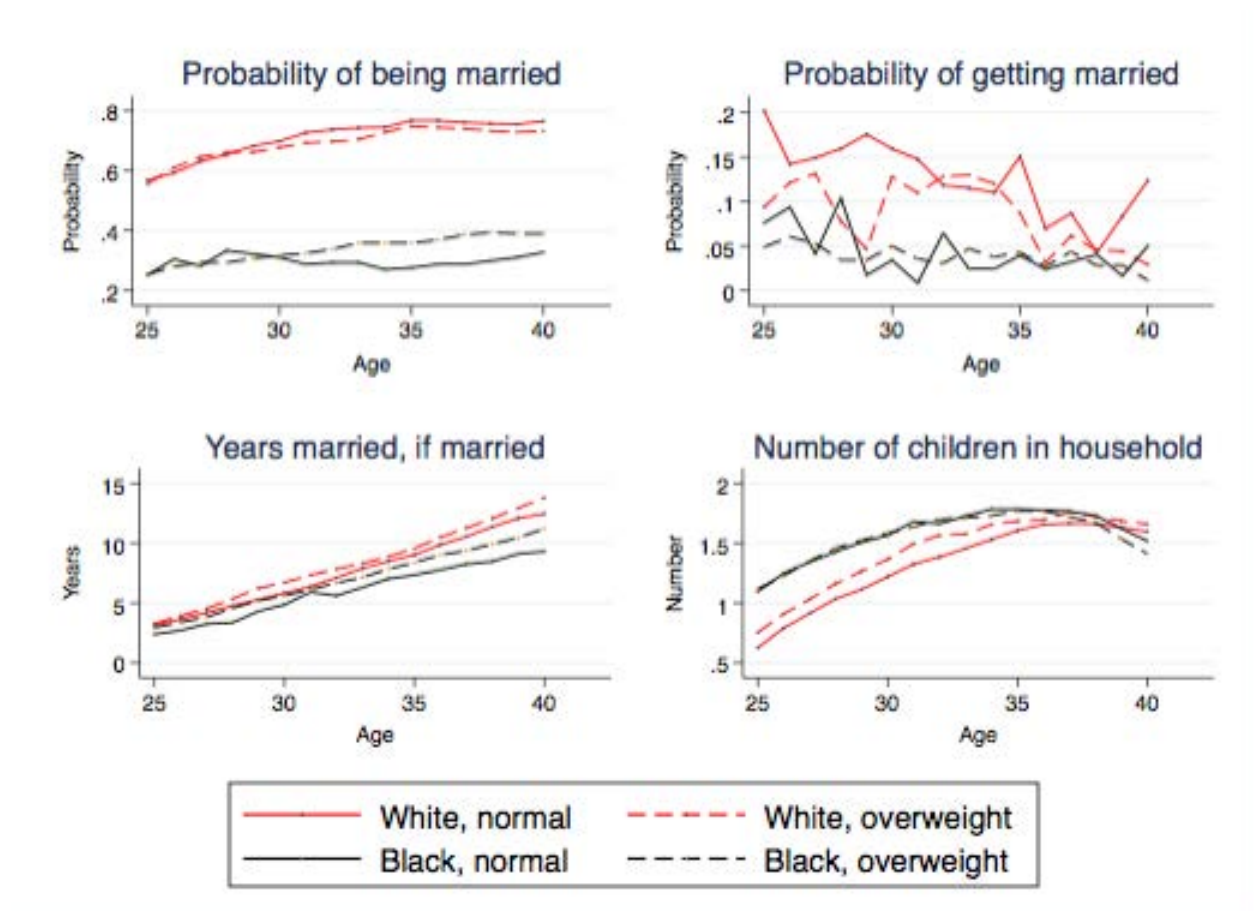
Figure 3: Marriage and Children Statistics by Age, Race, and Body Mass

4 Empirical Framework

Our empirical framework is motivated by a theoretical model of life-cycle decisionmaking and health production: per-period decisions regarding education, employment, marriage, and children with uncertain future wages and body mass transitions. The theory allows us to derive choice probabilities that we approximate using dynamic, non-linear functions of the information available (to an individual) at the beginning of each decisionmaking period. The theory also describes how these period-by-period decisions affect the production of health (or evolution of body mass) over time and the implied distribution of wages. We then estimate the approximated demand and production equations jointly, and allow for both permanent and time-varying individual unobserved heterogeneity (UH) through the use of discrete factor random effects (Heckman and Singer, 1983 and Mroz, 1999). Rather than estimate the conditional mean of wages and body mass, we estimate the conditional density of these observed outcomes using a conditional density estimation (CDE) technique (Gilleskie and Mroz, 2004).

This multi-faceted empirical approach enables us to consider aspects of the problem that may not be computationally-feasible with full-solution and estimation of the dynamic optimization problem. We 1.) include several endogenous explanatory variables that provide a detailed description of the histories of individual decisions over the life cycle, 2.) consider the continuous evolution of body mass rather than discrete categories, and 3.) explore the possibility that variables of interest may have different marginal effects at different points of support of the wage and body mass distributions. We seek to evaluate the causal pathways of body mass on wages of women (either directly or indirectly through other life-cycle determinants) and to distinguish the contemporaneous effect from that generated by life-cycle behaviors influenced by body mass.

4.1 Timing and Notation

In each year t after age 18, an individual obtains a wage offer drawn from the population distribution of wages. The individual also observes her spouse's wage if married.[14] The wage a female would receive if she worked, w_t, depends on education and employment experience, as described by Mincer (1974). We argue that these human capital variables may not fully capture a person's productivity at work. Her productivity, or effort, might be influenced by her health or her time demands. Thus, we include body mass (Cawley, 2004), marital status (Korenman and Neumark, 1991), and number of children in the household entering period t as determinants of the wage distribution.[15]

Each period the individual then jointly decides 1.) whether or not to attend school, 2.) whether to be non-employed, part-time employed, or full-time employed, 3.) whether or not to be married, and 4.) whether or not to acquire (or lose) a child (or children) living in the household. We let indicator variables (lower case) define the alternatives. That is, $s_t = 1$ if the individual attends school in period t and $s_t = 0$, otherwise. The employment indicator takes on the value $e_t = 0$ if the individual does not work in period t, $e_t = 1$ if he works part time (less than 1375 hours per year), and $e_t = 2$ if he works full time (1375 hours or more per year). We indicate the marriage decision by $m_t = 1$ if the individual is married in year t, and $m_t = 0$, otherwise. The number of children in the household may increase or decrease due to pregnancy (singleton or multiples), marriage, divorce, age of child, or child mortality.[16] The variable $k_t = 0$ indicates that there is no change in the number of children

---

[14]Spouse income is treated as exogenous. While we model the marriage decision, which may be influenced by body mass, we do not model the correlation between an individual's body mass and her potential spouse's income (Chiappori, et al., 2012).

[15]We recognize that body mass is an imperfect measure of health, and may not fully capture the role that health plays in one's productivity. Instead, or additionally, it may signal expected medical care expenditures (which may be compensated by the employer and hence affect offered wages) or reflect an employer's tastes based on appearance. We cannot distinguish these roles in estimation.

[16]Some reasons for a change in number of children in the household are available in the data. For example, the survey records whether a child was born, died, was adopted, etc. Additionally, a comparison of the number of children in the household from year to year reveals additions and losses of more than one child in a non-trivial number of cases. The changes could be due to birth by the sample female as well as children acquired (or lost) through other channels (e.g., adoption, marriage, divorce, aging out, and death). Because of the large number of ways household size can change, we are motivated to focus only on the change in number of children in the household rather than the channel. A larger household size, regardless of channel, places demands on time, finances, and energy, which may impact per-period decisionmaking.

in the household in year t, and values of $k_t = 1$ and $k_t = -1$ indicate that at least one child is acquired or lost in year t.

These yearly decisions (i.e., school attendance, employment, marriage, and children) produce stock variables upon entering period t that summarize the history of the decisions. The vectors of these history variables (upper case) are denoted: educational history (S_t), work experience (E_t), marital history (M_t), and household child accumulation (K_t). These vectors include both duration values as well as indicators of behavior or outcomes in the last period $(t-1)$ and polynomials of the continuous values. These histories up to period t influence the value of each period t schooling, employment, marriage, and child alternative.[17] We allow the vector of variables that capture body mass history up to the beginning of a period (B_t) to affect current schooling, employment, marital status and child accumulation decisions.

Other information known at the beginning of year t that may influence preferences, constraints, or expectations includes exogenous individual characteristics (e.g., age, gender, race, urbanicity, region) and exogenous aggregate business cycle indicators (e.g., calendar year trend variables) denoted by the vector X_t. Variables that capture known price and supply conditions such as average tuition, wage and employment information, welfare amounts, restaurant sales, and costs of different food items are denoted by the vector P_t. We use unique superscripts to denote, for descriptive purposes only, the variables in P_t that are related to particular endogenous behaviors and outcomes; that is, $P_t = (P_t^S, P_t^E, P_t^M, P_t^K, P_t^B)$. We recognize, however, that the optimal demand functions for each behavior may depend on own and cross prices. The vector of individual information known at the beginning of period t is denoted $\Omega_t = (B_t, S_t, E_t, M_t, K_t, X_t, P_t)$.

As forward-looking decisionmakers, the individuals face uncertainty about future wages. Wages depend on educational attainment, work experience, and marital and child histories. In addition to these human capital variables, wages may also be influenced by body mass, which explains variation in wages caused by health impacts on productivity or employer preferences about appearance. Future body mass is also uncertain. Having made the

---

[17]We are not explicit about how these variables may impact preferences or the budget constraint because we do not estimate the primitive parameters of the decisionmaking optimization problem. However, these variables, because they are a part of the information set when evaluating alternatives, are arguments of the non-linear demand functions resulting from optimization.

"beginning-of-period" decisions (i.e., education, employment, marriage, and child accumulation) that affect the per-period budget and time constraints, we assume that individuals then allocate income and time to the daily activities of caloric intake and caloric expenditure (i.e., eating and exercising) denoted ci_t^* and ce_t^*. Theoretically, the density that describes possible values of next period's body mass, B_{t+1}, depends on body mass this period and these daily caloric input and output behaviors during the period. Unfortunately, these latter activities are not observed in our data. The daily input decisions depend on market and time prices of food and exercise, P_t^B, as well as the less frequently changed behaviors regarding education, employment, marital, and child accumulation. Substituting in the determinants of caloric intake and expenditure, expectations of body mass transition are, indirectly, a function of the period t behaviors, the resulting income (own wages and spouse's wages if married), and relevant prices (P_t^B).

The timeline below depicts the observed and unobserved outcomes per period (e.g., one year), and the exogenous and endogenous information available at the beginning of each period.
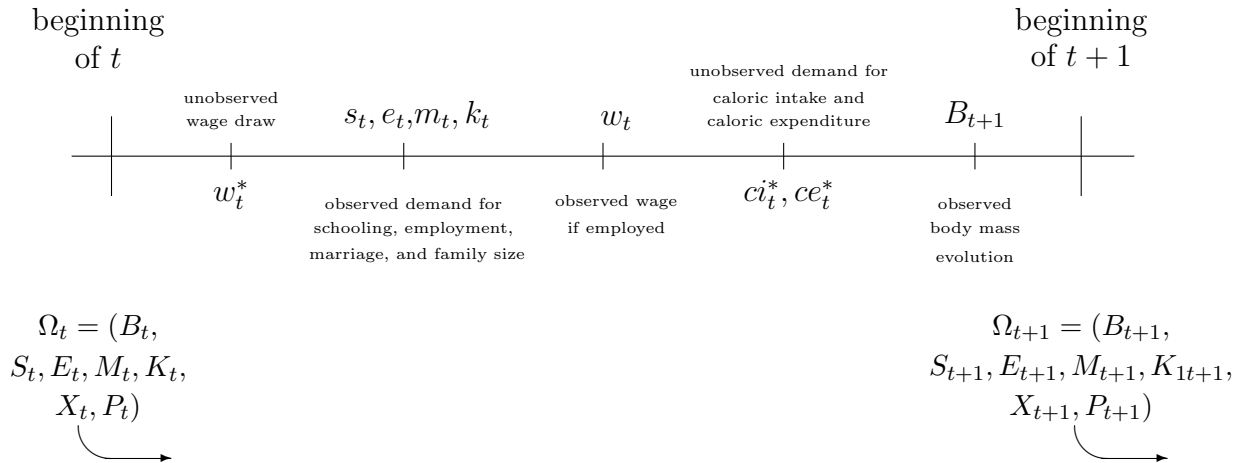
Variables that explain the observed outcomes are described in Tables 2, 3 and 4 for our research sample. The endogenous variables are summarized over all person-observations in the sample over the 1984-2002 period (Table 2). We begin by describing the variables that capture the history of one's body mass, B_t. In estimation, we use moments of the continuous index of body mass (BMI) rather than categories of adiposity (i.e., underweight, normal weight, overweight, and obese). We report summary statistics for both for informational

purposes. We also use time-varying indicators of ever overweight or ever obese to capture the contemporaneous effects of past weight issues on current behaviors and wages conditional on current BMI. Finally, we include a variable representing how different a person's BMI is from women of the same race in each year. This measure may reflect social capital, as deviations from the norm in body mass among a specific group capture social perceptions of effort or commitment.[18] An individual's education history, S_t, is captured by enrollment last period to capture state dependence and the number of years of schooling attainment. We avoid having to include additional equations for degree attainment (available in the NLSY data) by constructing indicators of at least 12 years of schooling and at least 16 years of schooling to approximate high school and college degrees, respectively. An indicator for the 13th year of schooling captures potential "freshman fifteen" effects on weight gain. The employment history vector, E_t, includes indicators of previous employment states and years of experience, separated by full and part time status. The marital history vector, M_t, includes the previous marriage state, years married if married in the previous period, and years single if ever married and single in the previous period. Finally, the variables summarizing the history of children in the household, K_t, includes the number of children in the household entering the period, an indicator of whether that number increased or decreased in the previous period (e.g., whether the female gained or lost a child), and an indicator of whether a child addition was within the last 5 years. Because nearly 90 percent of child additions are births, we interpret this variable as an indicator of having a child in the household who is less than 6 years of age.

The exogenous time-invariant individual variables are summarized for the 3213 females in our research sample in 1984, while the exogenous time-varying variables are summarized over all person-years (Table 3). Own citizenship and parental information are used to identify initially-observed values of endogenous variables (discussed in Section 4.6). Moments and interactions of variables are included in estimation. While we do not allow for a structural shift in variables per- and post- the 1993 Cook vs. Rhode Island federal policy change, we

---

[18]Teh NLSY79 does not include, ever wave, other measures of health that would allow us to measure variations in wages by body mass conditional on other measured health indicators (i.e., perhaps a beauty effect vs. a health effect).

Table 2

Description of Endogenous Individual Explanatory Variables

Variable name	White (N=1,951) Mean	Std Dev	Black (N=1,262) Mean	Std Dev
Endogenous individual variables over all person-years				
Body mass B_t				
BMI in t	23.92	5.06	26.56	6.10
Underweight: BMI in t $\leq$ 18.5	0.05	0.22	0.03	0.16
Normal weight: $18.5 \leq$ BMI in t < 25	0.65	0.48	0.46	0.50
Overweight: $25 \leq$ BMI in t < 30	0.19	0.40	0.27	0.44
Obese: BMI in t ≥ 30	0.11	0.31	0.24	0.43
Ever overweight prior to t	0.39	0.49	0.57	0.50
Ever obese prior to t	0.14	0.35	0.28	0.45
Weight in pounds at t	142.27	31.14	155.88	36.44
Height in inches at t	64.66	2.59	64.26	2.90
Standardized deviations from mean BMI in t	0.00	1.00	0.00	1.00
Education history S_t				
Enrolled in t-1	0.09	0.29	0.08	0.27
Years enrolled in school entering t	13.23	1.95	12.70	1.73
Years enrolled < 12 entering t	0.09	0.28	0.17	0.38
Years enrolled ≥ 12 entering t	0.91	0.28	0.83	0.38
Years enrolled ≥ 16 entering t	0.15	0.36	0.07	0.26
Freshmen year of college in t	0.02	0.13	0.02	0.13
Employment history E_t				
Employed in t-1	0.85	0.36	0.79	0.41
Employed full time in t-1	0.69	0.46	0.71	0.45
Employed part time in t-1	0.31	0.46	0.29	0.45
Years employed entering t	10.34	5.47	8.94	5.80
Years full time employed entering t	6.42	5.15	5.43	5.31
Years part time employed entering t	3.92	3.00	3.51	2.70
Marital history M_t				
Married in t-1	0.63	0.48	0.31	0.46
Years married entering t if married in t-1	8.00	5.21	6.64	4.87
Years single entering t if single in t-1	1.64	3.15	1.67	3.56
Child history K_t				
Number of children in hh entering t	1.12	1.16	1.42	1.29
Acquire any children in hh in t-1	0.09	0.28	0.09	0.28
Lose any children from hh in t-1	0.02	0.12	0.03	0.17
Any children under age 6 in t-1	0.35	0.48	0.39	0.49

do allow for a cubic in the time trend, which may capture this change if it is statistically relevant.

Table 3

Description of Exogenous Individual and Aggregate Explanatory Variables

Variable name	White (N=1,951)		Black (N=1,262)	
	Mean	Std Dev	Mean	Std Dev
Time-invariant individual variables in year 1984				
AFQT score minus median	19.23	25.47	-11.57	19.81
AFQT score missing	0.03	0.16	0.02	0.12
Non-US citizenship at birth	0.03	0.15	0.03	0.17
Non-US citizenship at age 14	0.01	0.07	0.01	0.10
Mother is non-US citizen	0.04	0.20	0.03	0.17
Father is non-US citizen	0.14	0.35	0.03	0.16
Years of education of mother	12.06	2.38	10.81	2.55
Mother's education missing	0.03	0.18	0.07	0.25
Years of education of father	12.39	3.16	10.23	3.52
Father's education missing	0.08	0.27	0.28	0.45
Time-varying individual variables over all person-years				
Age in years	30.28	6.11	30.19	6.08
Rural residence	0.26	0.44	0.17	0.37
Residence type missing	0.05	0.23	0.04	0.20
Northeast region	0.20	0.40	0.14	0.35
North central region	0.34	0.47	0.18	0.39
West region	0.16	0.37	0.07	0.25
South region (omitted category)	0.30	0.46	0.60	0.49
State of residence missing	0.01	0.08	0.01	0.10
Trend and Interval Variables				
	Mean	Std Dev	Min	Max
Time trend (1=1984)	9.16	5.40	1	19
Interval measure for density estimation	-1.42	0.68	-2.20	0.00

The exogenous price and supply-side variables vary by state or county and time, and are described in Table 4.[19]   While there are several sources of statistical identification in

---

[19]We obtain variation at the state, county (if available), and year levels. Average tuition rates are from the Higher Education General Information Survey (`www.icpsr.umich.edu/icpsrweb/ICPSR/series/30`) and vary by state and year. Employment data and retail, food, and restaurant sales data are from the Woods and Poole Economics Database (`www.woodsandpoole.com`) and vary by state and county and year. We use state and year level variation in the average Aid to Families with Dependent Children (AFDC) payment (Bernal and Keane, 2010). Yearly population data are from the Census Bureau

our model (discussed in Section 4.6), these exogenous time-varying variables serve as classic identification variables since they are included in theoretically-relevant demand (behavior) equations and excluded from the body mass production function conditional on included endogenous behaviors (as described in more detail in the next section).

(`www.census.gov/popest/data/historical`) with total population at the county level and sex ratios at the state level. Food, alcohol, and cigarette prices, from the Council for Community and Economic Research (formerly the American Chamber of Commerce Research Association or ACCRA), are aggregated to the county level by year (`www.c2er.org`). The food price consists of weighted prices of frying chicken (per pound), chunk light tuna (5 ou.), half gallon of whole milk, a dozen eggs, margarine (1 lb.), grated parmesan cheese (8 ou.), potatoes (10 lb.), bananas (per pound), iceberg lettuce, and white bread. The junk food price is the average price of a McDonald's quarter pound cheeseburger, a thin crusted cheese pizza at Pizza Hut or Pizza Inn, and fried chicken at KFC or Church's.

Table 4
Description of Exogenous Price and Supply-Side Variables

Variable name	Mean	Std Dev	Min	Max
Schooling variables P_t^S				
Two year college semester tuition in 000s	1.483	0.842	0.127	5.027
Four year college semester tuition in 000s	2.338	0.986	0.339	6.868
Graduate school semester tuition in 000s	2.644	1.139	0.518	7.076
Employment variables P_t^E				
Unemployment rate	6.86	2.26	1.90	17.70
Total employment per capita	0.58	0.11	0.38	1.34
Ratio of manuf empl to total empl	0.13	0.05	0.02	0.26
Ratio of service empl to total empl	0.28	0.05	0.18	0.48
Total earnings per employee	41.18	8.08	27.41	77.16
Ratio of manuf earnings to total earnings	0.18	0.08	0.02	0.39
Ratio of service earnings to total earnings	0.24	0.05	0.13	0.42
Marriage and Children variables P_t^M and P_t^K				
Total population in 000,000s	51.07	56.35	4.54	350.25
White gender ratio: male, 20-60/female, 15-50	1.05	0.07	0.92	1.64
Black gender ratio: male, 20-60/female, 15-50	0.89	0.09	0.75	1.79
Household income in 000s	59.62	12.29	36.97	106.29
AFDC per month per family of four in 00s	5.23	1.96	1.47	11.87
Body Mass variables P_t^B				
Mean price of food	1.85	0.15	1.51	2.76
Mean price of junk food	4.69	0.32	3.57	6.79
Mean price of carton of cigarettes	19.59	6.67	10.03	47.64
Mean price of 6-pack of beer	4.87	1.04	3.46	8.18
Mean price of bottle of wine	6.28	0.90	3.93	10.47
Mean price of liter of liquor	17.41	4.25	8.63	26.20
Ratio of food sales to total retail sales	0.18	0.03	0.08	0.25
Ratio of restaurant sales to total retail sales	0.10	0.03	0.06	0.32
Indicator of missing price data	0.06	0.24	0.00	1.00

Note: Data presented are means over 50 states and the District of Columbia for the years 1984-2002. Dollar amounts are in year 2000 dollars.

4.2 Behaviors: Schooling, Employment, Marriage, Children

Below we describe the dynamic demand functions derived from the joint schooling, employment, marriage, and children decisions and represented by probabilities of the observed outcomes. While theory suggests that these decisions are jointly made (among 36 mutually exclusive combinations of the four behaviors), we empirically model the derived demand as four equations and allow the equations to be correlated through time-invariant and time-varying observed and unobserved heterogeneity. The error terms that capture the unobserved determinants of each equation e, u_t^e, are decomposed into a permanent individual component (μ), a time-varying individual component (ν_t), and an idiosyncratic component (ϵ_t); specifically, $u_t^e = \rho^e \mu + \omega^e \nu_t + \epsilon_t^e$. Factor loadings ($\rho$ and ω) in each equation (and for each outcome in each equation) indicate the relative importance of the associated type of heterogeneity for that equation or outcome. The serially-uncorrelated idiosyncratic terms (ϵ_t^e) are Extreme Value distributed each period, producing logit and multinomial logit probabilities of observed decision outcomes.

The log odds ratio of being enrolled in school in period t ($s_t = 1$) relative to not being in school ($s_t = 0$) is

$$\ln\left[\frac{p(s_t = 1)}{p(s_t = 0)}\right] = \alpha' f^S(B_t, S_t, E_t, M_t, K_t, X_t, P_t^S, P_t^E, P_t^M, P_t^K, P_t^B) + \rho^S \mu + \omega^S \nu_t. \quad (1)$$

The log odds ratio of being non-employed ($e_t = 0$) or employed part-time ($e_t = 1$) relative to being employed full-time ($e_t = 2$) in period t is

$$\ln\left[\frac{p(e_t = d)}{p(e_t = 2)}\right] = \beta_d' f^E(B_t, S_t, E_t, M_t, K_t, X_t, P_t^S, P_t^E, P_t^M, P_t^K, P_t^B) + \rho_d^E \mu + \omega_d^E \nu_t \quad d = 0, 1. \quad (2)$$

The log odds ratio of being married in period t ($m_t = 1$) relative to not being married ($m_t = 0$) is

$$\ln\left[\frac{p(m_t = 1)}{p(m_t = 0)}\right] = \gamma' f^M(B_t, S_t, E_t, M_t, K_t, X_t, P_t^S, P_t^E, P_t^M, P_t^K, P_t^B) + \rho^M \mu + \omega^M \nu_t. \quad (3)$$

The log odds ratio of a change in the number of children in the household by at least j children in period t ($k_t = j$) relative to no change ($k_t = 0$) is

$$\ln\left[\frac{p(k_t = j)}{p(k_t = 0)}\right] = \delta_j' f^K(B_t, S_t, E_t, M_t, K_t, X_t, P_t^S, P_t^E, P_t^M, P_t^K, P_t^B) + \rho_j^K \mu + \omega_j^K \nu_t \quad j = -1, 1. \quad (4)$$

Note that these four probabilities are functions of the same vector of explanatory variables since the decisions that produce the observed outcomes are jointly made. Additionally, because the schooling, employment, marriage, and child accumulation decisions (s_t, e_t, m_t and k_t) are made jointly at the beginning of the period, all the variables representing supply side conditions ($P_t^S, P_t^E, P_t^M, P_t^K$) appear in each of Equations 1-4 in order to capture own- and cross-price effects. Similarly, individuals are forward looking and anticipate making optimal daily caloric intake and caloric expenditure decisions within the period, and these decisions depend on the known exogenous vector P_t^B. The equations are also correlated through both the permanent unobserved individual heterogeniety (μ) and the time-varying unobserved individual heterogeneity (ν_t), which enter each equation with different effects (ρ^e and $\omega^e, e = S, E, M,$ and K).

4.3   Body Mass Evolution

We assume that body mass transitions are captured by a production function that depends on caloric intake, ci_t^*, and caloric expenditure , ce_t^*: behaviors chosen by an individual but unobserved in our data. Specifically,

$$
\begin{aligned}
B_{t+1} &= b(B_t, ci_t^*, ce_t^*, X_t, u_t^B) \\
&\simeq \phi' f^B(B_t, S_{t+1}, E_{t+1}, M_{t+1}, K_{t+1}, w_t, P_t^B, X_t) + \rho^B \mu + \omega^B \nu_t + \epsilon_t^B .
\end{aligned} \tag{5}
$$

After substituting in the determinants of caloric intake and expenditure, body mass transition at the end of the period is a function of the observed endogenous decisions during the period, reflected by the vector of updated $(t+1)$ history variables.[20] Conditional on the observed behavior during period t, only the supply-side variables (P_t^B) that affect body mass inputs affect the body mass transition at the end of period t. Note that the unobserved permanent and time-varying individual UH that affects the schooling, employment, marriage, and children demand equations also influences body mass at the end of period t.

---

[20]Theory suggests that non-earned income (including spouse income where marital status is modeled as endogenous) and earned income (e.g., wages which are modeled as endogenous) should be included in the body mass evolution equation since the determinants of caloric intake and expenditure include one's income. However, we found the total household income measures in the NLSY to be fairly noisy and frequently missing. Rather than include own wages only, we have chosen to replace wages with their reduced form determinants in estimation of the body mass equation.

4.4 Distribution of Wages

Theory states that individuals receive a wage offer each period and, conditional on this wage, they evaluate the value of each set of alternative behaviors and choose the one that maximizes their discounted lifetime utility. Future wages are uncertain, so individuals must evaluate future utility unconditional on future wages (i.e., integrate over the distribution of wages). Empirically, we (the econometricians) do not observe wage offers. We only observe her accepted wage, conditional on choosing employment in period t. It is appropriate to address the potential bias in the marginal effects of explanatory determinants of the unconditional wage distribution by modeling the UH that impacts selection into employment (Equation 2) as well as wages. According to Mincer's seminal work, wages are a function of observed human capital, typically captured by educational attainment and work experience. Wages may also vary by productivity differences at the same levels of education and experience. Productivity is generally not observed, and certainly difficult to measure. We might imagine that productivity is influenced by employee characteristics as well as employer or job characteristics.[21] In particular, we allow productivity at work to differ by an employee's marital status and number of children in the household. Productivity may also be influenced by one's physical health; our only health measure is body mass. Significance of variation in body mass in explaining variation in wages may also reflect preferences by employers.

For these reasons our period t wage equation includes variables summarizing the history of education, employment, marriage, and children as well as body mass entering the current period. Each of these variables is endogenous, and we address potential bias stemming from the endogeniety of these wage determinants by modeling the UH that impacts the determinants (Equations 1-5) as well as the wages. We also include as wage determinants characteristics of the demand-side of the employment market at the state level (denoted P_t^E), aggregate trends (captured by a cubic function of the continuous year variable), and geographical region-year interactions to pick up variation in skill prices over locations and

---

[21]We observe an individual's occupation if she works, but few other employer characteristics. We could model occupational choice and include occupation indicators in the wage equation. Because this would add a large multinomial choice equation and many parameters, we do not jointly model occupation. We also do not include occupation indicators and treat them as exogenous. Rather, theory tells us that the occupation demand function is a function of the same variables that determine each of the jointly chosen behaviors that we do model. Hence, we assume that the endogenous human capital and productivity histories explain occupation.

time. If we wished to explain how the mean of the wage distribution varied by its determinants, we could estimate log wages conditional on being employed as is typical in the labor literature. That is, observed log wages in period t, conditional on being employed in period t, would be specified as

$$\ln(w_t \mid e_t \neq 0) = \eta' f^W(B_t, S_t, E_t, M_t, K_t, X_t, P_t^E) + \rho^W \mu + \omega^W \nu_t + \epsilon_t^W \tag{6}$$

where ϵ_t^W is a serially-uncorrelated error term. Rather than use OLS to explain average log wages, we estimate the entire density of wages conditional on explanatory variables (explained in the next section below).

4.5 Conditional Density Estimation of Wages and Body Mass

As explained above, future wages and body mass are uncertain when individuals are making their period t decisions. Thus, we seek to estimate the marginal effects of theoretically-relevant variables on the density of these outcomes rather than simply the first moment, or expected value. To do so, we estimate the density of wages (and of body mass) conditional on observable and unobservable variables. We use the conditional density estimation technique (Gilleskie and Mroz, 2004), which allows the effect of an explanatory variable to be different at different points of support of the wage (body mass) distribution.

Specifically, we begin by splitting the sample observations into K intervals. That is, we discretize the support of the dependent variable. We do this by choosing boundary points that place an equal number of observations in each interval (i.e., $\frac{1}{k}$th of the sample falls into each bin or interval) such that w_k is the wage that represents the kth percentile of the ordered wages. Figure 4 displays the histogram of wages with cut points associated with $K = 10$.

The probability of an individual's wage offer being from the kth interval (bounded by cutoff wages w_{k-1} and w_k), conditional on the explanatory variables x, is

$$p[w_{k-1} \leq W \leq w_k | x] = \int_{w_{k-1}}^{w_k} f(w|x) dw \ .$$

28
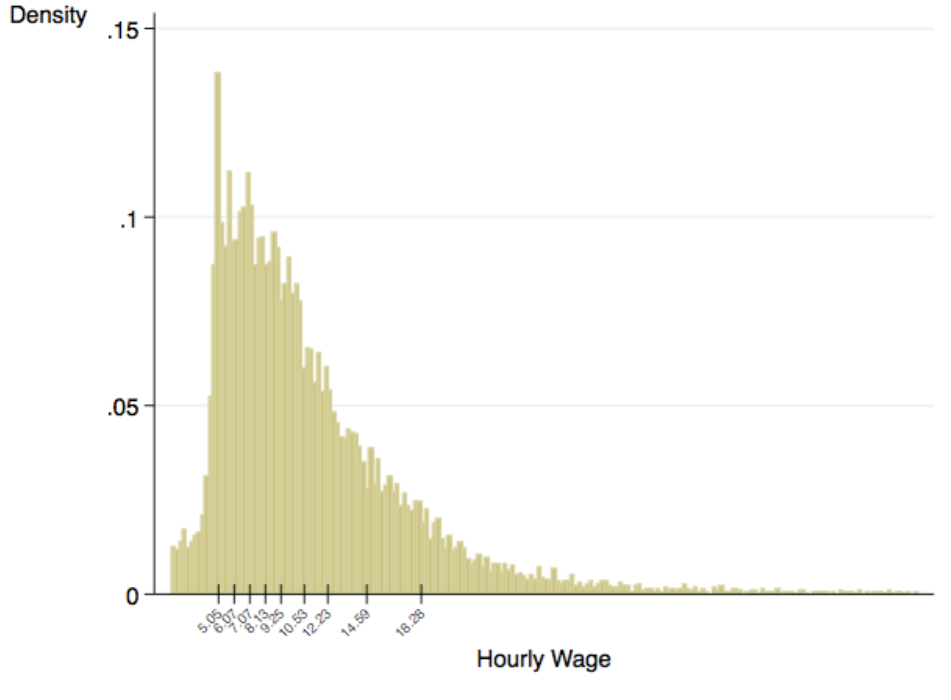
Figure 4: Histogram of Observed Wages with Interval Cut Points

We can define the discrete time hazard function capturing the probability of being in the kth interval conditional on not being in a previous interval as

$$\lambda(k,x) = p[w_{k-1} \leq W \leq w_k | x, W \geq w_{k-1}] = \frac{\int_{w_{k-1}}^{w_k} f(w|x)dw}{1 - \int_{w_0}^{w_{k-1}} f(w|x)dw} \; .$$

Using this hazard representation, we re-write the probability of a wage falling in the kth interval as

$$p[w_{k-1} \leq W \leq w_k | x] = \lambda(k,x) \prod_{j=1}^{k-1} [1 - \lambda(j,x)] \; .$$

We can recover any moment of the distribution once we have estimated this hazard function (which allows us to form the density). For example, the expected value of the wage conditional on the explanatory variables is

$$E[W|x] = \sum_{k=1}^{K} \overline{w}(k|K)\lambda(k,x) \prod_{j=1}^{k-1} [1 - \lambda(j,x)]$$

where $\overline{w}(k|K)$ is the mean of the observed wages, w, within the kth interval.[22]

---

[22]Note that it is important to estimate the employment outcome jointly with the conditional density of wages (as we do) in order to account for selection that may bias the estimated marginal effects of explanatory variables when we only observe the accepted wages of employed individuals.

It remains to estimate the conditional hazard, $\lambda(k, x)$. First, we replicate each of the person-year observations $K - 1$ times. We then define a dependent variable indicating whether or not the wage of a particular person in year t falls into each of the $K - 1$ intervals. (The indicator is undefined for intervals greater than the observed wage.) We know that the unconditional probability of an observed wage being in the kth interval given that it is not in an earlier interval is $\frac{1}{K-(k-1)}$ based on the way we created the intervals. Next, we define a transformation of the interval counter as $i_k = -\ln(K - k)$ for $k < K$. It follows that the logit probability for each indicator i_k, $\lambda(k, i) = logit(i_k) = \frac{e^{i_k}}{1+e^{i_k}}$, gives us the unconditional probability of the observed wage being in the kth interval (given that it is not in a previous interval) exactly because of the way we have defined the interval counter transformation. To estimate the conditional hazard (i.e., conditional on variables x), we interact the transformed interval counter (i.e., the variable i_k) fully with the wage determinants.

The hazards for wages and body mass are jointly estimated with other equations in our full model (i.e., Equations 1-4 and the attrition and initial conditions equations) and correlated through the permanent and time-varying UH whose distributions we also estimate. Once the conditional hazard is estimated we can use it to recover the mean wage and body mass conditional on the observables (and unobservables, in our case). The marginal effects of the explanatory variables on the mean wage (body mass) account for differences in effects at different points of support of the wage (body mass) distribution.

4.6   Attrition and Initial Condition Equations

In order to correctly estimate the distribution of permanent and time-varying UH, we must account for the fact that individuals attrit from our research sample over time. We include, in the jointly estimated set of equations, an equation for the probability of attrition at the end of the period. This probability depends on updated history variables, reflecting the period t decisions, and the UH.

Our analysis of individual behavior begins in 1984 when some of the endogenous variables are non-zero. We cannot explain the variation in these observed initial conditions using the dynamic specification of behaviors described above because we do not model the history of one's decisions prior to 1984 (i.e., we have restricted the model to explaining behavior

after age 18). For this reason we specify reduced-form equations that depend on exogenous characteristics, including variables that do not influence the subsequent per-period outcomes (conditional on the endogenous history variables). These initial condition equations are jointly estimated with the subsequent dynamic demand and outcome equations and correlated through the permanent individual UH. The endogenous state variables entering the first period (1984), when individuals are 18-26 years old, include: years of schooling (continuous variable: 7-16), marital state (logit variable: 0,1), years married if married (continuous variable: 1-4), years single if single (continuous variable: 1-4), number of children (continuous variable: 0-7), employment state (multinomial logit variable: full-time, part-time, not employed), years of full time experience (continuous variable: 0-5), years of part time experience (continuous variable: 0-5), and BMI (continuous variable). Identification of these initial conditions is achieved by inclusion of exogenous time-varying variables (P_t when $t = 0$) and other individual characteristics (e.g., mother's and father's years of education and indicators of mother's, father's, and child's (at birth and age 14) non-U.S. citizenship).

Additionally, we observe wages of those individuals who are employed in period t. However, in the NLSY data, wages are missing for some employed individuals. In a typical year, about 6 percent of wages are missing. We include an equation to represent any non-randomness in missing wages. In 1995, 1997, 1999, and 2001, however, wages are missing for all individuals in the sample because they were not surveyed in these years.[23] Table 5 below summarizes the jointly estimated set of equations and their determinants.

5 Empirical Results

5.1 Replication of Previous Literature

In order to convince the reader that our preferred model addresses omissions in the literature (namely dynamics and UH bias correction), we first estimate models similar to those in the

---

[23]We are able to construct values for demand behaviors during these non-surveyed years given observed behaviors in surveyed years. For example, while we do not observe whether someone was married in 1995, we do observe their marital status in 1994 and 1996. We make assumptions during years with missing data for only a small number of person-years. If the assumptions required to fill in these missing years are too strict we delete that individual from our research sample.

Table 5
Summary of Equation Specifications in Jointly Estimated System

Outcome	Estimator	Explanatory Variables		Unobserved Heterogeneity
		Endogenous	Exogenous	
Enrolled s_t	logit	B_t, S_t, E_t, M_t, K_t	$X_t, P_t^S, P_t^E, P_t^M, P_t^K, P_t^B$	$\rho^S \mu, \omega^S \nu_t, \epsilon_t^S$
Employed e_t	mlogit	B_t, S_t, E_t, M_t, K_t	$X_t, P_t^S, P_t^E, P_t^M, P_t^K, P_t^B$	$\rho^E \mu, \omega^E \nu_t, \epsilon_t^E$
Married m_t	logit	B_t, S_t, E_t, M_t, K_t	$X_t, P_t^S, P_t^E, P_t^M, P_t^K, P_t^B$	$\rho^M \mu, \omega^M \nu_t, \epsilon_t^M$
$\triangle$ Kids in hh k_t	mlogit	B_t, S_t, E_t, M_t, K_t	$X_t, P_t^S, P_t^E, P_t^M, P_t^K, P_t^B$	$\rho^K \mu, \omega^K \nu_t, \epsilon_t^K$
Wage if emp w_t	CDE	B_t, S_t, E_t, M_t, K_t	X_t, P_t^E	$\rho^W \mu, \omega^W \nu_t, \epsilon_t^W$
Body Mass B_{t+1}	CDE	B_t $S_{t+1}, E_{t+1}, M_{t+1}, K_{t+1}$	X_t, P_t^B	$\rho^B \mu, \omega^B \nu_t, , \epsilon_t^B$
Attrition A_{t+1}	logit	B_{t+1} $S_{t+1}, E_{t+1}, M_{t+1}, K_{t+1}$	X_t	$\rho^A \mu, \omega^A \nu_t, \epsilon_t^A$
Initially observed state variables	2 logit 7 ols		X_1, P_1, Z_1	$\rho^I \mu, \epsilon_1^I$

literature using our research sample. Table 6 presents coefficient estimates and marginal effects of BMI on wages as we consider different right hand side specifications (Models 1-3). That is, Model 1 is a very sparse model of log wages on BMI and several demographic variables. We find that a contemporaneous reduction in body mass from overweight status (BMI = 27.5) to normal weight status (BMI = 24) results in a \$0.31 increase in hourly wages for white women and a \$0.12 increase in the hourly wages of black women, where increases for black women are statistically significant.[24] Put differently, a five percent decrease in BMI (i.e., equivalent to about a seven pound annual weight loss for a 5' 4" female) leads to an \$0.11 increase in hourly wages for white females and a \$0.02 increase in hourly wages for black females.[25] When we include (in Model 2) variables representing human capital (but treat them as exogenous), the marginal effect of moving from overweight to normal weight is a \$0.27 and \$0.14 increase in wages, respectively. (A five percent decrease in BMI

---

[24]If we do not control for deviations from the mean BMI by race, BMI impacts on wages are statistically significant for both races (as the literature reports with this simple biased specification).

[25]To report marginal effects in level dollars we transform from logs to levels assuming normality and homoskedasticity of the wage error term.

results in an estimated $0.09 and $0.04 increase, respectively.) When we control for demand side wage shifters (Model 3) the marginal effect of body mass changes very little. In the latter two cases, the impact of body mass on the wages of white and black women are less precisely estimated. Model 4 of Table 6 presents the estimates and marginal effects when the model is estimated with individual fixed effects. These fixed effects purge the coefficient estimates of bias associated with permanent individual UH that might be correlated with wages and body mass (as well as education, work experience, marriage, and children). The resulting marginal effect of body mass on wages of white women is much smaller ($0.07) and is significantly estimated at the one percent level. The effect of a reduction in weight on wages of black women becomes negative, costing $0.05 (for a reduction from overweight to normal weight).[26] We find that the large reduction in impact after controlling for permanent heterogeneity is consistent with the literature. However, the results after introduction of a variable to capture deviation from the average BMI of individuals of one's same race suggest that these smaller impacts are statistically significant.

We also calculate the marginal effects using a quantile regression (QR) estimator in order to determine whether the marginal effect of a change in BMI is different at different levels of wages, as was examined by Johar and Katayama (2012).[27] Displayed in Table 7, we find that a reduction from overweight to normal weight among females leads to a smaller average increase in hourly wages in the bottom quartile of the distribution than in higher quartiles. On average, the wage increases are $0.26 and $0.15 for white and black women, respectively. These findings are qualitatively similar to those of Johar and Katayama; they are larger because we do not account for the endogeneity of BMI in these exploratory replications with our data. The conditional density estimation (CDE) approach is more general and flexible than the QR approach, and is easier to use in more complicated models that involve correction for selection (into employment) and endogeneity (of relevant explanatory variables such as BMI, education, experience, and productivity proxies). Estimating a single equation for wages using the CDE approach, we find average wage effects of the over- to normal-weight

---

[26]A five percent decrease in weight leads to an increase in wages of white women on $0.02 and a decrease in wages of black women of almost $0.04.

[27]These authors account for the endogeneity of BMI using an IV approach. They control for, but treat as exogenous, educational attainment, type of employment, work experience, marital status, and number of biological children. They do not control for selection into employment.

Table 6

Replicated Results from Literature: Estimated Effects of Body Mass on Wages

Variable	Model 1		Model 2		Model 3		Model 4	
BMI in t	-0.008	(0.006)	-0.008	(0.005)	-0.009	(0.005) **	-0.016	(0.005) ***
BMI in t × Black race	-0.015	(0.009) *	-0.007	(0.008)	-0.006	(0.008)	0.002	(0.008)
White: std devs from mean BMI	-0.002	(0.026)	0.004	(0.023)	0.015	(0.023)	0.073	(0.028) ***
Black: std devs from mean BMI	0.117	(0.039) ***	0.067	(0.035) **	0.071	(0.034) **	0.097	(0.032) ***
Estimation Method:	OLS on $\ln W_t$ clustered std err		OLS on $\ln W_t$ clustered std err		OLS on $\ln W_t$ clustered std err		OLS on $\ln W_t$ clustered std err fixed effects	
Model Includes:	X_t, B_t		X_t, B_t S_t, E_t, M_t, K_t		X_t, B_t S_t, E_t, M_t, K_t P_t^e		X_t, B_t S_t, E_t, M_t, K_t P_t^e	
R-squared	0.28		0.40		0.42		0.34	

Marginal Effect ($) of a change from Overweight to Normal Weight (at the point estimates)

White	0.31		0.27		0.26		0.07	
Black	0.12		0.14		0.14		-0.05	

Note: Coefficients are reported in the top panel; marginal effects are simulated in order
to account for interactions. Models 1-3 treat endogenous histories as exogenous.
All models include a cubic time trend and time trend * regional location interactions.

reduction to be \$0.22 and \$0.14. With CDE (estimated coefficients not shown in table), we find that the coefficients on BMI variables interacted with the continuous wage interval indicator are statistically significant for whites and blacks, suggesting that body mass has different effects across the support of the distribution of wages. These findings are consistent with the reported marginal effects from the QR framework above. It remains to explore how these results change when additional controls for selection, endogeneity, and measurement error are incorporated. Importantly, our multiple equation dynamic framework allows body mass to influence behaviors, as individuals age, that also determine wages, allowing us to capture the contemporaneous and life-cycle effects of an evolving body mass on wages.

Table 7
Estimated Effects of Body Mass across the Support of Wages

Variable	Quantiles from Quantile Regression (QR)					
	25th		50th		75th	
BMI in t	-0.106	(0.039) ***	-0.190	(0.041) ***	-0.240	(0.066) ***
BMI in t × Black race	0.019	(0.050)	-0.027	(0.050)	-0.065	(0.057)
White: std devs from mean BMI	0.289	(0.184)	0.576	(0.186) ***	0.715	(0.313) **
Black: std devs from mean BMI	0.343	(0.161) **	1.076	(0.151) ***	1.526	(0.268) ***

Model
Includes
$$X_t, B_t$$
$$S_t, E_t, M_t, K_t$$
$$P_t^e$$

Calculated Contemporaneous Effect
of Change from Overweight to Normal Weight (at the point estimates)

QR Average	White: 0.26	Black: 0.15
CDE Average	White: 0.22	Black: 0.14

Note: Coefficients are reported in the top panel; marginal effects are simulated
in order to account for interactions. Models treat endogenous histories as exogenous.
Models include a cubic time trend and time trend * regional location interactions.
Note: QR = quantile regression; CDE = conditional density estimator

5.2 Results from our Preferred Model

Graphical Fit of Model's Ability to Capture Dynamic Behavior

Appendix Tables A1-A6 present coefficient estimates from our dynamic multiple equation model that accounts for both permanent and time-varying individual UH and allows for different effects across the distributions of body mass and wages.[28] We begin by depicting, in Figure 6 and 7, outcomes from our data generating process (i.e., the jointly-estimated wage, demand, and production equations) with the observed data. We simulate behavior from the first period forward, using the simulated values of endogenous explanatory variables updated each period based on the estimated dynamic equations. The updated simulations from the model with heterogeneity suggest that the model captures well the dynamics associated with the modeled behaviors.[29]

Calculation of Model Implications

Examination of the coefficient estimates is not sufficient for understanding the dynamic role of body mass on wages over time. In order to calculate the contemporaneous and dynamic marginal effects of variables of interest on outcomes of interest, we simulate behavior over time using the estimated parameters on observed variables, the estimated distribution of UH, and assumed distributions of the remaining i.i.d unobservables. For example, we might desire to know how an improvement in body mass affects wages. But, is that a one time decrease in body mass at age 18? Or a decrease each year? And do we care about the average effect on wages over the life cycle, or at age 18, or at age 40? Any change in body mass at a particular age will have an effect on wages at every subsequent age.

---

[28]Estimates for the initial conditions equations are available from the authors. We have verified that the variables we include in the initial condition equations as identifying variables are statistically significant and can be excluded from the main equations conditional on the inclusion of the endogenous variables. Similarly, to test our exclusions restrictions in the dynamic equations, we have estimated a model where we include variables in equations for which theory suggests they should be excluded. Tests of statistical significance confirm that these identifying variables can be excluded after conditioning on the endogenous variables that they do explain.

[29]Likelihood ratio tests suggest that a model with six estimated discrete mass points for the permanent UH distribution [0.00 (0.02), 0.27 (0.10), 0.45 (0.19), 0.61 (0.34), 0.79 (0.27), 1.00, (0.08)] and three estimated discrete mass points for the time-varying UH distribution [0.00 (0.08), 0.60 (0.92), 1.00 (0.01)] provides the best statistical fit. Mass point values are listed with probability weights in parentheses.
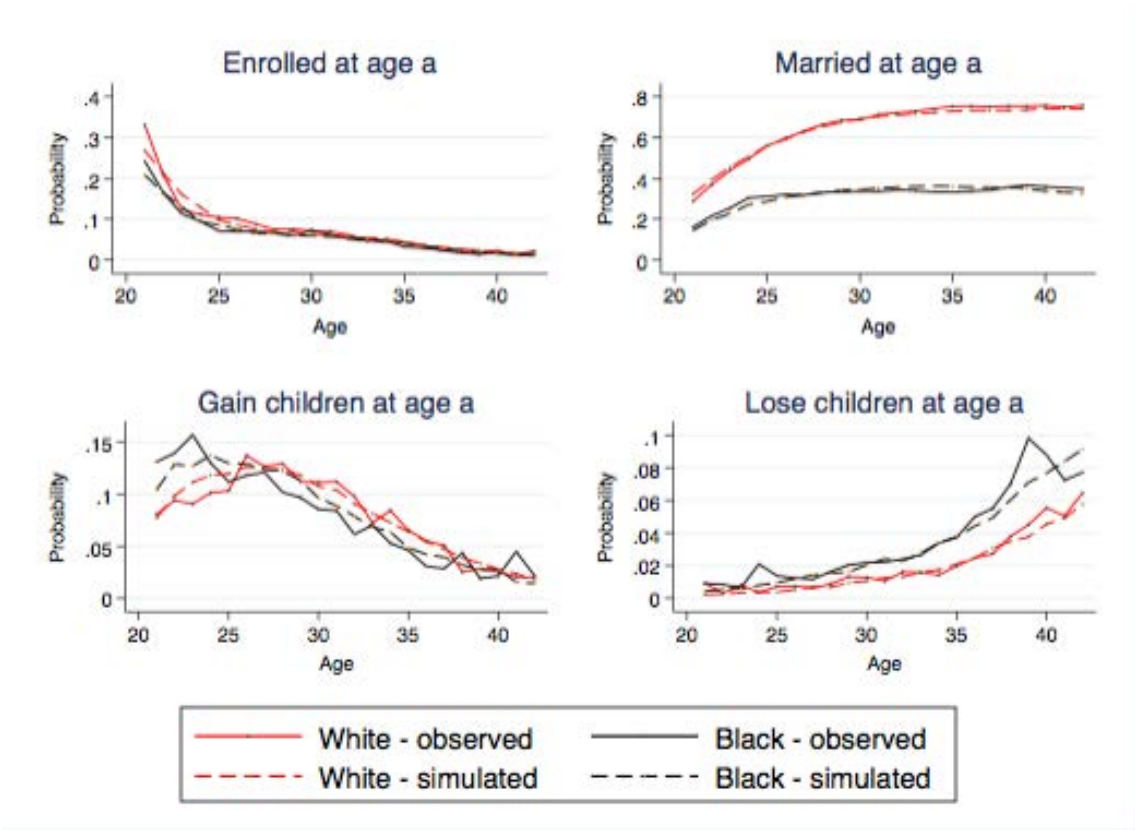
Figure 5: Comparison of Observed Data to Model Predictions
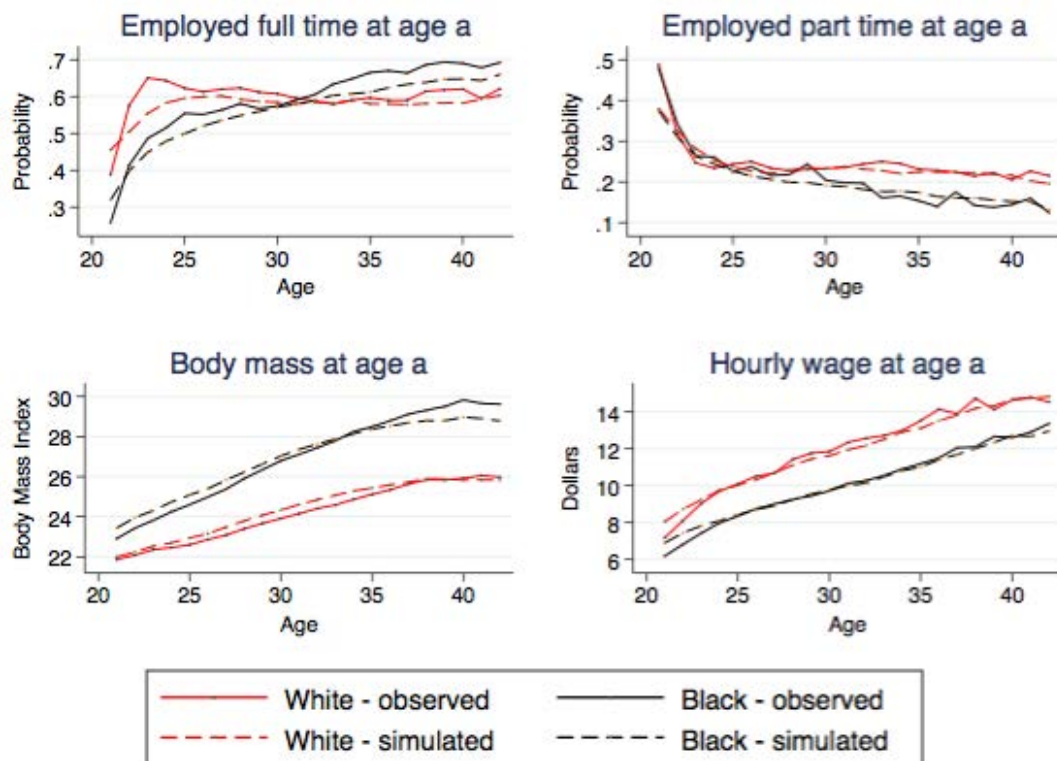
Figure 6: Comparison of Observed Data to Model Predictions

Since our goal is to understand the dynamic role of body mass on wages, we focus our attention here on understanding these marginal effects, and we encourage the reader to take note of the statistical significance of estimated parameters (available in Tables A1-A6). Note that we report impacts on both the conditional (on employment) and unconditional wage distribution (while also conditioning on observed wage determinants). First, we calculate the effect of an improvement in body mass from overweight (BMI=27.5) to normal weight (BMI=24) on contemporaneous wages. That is, if one's body mass currently was normal weight rather than overweight (all else equal), how might the hourly wage distribution differ? We call this the *contemporaneous effect*. This effect does not include the effect on human capital formation and accumulation (which also affect wages this year) of the changes in evolutionary body mass that may have precipitated the current change. Figure 8 plots the estimated wage differences across deciles for our preferred model with UH (solid lines), which addresses endogeneity, selection, and measurement error; it also displays the associated results for a model without UH (dotted lines), which represents biased impacts. The no UH estimates reflect gains similar to those described previously by our replications of models used in the literature. Our preferred model that includes UH suggests that, on average, the unbiased contemporaneous wage impacts are smaller; white women experience a $0.07 increase in hourly wages and black women experience a small decrease of $0.02. These averages reflect increases in wages at all deciles for whites while blacks experience reductions in wages that fall in the first six deciles, but improvements at higher wages.

We can compare the contemporaneous effect of body mass on wages with that of human capital and productivity variables. One year increases in education lead to contemporaneous increases in hourly wages of $0.79 and $0.66 for white and black females, respectively. One additional year of full time employment experience leads to increases of $0.26 and $0.21, respectively. The associated reduction of body mass by one BMI point leads to increases in contemporaneous wages of $0.05 for white women and decreases of $0.01 for black women, on average.

Now we want to calculate the *dynamic effect* (or life-cycle effect) of an improvement in body mass from overweight to normal weight. That is, we simulate schooling, employment,
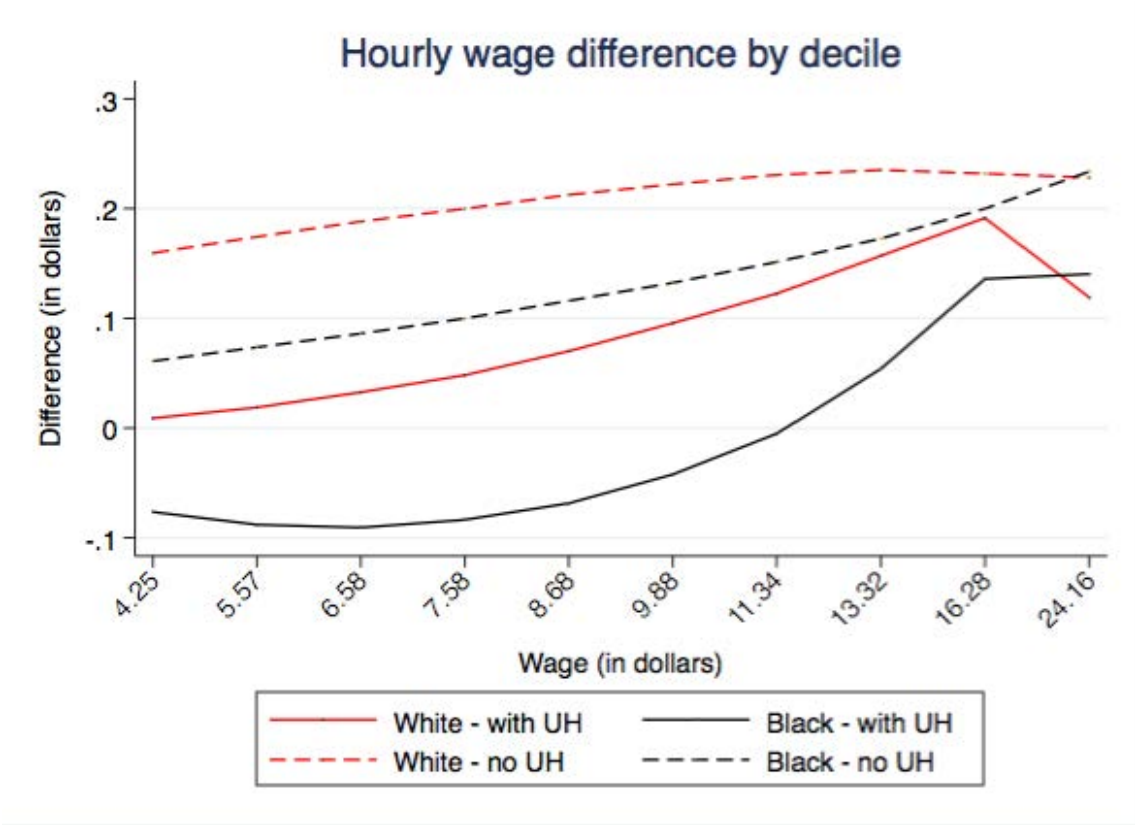
Figure 7: Contemporaneous Effect of Body Mass Improvement on Wages

marriage, and family size behaviors of individuals if they were normal or overweight through-out their lives (ages 18 through 45) and compare the predicted wages for each body mass type each period. Note that this calculation includes both the change in contemporaneous wages (via body mass) as well as the changes in wages caused by the effects of body mass on human capital (via education and experience) and productivity (via marriage and children).

Figure 9 depicts the average differences in wages associated with an improvement in body mass at different deciles of the wage distribution contemporaneously (solid line) and including the indirect, life-cycle impacts (dotted line). The life-cycle effect for white females is an average increase in hourly wages of $0.04. Black females experience a decline of $0.07 on average. Recall that the contemporaneous effects averaged $0.07 and $-0.02 for white and black females, respectively. These results suggest that the permanent reduction in body mass (from overweight to normal weight) led to changes in human capital and productivity behaviors over the life cycle that generated additional contributions to wages that reflect a smaller total life-cycle body mass impact. In general, a weight improvement results in the greatest nominal wage increase for white and black women in the highest deciles of wages. However, black women are likely to experience a wage reduction in the wage deciles where they are more likely to be observed.

For comparison, the life-cycle impact of a one year increase in schooling is $1.16 and $1.03 for white and black women, respectively (Figure 9). Increases are larger at higher wages. A one year increase in full-time work experience increases wages by $0.40 and $0.41 for white and black women. The life-cycle effect of a *one unit* decrease in BMI on wages, on average, is $0.04 for white women and $-0.01 for black women.

Given our rich model, we can explicitly trace the avenues through which this simulated reduction in body mass suggests differences in the contemporaneous and life-cycle wage ef-fects. We observe in Figure 10 that the marginal effects of a body mass improvement on school attendance by age is positive during typical schooling ages (i.e., under age 28). The health improvement increases part-time employment over the life cycle, but reduces full-time employment (with particularly stronger effects on black women). Body mass improvements increase marriage rates at younger ages for white women more so than that for older white

41
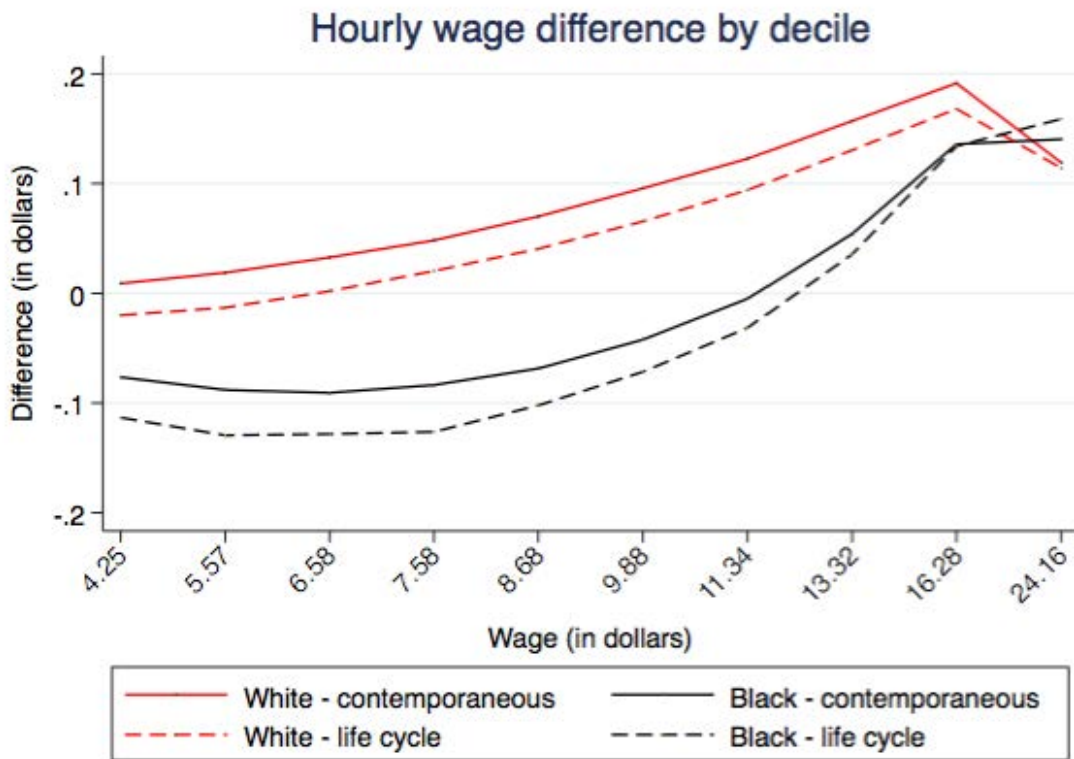
Figure 8: Life-Cycle Effect of Body Mass Improvement on Wages
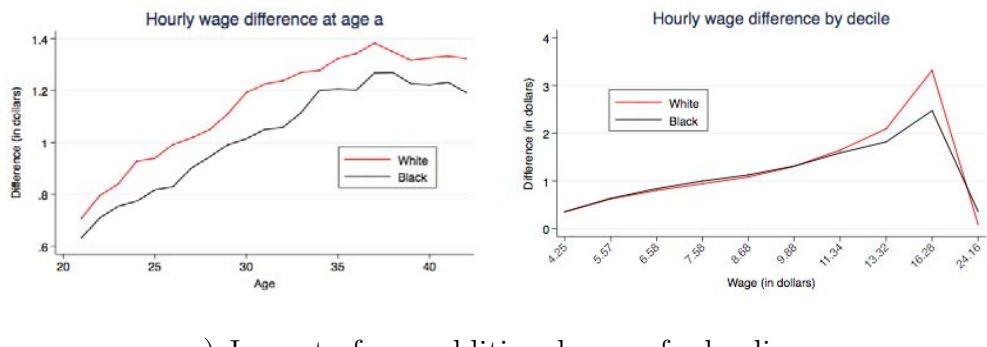
a.) Impact of one additional year of schooling
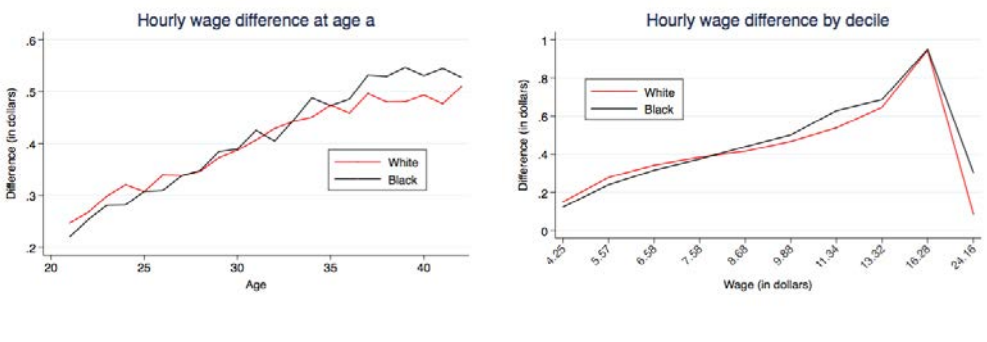
b.) Impact of one additional year of full-time work experience

Figure 9: Life-Cycle Effects of Human Capital on Wages, by Age and Decile

women. Black women, however, while experiencing increases in marriage rates with a re-
duction in body mass, see the largest gains after their late twenties. Reductions in body
mass lead to lower rates of adding children to the household for white women, while black
women's rate of child accumulation, while initially lower, actually increases slightly after age
28.
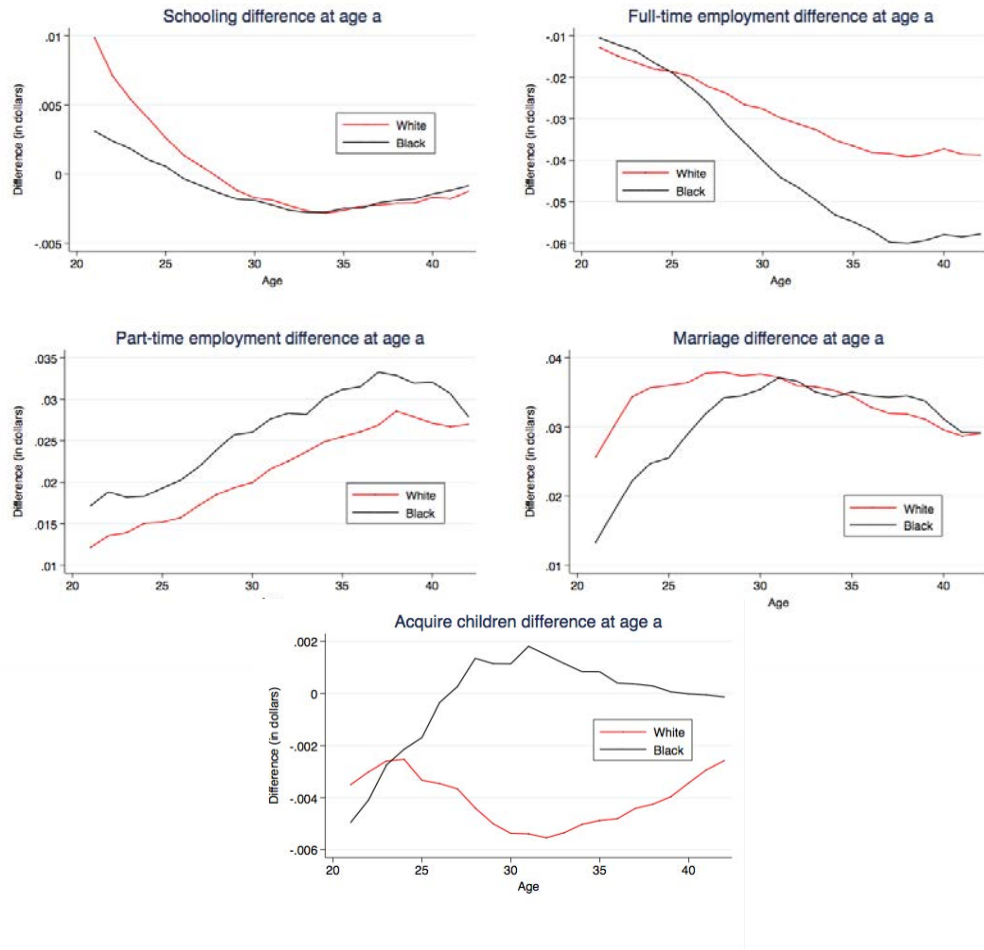
Figure 10: Life-Cycle Effects of Body Mass Improvement on Behaviors, by Age

Figure 11 displays the simulated wage differences of a permanent body mass improve-
ment from overweight to normal weight at every age over the life cycle on average wages.
Panel a. depicts wages unconditional on employment and panel b. conditions on being em-
ployed full- or part-time in both body mass simulations (i.e., normal and overweight).

To quantify these distributional impacts of wages across the life cycle, we examine the
impact of different life-cycle body mass trajectories on wages between the ages of 25 and

a.) Unconditional on Employment   b.) Conditional on Being Employed
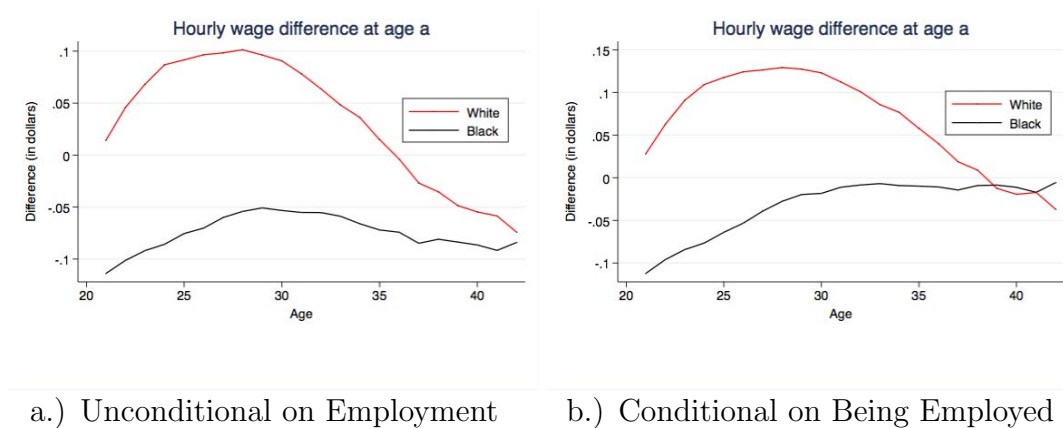
Figure 11: Effect of Body Mass Improvement on Average Wages by Age

40.[30] First, for each individual in the estimation sample replicated 50 times, we simulate their behavior each year regarding education, employment, marriage, and children, as well as body mass and wages. We call this our baseline, which represents the wage outcomes associated with endogenous behaviors and body mass evolution over the life cycle. We then force a body mass trajectory on each replicated individual. For example, we may force the individual to be normal weight over the age range (e.g., age 18 to 45). Table 10 describes the differences in average wages and total earnings over the ages of 25 through 40 for each of the imposed body mass trajectories.

In the baseline model simulation, body mass and behaviors are determined by the model estimates endogenously and updated each period. Thus, body mass is allowed to fluctuate as individuals age. In the simulations that follow, we force body weight to be a specified level each period, but continue to let behaviors and wages be determined endogenously according to the dynamic model. The average of the sum of wage draw differences when body mass is normal (BMI=24) at age 25 through age 40 relative to the baseline (when body mass is freely determined at each age) is negative.[31] However, what is relevant for our analysis is the

---

[30]Ostbye, et al. (2011) use the NLSY79 data to classify individuals into groups characterized by their BMI trajectory. These epidemiologists identify four trajectory groups: normal weight, overweight, late adulthood obesity, and early adulthood obesity. Grouping individuals together by gender, race, and cohort, they report membership rates of 35.0, 41.2, 19.7, and 4.2 percent for the four trajectory groupings, respectively. Males, blacks and those in the younger age cohort (aged 18 to 20 years in 1981) were more likely to be in the overweight, late overweight, and early overweight trajectory groups relative to the normal weight group.

[31]Note that the wage values in these calculations are draws from the wage distribution conditional on an individual's histories of human and health capital, regardless of employment decision (column one). In column two we average the sum of the wage differences conditional on choosing to work.

Table 10
Cumulative Average Wage and Total Earnings Differences by Simulated Body Mass Trajectory

	Average Cumulative Difference in Hourly Wages		Average Cumulative Difference in Total Earnings	
	unconditional on employment	conditional on employment	unconditional on employment	conditional on employment
White Females				
Baseline(weight determined by model)	12.60	13.42	403,301	383,294
Always normal weight	-1.97	-1.64	-3,936	-3,273
Always overweight	-2.80	-2.69	-5,604	-5,379
Always obese	-3.90	-3.84	-7,810	-7,681
Early normal weight	-2.20	-1.88	-4,410	-3,760
Early overweight	-3.45	-3.17	-6,892	-6,344
Always overweight vs always normal weight	-0.83	-1.04	-1,668	-2,083
Always obese vs always normal weight	-1.93	-2.17	-3,874	-4,341
Black Females				
Baseline(weight determined by model)	10.73	11.78	343,411	343,674
Always normal weight	-0.34	-0.04	-678	-75
Always overweight	0.68	0.32	1,360	635
Always obese	0.28	-0.13	556	-252
Early normal weight	-0.04	-0.19	-88	-380
Early overweight	0.52	0.01	1,045	11
Always overweight vs always normal weight	1.01	0.37	2,038	732
Always obese vs always normal weight	0.62	-0.09	1,234	-170

Note: Always refers to body mass from age 18 through age 40. Early refers to body mass up to age 25.
Total earnings calculations are based on full time (40 hours/week), full year (50 weeks per year)
employment over the 16 years between age 25 through age 40.

46

size of this difference when body mass is simulated to be overweight (BMI = 27.5) or obese (BMI = 32) relative to that difference when body mass is normal. For white females, the loss in hourly wages grows as body mass increases. Translated into total earnings over ages 25 through 40, we estimate a $1,668 reduction over 16 years if a white female is overweight versus normal weight throughout that age span (row three minus row two). If she is obese, she loses $3,874 (row four minus row two). These figures (column three) are based on the wage draws she receives. In column four, we report analogous values conditional on working (i.e., taking the offer).

Being normal weight (or overweight) from age 18 to age 25 (and then letting body mass evolve according to the model) vs being normal weight until age 40 is an additional penalty of $474 (row five minus row two) and $2,956 (row six minus row two), respectively. Finally, for white females, the reduction in total potential earnings (i.e., based on offered wages each period) for overweight (obese) body mass relative to normal body mass (rows seven and eight, respectively) from age 18 to 40 is $1,668 ($3,874). These wage penalties associated with body mass can reach as high as two percent of total earnings over this age range.

The lower panel of Table 10 reports these numbers for black females. The positive average wage difference between always overweight (row three) and always normal weight (row two) reflects the higher wages among overweight women, which characterizes black females in this age group. However, being obese (row five) and black results in lower wages than being overweight (row four) and black. Obesity among black women who choose to work amounts to a $170 loss in total earnings when she chooses to be employed (row eight). The wage penalties (or gains) for black women are tiny relative to total earnings over this age range. Recall, however, that these figures represent averages over the entire wage distribution. The wage penalties at the high end of the wage distribution are large and significant for both white and black women.

6   Conclusion

We find significant differences in the contemporaneous effect and the dynamic effect of body mass on wages, both across females of different races and over the distribution of wages. We highlight several differences between our model and those models frequently used in this literature: we use a richer incorporation of body mass, we account for selection into employment, we model endogenous body mass transitions, we jointly model endogenous variables (related to the history of schooling, employment, marriage, and children) that impact wages over time, we use random effects vs fixed effects to capture permanent heterogeneity which allows us to retain non-time-varying wage determinants, we include both permanent and time-varying individual unobserved heterogeneity, and we estimate the impact of determinants on the density of wages and body mass rather than simply the mean. Our model enables us to disentangle these two effects of interest. It also allows us to understand the mechanisms over the life cycle that contribute to these differences.

References

Averett, S., L. Argys, and J. Sorkin. 2013. "In Sickness and In Health: An Examination of Relationship Status and Health Using Data from the Canadian National Public Health Survey." *Review of Economics of the Household* 11(4): 599-633.

Averett, S. and S. Korenman S. 1996. "The economic reality of the beauty myth." *Journal of Human Resources* 31: 304-330.

Baum, C.L. and W.F. Ford. 2004. "The wage effects of obesity: a longitudinal study." *Health Economics* 13: 885-899.

Bernal, R. and M. Keane. 2011. "Child Care Choices and Childrens Cognitive Achievement: The Case of Single Mothers." *Journal of Labor Economics* 29(3): 459-512.

Bhattacharya, J. and M.K. Bundorf. 2005. "The incidence of the health care costs of obesity". NBER Working Paper # 11303.

Burkhauser, R. and J. Cawley. 2008. "Beyond BMI: The value of more accurate measures of fatness and obesity in social science research." *Journal of Health Economics* 27(2): 519-529.

Carpenter, C.S. 2006. "The effects of employment protection for obese people." *Industrial Relations* 45(3): 393-415.

Cawley, J. 2000. "An instrumental variables approach to measuring the effect of body weight on employment disability." *Health Services Research* 35(5): 1159-1179.

Cawley, J. 2004. "The Impact of Obesity on Wages." *The Journal of Human Resources* 39(2): 451-474.

Cawley, J. and C. Ruhm. 2012. "The Economics of Risky Health Behaviors." in <u>Handbook of Health Economics</u> Vol. 2 Book Series: Handbooks in Economics, edited by Pauly, M., McGuire, T., and Barros, P., p. 95-199.

Chen, A. 2012. "When Does Weight Matter Most?" *Journal of Health Economics* 31(1): 285-295.

Chiappori, P., S. Oreffice, and C. Quintana-Domeque. 2012. "Fatter Attraction: Anthropometric and Socioeconomic Matching on the Marriage Market." *Journal of Political Economy* 120(4): 659-695.

Conley, D. and G. Glauber. 2005. "Body mass and economic status." NBER working paper # 11343.

Crosnoe, R. and C. Muller. 2004. "Body mass index, academic achievement, and school context: examining the educational experience of adolescents at risk of obesity." *Journal of Health and Social Behavior* 45(4): 393-407.

Gilleskie, D. and T.A. Mroz. (2004)."A Flexible Approach for Estimating the Effects of Covariates on Health Expenditures." *Journal of Health Economics* 23(2): 391-418.

Grossman, M. 2009. "Effects of Weight on Children's Educational Achievement". *Economics of Education Review* Forthcoming.

Han E., E.C. Norton, and S.S. Stearns. 2009. "Weight and wages: Fat versus lean paychecks." *Health Economics* 18(5): 535-548.

Han E., E.C. Norton, and L.M. Powell. 2009. "Direct and indirect effects of teenage body weight on adult wages." NBER working paper #15027.

Han, E., E. Norton, and L. Powell. 2011. "Direct and Indirect Effect of Body Weight on Adult Wages." *Economics and Human Biology* 9(4): 381-392.

Harper B. 2000. "Beauty, Structure and the Labour Market: A British Cohort Study." *Oxford Bulletin of Economics and Statistics* 62: 771-800.

Heckman, James, Rodrigo Pinto, and Peter Savelyev. 2013. "Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes." *American Economic Review* 103(6): 2052-86.

Heckman, J. and B. Singer. 1983. "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data." *Econometrica* 52: 271-320.

Johansson, E., P. Bocherman, U. Kiiskinen, M. Heliovaara. 2009. "Obesity and Labour Market Success in Finland: The Difference between Having a High BMI and being Fat." *Economics and Human Biology* 7: 36-45.

Johar, M. and H. Katayama. 2012. "Quantile Regression Analysis ff Body Mass and Wages." *Health Economics* 21(5): 597-611.

Kaestner, R., M. Grossman, and B. Yarnoff. 2009. "Effects of Weight on Adolescent Educational Achievement". NBER working paper #14994.

Korenman, S. and D. Neumark. 1991. "Does Marriage Really Make Men More Productive?" *The Journal of Human Resources* 26(2): 282-307.

Lin, W., K. McEvilly, and J. Pantano. 2014."Obesity and Sex Ratios in the U.S.", working paper.

Malcolm, M. and K. Ilker. 2014. "Selection Works Both Ways: BMI and Marital Formation Among Young Women." *Review of Economics of the Household*, forthcoming; published online (DOI 10.1007/s11150-014-9247-8).

Mincer, J. 1974. Schooling, Experience, and Earnings. NBER and Columbia Press.

Morris, S. 2007. "The Impact of Obesity on Employment." *Labour Economics* 14: 413-433.

Mroz, T.A. 1999. "Discrete Factor Approximation in Simultaneous Equation Models: Estimating the Impact of a Dummy Endogenous Variable on a Continuous Outcome." *Journal of Econometrics* 92: 233-74.

Norton, E.C. and E. Han. 2008. "Genetic Information, Obesity, and Labor Market Outcomes." *Health Economics* 17: 1089-1104.

Ostbye, T., R. Malhotra, and L. R. Landerman. 2011. "Body Mass Trajectories Through Adulthood: Results from the National Longitudinal Survey of Youth 1979 Cohort (19812006)" *International Journal of Epidemiology* (2011) 40 (1): 240-250.

Pagan, J.A. and A. Davila. 1997. "Obesity, Occupational Attainment, and Earnings." *Social Science Quarterly* 78(3): 756-770.

Paraponaris, A., B. Saliba, and B. Ventelou. 2005. "Obesity, weight status and employability: empirical evidence from a French National Survey." *Economics and Human Biology* 3(2): 241-258.

Pinkston, J. 2015. "The Dynamic Effects of Obesity on the Wages of Young Workers." working paper.

Sabia, J.J. 2007. "The Effect of Body Weight on Adolescent Academic Performance." *Southern Economic Journal* 73(4): 871-900.

Sabia, J.J. and D. Rees. 2012. "Body Weight and Wages: Evidence from Add Health." *Economics and Human Biology* 10(1): 14-19.

Wada, R. and E. Tekin. 2010. "Body composition and Wages." *Economics and Human Biology* 8(2):242-254.

Table A1

Estimation Results: Wages (Conditional Density Estimation)

Variable name	Coeff	Std Error	
Constant	10.221	0.963	***
BMI in t	-0.231	0.047	***
BMI in t × Black race	0.193	0.061	***
White: standardized deviations from mean BMI in t	1.286	0.247	***
Black: standardized deviations from mean BMI in t	0.465	0.365	
Height in inches (- 64 inches)	-0.077	0.014	***
Height in inches (- 64 inches) squared/100	0.030	0.033	
Interval × BMI in t	-0.163	0.028	***
Interval × BMI in t × Black race	0.118	0.038	***
Interval × White: standardized deviations from mean BMI in t	0.850	0.136	***
Interval × Black: standardized deviations from mean BMI in t	0.447	0.194	**
Interval × Height in inches (- 64 inches)	-0.024	0.007	***
Interval × Height in inches (- 64 inches) squared/100	-0.001	0.017	
Enrolled in t-1	0.258	0.136	*
Enrolled in t-1 × Black race	0.141	0.335	
Years enrolled in school entering t	-0.769	0.050	***
Years enrolled in school entering t × Black race	0.206	0.084	**
Years enrolled ≥ 12 entering t	1.196	0.319	***
Years enrolled ≥ 12 entering t × Black race	-2.754	0.534	***
Years enrolled ≥ 16 entering t	0.943	0.196	***
Years enrolled ≥ 16 entering t × Black race	-0.521	0.407	
Employed in t-1	0.417	0.534	
Employed in t-1 × Black race	0.731	0.656	
Employed part time in t-1	-0.807	0.117	***
Employed part time in t-1 × Black race	0.072	0.258	
Years employed entering t	-0.175	0.025	***
Years employed entering t × Black race	-0.087	0.039	**
Years part time employed entering t	-0.008	0.018	
Years part time employed entering t × Black race	-0.079	0.031	***
Married in t-1	-0.218	0.221	
Married in t-1 × Black race	0.313	0.603	
Years married entering t if married in t-1	0.046	0.017	***
Years married entering t if married in t-1 × Black race	-0.041	0.048	
Years single entering t if single in t-1	0.045	0.023	**
Years single entering t if single in t-1 × Black race	-0.094	0.031	***
Number of children in hh entering t	0.136	0.071	*
Number of children in hh entering t × Black race	0.020	0.131	
Acquire any children in hh in t-1	0.025	0.239	
Acquire any children in hh in t-1 × Black race	-0.052	0.620	
Lose any children from hh in t-1	-0.052	0.478	
Lose any children from hh in t-1 × Black race	-0.038	0.599	
Any children under age 6 in t-1	-0.079	0.178	
Any children under age 6 in t-1 × Black race	0.114	0.391	

- continues on next page

Table A1 - continued
Estimation Results: Wages (Conditional Density Estimation)

Variable name	Coeff	Std Error	
Interval × Enrolled in t-1	-0.104	0.081	
Interval × Enrolled in t-1 × Black race	0.112	0.188	
Interval × Years enrolled in school entering t	-0.147	0.028	***
Interval × Years enrolled in school entering t × Black race	0.143	0.048	***
Interval × Years enrolled ≥ 12 entering t	0.477	0.166	***
Interval × Years enrolled ≥ 12 entering t × Black race	-1.267	0.272	***
Interval × Years enrolled ≥ 16 entering t	0.429	0.116	***
Interval × Years enrolled ≥ 16 entering t × Black race	-0.188	0.235	
Interval × Employed in t-1	0.913	0.276	***
Interval × Employed in t-1 × Black race	0.132	0.330	
Interval × Employed part time in t-1	-0.794	0.066	***
Interval × Employed part time in t-1 × Black race	0.071	0.139	
Interval × Years employed entering t	-0.001	0.013	
Interval × Years employed entering t × Black race	-0.034	0.021	*
Interval × Years part time employed entering t	-0.061	0.010	***
Interval × Years part time employed entering t × Black race	-0.024	0.018	
Interval × Married in t-1	-0.153	0.130	
Interval × Married in t-1 × Black race	0.208	0.340	
Interval × Years married entering t if married in t-1	0.020	0.010	*
Interval × Years married entering t if married in t-1 × Black race	-0.024	0.027	
Interval × Years single entering t if single in t-1	0.012	0.014	
Interval × Years single entering t if single in t-1 × Black race	-0.043	0.019	**
Interval × Number of children in hh entering t	-0.032	0.042	
Interval × Number of children in hh entering t × Black race	0.054	0.073	
Interval × Acquire any children in hh in t-1	0.066	0.140	
Interval × Acquire any children in hh in t-1 × Black race	-0.092	0.345	
Interval × Lose any children from hh in t-1	-0.206	0.273	
Interval × Lose any children from hh in t-1 × Black race	0.026	0.331	
Interval × Any children under age 6 in t-1	-0.034	0.105	
Interval × Any children under age 6 in t-1 × Black race	0.124	0.221	

- continues on next page

Table A1 - continued
Estimation Results: Wages (Conditional Density Estimation)

Variable name	Coeff	Std Error	
Black race	-4.348	1.134	***
Age in years - 18	0.096	0.026	***
Black × age	0.045	0.041	
Rural residence	0.124	0.083	
Residence type missing	-0.793	0.239	***
Northeast region	-0.589	0.178	***
North central region	-0.351	0.176	**
West region	-0.566	0.174	***
AFQT score minus median	-0.006	0.003	*
AFQT score squared	-0.001	0.001	
AFQT score cubed	0.000	0.000	
AFQT score missing	-1.193	0.335	***
Interval × Black race	-3.225	0.865	***
Interval × Age in years - 18	0.040	0.013	***
Interval × Black × age	0.014	0.022	
Interval × Rural residence	-0.012	0.048	
Interval × Residence type missing	-0.457	0.137	***
Interval × Northeast region	0.066	0.086	
Interval × North central region	-0.068	0.079	
Interval × West region	-0.086	0.076	
Interval × AFQT score minus median	0.002	0.002	
Interval × AFQT score squared	-0.001	0.001	
Interval × AFQT score cubed	0.000	0.000	
Interval × AFQT score missing	-0.431	0.196	**
Unemployment rate	0.058	0.013	***
Total employment per capita	2.139	0.567	***
Ratio of manuf empl to total empl	-1.621	2.286	
Ratio of service empl to total empl	11.515	1.692	***
Total earnings per employee	-0.035	0.009	***
Ratio of manuf earnings to total earnings	4.397	1.668	***
Ratio of service earnings to total earnings	-14.934	1.625	***
Interval × Unemployment rate	0.014	0.008	*
Interval × Total employment per capita	0.999	0.341	***
Interval × Ratio of manuf empl to total empl	-2.086	1.303	
Interval × Ratio of service empl to total empl	-6.880	1.174	***
Interval × Total earnings per employee	0.013	0.005	***
Interval × Ratio of manuf earnings to total earnings	1.825	0.966	*
Interval × Ratio of service earnings to total earnings	1.763	1.086	
Time trend (1=1984)	-0.282	0.031	***
Time trend squared/100	1.227	0.317	***
Time trend cubed/1000	-0.216	0.106	**
Time trend × Northeast region	-0.019	0.007	***
Time trend × North central region	-0.026	0.007	***
Time trend × West region	-0.008	0.008	
Interval measure for density estimation	7.504	0.700	***
Interval squared	0.472	0.151	***
Interval cubed	0.313	0.043	***
UH Factor loading: permanent	7.478	0.149	***
UH Factor loading: time-varying	8.294	0.682	***

Table A2

Estimation Results: Employment Status (relative to full time)

Variable name	Part time			Non-employed		
	Coeff	Std Err		Coeff	Std Err	
Constant	2.286	1.010	**	-5.875	2.245	***
BMI in t	0.269	0.127	**	0.200	0.141	
BMI squared/100 in t	-1.009	0.455	**	-0.819	0.541	
BMI cubed/1000 in t	0.107	0.048	**	0.096	0.056	*
BMI in t × Black race	-0.543	0.137	***	-0.325	0.166	*
BMI squared/100 in t × Black race	1.462	0.487	***	0.646	0.631	
BMI cubed/1000 in t × Black race	-0.140	0.054	***	-0.058	0.070	
Ever overweight prior to t	-0.048	0.048		0.000	0.073	
Ever obese prior to t	0.063	0.080		0.108	0.124	
White: standardized deviations from mean BMI in t	0.123	0.154		0.060	0.257	
Black: standardized deviations from mean BMI in t	0.430	0.173	***	0.593	0.309	*
Height in inches (- 64 inches)	-0.015	0.005	***	-0.013	0.008	
Height in inches (- 64 inches) squared/100	-0.004	0.011		0.013	0.015	
Enrolled in t-1	0.067	0.052		-0.610	0.090	***
Years enrolled in school entering t	-0.124	0.017	***	-0.114	0.026	***
Years enrolled ≥ 12 entering t	0.092	0.060		-0.159	0.085	*
Years enrolled ≥ 16 entering t	0.048	0.073		0.219	0.121	*
Employed in t-1	-2.894	0.073	***	-6.750	0.099	***
Employed part time in t-1	2.188	0.034	***	3.287	0.073	***
Years employed entering t	-0.080	0.006	***	-0.119	0.009	***
Years part time employed entering t	0.125	0.006	***	0.029	0.010	***
Married in t-1	0.367	0.047	***	0.515	0.076	***
Years married entering t if married in t-1	-0.001	0.005		-0.023	0.008	***
Years single entering t if single in t-1	0.005	0.007		-0.006	0.011	
Number of children in hh entering t	-0.015	0.018		-0.067	0.026	***
Acquire any children in hh in t-1	0.022	0.056		0.239	0.075	***
Lose any children from hh in t-1	0.033	0.112		0.033	0.189	
Any children under age 6 in t-1	0.288	0.040	***	0.490	0.060	***

- continues on next page

Table A2 – continued
Estimation Results: Employment Status (relative to full time)

Variable name	Part time			Non-employed		
	Coeff	Std Err		Coeff	Std Err	
Black race	6.113	1.080	***	4.571	1.360	***
Age in years - 18	-0.305	0.045	***	-0.075	0.061	
Age - 18 squared/100	0.267	0.038	***	0.136	0.052	***
Age - 18 cubed/1000	-0.063	0.009	***	-0.035	0.013	***
Rural residence	-0.021	0.040		0.057	0.058	
Residence type missing	0.129	0.093		0.143	0.149	
Northeast region	0.326	0.136	**	0.365	0.208	*
North central region	0.171	0.143		0.165	0.147	
West region	0.097	0.095		-0.024	0.192	
AFQT score minus median	0.000	0.001		-0.004	0.002	**
AFQT score squared	0.001	0.000	***	0.003	0.001	***
AFQT score cubed	0.000	0.000		0.000	0.000	*
AFQT score missing	0.265	0.143	*	0.260	0.253	
Two year college semester tuition in 000s	0.061	0.053		0.037	0.060	
Four year college semester tuition in 000s	-0.081	0.050		-0.121	0.066	*
Graduate school semester tuition in 000s	0.042	0.036		0.021	0.062	
Unemployment rate	0.004	0.010		0.049	0.013	***
Total employment per capita	-2.151	0.710	***	-3.669	0.869	***
Ratio of manuf empl to total empl	0.028	0.959		-0.960	0.947	
Ratio of service empl to total empl	-1.954	0.944	**	-1.621	1.051	
Total earnings per employee	-0.005	0.010		-0.005	0.012	
Ratio of manuf earnings to total earnings	0.000	0.908		0.304	0.838	
Ratio of service earnings to total earnings	1.092	0.927		2.356	1.054	**
Total population in 000,000s	-0.004	0.001	***	-0.007	0.002	***
Total population in 000,000s squared/100	0.001	0.000	***	0.002	0.001	***
White gender ratio: male, 20-60/female, 15-50	-0.042	0.751		0.648	0.795	
Black gender ratio: male, 20-60/female, 15-50	0.512	0.762		1.313	0.856	***
AFDC per month per family of four in 00s	-0.014	0.018		-0.031	0.029	
Household income in 000s	0.010	0.006	*	0.016	0.008	**
Mean price of liter of liquor	0.032	0.009	***	0.035	0.013	***
Mean price of 6-pack of beer	0.031	0.047		-0.075	0.067	*
Mean price of bottle of wine	-0.041	0.026		-0.019	0.040	
Mean price of carton of cigarettes	0.004	0.010		-0.004	0.014	
Mean price of food	-0.262	0.408		-0.264	0.771	
Mean price of junk food	0.012	0.176		-0.168	0.297	
Ratio of food sales to total retail sales	-0.551	1.005		-2.069	1.096	***
Ratio of restaurant sales to total retail sales	8.751	1.141	***	11.542	1.439	***
Indicator of missing state of residence	-0.334	0.796		1.348	0.920	
Indicator of missing price data	0.143	0.820		-1.189	0.867	
Time trend (1=1984)	0.049	0.063		-0.006	0.080	
Time trend squared/100	-1.318	0.727	*	-0.816	0.894	
Time trend cubed/1000	0.512	0.252	**	0.520	0.311	**
UH Factor loading: permanent	0.615	0.085	***	0.677	0.162	***
UH Factor loading: time-varying	-1.682	0.177	***	16.538	2.200	***

Table A3

Estimation Results: School Enrollment and Marital Status

Variable name	Enrolled			Married		
	Coeff	Std Err		Coeff	Std Err	
Constant	-2.625	1.017	***	-7.606	0.995	***
BMI in t	-0.285	0.141	**	0.351	0.146	**
BMI squared/100 in t	1.166	0.568	**	-0.714	0.674	
BMI cubed/1000 in t	-0.122	0.064	*	0.055	0.075	
BMI in t × Black race	0.411	0.137	***	-0.094	0.147	
BMI squared/100 in t × Black race	-1.385	0.540	***	0.313	0.673	
BMI cubed/1000 in t × Black race	0.143	0.063	**	-0.032	0.080	
Ever overweight prior to t	-0.007	0.077		-0.225	0.068	***
Ever obese prior to t	-0.124	0.152		-0.324	0.148	**
White: standardized deviations from mean BMI in t	-0.303	0.262		-0.284	0.357	
Black: standardized deviations from mean BMI in t	-0.255	0.335		-0.337	0.412	
Height in inches (- 64 inches)	-0.004	0.008		-0.001	0.007	
Height in inches (- 64 inches) squared/100	0.005	0.017		-0.020	0.016	
Enrolled in t-1	2.731	0.050	***	-0.400	0.072	***
Years enrolled in school entering t	0.207	0.024	***	0.114	0.023	***
Years enrolled ≥ 12 entering t	-0.251	0.105	**	-0.003	0.081	
Years enrolled ≥ 16 entering t	-0.319	0.090	***	-0.161	0.097	*
Employed in t-1	-0.272	0.085	***	0.004	0.077	
Employed part time in t-1	0.258	0.060	***	0.076	0.055	
Years employed entering t	0.023	0.012	*	0.030	0.009	***
Years part time employed entering t	-0.016	0.012		-0.020	0.009	**
Married in t-1	-0.157	0.072	**	4.686	0.060	***
Years married entering t if married in t-1	-0.005	0.010		0.102	0.008	***
Years single entering t if single in t-1	0.021	0.012	*	0.007	0.009	
Number of children in hh entering t	0.149	0.035	***	0.015	0.025	
Acquire any children in hh in t-1	-0.222	0.100	**	0.157	0.082	*
Lose any children from hh in t-1	0.264	0.913		-0.075	0.150	
Any children under age 6 in t-1	-0.383	0.071	***	0.025	0.060	

Variable name	Enrolled			Married		
	Coeff	Std Err		Coeff	Std Err	
Black race	-3.771	1.009	***	0.193	1.004	
Age in years - 18	-0.329	0.058	***	0.047	0.057	
Age - 18 squared/100	0.178	0.055	***	-0.076	0.049	
Age - 18 cubed/1000	-0.039	0.015	***	0.017	0.012	
Rural residence	-0.132	0.065	**	0.156	0.056	***
Residence type missing	0.046	0.145		0.055	0.141	
Northeast region	0.039	0.194		0.063	0.177	
North central region	-0.050	0.170		-0.089	0.155	
West region	0.290	0.144	**	-0.224	0.147	
AFQT score minus median	0.013	0.002	***	0.002	0.002	
AFQT score squared	-0.002	0.001	***	0.000	0.001	
AFQT score cubed	0.000	0.000		0.000	0.000	
AFQT score missing	-0.012	0.358		0.283	0.160	*
Two year college semester tuition in 000s	-0.026	0.080		0.040	0.067	
Four year college semester tuition in 000s	0.036	0.078		-0.007	0.066	
Graduate school semester tuition in 000s	-0.031	0.063		0.009	0.063	
Unemployment rate	-0.010	0.014		-0.006	0.013	
Total employment per capita	0.676	0.972		-0.772	0.849	
Ratio of manuf empl to total empl	-1.553	0.987		0.625	0.987	
Ratio of service empl to total empl	-2.945	1.048	***	0.835	1.042	
Total earnings per employee	0.031	0.013	**	0.002	0.013	
Ratio of manuf earnings to total earnings	1.777	0.964	*	-1.932	0.962	**
Ratio of service earnings to total earnings	3.559	1.031	***	-3.456	1.026	***
Total population in 000,000s	0.000	0.002		0.000	0.002	
Total population in 000,000s squared/100	0.000	0.001		0.000	0.001	
White gender ratio: male, 20-60/female, 15-50	1.246	0.953		1.565	0.945	*
Black gender ratio: male, 20-60/female, 15-50	1.577	0.900	*	1.620	0.893	*
AFDC per month per family of four in 00s	-0.011	0.027		0.004	0.025	
Household income in 000s	-0.027	0.009	***	-0.004	0.008	
Mean price of liter of liquor	-0.003	0.012		0.000	0.012	
Mean price of 6-pack of beer	-0.136	0.081	*	-0.059	0.070	
Mean price of bottle of wine	-0.005	0.040		-0.022	0.036	
Mean price of carton of cigarettes	0.059	0.015	***	-0.001	0.014	
Mean price of food	0.064	0.905		0.089	0.863	
Mean price of junk food	-0.250	0.302		0.039	0.322	
Ratio of food sales to total retail sales	2.379	1.037	**	-1.474	0.998	
Ratio of restaurant sales to total retail sales	-4.876	1.163	***	-1.486	1.009	
Indicator of missing state of residence	1.667	0.917	*	-0.075	0.948	
Indicator of missing price data	-0.833	0.944		-0.017	0.887	
Time trend (1=1984)	-0.019	0.075		-0.003	0.076	
Time trend squared/100	1.394	0.932		-0.195	0.875	
Time trend cubed/1000	-0.904	0.354	***	0.071	0.309	
UH Factor loading: permanent	-0.034	0.120		-0.426	0.110	***
UH Factor loading: time-varying	-0.415	0.143	***	-0.127	0.145	

Table A4
Estimation Results: Child Accumulation (relative to none)

Variable name	Gain			Lose		
	Coeff	Std Err		Coeff	Std Err	
Constant	-5.865	1.009	***	0.851	1.031	
BMI in t	0.408	0.145	***	-0.298	0.176	*
BMI squared/100 in t	-0.817	0.675		0.845	0.763	
BMI cubed/1000 in t	0.053	0.076		-0.081	0.083	
BMI in t × Black race	-0.267	0.138	*	0.417	0.166	***
BMI squared/100 in t × Black race	0.571	0.674		-1.217	0.765	
BMI cubed/1000 in t × Black race	-0.044	0.082		0.127	0.086	
Ever overweight prior to t	-0.099	0.059	*	0.075	0.125	
Ever obese prior to t	-0.226	0.136	*	-0.023	0.144	
White: standardized deviations from mean BMI in t	-0.173	0.369		0.039	0.536	
Black: standardized deviations from mean BMI in t	0.050	0.387		-0.284	0.642	
Height in inches (- 64 inches)	0.002	0.007		-0.003	0.013	
Height in inches (- 64 inches) squared/100	-0.008	0.014		-0.047	0.028	*
Enrolled in t-1	-0.762	0.081	***	0.318	0.230	
Years enrolled in school entering t	0.046	0.022	**	-0.254	0.067	***
Years enrolled ≥ 12 entering t	-0.167	0.071	**	-0.103	0.191	
Years enrolled ≥ 16 entering t	-0.061	0.090		-0.162	0.557	
Employed in t-1	-1.169	0.062	***	0.369	0.112	***
Employed part time in t-1	0.727	0.050	***	-0.101	0.091	
Years employed entering t	0.049	0.009	***	-0.002	0.010	
Years part time employed entering t	-0.022	0.009	**	0.010	0.013	
Married in t-1	1.517	0.056	***	-0.142	0.117	
Years married entering t if married in t-1	-0.088	0.008	***	-0.033	0.010	***
Years single entering t if single in t-1	0.046	0.010	***	0.005	0.010	
Number of children in hh entering t	-0.363	0.026	***	0.548	0.030	***
Acquire any children in hh in t-1	-0.484	0.062	***	0.514	0.105	***
Lose any children from hh in t-1	1.035	0.104	***	0.308	0.162	*
Any children under age 6 in t-1	0.548	0.051	***	0.173	0.083	**

- continues on next page

Variable name	Gain			Lose		
	Coeff	Std Err		Coeff	Std Err	
Black race	4.799	1.008	***	-3.576	0.998	***
Age in years - 18	-0.006	0.058		-0.308	0.118	***
Age - 18 squared/100	-0.010	0.052		0.329	0.084	***
Age - 18 cubed/1000	-0.008	0.014		-0.066	0.019	***
Rural residence	0.038	0.047		0.166	0.084	**
Residence type missing	0.089	0.124		0.357	0.162	**
Northeast region	0.221	0.147		-0.057	0.500	
North central region	0.153	0.130		-0.064	0.309	
West region	0.129	0.134		-0.410	0.443	
AFQT score minus median	0.000	0.002		-0.003	0.003	
AFQT score squared	0.001	0.001	**	-0.002	0.001	**
AFQT score cubed	0.000	0.000	*	0.000	0.000	
AFQT score missing	0.245	0.121	**	-0.363	0.691	
Two year college semester tuition in 000s	0.094	0.056	*	-0.013	0.020	
Four year college semester tuition in 000s	-0.087	0.059		1.154	1.019	
Graduate school semester tuition in 000s	0.067	0.055		-4.953	1.807	***
Unemployment rate	-0.002	0.012		-10.189	2.294	***
Total employment per capita	-0.659	0.800		-0.003	0.019	
Ratio of manuf empl to total empl	0.061	0.950		0.858	1.352	
Ratio of service empl to total empl	1.954	1.045	*	9.244	1.956	***
Total earnings per employee	-0.005	0.011		0.092	0.095	
Ratio of manuf earnings to total earnings	-0.931	0.883		-0.024	0.114	
Ratio of service earnings to total earnings	-3.695	1.022	***	-0.032	0.079	
Total population in 000,000s	-0.002	0.002		-0.002	0.003	
Total population in 000,000s squared/100	0.001	0.001	*	0.001	0.001	
White gender ratio: male, 20-60/female, 15-50	0.044	0.817		1.519	0.966	
Black gender ratio: male, 20-60/female, 15-50	-0.967	0.823		0.375	0.936	
AFDC per month per family of four in 00s	-0.042	0.023	*	0.018	0.060	
Household income in 000s	0.005	0.007		-0.019	0.010	*
Mean price of liter of liquor	-0.017	0.010		0.006	0.021	
Mean price of 6-pack of beer	-0.030	0.067		0.032	0.067	
Mean price of bottle of wine	-0.036	0.032		-0.033	0.087	
Mean price of carton of cigarettes	0.005	0.012		0.028	0.017	
Mean price of food	-0.121	0.789		0.336	0.610	
Mean price of junk food	0.004	0.288		0.072	0.297	
Ratio of food sales to total retail sales	-4.344	1.023	***	-0.419	1.119	
Ratio of restaurant sales to total retail sales	2.431	1.179	**	-5.268	1.717	***
Indicator of missing state of residence	-3.002	0.879	***	-1.817	0.957	*
Indicator of missing price data	-0.730	0.891		1.985	0.972	**
Time trend (1=1984)	0.211	0.075	***	-0.085	0.117	
Time trend squared/100	-2.710	0.884	***	-0.264	0.953	
Time trend cubed/1000	0.798	0.321	***	0.082	0.306	
UH Factor loading: permanent	-0.352	0.105	***	0.812	0.384	**
UH Factor loading: time-varying	-0.139	0.116		0.353	0.495	

Table A5

Estimation Results: Body Mass Index at $t+1$ (Conditional Density Estimation)

Variable name	Coeff	Std Err	
Constant	-153.183	5.403	***
BMI in t	13.467	0.447	***
BMI squared/100 in t	-35.735	1.197	***
BMI cubed/1000 in t	2.757	0.104	***
BMI in t × Black race	3.279	0.510	***
BMI squared/100 in t × Black race	-10.422	1.441	***
BMI cubed/1000 in t × Black race	1.102	0.132	***
Ever overweight prior to t	-0.091	0.191	
Ever obese prior to t	-0.854	0.136	***
Height in inches (- 64 inches)	-0.012	0.008	
Height in inches (- 64 inches) squared/100	-0.073	0.017	***
Interval × BMI in t	5.656	0.189	***
Interval × BMI squared/100 in t	-11.353	0.540	***
Interval × BMI cubed/1000 in t	0.695	0.051	***
Interval × BMI in t × Black race	0.724	0.197	***
Interval × BMI squared/100 in t × Black race	-1.901	0.591	***
Interval × BMI cubed/1000 in t × Black race	0.161	0.062	***
Interval × Ever overweight prior to t	0.115	0.115	
Interval × Ever obese prior to t	-0.611	0.124	***
Interval × Height in inches (- 64 inches)	-0.011	0.005	**
Interval × Height in inches (- 64 inches) squared/100	-0.037	0.011	***
Enrolled in t	-0.067	0.121	
Years enrolled in school at end of t	-0.002	0.017	
Freshmen year of college in t	0.393	0.822	
Employed in t	0.125	0.066	*
Employed part time in t	-0.060	0.060	
Married in t	-0.235	0.077	***
Years married at end of t if married in t	0.017	0.007	**
Years single at end of t if single in t	-0.001	0.008	
Number of children in hh at end of t	-0.003	0.026	
Acquire any children in hh in t	0.154	0.092	*
Lose any children from hh in t	-0.284	0.401	
Any children under age 6 in t-1	0.143	0.061	**
Interval × Enrolled in t	-0.072	0.069	
Interval × Years enrolled in school at end of t	-0.012	0.010	
Interval × Freshmen year of college in t	0.260	0.442	
Interval × Employed in t	0.030	0.040	
Interval × Employed part time in t	0.003	0.036	
Interval × Married in t	0.019	0.046	
Interval × Years married at end of t if married in t	-0.001	0.004	
Interval × Years single at end of t if single in t	0.003	0.006	
Interval × Number of children in hh at end of t	-0.016	0.016	
Interval × Acquire any children in hh in t	0.050	0.055	
Interval × Lose any children from hh in t	-0.203	0.248	
Interval × Any children under age 6 in t-1	0.031	0.037	

- continues on next page

Variable name	Coeff	Std Err	
Black race	-34.480	5.814	***
Age in years - 18	0.017	0.010	*
Black × age	0.012	0.009	
Rural residence	0.034	0.059	
Residence type missing	0.014	0.126	
Northeast region	-0.019	0.109	
North central region	0.002	0.084	
West region	0.110	0.136	
AFQT score minus median	0.000	0.002	
AFQT score squared	0.000	0.001	
AFQT score cubed	0.000	0.000	
AFQT score missing	0.071	0.260	
Interval × Black race	-8.823	2.229	***
Interval × Age in years - 18	0.009	0.006	
Interval × Black × age	0.004	0.005	
Interval × Rural residence	0.019	0.035	
Interval × Residence type missing	0.057	0.075	
Interval × Northeast region	0.005	0.057	
Interval × North central region	0.076	0.041	*
Interval × West region	0.120	0.066	*
Interval × AFQT score minus median	-0.001	0.001	
Interval × AFQT score squared	0.000	0.000	
Interval × AFQT score cubed	0.000	0.000	
Interval × AFQT score missing	0.006	0.151	
Mean price of liter of liquor	0.026	0.009	***
Mean price of 6-pack of beer	-0.045	0.043	
Mean price of bottle of wine	0.064	0.038	*
Mean price of carton of cigarettes	0.014	0.008	*
Mean price of food	0.162	0.700	
Mean price of junk food	-0.215	0.266	
Ratio of food sales to total retail sales	1.023	0.895	
Ratio of restaurant sales to total retail sales	-1.440	0.890	
Indicator of missing state of residence	0.457	0.683	
Indicator of missing price data	0.409	0.808	
Interval × Mean price of liter of liquor	0.016	0.005	***
Interval × Mean price of 6-pack of beer	-0.037	0.027	
Interval × Mean price of bottle of wine	0.034	0.022	
Interval × Mean price of carton of cigarettes	0.002	0.005	
Interval × Mean price of food	0.338	0.390	
Interval × Mean price of junk food	-0.102	0.160	
Interval × Ratio of food sales to total retail sales	0.838	0.614	
Interval × Ratio of restaurant sales to total retail sales	-1.555	0.647	**
Interval × Indicator of missing state of residence	0.184	0.402	
Interval × Indicator of missing price data	0.618	0.486	
Time trend (1=1984)	-0.155	0.028	***
Time trend squared/100	1.539	0.326	***
Time trend cubed/1000	-0.457	0.116	***
Time trend × Northeast region	0.005	0.005	
Time trend × North central region	0.008	0.004	**
Time trend × West region	0.007	0.006	
Interval measure for density estimation	-79.514	2.222	***
Interval squared	3.148	0.337	***
Interval cubed	2.260	0.086	***
UH Factor loading: permanent	-0.059	0.040	
UH Factor loading: time-varying	-0.134	0.062	**

Table A6
Estimation Results: Attrition at end of t + 1

Variable name	Coeff	Std Err	
Constant	-2.444	0.394	***
BMI in t	-0.018	0.012	
BMI in t × Black race	0.019	0.012	
Ever overweight prior to t	0.056	0.083	
Ever obese prior to t	-0.244	0.121	**
Height in inches (- 64 inches)	0.013	0.011	
Height in inches (- 64 inches) squared/100	-0.030	0.025	
Enrolled in t	0.814	0.088	***
Employed in t	-0.709	0.077	***
Employed part time in t	0.037	0.083	
Married in t	-0.049	0.065	
Number of children in hh at end of t	-0.069	0.032	**
Acquire any children in hh in t	-0.334	0.126	***
Lose any children from hh in t	-0.265	0.566	
Any children under age 6 in t-1	-0.229	0.076	***
Black race	-0.267	0.320	
Age in years - 18	0.172	0.079	**
Age - 18 squared/100	-0.117	0.066	*
Age - 18 cubed/1000	0.023	0.016	
Rural residence	0.001	0.073	
Residence type missing	-0.818	0.210	***
Northeast region	-0.023	0.083	
North central region	-0.050	0.075	
West region	0.190	0.089	**
AFQT score minus median	0.003	0.002	
AFQT score squared	-0.003	0.001	***
AFQT score cubed	0.000	0.000	*
AFQT score missing	0.406	0.382	
Indicator of missing state of residence	0.981	0.401	**
Time trend (1=1984)	-0.204	0.078	***
Time trend squared/100	2.543	0.942	***
Time trend cubed/1000	-0.700	0.323	**
UH Factor loading: permanent	-0.355	0.211	*
UH Factor loading: time-varying	-0.468	0.335	