Human Capital and Economic Opportunity:
A Global Working Group

Working Paper Series

Working Paper No. 2012-018

Caught in the Bulimic Trap? Persistence and State Dependence of Bulimia Among Young Women

John Ham
Daniela Iorio
Michelle Sovinsky

September, 2012

Human Capital and Economic Opportunity Working Group
Economic Research Center
University of Chicago
1126 E. 59th Street
Chicago IL 60637
humcap@uchicago.edu

Caught in the Bulimic Trap?
Persistence and State Dependence of Bulimia Among Young Women

John C. Ham, Daniela Iorio, and Michelle Sovinsky[1]

August 15, 2012

First version: December 2008

Abstract

Eating disorders are an important and growing health concern, and bulimia nervosa (BN) accounts for the largest fraction of eating disorders. Health consequences of BN are substantial and especially serious given the increasingly compulsive nature of the disorder. However, remarkably little is known about the mechanisms underlying the persistent nature of BN. We use data from a unique panel data set, the National Heart, Lung, and Blood Institute Growth and Health Study, which was conducted for ten years on young women aged 9-10 at the start of the survey (in 1987). Using instrumental variable techniques, we document that unobserved heterogeneity plays a role in the persistence of BN, but, strikingly, up to two-thirds of this persistence is due to true state dependence. Our findings have important implications for public policy since they suggest that the timing of policy is crucial: preventive educational programs should be coupled with more intense (rehabilitation) treatment at the early stages of bingeing and purging behaviors. Our results are robust to different model specifications and identifying assumptions.

Keywords: Bulimia Nervosa, Demographics, State Dependence, Instrumental Variables, and Dynamic Panel Data Estimation

JEL Codes: I12, I18

---

[1] Ham is at the University of Maryland, IFAU, IFS, IRP (Madison), and IZA; Iorio is at Universitat Autonoma de Barcelona and Barcelona GSE; Sovinsky is at the the University of Zurich, CEPR, and the Becker Friedman Institute. Corresponding author is Sovinsky (email: michelle.sovinsky@gmail.com). We thank the editor and referees for comments which greatly improved the paper. We thank Lynne Casper, James Heckman, Geert Ridder, Seth Sanders, Duncan Thomas and seminar participants at Alicante, Arizona, Chicago, Erlangen-Nuremberg, John Hopkins, IMT-Lucca, RAND, University of Southern California, Society of Economic Dynamics Meetings (Istanbul), the Econometrics Society Meetings (San Francisco), and the Econometrics Society European Meetings (Barcelona) for helpful comments. We are grateful to the National Science Foundation, the Claremont McKenna Lowe Institute for Political Economy, the USC College of Letters, Arts and Sciences, the Ministerio de Educación y Ciencia (SEJ2006-00712), Ministerio de Ciencia y Tecnología (SEJ2006-00538), the Barcelona GSE, and the government of Catalonia for financial support. Any opinions in this paper are those of the authors and do not necessarily reflect the views of the NSF.

1 Introduction

In the United States, eating disorders are more common than Alzheimer's disease – as many as 10 million people have an eating disorder (ED) compared to 4.5 million with Alzheimer's (National Eating Disorders Association, 2008). Bulimia nervosa (BN), which disproportionately impacts women, is the most common form of an ED.[1] In the past decade, 6 to 8.4% of female adolescents engaged in purging behaviors (National Youth Risk Behavior, 2005). Females who engage in BN typically start when they are in their teens or early twenties; however, the onset age appears to be dropping. Children are reporting bulimic behaviors at ever younger ages, where the behavior is increasingly seen in children as young as 10 (Cavanaugh and Ray, 1999).

Bulimia is characterized by recurrent episodes of "binge-eating" followed by compensatory purging.[2] There are serious health consequences from these binge and purge cycles, including electrolyte imbalances that can cause irregular heartbeats, heart failure, inflammation and possible rupture of the esophagus from frequent vomiting, tooth decay, gastric rupture, muscle weakness, anemia, and malnutrition (American Psychiatric Association, 1993). The impact on adolescents and children is even more pronounced due to irreversible effects on physical development and emotional growth.[3]

Our work is motivated by evidence that bulimics persist in their behaviors (Keel et al., 2005), which may have long-run effects on health outcomes and human capital accumulation. One possible reason that individuals may persist in BN is that starving, bingeing, purging, and exercise increase $\beta-$endorphin levels, resulting in the same chemical effect as that delivered by opiates. Along these lines, Bencherif et al. (2005) compare women with BN to healthy women of the same age and weight. They scan their brains using positron emission tomography after injection with a radioactive compound that binds to opioid receptors. The opioid receptor binding in bulimic women was lower than in healthy women in the area of the brain involved in processing taste, as well as the anticipation and reward of eating. This reaction has been found in other behaviors that exhibit substantial persistence, such as drug addiction and gambling. Finally, some studies in the biological literature suggest that there may be a genetic component to BN beyond the production of opioids (Bulik et al., 2003).

[1] Approximately 80% of BN patients are female (Gidwani, 1997).

[2] Binge-eating is the consumption of an unusually large amount of food (by social comparison) in a two-hour period accompanied by a loss of control over the eating process. Compensatory behavior includes self-induced vomiting, misuse of laxatives, diuretics, or other medications, fasting, or excessive exercise. BN is identified with frequent weight fluctuations.

[3] Irreversible risks include pubertal delay or arrest and impaired acquisition of peak bone mass resulting in growth retardation and increased risk of osteoporosis (Society for Adolescent Medicine, 2003).

It is has not been examined whether the persistence in BN is due to individual heterogeneity (i.e., some girls have persistent traits that make them more prone to bulimic behavior, but they are not influenced by past experience) or true state dependence (i.e., past BN behavior is an important determinant of current BN behavior) (Heckman, 1981). In this paper we exploit longitudinal data on individuals' history of bulimic behavior and time-changing explanatory variables to separate state dependence from individual heterogeneity in BN persistence. We find that up to two-thirds of BN persistence is due to true state dependence. Also, the impact of past behavior on current behavior is four-fold higher among African American girls, and girls from low income households exhibit the highest persistence.

These findings have important policy implications. Since true state dependence is the most important cause of persistence in BN, it is reasonable to expect that the longer an individual experiences BN, the less responsive she will be to policy aimed at combatting the behavior. In this respect the timing of policy intervention is crucial: preventive educational programs aimed at instructing girls about the deleterious health effects of BN, as well as treatment interventions, will be most effective if provided in the early stages.[4] Moreover, since the role of state dependence is not the same across racial and income groups, early intervention should pay special attention to African Americans and girls from low-income families. Second, making the case for BN exhibiting positive state dependence would help put those exhibiting BN on equal footing (from a treatment reimbursement perspective) with individuals abusing drugs or alcohol. In some states this is a current policy issue, since in several states treatment for alcoholism and drug addiction is covered but ED treatment is not covered or is covered less generously.[5] In fact, only 6% of people with bulimia receive mental health care (Hoek and van Hoeken, 2003), while a majority of states cover treatment for alcoholism and drug addiction (Center for Mental Health Services, 2008.)[6] Finally, there are potential long-run implications of ED behaviors on educational attainment given that eating disorders impact health outcomes.

[4] Our policy suggestions are consistent with recent findings in the psychiatric literature. For instance, Reas et al. (2000) report that the BN recovery rate is close to 80% if treatment is given within the first 5 years, but falls to 20% if treatment is delayed until after 15 years. This does not mean that current BN has a positive causal effect on future BN, however. Women treated in the first 5 years likely consist of both "casual" and "hard core" bulimics, while women treated after 15 years are only "hard core" bulimics.

[5] Recently the Mental Health Parity Act of 2008 was implemented (in 2010). The act requires large employer-provided insurance policies that cover mental health issues to cover them at the same level as they cover other medical issues. Note that the Act does not require policies to cover mental health issues per se. Also, policies that do offer mental health benefits do not have to cover every mental health issue (HR 6983: Wellstone Mental Health Parity and Addiction Equity Act of 2008). State mental health parity laws apply to privately insured plans offered through an employer. These laws vary significantly from state to state.

[6] Daly (2008) found that typical coverage by insurance companies for EDs failed to provide adequate reimbursement for the most basic treatment as recommended by the American Psychiatric Association.

Recent work has shown that poor child health and nutrition reduces time in school and learning during that time. These findings suggest that policies aimed at improving health early in the process could also serve to improve educational attainment.[7]

In order to investigate the persistence of BN, we estimate dynamic linear, Tobit, Ordered Probit, and Probit models that address the limited dependent nature of our measures of bulimic behavior. Our control variables are demographic variables and time-changing measures of perfectionism, distrust, and feelings of ineffectiveness, as well as a poor body image in some specifications. The time-changing control variables enable us to allow for endogenous past behavior. However, we also allow for the possibility that time-changing personality indices are correlated with an unobserved time constant individual effect since, for example, some medical studies have found that genetic factors may play a role in BN incidence (Lilenfeld et al., 1998; Bulik et al., 2003). Our approach of allowing personality traits to impact bulimic outcomes is in the same spirit as the literature on the impact of non-cognitive skills and personality traits on economic outcomes (e.g., Borghans et al., 2008). We also consider weak IV and overidentifying restrictions test. Our restrictions pass these tests, and our estimates are robust to different estimation methods and identifying assumptions.

The outline of the paper is as follows. In section 2 we present a literature overview. In section 3 we describe the data and present basic statistics on BN persistence. In section 4 we present our methodology and discuss identification, while in section 5 we present our results. We conclude in section 6.

2 Literature Review and Background

In the social science literature, there are three papers on bingeing or purging behaviors. Hudson et al. (2007) and Reagan and Hersch (2005) focus on the prevalence of various types of ED behaviors among women and men. In a companion paper, Ham, Iorio, and Sovinsky (2011, hereafter HIS), we use data from the National Heart, Lung, and Blood Institute Growth and Health Study (hereafter NHLBI) to examine which adolescent females are most at risk for BN in a multivariate framework. The NHLBI Growth and Health survey was an epidemiological study conducted by Striegel-Moore et al. (2000); they examined univariate correlations between BN and race and between BN and parental education. HIS find that African-Americans are more likely than Whites to exhibit bulimic behaviors (consistent with Striegel-Moore et al., 2000) and that these effects remain after controlling for the education of the parent, family income,

[7] See the Handbook of Development Economics Chapter, "The Impact of Child Health and Nutrition on Education in Less Developed Countries," (Glewwe and Miguel, 2008) and references therein.

and personality traits. However, HIS find a more subtle pattern from the interaction of income class and race: low and middle income African American girls, and low income White girls, are at substantially higher risk of bulimic behaviors than girls from other race-income groups.

The work in this paper differs from previous studies in the economics and epidemiology literatures along many important dimensions. First, we consider dynamic aspects of BN and distinguish between persistence due to individual heterogeneity and true state dependence, where we allow for racial and income differences in persistence. Furthermore, given that genetic factors may contribute to BN, persistence due to individual heterogeneity may be important. Our investigation of the relative roles of state dependence and individual heterogeneity is related to the existing empirical literature on this issue in other contexts (see, e.g., Labeaga and Jones, 2003; Gilleskie and Strumpf, 2005; for a survey see Chaloupka and Warner, 2000).

The large and growing literature on obesity is related to our work in the broad sense that it pertains to food consumption, but is otherwise unrelated given that women suffering from BN are characterized by average body weight (Department of Health and Human Services, 2006). Our work is also related to the growing literature using economic identification strategies and appropriate econometric methods to investigate public health issues, (see, e.g., Adams, et al., 2003; Engers and Stern, 2002; Heckman, et al., 2007; Hinton, et al., 2010; and Smith, 2007). Finally, our work is different from previous research in the economics and epidemiology literature on habit formation in that we consider nonlinear and fixed effects estimators appropriate for limited dependent variables.

3 Data

We use data from the National Heart, Lung, and Blood Institute Growth and Health Study, a survey of 2379 girls from schools in Richmond, California and Cincinnati, Ohio, and from families enrolled in a health maintenance organization in the Washington, DC area.[8] The survey was conducted annually for ten years and contains substantial demographic and socioeconomic information such as age, race, parental education, and initial family income (in categories) as well as questions on BN behavior. The latter were first asked in 1990, when the girls were aged 11-12 (which was wave 3) and subsequently asked in waves 5, 7, 9, and 10. We present descriptive statistics in Table 1. We include clustered standard errors of the mean to account for the

[8] The data do not report the location of the participant due to confidentiality concerns. Schools were selected to participate in the study based on census tract data with approximately equal fractions of African American and White children where there was the least disparity in income and education between the two ethnic groups. The majority of the cohort was randomly drawn from families with nine (or ten) year-old girls that participated in the Health Maintenance Organization (HMO). A small percentage was recruited from a Girl Scout troop located in the same geographical area as the HMO population.

fact that for all demographic variables (except age) we have one observation per person, while for the other variables we have multiple observations per person. The survey is an exogenously stratified sample, designed to be approximately equally distributed across race, income, and (highest educated) parental education level, as the descriptive statistics in Table 1 confirm.

Table 1: Descriptive Statistics

	Mean	Standard Deviation	Clustered Standard Error of Mean	Number of Waves
Age	14.363	2.991	0.014	All 10
White	0.480	0.499	0.010	1
Parents High School or Less	0.255	0.436	0.009	1
Parents Some College	0.393	0.488	0.010	1
Parents Bachelor Degree or More	0.352	0.477	0.010	1
Income less than $20,000	0.318	0.466	0.010	1
Income in [$20000, $40000]	0.315	0.465	0.010	1
Income more than $40,000	0.367	0.482	0.010	1
ED-BN Index	1.279	2.682	0.039	3,5,7,9,10
Clinical Bulimia	0.022	0.145	0.002	3,5,7,9,10
Body Dissatisfaction Index*	8.039	7.554	0.131	3,5,7,9,10
Distrust Index**	3.589	3.492	0.056	3,5,9,10
Ineffectiveness Index***	2.752	3.915	0.063	3,5,9,10
Perfectionism Index****	6.468	3.290	0.052	3,5,9,10

Notes: Income is in 1988$; * ranges from 0 to 27 (maximal dissatisfaction); ** ranges from 0 to 21 (maximal distrust); *** ranges from 0 to 29 (maximal ineffectiveness); **** ranges from 0 to 18 (maximal perfectionism). See Appendix for more detailed description of the variables.

The questions regarding bulimic behaviors were developed to be easy to understand by young respondents and to be consistent with diagnostic criteria for BN.[9] In particular, for each respondent the data contain an Eating Disorders Inventory index developed by a panel of medical experts, which was designed to assess the psychological traits relevant to bulimia (Garner

[9] Clinical criteria for BN, according to the Diagnostic and Statistical Manual of Mental Disorders fourth edition (American Psychiatric Assocation, 2000a), require the cycle of binge-eating and compensatory behaviors occur at least two times a week for three months or more and that the individual feel a lack of control during the eating episodes. Due to data restrictions, we cannot examine the prevalence of anorexia nervosa.

et al., 1983). Thus, a major advantage of these data is that all sample participants are evaluated regarding BN behaviors, and a BN eating disorder index is developed for each participant independent of any diagnoses or treatment they have received. The survey reports an Eating Disorders Inventory Bulimia subscale for each respondent (hereafter the ED-BN index), which measures degrees of her behavior associated with BN. The ED-BN index is constructed based on the subjects' responses ("always"=1, "usually"=2, "often"=3, "sometimes"=4, "rarely"=5, and "never"=6) to seven items: 1) I eat when I am upset; 2) I stuff myself with food; 3) I have gone on eating binges where I felt that I could not stop; 4) I think about bingeing (overeating); 5) I eat moderately in front of others and stuff myself when they are gone; 6) I have the thought of trying to vomit in order to lose weight, and 7) I eat or drink in secrecy. A response of 4-6 on a given question contributes zero points to the ED-BN index; a response of 3 contributes 1 point; a response of 2 contributes 2 points; and a response of 1 contributes 3 points. The ED-BN index is the sum of the contributing points and ranges from 0 to 21 in our data. For instance, if a respondent answers "sometimes" to all questions, her ED-BN index will be zero. We have only the aggregate score, not the answers to individual questions. As Table 1 indicates, the mean ED-BN index is 1.2.

A higher ED-BN score is indicative of more BN related problems that are characterized by uncontrollable eating episodes followed by the desire to purge. According to the team of medical experts that developed the index (Garner et al., 1983), a score higher than 10 indicates that the girl is very likely to have a clinical case of BN. The quantitative interpretation in terms of who is perceived to be suffering from clinical BN is motivated by results from surveys among women diagnosed with BN (by the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) criteria): the average ED-BN index among this subsample was 10.8.[10] For this reason, we will refer to a value of the ED-BN index of greater than 10 as clinical bulimia for the remainder of the paper. The ED-BN index is widely used in epidemiological and ED studies (Rush et al., 2008). As shown in Table 1, approximately 2.2% of the girls (who are 14 years old on average) have a case of clinical BN, which is close to the national average reported from other sources.[11] However, in estimating some, but not all, of our models, we will exploit the fact that we know the numerical value of the index rather than simply whether it is greater than 10; this tends to result in an efficiency gain but does not change the basic nature of our results.

The NHLBI Growth and Health survey also contains questions used to construct four other indices based on psychological criteria. These indices were developed by a panel of medical experts (see Garner et al., 1983 for a discussion of the association of these personality

[10] See Garner et al. (1983) for more details on the development and validation of the ED-BN index.

[11] See for instance, Hudson et al. (2007) and National Eating Disorders Association (2008).

traits with EDs). The four additional indices measure a respondent's potential for personality traits/disorders, and below we refer to these indices collectively as the "personality indices." The first index is a measure of each girl's dissatisfaction with her body. This index is reported every year and is a sum of the respondents' answers to nine items intended to assess satisfaction with size and shape of specific parts of the body. Hereafter we refer to it as the body dissatisfaction index. We also use three additional indices based on psychological criteria, measuring tendencies toward: perfectionism (hereafter the perfectionism index), feelings of ineffectiveness (hereafter the ineffectiveness index), and interpersonal distrust (hereafter the distrust index). These indices are available in waves 3, 5, 9, and 10 and thus overlap with the ED-BN index availability, with the exception that the ED-BN index is also available in wave 7. For ease of exposition, we provide details on the questions used to form the personality indices in Appendix A. In all cases we do not have the responses to the questions used to construct the score, just the aggregated index, where a higher score indicates a higher level of the personality trait.

Table 2: Mean of ED-BN Index and Incidence of Clinical Bulimia by Characteristics

Variable	ED-BN Index			Clinical Bulimia (BN)		
	Mean	Standard Deviation	Clustered Std. Error	Mean	Standard Deviation	Clustered Std. Error
Years:						
1989	1.814	3.287	0.070	0.038	0.191	0.004
1991	1.610	3.021	0.067	0.033	0.178	0.004
1993	1.098	2.342	0.054	0.014	0.117	0.003
1995	0.860	2.054	0.046	0.008	0.092	0.002
1996	0.955	2.279	0.050	0.013	0.113	0.002
White	1.042	2.437	0.051	0.017	0.130	0.002
African American	1.498	2.873	0.058	0.026	0.158	0.003
Parents High School or Less	1.648	3.136	0.096	0.033	0.178	0.005
Parents Some College	1.325	2.682	0.060	0.020	0.141	0.003
Parents Bachelor Degree or More	0.973	2.278	0.055	0.015	0.122	0.002
Household Income (in 1988$):						
Income less than $20,000	1.721	3.146	0.086	0.033	0.179	0.004
Income in [$20000, $40000]	1.198	2.633	0.072	0.021	0.144	0.003
Income more than $40,000	0.982	2.245	0.053	0.013	0.112	0.002

Correlations of ED-BN Index and Clinical Bulimia with Personality Characteristics

Personality Characteristic Index	ED-BN Index	Clinical Bulimia (BN)
Body Dissatisfaction Index	0.221	0.114
Distrust Index	0.213	0.107
Ineffectiveness Index	0.439	0.274
Perfectionism Index	0.229	0.145

Notes: The top panel reports clustered (by individual) standard errors of the mean. All correlations in the bottom panel are significant at the 1% level.

Table 2 shows the univariate relationship between the demographic variables, the ED-BN index (columns 1-3), and BN incidence (columns 4-6). Again, in each case we cluster the standard errors (by individual) for the means. The results indicate that as the girls age, both the ED-BN index and BN incidence fall. A notable point, which we examine in detail in our companion paper (HIS 2011), is that African American girls have a statistically significant higher ED-BN index and incidence of clinical BN than White girls. Furthermore, both the ED-BN index and the incidence of clinical BN decrease as (the highest educated) parental education and family income increase, and again these differences are statistically significant at standard confidence levels. These results suggest that BN is more problematic among African American girls, girls from low income families, and girls from families with low parental education. As we discuss in HIS, these findings are not due to an incorrect interpretation of what the ED-BN index measures, i.e., the possibility that it might capture obesity (binge eating) instead of BN behaviors. Neither do these findings disappear once we condition on the other demographic variables or personality indices. The bottom panel of Table 2 shows that both the ED-BN Index and BN incidence are correlated with the indices measuring personality traits.

4 Empirical Models

In this section we describe our procedure to obtain a non-experimental estimate of the role of state dependence in the high degree of persistence in bulimic behavior (i.e., BN behavior in the past has a causal effect on BN behavior this period). From an evaluation point of view, it would be ideal to randomly assign individuals to the state in question in a baseline period, and then see how this assignment affects their presence in the state relative to a randomly chosen control group. In this way we could observe their persistence in the state, which would be solely due to state dependence. Of course, ethical considerations immediately rule out this approach, so we turn to other methods to distinguish the role of state dependence in persistence as opposed to that due to observed and unobserved heterogeneity (i.e., some girls have persistent traits that make them more prone to bulimic behavior). We first consider a linear regression framework, since it allows an extended discussion of identification issues, which arise in any non-experimental estimation of this type. We then consider limited dependent variable models to estimate state dependence in bulimic behavior.

We consider four model specifications: i) a linear regression structure that treats a zero value of the ED-BN index as lying on the regression line; ii) a Tobit structure for the ED-BN index; iii) a linear probability model (LPM) for the incidence of clinical BN (i.e., a value for the ED-BN index greater than 10) and iv) a Probit model.

4.1 Linear Model

We begin with the regression model and consider our most basic specification

$$y_{it} = \alpha_0 + \alpha_1 y_{it-1} + \widetilde{\delta}_i + v_{it}, \tag{1}$$

where y_{it-1} is the lag of the observed value of the ED-BN index, $\widetilde{\delta}_i$ are (unobserved) individual-specific random effects, and v_{it} is an uncorrelated (over time) error term. We drop the year dummies for ease of exposition.[12] The least squares estimate of α_1 will reflect both observed and unobserved heterogeneity as well as true state dependence. To account for observed heterogeneity, we include current explanatory variables X_{it} to obtain

$$y_{it} = \gamma_0 + \gamma_1 y_{it-1} + \gamma_2 X_{it} + \delta_i + v_{it}. \tag{2}$$

In our application X_{it} will consist of some or all of the current level of the personality characteristics (henceforth CPC) and the demographic variables (ethnicity, income, and the highest education of the parents) and in our basic model we assume that they are uncorrelated with δ_i and with v_{it}. We now consider issues related to identification to ensure that our estimate of γ_1 reflects only true state dependence.

4.1.1 Identification

Identification is an important and difficult issue in the estimation of dynamic models since they often do not lend themselves to using experimental data to estimate the parameters of interest. Researchers generally face a number of options for achieving identification, none of which may be totally convincing on its own. Therefore, we consider a number of identification strategies to see whether our results are robust to changing the identification strategies. Our first approach is to treat δ_i as a random effect uncorrelated with X_{it}, and to use the time-changing components of X_{it-1} (i.e., the lagged personality characteristics, henceforth LPC) as excluded IV for the endogenous lagged dependent variable.[13] Consider the case where we use only one lag of the personality characteristics as IV. Our approach will not produce consistent estimates of γ_1 if X_{it-1} are weak instruments, i.e., $\pi_2 \to 0$ as $N \to \infty$ in the first stage equation,

$$y_{it-1} = \pi_0 + \pi_1 X_{it} + \pi_2 X_{it-1} + e_{it-1}. \tag{3}$$

[12] If we add time dummies, the only real change is that age becomes very insignificant.

[13] An alternative identification strategy, which we did not investigate, is offered by Lewbel (2007). He shows that one does not need exclusion restricitons if one is willing to assume that the variance in the first stage error term differs across individuals and depends on observable characteristics while the covariance between the first stage and second stage error terms is constant.

Standard tests indicate that in our study X_{it-1} are not weak instruments in the sense that they affect y_{it-1} conditional on X_{it} (see Table 4).[14] Thus, the validity of our identification strategy, conditional on treating δ_i as a random effect uncorrelated with X_{it}, rests on whether it is reasonable to assume that the LPC X_{it-1} affect y_{it} only through y_{it-1}. Suppose that this is not true in our data, and that the correct specification is

$$y_{it} = \gamma_0 + \gamma_1 y_{it-1} + \gamma_2 X_{it} + \gamma_3 X_{it-1} + \delta_i + v_{it}. \tag{4}$$

However, if equation (4) holds, we expect the overidentifying test for equation (2) to fail.

To see this consider a "reduced form" version of equation (2) for current BN behavior

$$y_{it} = \rho_0 + \rho_1 X_{it} + \rho_2 X_{it-1} + e_{it}. \tag{5}$$

The overidentifying restriction test considers the null hypothesis $\rho_2 = \gamma_1 \pi_2$, which we would not expect to hold if equation (4) is the correct model. We do not fail these tests, and thus we conclude that the data suggest that X_{it-1} affects y_{it} only through y_{it-1}.[15]

Finally, one may be concerned that δ_i is correlated with X_{it}. An extreme version of this issue has been raised in the medical literature, where, as noted above, it is hypothesized that X_{it}, X_{it-1} and y_{it} are a function of a single unobserved factor, plus a random noise. To consider this, let

$$y_{it} = \alpha_i + v_{it},$$

where α_i is *iid* across i and has mean 0 and variance σ_α^2, v_{it} is *iid* across i and t with mean 0 and variance σ_v^2, and $E(\alpha_i, v_{i't}) = 0$ for all i, i' and t. Further, assume that personality characteristic k, X_{kit}, is determined by

$$X_{kit} = \phi_k \alpha_i + e_{kit}, \; k = 1, ..., K,$$

where $E(v_{it} e_{ki\tau}) = 0$ and $E(e_{ki\tau} e'_{k'i'\tau}) = 0$ for all i, i', t, τ and $k \neq k'$. Given the true value of each β_k is zero, we can consider the regression

$$y_{it} = \beta X_{it} + \alpha_i + v_{it},$$

where α_i is treated as a random effect uncorrelated with X_{it}. However, the least squares coefficients are biased, i.e. $E(\hat{\beta}) \neq 0$, because

$$E[X_{kit}(\alpha_i + v_{it})] = E[(\phi_k \alpha_i + e_{kit})(\alpha_i + v_{it})] = \phi_k \sigma_\alpha^2, \quad k = 1, ..., K.$$

---

[14] In the presence of weak instruments, a natural response is to include y_{it-2} as an IV, which requires the nontrivial assumption that v_{it} is independent over time or, at worst, follows an MA(1) process.

[15] This, of course, assumes that the overidentifying tests are not passed simply because of a lack of power.

If we first difference the equations for y_{it} and X_{it} we obtain

$$\Delta y_{it} = \beta \Delta X_{it} + \Delta v_{it}, \tag{6}$$

where Δ represents the first-difference operator. Now the least squares coefficients are unbiased, i.e. $E(\hat{\beta}) = 0$, because

$$E[\Delta X_{kit} \Delta v_{it},] = E[\Delta e_{kit} \Delta v_{it}] = 0 \text{ for all } k = 1, .., K.$$

To investigate the single factor hypothesis, we estimate equation (6) and test the null hypothesis $\beta = 0$ for each specification considered below. We decisively reject the null hypothesis $\beta = 0$ in all cases and thus conclude that the single factor model is not appropriate in our application.[16]

We next consider a specification of our general model given by equation (2) where it is appropriate to treat δ_i as a fixed effect (FE). As is well known, care must be exercised when estimating FE dynamic models. To obtain consistent estimates, we follow Arellano and Bond (1991; hereafter AB) and eliminate the FE by first differencing equation (2) to obtain

$$\Delta y_{it} = \beta_0 + \beta_1 \Delta y_{it-1} + \beta_2 \Delta X_{it} + \Delta v_{it}, \tag{7}$$

where Δ represents the first-difference operator. We consider two cases. First, we assume that v_{is} is independent of X_{it} for any t, s conditional on δ_i, i.e., X_{it} is strictly exogenous (Wooldridge, 2002, p. 253). Under this assumption we can treat ΔX_{it} as exogenous in equation (7), and X_{it-1} as excluded IV, i.e., ΔX_{it} acts as its own instrument. However, often these will be weak IVs, and this indeed is a problem in our application. AB consider this problem and suggest that researchers also use y_{it-2} as an IV. Note that the lag of the dependent variable will be a valid IV as long as v_{it} is independent over time. AB stress the importance of specification tests in using this assumption for identification. Specifically, one can test the null hypothesis that v_{it} is independent over time, as well as the null hypothesis that the overidentifying restrictions hold. We find we do not reject either of these null hypotheses.[17]

AB note that the use of y_{it-2} as an IV allows one to make weaker assumptions on the X_{it}. For example, there may be feedback effects from v_{it} to future values of X_{it}, and in this case strict

[16] When we do not include body dissatisfaction in the personality characteristics, the Wald statistics for the null hypothesis $\beta = 0$ when we use (do not use) the interpolated data are 190.652 (128.498), which are both much bigger than any reasonable critical value for $\chi^2(3)$. When we include body dissatisfaction in the personality characteristics, the Wald statistics when we use (do not use) the interpolated data are 232.850 (145.423), which are both much bigger than any reasonable critical value for $\chi^2(4)$.

[17] Again we need to add the caveat that we may not reject these null hypotheses simply because of a lack of power.

exogeneity would no longer hold. To address this potential issue, we assume only sequential exogeneity i.e., that v_{is} is independent of X_{it} only for $s \geq t$ conditional on δ_i (Wooldridge, 2002, p. 299). Under the sequential exogeneity assumption, we estimate the parameters of equation (7) by 2SLS while also treating ΔX_{it} as endogenous; we use y_{it-2} and X_{it-1} as our excluded IV. We find that for this specification we also cannot reject the null hypothesis that v_{it} is independent over time, nor can we reject the overidentifying assumptions.[18] Below we find that these different approaches produce similar estimates of true state dependence, presumably increasing the confidence readers can place in our estimates.

4.2 Tobit Model

For the Tobit model, we start by considering the simplest latent variable equation

$$y^*_{it} = \lambda_0 + \lambda_1 y_{it-1} + \widetilde{\mu}_i + e_{it}, \tag{8}$$

where $\widetilde{\mu}_i$ are (unobserved) individual-specific random effects and e_{it} is an uncorrelated (over time) error term, both of which are normally distributed. The estimate of λ_1 will capture observed and unobserved heterogeneity and true state dependence. To account for observed heterogeneity, we add explanatory variables X_{it} to obtain

$$y^*_{it} = \theta_0 + \theta_1 y_{it-1} + \theta_2 X_{it} + \mu_i + e_{it}, \tag{9}$$

where the estimate of θ_1 will reflect unobserved heterogeneity and true state dependence. To capture only the latter, we consider the Wooldridge (2005) dynamic correlated random effects Tobit model based on Chamberlain (1984), and assume that

$$\mu_i = \varphi_3 \overline{X}_i + \varphi_4 y_{i0} + c_i, \tag{10}$$

where $\overline{X_i}$ denotes the mean value of the explanatory variables, y_{i0} the initial condition, and c_i an individual specific error term. We now have

$$y^*_{it} = \varphi_0 + \varphi_1 y_{it-1} + \varphi_2 X_{it} + \varphi_3 \overline{X}_i + \varphi_4 y_{i0} + c_i + e_{it}. \tag{11}$$

We estimate the model by following Wooldridge (2005) in assuming strict exogeneity for the X_{it} (with respect to e_{it}) and then using MLE; in this case the estimate of φ_1 reflects only true

[18] To allow for genetic factors to play different roles at different ages, we estimated a model with both a fixed effect and a fixed effect interacted with a trend. In a linear model this leads to the Heckman-Hotz (1989) random growth model, which can be estimated by double-differencing and using IV procedures. However, given our limited number of observations, this model was too rich for our data; we obtained a coefficient roughly the size of our other IV estimates, but its standard error was so large that it was not statistically significant.

state dependence. Restricting the initial condition to depend on the initial observation of the ED-BN index is less of a problem in our sample because we have data on the respondents when they are young, and hence it seems reasonable to assume that y_{i0} captures initial conditions.

As a robustness check we also estimate a dynamic Probit model (using the Wooldridge procedure) and a dynamic LPM for the incidence of the ED-BN index being greater than 10. For the LPM, we proceed in a manner analogous to the linear regression model, and for the Probit model, we proceed in a manner analogous to the Tobit. See Appendix B for details.

5 Empirical Results

5.1 Results for the Linear Model

Table 3 contains our parameter estimates for the linear model. In column (1) we consider a model where the only explanatory variable is the (assumed to be exogenous) lagged dependent variable; its coefficient is estimated at 0.44 and, not surprisingly, it is very statistically significant. Regarding the effect of past ED-BN experience on current behavior, the coefficient can be interpreted as an elasticity since we would expect the mean of a variable and its lag to be equal. We obtain a relatively large estimate of the elasticity of 0.44. To look at the magnitude of the coefficient in another way, an individual with a lagged ED-BN index of 5 would have a current ED-BN index over two points higher than someone with a lagged index of 0; this difference is almost 150% of the mean value of the ED-BN index. After we add the demographic variables in column (2) and the personality indices in column (3), the lag coefficient drops to 0.421 and 0.35, respectively, and is insensitive to including body dissatisfaction in column (4). These results demonstrate substantial persistence in BN behavior that can be due to both unobserved heterogeneity and true state dependence.

To focus on the latter, we first assume the individual effect in equation (2) is uncorrelated with X_{it}. As noted above, in this case researchers can use X_{it-1} as IV as long as they are not weak. Fortunately, in our case X_{it-1} are not weak instruments, and thus we do not need to add y_{it-2} as an IV, which would require restrictions on the covariance of v_{it} over time for the same individual. Thus in columns (5) to (8), we estimate equation (2) while treating the lagged dependent variable as endogenous and use X_{it-1} as the excluded IV. Specifically, in columns (5) we exclude body dissatisfaction from the first and second stage equations, while in column (6) we include body dissatisfaction. Columns (5) and (6) both report a lagged coefficient of approximately 0.2, suggesting that over half the variation in persistence attributed to unobserved heterogeneity and state dependence is actually due to the latter. The coefficient estimate of 0.2 suggests an elasticity of 0.2 for the effect of lagged BN on current behavior. To put this another way, the expected ED-BN index for someone who has a lagged value of the ED-BN index equal to 5 compared to someone who has a lagged value of 0 would be higher by 1, approximately 80% of the mean value of 1.2. [19]

Table 3: Linear Regression Estimates of the Persistence of ED-BN Index

Variables					Two Stage Least Squares				Arellano-Bond			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Lagged ED-BN Index	0.444***	0.421***	0.355***	0.349***	0.190***	0.188***	0.149***	0.131***	0.192***	0.189***	0.177***	0.172***
	(0.028)	(0.009)	(0.027)	(0.027)	(0.062)	(0.059)	(0.048)	(0.046)	(0.038)	(0.038)	(0.042)	(0.041)
White		-0.028	-0.038	-0.081	-0.105	-0.174	-0.134*	-0.201***				
		(0.056)	(0.085)	(0.084)	(0.123)	(0.121)	(0.084)	(0.084)				
Age		-0.051***	-0.051***	-0.063***	-0.021	-0.032	-0.065***	-0.080***	-0.092***	-0.114***	-0.068***	-0.084***
		(0.013)	(0.016)	(0.016)	(0.022)	(0.022)	(0.018)	(0.018)	(0.117)	(0.118)	(0.015)	(0.017)
Parents Some College		-0.045	0.073	0.073	0.017	-0.006	-0.066	-0.089				
		(0.070)	(0.101)	(0.101)	(0.154)	(0.153)	(0.097)	(0.100)				
Parents Bachelor Degree or more		0.007	0.122	0.131	-0.009	-0.011	-0.035	-0.040				
		(0.801)	(0.110)	(0.110)	(0.167)	(0.167)	(0.105)	(0.108)				
Income in [$20000, $40000]		-0.196***	-0.236**	-0.238**	-0.524***	-0.539***	-0.240***	-0.248***				
		(0.071)	(0.102)	(0.102)	(0.154)	(0.154)	(0.097)	(0.100)				
Income more than $40,000		-0.284***	-0.207**	-0.221**	-0.463***	-0.486***	-0.288***	-0.296***				
		(0.077)	(0.104)	(0.103)	(0.159)	(0.159)	(0.094)	(0.096)				
Distrust Index			-0.019	-0.018	-0.040**	-0.041**	-0.002	-0.002	-0.002	-0.006	-0.018	-0.016
			(0.014)	(0.014)	(0.019)	(0.019)	(0.015)	(0.015)	(0.019)	(0.019)	(0.039)	(0.039)
Ineffectiveness Index			0.205***	0.188***	0.258***	0.229***	0.230***	0.206***	0.178***	0.158***	0.169***	0.149***
			(0.020)	(0.020)	(0.029)	(0.028)	(0.022)	(0.021)	(0.026)	(0.027)	(0.034)	(0.032)
Perfectionism Index			0.097***	0.095***	0.129***	0.125***	0.096***	0.093***	0.123***	0.120***	0.120***	0.121***
			(0.013)	(0.013)	(0.019)	(0.019)	(0.013)	(0.013)	(0.020)	(0.020)	(0.029)	(0.028)
Body Dissatisfaction Index				0.027***		0.040***		0.036***		0.050***		0.041**
				(0.005)		(0.008)		(0.005)		(0.011)		(0.019)
Constant	0.597***	0.161***	0.592*	0.657**	0.515	0.538	1.138***	1.233***	1.154***	1.190***	0.828**	0.794**
	(0.037)	(0.225)	(0.304)	(0.303)	(0.379)	(0.375)	(0.330)	(0.333)	(0.315)	(0.316)	(0.353)	(0.350)
Interpolated Indices	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Autocorrelation Test									0.495	0.570	0.495	0.570
									(0.620)	(0.568)	(0.620)	(0.568)
First Difference	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Sample Size	4151	3938	3938	3928	2285	2273	5426	5384	3612	3586	3612	3586

Notes: Standard errors robust to heteroskedasticity and intra-group correlation are reported in parenthesis. * indicates significant at 10%; ** significant at 5%; *** significant at 1%. Instruments are one-period lags of: all personality indices in columns (6) and (8); all indices excluding body dissatisfaction in columns (5) and (7). In columns (9) and (10), in the difference equation, instruments are two-period lags of the ED-BN index and the first difference of personality indices. In columns (11) and (12) instruments are two period lags of the ED-BN index and the available lags of the personality characteristics. Regarding the autocovariance test in the AB specifications, we fail to reject the null hypothesis of no autocorrelation in the idiosyncratic error term in all specifications. Columns (7)-(12) use interpolated values of personality indices in wave 7.

Our sample size is limited by the fact that the personality indices are not available in wave 7, and this limitation is especially important in our AB analysis.[20] However, we can increase our sample size if we assume that the personality index values vary smoothly from wave 5 to 9,

[19] Some girls in our sample may receive treatment once they begin bulimic behavior, although we cannot identify who they are. If this treatment is even partially effective, it will reduce the degree of true state dependence, so our estimates are lower bounds on the degree of true state dependence in untreated BN.

[20] Specifically, in the AB analysis we lose the independent variables ΔX_{it} when the dependent variable is $y_{i9} - y_{i7}$ and when the dependent variable is $y_{i10} - y_{i9}$.

and use interpolated values wave 7, which doubles our sample size.[21] The 2SLS estimates of our basic model using the imputed data (with and without body dissatisfaction) are in columns (7) and (8). Comparing the results in columns (7) and (8) to those in columns (5) and (6), respectively, indicates that using the imputed data diminishes the role of true state dependence by about one-fifth, but that the coefficient on the lagged value is still highly significant.[22] The interpolated indices also allow us to use X_{t-1} and X_{t-2} as instruments. When we do this, we obtain estimated coefficients (standard errors) of 0.252 (0.071) and 0.177 (0.066), respectively for columns 7 and 8 of Table 3.

As is standard practice, we consider two diagnostics for our 2SLS estimates in columns (5) to (8). Table 4 presents the reduced form estimates to investigate the issue of weak instruments. There will be heteroskedasticity in the first-stage regression equation for a censored dependent variable; therefore, the widely used rule of thumb for the first-stage F-statistic of excluded instruments (from Staiger and Stock (1997) and Stock and Yogo (2005)) will be inappropriate. Instead, we use the conjecture by Hansen, Hausman, and Newey (2008) that in the presence of heteroskedasticity in the first-stage equation, the Wald statistic for the null hypothesis that the excluded instruments are zero in the first stage, minus the number of instruments, should be greater than 32. Note first that we pass the weak IV test in all specifications, and that the perfectionism, ineffectiveness, and body dissatisfaction (when used) indices are always individually significant, suggesting that they are not simply driven by a single (genetic) factor.[23]

Further, when we consider the instruments on an individual basis, we pass the weak IV test for the perfectionism, ineffectiveness, and body dissatisfaction indices.[24]

[21] When we use the interpolated indices we obtain a lagged ED-BN index coefficient of 0.327(0.022) and 0.323(0.022), for columns 3 and 4, respectively. These estimates indicate that the results are very robust to the use of interpolated indices.

[22] We also investigate whether the results are robust when we control for depression. We have self-reported information on depression in two waves. Using this subsample, we estimate the model with and without depression. The coefficient of the lagged ED-BN index is virtually the same and statistically significant in both cases.

[23] We also consider only the perfectionism and ineffectiveness indices as IV in column (7) of Table 4. We obtain an estimate for the lagged coefficient of the ED-BN index of 0.163, suggesting that the results are robust to the exclusion of the distrust index (which is the only IV that is not significant in the first stage results).

[24] We present the additional first-stage estimates in Table B1 in Appendix B.

Table 4: First Stage Estimates for Table 3

	(1)	(2)	(3)	(4)
		Estimates Corresponding to Columns (5)-(8) of Table 3		
Instruments for Lagged ED-BN Index				
Lagged Perfectionism Index	0.154 ***	0.154 ***	0.165 ***	0.165 ***
	(0.019)	(0.019)	(0.014)	(0.014)
Lagged Ineffectiveness Index	0.262 ***	0.228 ***	0.250 ***	0.220 ***
	(0.018)	(0.019)	(0.013)	(0.014)
Lagged Distrust Index	0.017	0.013	-0.002	-0.006
	(0.020)	(0.020)	(0.015)	(0.015)
Lagged Dissatisfaction Index		0.060 ***		0.053 ***
		(0.011)		(0.007)
Other Regressors				
White	-0.221 *	-0.194	-0.249 ***	-0.282 ***
	(0.130)	(0.130)	(0.080)	(0.080)
Age	-0.060 **	-0.083 ***	-0.078 ***	-0.106 ***
	(0.027)	(0.027)	(0.018)	(0.019)
Parents Some College	-0.181	-0.212	-0.171 *	-0.198 **
	(0.155)	(0.155)	(0.095)	(0.095)
Parents Bachelor Degree	-0.407 **	-0.428 **	-0.266 **	-0.276 ***
or More	(0.175)	(0.174)	(0.107)	(0.107)
Income in [$20000, $40000]	0.026	-0.021	-0.227 **	-0.231 **
	(0.159)	(0.158)	(0.096)	(0.095)
Income more than $40,000	0.013	-0.041	-0.248 **	-0.263 ***
	(0.171)	(0.170)	(0.103)	(0.102)
Distrust Index	0.040 **	0.051 ***	0.023	0.031 **
	(0.019)	(0.019)	(0.015)	(0.015)
Ineffectiveness Index	0.053 ***	0.051 ***	0.032 **	0.028 **
	(0.017)	(0.018)	(0.013)	(0.014)
Perfectionism Index	0.005	0.005	-0.019	-0.020
	(0.018)	(0.018)	(0.014)	(0.014)
Body Dissatisfaction Index		-0.020 *		-0.012 *
		(0.010)		(0.006)
Constant	0.619	0.829 *	1.350 ***	1.640 ***
	(0.453)	(0.452)	(0.327)	(0.328)
Hansen, Hausmann and Newey Weak IV Test Statistic	143	165	222	265
Overidentification Test	1.796	2.005	2.736	3.096
	(0.407)	(0.571)	(0.213)	(0.407)
Interpolated Values	No	No	Yes	Yes
Sample Size	2285	2273	5426	5384

Notes: Standard errors robust to heteroskedasticity and intra-group correlation are reported in parenthesis. * indicates significant at 10%; ** significant at 5%; *** significant at 1%. Regarding the weak IV test, Hansen, Hausman, and Newey (2008) suggest that, in the presence of heteroskedasticity in the first stage equation, the test statistic should be greater than 32. Regarding the overidentifying test, under the null hypothesis that the overidentifying restrictions are satisfied the test statistic should be distributed Chi-Squared (2) (Chi-Squared (3)) in columns (1) and (3) (columns (2) and (4)). The p-values are in parenthesis. The overidentification test is consistent with clustering.

Our second diagnostic pertains to the overidentification restrictions. We present a Wald statistic to test the overidentification restrictions that the instruments are valid, which is suitable with heteroskedasticity and clustering; here the critical value is $\chi^2(l)$, where l is the degree of overidentification. Intuitively, the test can be thought of as assuming that one of the instruments is valid, and then examining whether the other instruments have zero coefficients in the structural equation. Also, we specifically test the validity of body dissatisfaction as an instrument, conditional on the other personality indices being valid, by entering its lagged value as an explanatory variable in column (6) and testing whether its coefficient is significantly different from zero. As the p-values show, we can not reject the null hypothesis that the overidentifying restriction with respect to restricting lagged body dissatisfaction is valid. Thus, overall the diagnostics show that our instruments are not weak and the overidentifying restrictions, including that for body dissatisfaction in column (6), are not rejected.

The 2SLS estimates in columns (5) to (8) of Table 3 are consistent if we assume that v_{is} and δ_i are independent of X_{it} for all s, t. As noted above, to relax this assumption we also present the results using the AB approach of differencing before using 2SLS to allow for the personality indices to be correlated with δ_i. We first assume that the personality traits are strictly exogenous with respect to v_{it} in equation (2) (i.e., that the personality traits are uncorrelated with v_{is} at all s, t) In this case we treat ΔX_{it} as exogenous and use y_{it-2} and ΔX_{it} as excluded IV under the assumption that the v_{it} are independent over time. The results are in columns (9) and (10) of Table 3 when we exclude and include body dissatisfaction, respectively. The results in column (9) show a highly significant lag coefficient of around 0.19 and the coefficient estimates remain the same when we include body dissatisfaction as an explanatory variable in column (10).[25] The test of the null hypothesis of no serial correlation is essentially a test of the overidentifying restriction on the lagged dependent variable (after allowing for heteroskedasticity). From the bottom of columns (9) and (10) we see that we cannot reject the null hypothesis, indicating that values of the ED-BN index lagged two periods (or more) are valid instruments in the equations in first differences, and our AB estimates are consistent.

Next we relax the strict exogeneity restriction by assuming that the personality traits are sequentially exogenous in the sense that we only assume $E(X_{it}v_{is}) = 0$ for $t \leq s$ to allow for feedback from current v_{is} to future X_{it}. Note that relaxing strict exogeneity implies we must treat ΔX_{it} as endogenous in equation (7), and we use y_{it-2} and X_{it-1} as excluded IV in the first-differenced equation. The AB results for this case are in columns (11) and (12) when we exclude and include body dissatisfaction, respectively. Again, the test for serial correlation suggests that lagged two periods (or more) value of the ED-BN index is a valid instrument.[26] The coefficient of the lagged dependent variable is estimated at 0.18 in columns (11) and (12).

When carrying out IV estimation, it is not possible to test whether a model is identified (although it is possible to test over-identifying restrictions). However, the results from diag-

[25] Weak instruments are not an issue because of the lagged dependent variable.

[26] We also estimated a specification of the model in which we use both y_{it-2} and y_{it-3} as instruments. The results are robust and a serial correlation test shows that the IV are valid. Further, our results are similar when we change the number of lagged X to include as IV. All robustness checks are available upon request.

nostic and robustness checks help us to add support to the notion that our model specification and identifying assumptions are appropriate. The estimates obtained in columns (5)-(12) are robust to a number of different identification strategies in terms of our assumptions on the independence of the personality traits X_{it} with respect to δ_i and v_{it} in equation (2), and with respect to whether or not we include body dissatisfaction in the model. Further, in terms of diagnostics, each of the different specifications passes weak IV and overidentification tests. Note in particular that our results are robust to allowing for the possibility i) that personality indices are driven by a genetic component in δ_i, i.e., all personality traits are driven by one factor and ii) that there may be feedback from current shocks to future values of personality indices.

In summary, we find that there is substantial persistence in BN, and that about half of this persistence is due to true state dependence. Further, the magnitude of the effect suggests that state dependence is quite important. Finally, these results are robust to changes in the explanatory variables and identification strategy.

So far we have focused on models where state dependence is constant across race and income class. Table 5 presents 2SLS estimates describing the racial and income differences in the persistence of BN when we address the endogeneity of past behavior. We use interpolated values for wave 7 (since we are estimating a richer model) and exclude body dissatisfaction as an explanatory variable. To facilitate the comparison with these results, column (1) repeats the results of Table 3 column (7), where the lag is not interacted with race or income. In the remaining columns, we use the socioeconomic indicator of focus interacted with the lag of the perfectionism and ineffectiveness indices as IV. For example, in column (2) we allow the persistence to differ by race, where the IV are race interacted with the lagged personality indices. Column (2) indicates that much of the persistence in the overall sample is driven by the behavior of African American girls. Indeed, the estimate for persistence among Whites is very small and significant (0.05), while it is substantial and significant for African-Americans (0.21). In column (3), where we consider income differences in persistence, we observe that the strongest persistence is in low income families, as the estimated coefficient on the lagged behavior is significant and very large at 0.32 (given we are instrumenting and imputing personality indices). It falls to 0.17 for middle income families and is essentially zero for girls from high income families. These results show interesting race and income effects of BN persistence.[27]

[27] The data are not rich enough for a model with race-income interactions in the levels and in the persistence.

Table 5: Racial and Income Class Differences in the Persistence of BN

	Two-Stage Least Squares Estimates		
	(1)	(2)	(3)
White	-0.134*	0.058	-0.129*
	(0.084)	(0.093)	(0.069)
Age	-0.065***	-0.062***	-0.067***
	(0.018)	(0.016)	(0.016)
Parents Some College	-0.066	-0.066	-0.024
	(0.097)	(0.082)	(0.082)
Parents Bachelor Degree or More	-0.035	-0.052	-0.012
	(0.105)	(0.093)	(0.092)
Income in [$20000, $40000]	-0.240***	-0.226***	0.067
	(0.097)	(0.083)	(0.124)
Income more than $40000	-0.288***	-0.259***	0.255**
	(0.094)	(0.089)	(0.123)
Lagged ED-BN Index	0.149***	0.206***	0.318***
	(0.048)	(0.036)	(0.042)
Interaction with Lagged ED-BN Index:			
White		-0.146***	
		(0.050)	
Income in [$20000, $40000]			-0.145**
			(0.058)
Income more than $40000			-0.362***
			(0.057)
Sample Size	5426	5426	5426

Notes: Results in all columns are with interpolated values of the indices and include all control variables as in Table 3 column (7). Standard errors robust to heteroskedasticity and intra-group correlation are in parenthesis. * significant at 10%; ** at 5%; *** at 1%.

5.2 Results for the Tobit and other Nonlinear Models

The Tobit partial effect estimates are given in Table 6. Column (1) presents estimates where the only explanatory variable is the lagged dependent variable, and the estimated partial effect is 0.27. In columns (2) and (3) we control for observable heterogeneity by including demographic variables and personality indices without and with body dissatisfaction respectively. The partial effect of the lagged dependent variable falls to 0.20 in both columns (2) and (3).[28] In order to control for unobserved heterogeneity in columns (4) and (5), we include correlated random effects using the Wooldridge (2205) approach, where we exclude and include body dissatisfaction, respectively. The estimates of 0.19 and 0.18 of the partial effect of the lagged dependent

[28] We also estimated the model for column (3) using the interpolated data, and these results (not shown) were very close to those for the non-imputed data presented in column (3).

variable in these two columns capture true state dependence, and represent about two-thirds of BN persistence, estimated at 0.27 in column (1), which reflects observed heterogeneity, unobserved heterogeneity, and true state dependence. Further, the persistence estimates in columns (4) and (5) are approximately equal to those in columns (2) and (3) respectively, suggesting that state dependence plays a much larger role than unobserved heterogeneity.

The estimated partial effects from the Probit and LPM models are of the same sign as the linear and Tobit estimates (see Tables B2 and B3 in Appendix B), but fewer estimated coefficients are statistically significant. This is expected since the Probit and LPM use much less information per person. Indeed, our estimates illustrate the importance of not focusing only on whether an individual has BN for understanding the determinants of the disorder.

Table 6: Tobit Partial Effects Estimates for the Persistence of the ED-BN Index

	(1)	(2)	(3)	(4)	(5)
Lagged ED-BN Index	0.270***	0.200***	0.184***	0.190***	0.180***
	(0.013)	(0.012)	(0.009)	(0.013)	(0.013)
White		-0.077	-0.104**		
		(0.070)	(0.060)		
Age		-0.041***	-0.036***		
		(0.013)	(0.067)		
Parents Some College		0.096	0.035		
		(0.083)	(0.067)		
Parents Bachelor Degree		0.127	0.065		
or More		(0.095)	(0.079)		
Income in [$20000, $40000]		-0.224***	-0.160***		
		(0.076)	(0.065)		
Income more than $40,000		-0.169**	-0.160***		
		(0.086)	(0.065)		
Distrust Index		-0.007	-0.001	-0.015	-0.015
		(0.010)	(0.009)	(0.012)	(0.012)
Ineffectiveness Index		0.123***	0.118***	0.114***	0.099***
		(0.010)	(0.009)	(0.011)	(0.011)
Perfectionism Index		0.066***	0.060***	0.092***	0.044***
		(0.009)	(0.008)	(0.013)	(0.018)
Body Dissatisfaction Index			0.019***		0.033***
			(0.003)		(0.007)
Interpolated Indices	No	No	No	No	No
Chamberlain/Wooldridge Fixed Effects	No	No	No	Yes	Yes
Sample Size	4151	3938	3928	3938	3928

Notes: Standard errors robust to intra-individual correlation are reported in parenthesis.
* indicates significant at 10%; ** significant at 5%; *** significant at 1%.

6 Conclusions

This is the first study that quantifies the role of true state dependence and individual heterogeneity in bulimia nervosa among adolescent girls. We use a panel data set, the NHLBI Growth and Health Survey, that is uniquely suited for studying these issues. A major advantage of these data is that all sample participants were evaluated regarding bulimic behaviors for ten years, starting when they were young (aged 11-12 years), independent of any diagnoses or treatment they had received. For each respondent, the data contain i) an Eating Disorders Inventory index, developed by medical experts; ii) information on SES, and iii) information on time-changing personality traits.

Our use of these data produces a number of important results. First, and perhaps most importantly, we find that much of the persistence in bulimic behavior is due to true state dependence after controlling for individual heterogeneity, and that this result continues to hold when we allow for the possibility that the personality traits are correlated with an individual random effect (possibly driven by a genetic factor), and the possibility that there is feedback from the current shock in BN to future values of the personality indices. Indeed we find that up to two-thirds of the persistence in BN is due to the true state dependence, and that the past four years of behavior positively and significantly impact bulimic behavior in the current period.

Further, we show that African-Americans are more likely to persist in bulimic behavior relative to Whites. Indeed, the estimates suggest that the impact of past behavior on current behavior is four-fold higher among African-Americans. In addition, the strongest persistence (among income groups) is present in low income families.

Our results have several important policy implications. First, since state dependence plays an important role in BN persistence, it is reasonable to expect that the longer an individual experiences BN, the less responsive she will be to policy aimed at combatting it. In this respect it is important to instruct a wide range of young women on the deleterious effects of BN and the importance of getting help, especially at the initial stages of bulimic behaviors. In addition, to the extent that poor health is linked with lower educational attainment, policy aimed at combatting the onset of bulimic behaviors among young girls could also serve to improve educational attainment.

Finally, a number of aspects of BN behavior are consistent with medical criteria that define an addiction. According to the DSM-IV, in order to be classified as an addiction, a behavior or substance abuse must satisfy at least three of seven criteria in a given year: 1) experiencing a persistent desire for the substance or behavior or an inability to reduce or control its use; 2) use of the substance or behavior continuing despite known adverse consequences; 3) withdrawal; 4) tolerance (more is needed for the same effect); 5) taking a larger amount of the substance or taking the substance for a longer period, than was intended; 6) spending much time seeking or consuming the substance or recovering from its effects; and 7) use of the substance or behavior interfering with important activities.[29] It is straightforward to see that BN fulfills criterion 1 (inability to control its use) as one of the diagnostic criteria for BN involves loss of control over the eating process.[30] Regarding criterion 2, we document that young women persist in their behaviors. Due to data limitations we are not able to determine whether the respondents are aware of the negative consequences of their behavior; however, a number of the adverse health effects will be readily apparent to anyone who continues with BN behavior, such as inflamed and irritated esophagus, tooth decay, muscle weakness, gastric rupture, and anemia. In this sense the continued behavior is consistent with addiction criterion 2 (i.e., use continues despite known adverse consequences). There is separate scientific evidence of withdrawal symptoms (criterion 3) in laxative use, which is a purging behavior (Colton et al., 1998). Hence, while not conclusive, the evidence is suggestive that BN may satisfy at least some of the criteria of a medical addiction.

[29] Further, note that to be diagnosed with a physiological dependence it is necessary that either criterion 3 or 4 be met; thus, physiological dependence is neither necessary nor sufficient for the medical definition of addiction.

[30] Corwin and Grigson (2009) note that other diagnostic criteria for bingeing-related disorders approximate the DSM-IV criteria for addiction. These include binge-type consumption, (i.e., criterion 5); bingeing is followed by inappropriate compensatory behavior (i.e., criterion 2); bingeing occurs at least twice a week for 3 months (i.e., criterion 5). Their argument is not based on an empirical analysis, but rather on their interpretation of the relation between the DSM-IV addiction and BN criteria.

Appendix

A Data Variable Definitions

We describe the construction of the ED-BN index in the main text of the paper. The body dissatisfaction index is based on subject responses to nine items: 1) I think that my stomach is too big, 2) I think that my thighs are too large, 3) I think that my stomach is just the right size, 4) I feel satisfied with the shape of my body, 5) I like the shape of my buttocks, 6) I think my hips are too big, 7) I think that my thighs are just the right size, 8) I think that my buttocks are too large, 9) I think my hips are just the right size. This index ranges from 0 to 27, and responses are scored such that a higher score indicates greater dissatisfaction.[31]

The perfectionism index is based on subject responses to six items: 1) In my family everyone has to do things like a superstar; 2) I try very hard to do what my parents and teachers want; 3) I hate being less than best at things; 4) My parents expect me to be the best; 5) I have to do things perfectly or not to do them at all; 6) I want to do very well. The subjects are offered the same responses, and the responses are scored in the same way as the ED-BN index.

The distrust index is based on subject responses to seven items: 1) I tell people about my feelings; 2) I trust people; 3) I can talk to other people easily; 4) I have close friends; 5) I have trouble telling other people how I feel; 6) I don't want people to get to know me very well; and 7) I can talk about my private thoughts or feelings. The scoring rule is as follows: "always"=1, "usually"=2, "often"=3, "sometimes"=4, "rarely"=5, and "never"=6 in questions 5 and 6; and "always"=6, "usually"=5, "often"=4, "sometimes"=3, "rarely"=2, and "never"=1 in questions 1, 2, 3, 4, and 7. A response of 4-6 on a given question contributes zero points to the distrust index; a response of 3 contributes 1 point; a response of 2 contributes 2 points; and a response of 1 contributes 3 points. The distrust index is a sum of all contributing points.

The ineffectiveness index is based on subject responses to ten items: 1) I feel I can't do things very well; 2) I feel very alone; 3) I feel I can't handle things in my life; 4) I wish I were someone else; 5) I don't think I am as good as other kids; 6) I feel good about myself; 7) I don't like myself very much; 8) I feel I can do whatever I try to do; 9) I feel I am a good person; 10) I feel empty inside. The scoring rule is as follows: "always"=1, "usually"=2, "often"=3,

[31] The scoring rule is as follows: "always"=6, "usually"=5, "often"=4, "sometimes"=3, "rarely"=2, and "never"=1 in questions 3, 4, 5, 7, and 9 and "always"=1, "usually"=2, "often"=3, "sometimes"=4, "rarely"=5, and "never"=6 in questions 1, 2, 6, and 8. Again a response of 4-6 on a given question contributes zero points to the body image index; a response of 3 contributes 1 point; a response of 2 contributes 2 points; and a response of 1 contributes 3 points. The body image index is the sum of the contributing points.

"sometimes"=4, "rarely"=5, and "never"=6 in questions 1,2,3,4,5,7, and 10; and "always"=6, "usually"=5, "often"=4, "sometimes"=3, "rarely"=2, and "never"=1 in questions 6,8, and 9. A response of 4-6 on a given question contributes zero points to the ineffectiveness index; a response of 3 contributes 1 point; a response of 2 contributes 2 points; and a response of 1 contributes 3 points. The ineffectiveness index is a sum of all contributing points.

Table A1 provides more details on the variables used in the paper.

Table A.1: Variable Definitions

Variable	Description	Coding	Waves
ED-BN Index	Eating Disorders Bulimia Subscale	Categorical Variable; Range 0-21	3,5,7,9,10
Clinical Bulimia	Case of Clinical Bulimia	=1 if ED-BN Index >10; =0 Otherwise	3,5,7,9,10
Body Dissatisfaction Index	Measures Poor Body Image Concerns	Categorical Variable; Range 0-27	3,5,7,9,10
Perfectionism Index	Measures Driveness for Perfection	Categorical Variable; Range 0-18	3,5,9,10
Ineffectiveness Index	Measures Feelings of Ineffectiveness	Categorical Variable; Range 0-29	3,5,9,10
Distrust Index	Measures Interpersonal Distrust	Categorical Variable; Range 0-21	3,5,9,10
Age	Respondent Age		All 10
White	Respondent Race is White	=1 if Race is White; =0 if African American	1
Parents High School or Less	Highest Education of Parents	Dummy Variable Highest Education High School or Less	1
Parents Some College	Highest Education of Parents	Dummy Variable Highest Education Some College	1
Parents Bachelor Degree or More	Highest Education of Parents	Dummy Variable Highest Education College Degree or More	1
Income less than $20,000	Household income (in 1988$)	Dummy Variable Household Income is Less than $20,000	1
Income in [$20000, $40000]	Household income (in 1988$)	Dummy Variable Household Income is in Range [$20,000,$40,000]	1
Income more than $40,000	Household income (in 1988$)	Dummy Variable Household Income is Higher than $40,000	1

B Additional Regression Results

Table B1 presents the reduced form estimates to investigate the issue of weak instruments. There will be heteroskedasticity in the first-stage regression equation for a censored dependent variable; therefore, the widely used rule of thumb for the first-stage F-statistic of excluded instruments (from Staiger and Stock (1997) and Stock and Yogo (2005)) will be inappropriate. Instead, we use the conjecture by Hansen, Hausman, and Newey (2008) that in the presence of heteroskedasticity in the first-stage equation, the Wald statistic for the null hypothesis that the excluded instruments are zero in the first stage, minus the number of instruments, should be greater than 32. The estimates in Columns (1)-(3) consider the instruments on an individual basis. We pass the weak IV test for the perfectionism and ineffectiveness.[32]

[32] We also estimated the specification in column (6) of Table 4 using separate instruments including body dissatisfaction. The results are very similar and are available upon request.

Table B1: Additional First Stage Estimates for Table 3 (Persistence of ED-BN Index)

	Estimates For Specification from Column (5) of Table 3 using Separate Instruments		
	(1)	(2)	(3)
Instruments for Lagged ED-BN Index			
Lagged Perfectionism Index	0.212 ***		
	(0.020)		
Lagged Ineffectiveness Index		0.286 ***	
		(0.017)	
Lagged Distrust Index			0.108 ***
			(0.020)
Lagged Dissatisfaction Index			
Other Regressors			
White	-0.300 **	-0.350 ***	-0.388 ***
	(0.135)	(0.129)	(0.137)
Age	-0.105 **	-0.055 **	-0.074 ***
	(0.028)	(0.027)	(0.028)
Parents Some College	-0.205	-0.150	-0.182
	(0.163)	(0.157)	(0.165)
Parents Bachelor Degree	-0.626 ***	-0.350 **	-0.553 ***
or More	(0.184)	(0.177)	(0.185)
Income in [$20000, $40000]	-0.251	-0.038	-0.276 *
	(0.167)	(0.161)	(0.168)
Income more than $40,000	-0.180	-0.064	-0.182
	(0.180)	(0.173)	(0.182)
Distrust Index	0.063 ***	0.050 ***	0.034 *
	(0.019)	(0.019)	(0.020)
Ineffectiveness Index	0.170 ***	0.044 ***	0.171 ***
	(0.016)	(0.017)	(0.016)
Perfectionism Index	-0.020	0.063 ***	0.055 ***
	(0.019)	(0.017)	(0.018)
Body Dissatisfaction Index			
Constant	1.842 ***	1.231 ***	1.979 ***
	(0.458)	(0.441)	(0.475)
Weak IV Test Statistic*	135	452	28
Interpolated Values	No	No	No
Sample Size	2309	2303	2308

Notes: Standard errors robust to heteroskedasticity and intra-group correlation are reported in parenthesis. Regarding the weak IV test, Hansen, Hausman, and Newey (2008) suggest that, in the presence of heteroskedasticity in the first stage equation, the test statistic should be greater than 32.
* indicates significant at 10%; ** significant at 5%; *** significant at 1%.

The dynamic LPM and Probit model estimates are in Tables B2 and B3, respectively. These results suggest that the dynamic model is too rich for the zero-one data, since the IV regression coefficient on the lagged dependent variable is significant only if we difference the data and use the AB approach. Further, the Probit partial effects for the lagged incidence of BN are not significant once we include the fixed effects. The insignificant partial effects on the lagged incidence of BN in columns (4) and (5) have large confidence intervals; in other words, they are imprecisely estimated "zero" coefficients.

Table B2: Linear Probability Estimates of the Persistence of Clinical Bulimia

Variables				Two Stage Least Squares				Arellano-Bond	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Lagged Clinical Bulimia	0.196***	0.150***	0.149***	0.034	0.005	0.017	-0.008	0.093**	0.093**
	(0.043)	(0.041)	(0.041)	(0.090)	(0.062)	(0.089)	(0.060)	(0.050)	(0.051)
White		-0.005	-0.005	-0.007	-0.004	-0.009	-0.005	-0.017**	-0.019**
		(0.005)	(0.005)	(0.008)	(0.004)	(0.008)	(0.004)	(0.008)	(0.008)
Age		-0.002**	-0.003**	-0.002	-0.003***	-0.002	-0.003***		
		(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)		
Parents Some College		0.001	0.001	-0.004	-0.005	-0.004	-0.005		
		(0.006)	(0.006)	(0.010)	(0.005)	(0.010)	(0.005)		
Parents Bachelor Degree		0.006	0.006	0.002	-0.001	0.002	-0.001		
or More		(0.007)	(0.007)	(0.011)	(0.006)	(0.011)	(0.006)		
Income in [$20000, $40000]		-0.007	-0.007	-0.009	-0.008*	-0.010	-0.009*		
		(0.007)	(0.007)	(0.010)	(0.005)	(0.010)	(0.005)		
Income more than $40,000		-0.009	-0.009	-0.010	-0.012**	-0.011	-0.012**		
		(0.006)	(0.006)	(0.010)	(0.005)	(0.011)	(0.005)		
Distrust Index		-0.001	-0.001	-0.002*	-0.001	-0.002*	-0.000	-0.002**	-0.002**
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Ineffectiveness Index		0.008***	0.008***	0.011***	0.008***	0.010***	0.008***	0.006**	0.006**
		(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Perfectionism Index		0.003***	0.003***	0.005***	0.004***	0.005***	0.004***	0.003***	0.003***
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Body Dissatisfaction Index			0.001			0.001**	0.001**		
			(0.000)			(0.001)	(0.000)		
Constant	0.016***	0.023	0.024	0.010	0.031*	0.011	0.033*	0.027*	0.028*
	(0.002)	(0.020)	(0.020)	(0.026)	(0.017)	(0.027)	(0.017)	(0.016)	(0.016)
Interpolated Indices	No	No	No	No	No	Yes	Yes	Yes	Yes
First Difference	No	No	No	No	No	No	No	Yes	Yes
Sample Size	4151	3938	3928	2285	2273	5426	5384	3437	3411

Notes: Standard errors robust to heteroskedasticity and intra-group correlation are reported in parenthesis. NA denotes not applicable; * indicates significant at 10%; ** significant at 5%; *** significant at 1%. Instruments are one-period lags of: all personality indices in columns (5) and (7); all indices excluding body dissatisfaction in columns (4) and (6). Columns (6)- (9) use interpolated values of personality indices in wave 7.

Table B3: Probit Partial Effects for the Persistence of Clinical Bulimia

Variables	(1)	(2)	(3)	(4)	(5)
Lagged Clinical Bulimia	0.196***	0.074***	0.070***	0.017	0.017
	(0.044)	(0.025)	(0.024)	(0.016)	(0.016)
White		-0.009**	-0.011***	-0.005	-0.006*
		(0.004)	(0.004)	(0.003)	(0.003)
Age		-0.003***	-0.003***	-0.002**	-0.002***
		(0.001)	(0.001)	(0.001)	(0.001)
Parents Some College		0.001	0.000	0.001	0.000
		(0.005)	(0.005)	(0.004)	(0.004)
Parents Bachelor Degree or More		0.006	0.005	0.005	0.004
		(0.007)	(0.007)	(0.005)	(0.005)
Income in [$20000, $40000]		-0.007	-0.007	-0.005	-0.005
		(0.004)	(0.004)	(0.003)	(0.003)
Income more than $40,000		-0.008*	-0.008*	-0.004	-0.005
		(0.004)	(0.004)	(0.003)	(0.003)
Distrust Index		-0.001	-0.001	-0.001	-0.001
		(0.001)	(0.001)	(0.001)	(0.001)
Ineffectiveness Index		0.003***	0.003***	0.002***	0.002***
		(0.000)	(0.001)	(0.001)	(0.001)
Perfectionism Index		0.002***	0.002***	0.001**	0.001**
		(0.001)	(0.001)	(0.001)	(0.001)
Body Dissatisfaction Index			0.001***		0.001
			(0.001)	(0.007)	(0.001)
Chamberlain/Wooldridge Fixed Effects	No	No	No	Yes	Yes
Constant	-2.137***	-0.653*	-1.500***	-1.437***	-1.812***
	(0.050)	(0.385)	(0.429)	(0.442)	(0.576)
Sample Size	4151	3938	3938	3938	3928

Notes: Standard errors robust to intra-individual correlation are in parenthesis. * indicates significant at the 10% level; ** at 5%; *** at 1%.

References

Adams, P., M. Hurd, D.McFadden, A. Merrill, and T. Ribeiro (2003) "Healthy, Wealthy and Wise? Tests for Direct Causal Paths Between Health and Socioeconomic Status," *Journal of Econometrics* 112: 3-56.

American Psychiatric Association (APA) (1993) "Practice Guideline for Eating Disorders," *American Journal of Psychiatry* 150: 212-228.

American Psychiatric Association (APA) (2000a) Diagnostic and Statistical Manual of Mental Disorders. American Psychiatric Press: Washington DC. 4th Edition.

American Psychiatric Association (APA) (2000b) "Practice Guideline for the Treatment of Patients with Eating Disorders (Revision)," *American Journal of Psychiatry* 15: 1-39.

Anderson, T. and C. Hsiao (1981) "Estimation of Dynamic Models with Error Components," *Journal of the American Statistical Association* 76: 598-606.

Arellano, M. and S. Bond (1991) "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies* 58: 277-97.

Borghans, L., A. Duckworth, J. Heckman, and B. ter Weel (2008) "The Economics and Psychology of Personality Traits," *Journal of Human Resources* 43: 972-1052.

Bulik, C., B. Devlin, and S. Bacanu (2003) "Significant Linkage on Chromosome *10p* in Families with Bulimia Nervosa," *American Journal of Human Genetics* 72: 200-207.

Cavanaugh, C. and L. Ray (1999) "What we Know about Eating Disorders: Facts and Statistics," In Lemberg, Raymond and Cohn, Leigh (Eds) Eating Disorders: A Reference Sourcebook. Oryx. Press. Phoenix, AZ.

Center for Mental Health Services (2008) "Report on State Parity Laws," http://download.ncadi.samhsa.gov /Ken/pdf /SMA07-4228/ CMHS17a_StateMandates_web.pdf.

Chaloupka, F. and Warner (2000) "The Economics of Smoking," in Handbook of Health Economics, North Holland.

Chamberlain, G. (1984) "Panel Data," in Z. Griliches and M.D. Intriligator (eds.), Handbook of Econometrics, Volume 2. Amsterdam: North Holland, 1247-1318.

Daly, R. (2008) "Few States Act on Parity Improvements This Year," *Psychiatry News* 43: 2-25.

Department of Health and Human Services (2006) Agency for Healthcare Research and Quality, Management of Eating Disorders, Evidence Report/Technology Assessment, Number 135, 2006. AHRQ publication 06-E010. Senate Report accompanying the 2006 budget as reported by the Eating Disorders Coalition. URL: www.eatingdisorderscoalition.org.

Engers, M. and S. Stern (2002) “Long-Term Care and Family Bargaining,” *International Economic Review* 43(1): 73-114.

Garner, D., P. Marion, M. Olmstead, and J. Polivy (1983) “Development and Validation of a Multidimensional Eating Disorder Inventory for Anorexia Nervosa and Bulimia,” *International Journal of Eating Disorders* 2: 15-34.

Gidwani, G.P. and Rome, E.S. (1997) “Eating Disorders,” *Clinical Obstetrics and Gynecology* 40:601-615.

Gilleskie, D. and K. Strumpf (2005) “The Behavioral Dynamics of Youth Smoking,” *Journal of Human Resources* 40: 822-866.

Glewwe, P. and E. Miguel (2008) “The Impact of Child Health and Nutrition on Education in Less Developed Countries,” Handbook of Development Economics Volume 4, (eds.) T. Paul Schultz and John Strauss, Elsevier.

Ham, J., Iorio, D., and M. Sovinsky (2011) “On the Incidence of Bulimia Among Blacks and Whites," mimeo, University of Zurich.

Hansen, C., J. Hausman, and W. Newey (2008) “Estimation with Many Instrumental Variables,” *Journal of Business and Economic Statistics* 26: 398-422.

Heckman, J. (1981) “Heterogeneity and State Dependence,” in Sherwin Rosen (Eds), Studies in Labor Markets, Chicago: University of Chicago Press.

Heckman, J. and J. Hotz (1989) “Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training,” *Journal of the American Statistical Association* 84: 862-880 (with discussion).

Heckman, J., A. Basu, S. Navarro, and S. Urzua (2007) “Use of Instrumental Variables in the Presence of Heterogeneity and Self-Selection: An Application to Treatments of Breast Cancer Patients,” *Health Economics* 16: 1133-1157.

Hinton, I., J. Howell, E. Merwin, S. Stern, S. Turner, I. Williams, and M. Wilson (2010) “The Educational Pipeline for Health Care Professionals: Understanding the Source of Racial Differences,” *Journal of Human Resources* 45: 116-156.

Hoek, H.W., and van Hoeken, D. (2003) “Review of the Prevalence and Incidence of Eating Disorders,” *International Journal of Eating Disorders* 23: 383-396.

House of Representatives (2008) 6983, 110th Congress, Paul Wellstone and Pete Domenici Mental Health Parity and Addiction Equity Act of 2008

Hudson, J., E. Hiripi, H. Pope Jr, and R. Kesslre (2007) “The Prevalence and Correlates of Eating Disroders in the National Comorbidity Survey Replication,” *Biological Psychiatry* 61: 348-358.

Labeaga, J. and A. Jones (2003) “Individual Heterogeneity and Censoring in Panel Data Estimates of Tobacco Expenditures,” *Journal of Applied Econometrics* 18(2): 157-77.

Lewbel, A. (2007) "Using Heteroskedasticity to Identify and Estimate Mismeasured and Endogenous Regressor Models," Boston College Working Paper.

Lilenfeld, L., W. Kaye, and C. Greeno (1998) "A Controlled Family Study of Restricting Anorexia and Bulimia Nervosa: Comorbidity in Probands and Disorders in First-Degree Relatives," *Archives of General Psychiatry* 55: 603-10.

National Eating Disorders Association (2008) URL: www.nationaleatingdisorders.org. Accessed Dec 29, 2008.

National Youth Risk Behavior Survey (2005) URL: http:// www.cdc.gov /HealthyYouth /yrbs/index.htm.

Reagan, P. and J. Hersch (2005) "Influence of Race, Gender, and Socioeconomic Status on Binge Eating Frequency in a Population-Based Sample," *International Journal of Eating Disorders* 38: 252–256.

Reas, D., D.Williamson, C. Martin, and N. Zucker (2000) "Duration of Illness Predicts Outcome for Bulimia Nervosa: A Long-term Follow-up Study," *International Journal of Eating Disorders* 27: 428-34.

Rush, J., M. First, and D. Blacker (2008) Handbook of Psychiatric Measures. American Psychiatric Publishing: Arlington, VA, 2nd Edition.

Society for Adolescent Medicine (2003) "Eating Disorders in Adolescents: Position Paper of the Society for Adolescent Medicine," *Journal of Adolescent Health* 33: 496-503.

Smith, J. (2007) "The Impact of Social Economic Status on Health over the Life Course," *Journal of Human Resources* 42: 739-764.

Staiger, D., and J. Stock (1997) "Instrumental Variables Regression with Weak Instruments," *Econometrica* 65: 557-586.

Striegel-Moore, R., G. Schreiber, A. Lo, P. Crawford, E. Obarzanek, and J. Rodin (2000) "Eating Disorder Symptoms in a Cohort of 11 to 16-Year-Old Black and White Girls: The NHLBI Growth and Health Study," *International Journal of Eating Disorders* 27: 49-66.

Stock, J. and M. Yogo (2005) "Testing for Weak Instruments," Chapter 5, in Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg Andrews, D.W.K. and J.H. Stock eds. Cambridge University Press, Cambridge.

Wooldridge, J. (2002) Econometric Analysis of Cross Section and Panel Data. MIT Press.

Wooldridge, J. (2005) "Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Effects," *Journal of Applied Econometrics* 20: 39-54.