Depressive Symptoms and Endogenous Physical Activity:
An Ordered Probability Approach

Jun Zhang
Graduate Student

Steven T. Yen
Professor

Department of Agricultural and Resource Economics
The University of Tennessee
Knoxville, TN 37996-4518

Depressive Symptoms and Endogenous Physical Activity: An Ordered Probability Approach

Abstract

Depression is a serious mental disorder which affects more than 350 million people of all ages worldwide in the 2012 and physical activity is generally believed to be effective in combating depressive symptoms. This study investigates the effects of regular physical activity and socio-demographic factors on depressive symptoms for both men and women. Data for this study come from the 2011 Behavioral Risk Factor Surveillance System (BRFSS) and an ordered probability model with binary endogenous physical activity is developed to accommodate the ordinal nature of depression outcomes. Results suggest that physical activity is most beneficial for mild and moderate depressed individuals and the effect of regular physical activity is most notable on mild depressed females. In addition, socio-demographic factors are found to vary significantly between gender, and factors of age, income, race, education, employment status and recent mental health condition play important roles in affecting depressive symptoms.

Key words: Depressive symptoms, physical activity, switching probability model, treatment effect.

Introduction

Depression is a common mental disorder involving the brain and is commonly characterized by sadness, loss of interest or pleasure, feelings of guilt or low self-esteem, disturbed sleep or appetite, feelings of tiredness, and poor concentration. According to the World Health Organization (WHO 2012), more than 350 million people of all ages suffered from depression during 2012 in all parts of the world. It is estimated that 1 out of 20 people reported having an episode of depression in the previous year worldwide (Kessler et al. 2008), by 2020, depression will be the second leading cause of world disability (WHO 2001) and by 2030 it is expected to be the largest contributor to disease burden (WHO 2008).

Many people in developed countries suffer from depression and other diseases related to depression. In the years of 2006 and 2008, about 9% of Americans met the criteria for current depression and 3.4% for major depression (Centers for Disease Control and Prevention (CDC) 2010).[1] Depression was the third leading cause of disease burden worldwide and a leading cause of disability in high-income countries in 2004 (WHO 2008). Depression can adversely affect the outcome of common chronic conditions, such as arthritis, asthma, cardiovascular disease, cancer, diabetes, and obesity (Chapman et al. 2005), it can also result in increased work absenteeism, short-term disability, and decreased productivity (Goetzel et al. 2003).

Factors contributing to depression are complicated and include both biological and social factors. Some researchers attribute depression primarily to biological factors (e.g., Ranga and Krishnan 2002; Riso, Miyatake, and Thase 2002), while a number of other studies suggest that depression is mainly caused by social factors rather than biological factors (e.g., Jorm et al. 1997; Hansson et al. 2009). Known to contribute to depression are many traditional socio-demographic

---

[1] Current depression is defined as meeting the Behavioral Risk Factor Surveillance System (BRFSS) criteria for either major depression or "other depression" during the 2 weeks preceding the survey (CDC 2010).

factors such as marital status, gender, income and age (e.g., Addis 2008; De-Velde, Bracke, and Levecque 2010), and other factors such as physical activity level (e.g., Camacho 1991; Robertson 2012; Goodwin 2003; De-Moor et al. 2006).

A number of studies have investigated the relationship between physical activity and mental health (e.g., Farmer et al. 1988; Camacho et al. 1991; Goodwin 2003; De-Moor et al. 2006), and physical activity is also regarded as an important way to alleviate depressive symptoms (Salmon 2001; Mota-Pereira et al. 2011). One popular explanation for the relation between exercise and depression is based on the theory that exercise has a positive effect on depression due to an increased release of beta-endorphins following exercise. Endorphins are related to positive mood and thus enhance the sense of well-being (Craft et al. 2004). Another explanation, related to the theory of self-efficacy, is that exercise would increase the feeling of coping self-efficacy which is inversely related to depression (Craft 2005). However, Chalder et al. (2012) suggest that adding a physical activity intervention to usual care does not reduce symptoms of depression more than usual care alone. This finding challenges the current clinical guidance which recommends exercise to help combat depression (Babyak et al. 2000; Foley et al. 2008; Hoffman et al. 2011).

Although findings have not been entirely consistent, many studies suggest that physical activity or exercise could reduce symptoms of mild to moderate depression (e.g., Babyak et al. 2000; Foley et al. 2008; Mota-Pereira et al. 2011). Most of these studies rely on small and selected clinical samples which do not represent the general population. In addition, few studies have taken other socio-demographic factors into account when studying the relation between physical activity and depression; exceptions include Farmer et al. (1988) and Goodwin (2003). Analysis not accounting for other factors can be misleading if both physical activity and socio-

demographic factors affect the level of depressive symptoms simultaneously. To our knowledge, none of the previous studies have investigated the quantitative effects of physical activity and other socio-demographic factors simultaneously on the ordinal outcome of depressive symptoms. This study will fill this gap of knowledge. Findings can inform policy deliberation in the effort to improve the mental health and general health of targeted population.

Literature review

Physical activity and depression

The empirical literature on depression and physical activity has provided much evidence that physical activity is negatively related to the level of depressive symptoms, and many researchers employ a variety of methods and data from either surveys or clinical samples. Early examples of the research have shown solid evidence that physical activity is likely to affect the level of depressive symptoms. In the 1980s, Farmer et al. (1988) find a negative association in white individuals between physical activity and depressive symptoms by using the National Health and Nutrition Examination Survey and logistic regression models that include demographic variables of age, race, education, employment status, self-reported health, household income and length of follow-up. Camacho et al. (1991) use a method similar to Farmer et al. but a different dataset. Based on samples from 1965 to 1983 in the Alameda County, California, they find that men and women who report a low activity level at baseline have greater risk of depression than those who report a high activity level. This finding suggests that high activity level can indeed reduce the risk of depression in the long term.

Many recent studies on depression also confirm that physical activity or exercise can reduce the level of depressive symptoms. Goodwin (2003) estimates the impacts of self-reported

physical activity and socio-demographic factors on mental disorders by using samples from the National Comorbidity Survey. Results from logistic regressions indicate that regular physical activity is associated with a significantly decreasing likelihood of having current major depression. De-Moor et al. (2006) empirically show that regular exercisers are on average less depressed than non-exercisers by using large national samples from the Netherland. Hamer, Stamatakis, and Steptoe (2008) use data from the Scottish Health Survey to further supplement different types of physical activity in relation to mental health, and demonstrate strong associations between physical activity and the reduced odds of psychological distress. Sieverdes et al. (2012) focus on leisure time physical activities of men and divide individuals into categories according to time spent on physical activities per week. They find that men in median and high physical activity categories are 51% less likely to have depressive symptoms compared with men who do not participate in any physical activities.

In clinical research, physical activity is shown to be an effective treatment to alleviate mild and moderate depressive symptoms. Babyak et al. (2000) show that among individuals with major depressive disorder (MDD), after 4 months of treatment with exercise 60.4% of patients in the exercise group no longer meet the DSM-IV criteria for MDD.[2] Foley et al. (2008) find that both aerobic and stretching exercises are associated with significant decreases in severity of depression and increased in coping efficacy and episodic memory over 12 weeks. Mota-Pereira et al. (2011) suggest a 12 week exercise program of 30-45 minutes walks 5 times a week results in the improvement of all studied parameters of depression and this improvement is not due to social interaction. Hoffman et al. (2011) find that among clinical samples of depressed elder adults, 46% were fully remitted at the end of the original 4-month study treatment with exercise,

---

[2] DSM-IV is short for the Diagnostic and Statistical Manual of Mental Disorders, 4th edition (American Psychiatric Association 1994).

and 66% were fully remitted 1 year after the end of treatment. This finding suggests a lasting effect of physical activity in reducing depressive symptoms. Most recently, Chalder et al. (2012) use samples from 361 depressed adults age 18-69 and find no evidence that participants offered the physical activity intervention reported improvement in mood by the four month follow-up point compared with those in the usual care group. However, the result is questionable because the effect of exercise on depression was not tested and further, the samples are very small.

Socio-demographic factors and depression

Besides physical activity, socio-demographic factors are also found to play a role in affecting depression. Age is generally accepted as an important factor, but the relation between age and depression is not consistent in previous findings. Mirowsky and Ross (1992) suggest a U-shape relation between age and depression and find depression reaches its lowest level around age 45 with samples from the years of 1985 and 1990 in the United States (US). Kessler et al. (1992) show a similar relationship with samples from two national surveys of the US. Wade and Cairney (1997) find a steady decline across age groups after other socio-demographic factors are controlled for, by using Canadian samples. Schieman et al. (2002) reinforce the notion of negative linear relationship between age and depression with data from physically disabled and nondisabled residents respectively. Streiner et al. (2006) provide evidence of a linear decrease in all disorders after age 55 for men and women, for both people born in Canada and people who immigrated to Canada after age 18.

Gender is another important factor, and most earlier studies have concluded that women have higher risk than men of having depressive symptoms. Kessler et al. (1993) suggest that depressive disorders are more common in women, who have lifetime rates for major depressive episodes of 21.3%, compared with 12.7% in men. Using logistic regression, Goodwin and Gotlib

(2004) find that women are more likely to have major depression than men. De-Velde, Bracke, and Levecque (2010) estimate the gender difference in depression with large datasets of 23 European countries and their results indicate that women report higher levels of depression than men do in all countries. They also confirm that socio-demographic factors have strong association with depression in both men and women.

Plenty of previous studies suggest that income, race, education, marriage and employment status can affect depressive symptoms. Whooley et al. (2002) find low-income young adults more likely to have depressive symptoms than high-income young adults by using a sample of 5115 individuals age 18 to 30. Zimmerman and Katon (2005) report a negative relation between income and depression symptoms with Kernel regression for both men and women. Somervell et al. (1989) use large samples from 5 communities in the US to test the difference in major depression between white and black adults. Results show that in the 18-24 age group, white men have higher prevalence of depression than black men, while white women have lower prevalence of depression than black women. Bromberger et al. (2004) indicate that compared with white women, African American and Hispanic women have higher odds and Chinese woman have lower odds, of a CES-D score of 16 or higher.[3] Craig and Natta (1979) study the influence of education on depressive symptoms and find that less educated individuals are more likely to exhibit depressive symptoms. Jang et al. (2009) investigate the relation between marital status and depression with large samples for Korean individuals age 45 and above. Results reveal that both male and female who are divorced, separated or widowed have higher scores for depression than married individuals. Based on logistic regression with panel data, Dooley, Catalano, and Wilson (1994) find that unemployment increases the risk of depressive symptoms.

---

[3] CES-D is short for the Center for Epidemiologic Studies Depression Scale. See Radloff (1977).

Conceptual framework

The empirical model of this study is derived from the utility maximization framework where utility is specified as a function of the level of depressive symptoms. Assume that each individual's current level of depressive symptoms is $D = D(M, E; Z_1)$, which is determined by the current mental health status (M), recent physical activity or exercise (E) and socio-demographic variables (Z_1). Deriving utility from the level of depressive symptoms (D), current mental health status (M) and recent physical activity (E), each individual has a utility function

$$U = U(D(M, E; Z_1), M, E; Z_2) \tag{1}$$

where Z_2 is another set of socio-demographic variables. Under the assumption that health condition of each individual is restricted by age, this utility function is maximized subject to the health condition constraint

$$g(M, E, H_m, H_p) \leq f(A) \tag{2}$$

where $g(\cdot)$ is a function which reflects the current health status in numerical values, $f(\cdot)$ is a function of age which reflects the optimal possible health status of normal people at specific age (A) in numerical values, H_m is the recent mental health condition, and H_p is the recent physical health condition.[4] Solving the constrained utility maximization problem in equation (1) yields the optimal level of current level of depressive symptoms, current mental health status, recent physical activity and socio-demographic variables. The optimal level of depressive symptoms (D^*) can be denoted as

$$D^* = D^*(M^*(H_m, H_p, A; Z_1, Z_2), \ E^*(H_m, H_p, A; Z_1, Z_2); \ Z_2) \tag{3}$$

---

[4] In this study, we assume previous mental and physical health conditions, especially recent conditions will affect current health status.

Thus optimal level of depressive symptoms is a function of current mental health status and physical activity. Note that physical activity is an endogenous variable, as is the optimum level of current mental health status. Lacking the information of current mental health status in the dataset, recent mental and physical health conditions are used to express current mental health status indirectly, and equation (3) is rewritten as

$$D^* = \tilde{D}^*(H_m, H_p, A; \ Z_1, Z_2, E^*(H_m, H_p, A; \ Z_1, Z_2); \ Z_2) \tag{4}$$

Drawing on the optimal depression level of equation (4), one relevant empirical specification is the treatment effect model (Barnow, Cain, and Goldberger 1980). Another empirical approach is the switching regression model, which is a more generalized case of treatment effect model. With the endogenous variables of physical activity, recent mental health condition, age and other socio-demographic variables, optimal level of depressive symptoms of each individual i is expressed by a linear equation:

$$y_i = x_i'\beta + \gamma d_i + u_i \quad , \quad i = 1\ldots n \tag{5}$$

Where y_i denotes observed level of depressive symptoms, x_i is a vector containing exogenous variables such as recent mental health condition, age and other socio-demographic characteristics (with corresponding parameter vector β), d_i is an endogenous variable representing physical activity (with scalar parameter γ), and u_i is random error which reflects the unobservable.

Empirical model

To our knowledge, the empirical approaches used by most previous studies are either logistic regression or linear regression. With the large portion of zeros in the ordinal outcome variable of depression, linear regression will produce biased estimates and logistic regression

cannot reflect the ordinal depression level accurately. Facing such ordinal outcome problems,

recent studies in the field of applied economics provide more efficient and accurate procedures.

Yen, Shaw, and Yuan (2010) implement an ordered probability treatment model to study the

effect of ordinal cigarette smoking on ordinal health outcomes. Li and Tobias (2005) and Yen,

Bruce, and Jahns (2012) extend the treatment model to a switching ordered probability model to

investigate the effect of binary participation in the Supplemental Nutrition Assistance Program

on ordinal self-accessed health conditions. We adopted the approach of ordered probability

model with binary endogenous switching to accommodate the ordinality of dependent variable

and better differentiate the effect of seldom and regular exercise on the depression categories.

Switching regression models date back to Roy (1951) who was concerned with an

individual's decision between earning income as fisher or hunters, and have been extensively

used in economics. Unlike the ordered treatment effect model, in the switching probability model

the ordinal outcome y_i (current depression level) is modeled with two different processes. For

current application, consider a binary switching equation for endogenous variable d_i (physical

activity)

$$
\begin{aligned}
d_i &= 1 \quad \text{if} \quad z_i'\alpha + v_i > 0 \\
&= 0 \quad \text{if} \quad z_i'\alpha + v_i \le 0
\end{aligned}
\tag{6}
$$

and a set of ordered probability models for the current depression level, for the regular exerciser

$(d_i = 1)$ and seldom exerciser $(d_i = 0)$ regimes

$$
y_i^{(s)} = k \quad \text{if} \quad \xi_{k-1}^{(s)} < x_i'\beta^{(s)} + u_i^{(s)} < \xi_k^{(s)}, \quad k = 1 \ldots K; \quad s = 0,1
\tag{7}
$$

where $y_i^{(1)}$ denotes the outcome received by each individual with treatment state $(d_i = 1)$ and

$y_i^{(0)}$ denotes the outcome received by each individual without treatment state $d_i = 0$. Only one

outcome, denoted as y_i, is observed for each individual, and thus

$$y_i = d_i y_i^{(1)} + (1 - d_i) y_i^{(0)} \qquad (8)$$

In equations (6) and (7), z_i and x_i are vectors of explanatory variables for individual i, α and

$\beta^{(s)}$ are conformable parameter vectors, v_i and $u_i^{(s)}$ are random errors for each individual, and

$\xi_k^{(s)}$ are threshold parameter such that $\xi_0^{(s)} = -\infty$, $\xi_1^{(s)} = 0$, $\xi_K^{(s)} = \infty$, and $\xi_2^{(s)} \dots \xi_{K-1}^{(s)}$ are

estimable. Assume that the random error vector $[v_i, u_i^{(0)}, u_i^{(1)}]'$ are distributed as standard

trivariate normal distribution with zero means, unitary variances, and a finite correlation matrix:

$$\begin{bmatrix} v_i \\ u_i^{(0)} \\ u_i^{(1)} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho_{v0} & \rho_{v1} \\ \rho_{0v} & 1 & \rho_{01} \\ \rho_{1v} & \rho_{10} & 1 \end{bmatrix} \right) \qquad (9)$$

As in conventional switching probability models, the error correlation ρ_{01} between $u_i^{(0)}$ and $u_i^{(1)}$

is not identified and needs not be identified. The parameters in (6), (7) and (9) are estimated by

the maximum-likelihood (ML) procedure. To construct the likelihood contribution for the sample

observation, first define a bivariate standard normal cumulative function (CDF)

$_2(x, y, \rho) = \Pr(X \leq x, Y \leq y)$ with correlation ρ and marginal CDFs $_1(x) = \Pr(X \leq x)$ and

$_1(y) = \Pr(Y \leq y)$. Then given the distribution of error terms in (9) and information from

equation (6)-(8), the likelihood contributions for the two distinctive sample regimes $(d_i = 0)$ and

$(d_i = 1)$ can be obtained.

The model reduces to the treatment effect model (Yen, Shaw, and Yuan 2011) by

imposing parametric restrictions that all "slope" coefficients in $\beta^{(0)}$ and $\beta^{(1)}$ and the two error

correlations ρ_{v0} and ρ_{v1} are equal between the two regimes. In addition, imposing the restriction

that $\rho_{v0} = \rho_{v1} = 0$ reduces the model to one with exogenous switching, in which case, binary probit for the switching equation and ordered probit for each of the two regimes can be estimated separately. Tests for the above restrictions can be carried out by regular means, using likelihood-ratio (LR), Wald, or Lagrange multiplier (LM) test.

To facilitate interpretation of the effects on explanatory variables, marginal effects of explanatory variables on the probabilities of depression categories and treatment effect of physical activity on the depression categories are calculated. Specially, for each individual, the probabilities of being regular or seldom exercisers are

$$\Pr(d_i = 1) = \Phi_1(z_i'\alpha) \tag{10}$$

$$\Pr(d_i = 0) = 1 - \Phi_1(z_i'\alpha) \tag{11}$$

applying information above and equations (10) and (11), the probabilities of each depression category conditional on seldom exerciser and regular exerciser are

$$\Pr(y_i^{(0)} = k, d_i = 0 \mid d_i = 0) = \frac{\Phi_2(-z_i'\alpha, \xi_k^{(0)} - x_i'\beta^{(0)}, \rho_{v0}) - \Phi_2(z_i'\alpha, \xi_{k-1}^{(0)} - x_i'\beta^{(0)}, \rho_{v0})}{1 - \Phi_1(z_i'\alpha)} \tag{12}$$

$$\Pr(y_i^{(1)} = k, d_i = 1 \mid d_i = 1) = \frac{\Phi_1(\xi_k^{(1)} - x_i'\beta^{(1)}) - \Phi_1(\xi_{k-1}^{(1)} - x_i'\beta^{(1)})}{\Phi_1(z_i'\alpha)}$$
$$- \frac{\Phi_2(-z_i'\alpha, \xi_k^{(1)} - x_i'\beta^{(1)}, \rho_{v1}) - \Phi_2(z_i'\alpha, \xi_{k-1}^{(1)} - x_i'\beta^{(1)}, \rho_{v1})}{\Phi_1(z_i'\alpha)} \tag{13}$$

Marginal effects of each continuous (binary) explanatory variable can be derived by differentiating (differencing) equations (10), (11), (12), and (13). In addition, the treatment effect of physical activity on each depression category is

$$TE_k = \Pr(y_i = k \mid d_i = 1) - \Pr(y_i = k \mid d_i = 0), \quad k = 1 \ldots K \tag{14}$$

For statistical inference, standard errors of the marginal and treatment effects can be derived by the delta method (Papke and Wooldridge 2005).

Data and samples

Data for this study come from the 2011 Behavioral Risk Factor Surveillance System (BRFSS) collected by state health departments in collaboration with the Centers for Disease Control (CDC). The BRFSS is a state-based system of health surveys that collects information on health risk behaviors, and the 2011 BRFSS is the most recent large national survey which provides adequate information for depression and socio-demographic factors. After removing missing values for important variables, the pooled sample consists of 11,560 individuals age 18 to 99, of which 4,798 are males and 6,762 are females.

Dependent variable

The outcome variable is current depression level, which is measured by the eight-item self-reported Patient Health Questionnaire Depression Scale (PHQ-8). PHQ-8 covers eight of the nine criteria from the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) for diagnosis of major depressive disorder (CDC 2010). The ninth criterion in the DSM-IV is omitted because it is to access extreme depressive symptoms, such as suicide, which is beyond the scope of this study. PHQ-8 is one of the valid diagnostic and severity measures for depression in large clinic studies (e.g., Kroenke et al. 2009). Compared with the Center for Epidemiologic Studies Depression Scale (CES-D), the PHQ-9 (with additional suicide item than PHQ-8) is reliable and advantageous because it is just half the length of CES-D (Milette et al. 2010).

The depression section of the BRFSS questionnaire provides eight self-reported items which belong to the PHQ-8 system. Each depression level indicator was calculated based on the eight PHQ-8 items from BRFSS and the dependent variable which measures depression level is

denoted as PHQ-8. The resulting value for PHQ-8 is a non-negative integer ranging from 0 (no

depressive symptoms) to 4 (severe depressive symptoms). And depression is classified as major

depression (PHQ-8≥2) and other depression due to corresponding PHQ-8 scores (Kroenke et al.

2009). The pooled sample is restricted to individuals age >18 with a sample size of 11,560. The

frequencies for PHQ-8 are 8,802 (76.1%) for value 0, 1,702 (14.7%) for value 1, 578 (5.0%) for

value 2, 315 (2.7%) for value 3, and 163 (1.4%) for value 4. The sample size for men is 4,798

and corresponding frequencies for PHQ-8 are 3,834 (79.9%) for value 0, 597 (12.4%) for value 1,

207(4.3%) for value 2, 105 (2.2%) for value 3 and 55 (1.2%) for value 4. Among 6,762 female

individuals, the frequencies for PHQ-8 are 4,968 (73.5%) for value 0, 1105 (16.3%) for value 1,

371 (5.5%) for value 2, 201 (3.0%) for value 3 and 108 (1.6%) for value 4. Compared to males,

there are more females suffering from all levels of depressive symptoms, and the sample

statistics are consistent with previous research (e.g., Kessler et al. 1993; Goodwin and Gotlib

2004).

Endogenous variable

To better differentiate the effects of physical activity on depression level, physical

activity is specified as a binary variable such that physical activity = 1 for regular exercisers and

physical activity = 0 for seldom exercisers. The measurement for physical activity is drawn from

BRFSS questionnaire item "How many times per week or per month did you take part in this

activity during the past month". Regular exercisers are defined as those who did physical activity

or exercise at least 15 times during last month, while seldom exercisers are those who did less

than 15 times during last month. In the pooled sample of 11,560 individuals, about 39% are

regular exercisers. Considering gender difference and depression categories, table 1 and figure 1

provide the frequency distribution of physical activity and depression categories, which suggest regular exercisers are less likely to be depressed at each depression category.

Explanatory variables

Table 2 provides sample statistics and definitions of all explanatory variables. Socio-demographic variables include season, age, income, race, education, household composition (children), gender, home ownership, employment status, and marital status.

Recent mental health condition plays a notable role in affecting the current level of depressive symptoms. The measurement for the recent mental health condition is drawn from BRFSS questionnaire item "For how many days during the past 30 days was your mental health not good". The sample mean of not good mental health days in the last 30 days are 2.93 for males and 3.89 for females. About 3,513 (73.2%) male individuals and 4,298 (63.6%) female individuals reported excellent recent mental health conditions (0 mental health not good days). Men are less likely to have mental health problems than women.

The measurement for the recent physical health condition is drawn from BRFSS questionnaire item "For how many days during the past 30 days was your physical health not good?" The sample mean of not good physical health days in the last 30 days are 4.13 for men and 4.43 for women. About 3,161 (65.9%) men and 4,162 (61.6%) women reported excellent recent physical health conditions (0 physical health not good days).

In clinical research of depression, season is found to affect depression (e.g., Rosenthal et al. 1984; Harmatz et al. 2000), because mood is closely related to seasonal variation (Harmatz et al. 2000). In this study, each season is indicated by a binary variable. The pooled sample

frequencies are 0.25, 0.21, 0.27 and 0.27for the season of fall, winter, spring and summer respectively.

Age ranges from 18 to 99, with a mean of 54.3 for men and 54.5 for women. The mean of annual household income level is 5.84 (5 denotes annual household income between 25,000 to 35,000 and 6 denotes annual household income between 35,000 to 50,000) for the male sample and 5.44 for the female sample. About 70.9% of the male sample are white, compared to 72.0% of the female sample. The percentages of Hispanics are 21.9% for the male sample and 20.8% for the female sample

About 40% of the male sample have a bachelor's degree or above, compared to 38% of the female sample. As to household composition, one variable is used to measure the number of children under 18 years old, with a sample mean of 0.53 for the male sample and 0.58 for the female sample. About 78.6% of the male sample are home owners and the percentage for the female sample is 77.7%.

About 45.1% of the men are employed and 24.6% are retired, compared to 40.2% and 20.9% of the women. As to marriage status, 59.8% of the men are married and 14.4% are divorced. Among women, 50.5% are married and 17.7% are divorced.

Results and Discussion

An important issue in estimation is the identification of model parameters and endogenous effects. For instrumental variable estimation, at least one variable which is correlated with the endogenous variable, uncorrelated with the error term of the outcome equation, and does not affect the outcome equation is required in the switching equation for parameter identification. However, for ML estimation of current model, the nonlinear

identification criteria are met without exclusion restrictions owing to distributional assumption of the error term. Nonlinear functional form relying solely on distributional assumptions often fails to generate sufficient variation to identify model parameters which can be capricious. To avoid over-burdening the nonlinear functional forms for parameter identification, exclusion restrictions with different sets of explanatory variables in the switching (physical activity) and PQH-8 equations are imposed. Recent mental health condition is used solely in the PHQ-8 equation and recent physical health condition is only placed in the switching (physical activity) equation.

Tests for the switching probability model

Several empirical tests are performed to evaluate suitability of the switching probability model. Based on LR and Wald tests, the hypothesis of equal "slope" and error correlation coefficients between the two regimes is rejected for the male, female, and pooled samples (Table 3), suggesting that the switching probability model performs better than the treatment effect model in fitting the data. The next statistical test is for gender equality, viz., that all parameters are equal between men and women. This is carried out with a LR test, which is similar to Chow test in linear regression models. Specifically, define the log-likelihood values for the male, female, and pooled sample samples as $\log L_m, \log L_f,$ and $\log L_p$, with corresponding numbers of parameters $k_m, k_f,$ and k_p. Then, under the null hypothesis that parameters are equal between genders, the test statistics $LR = 2(\log L_m + \log L_f - \log L_p)$ is Chi-square distributed with $k_m + k_f - k_p$ degrees of freedom (df). For the switching probability model, the hypothesis of equal slope coefficients between male and female samples is rejected (LR = 117.82, df = 80, *p*-value = 0.0038), which suggests separate estimation of the model by gender-segmented samples.

ML estimates with gender-segmented samples

Table 4 presents ML estimates for the switching probability model with male and female samples. All threshold parameter estimates are positive and significant at the 1% level of significance or lower, which suggest that the ordered probability model (switching probability model) is successful in delineating the PHQ-8 categories for regular exercisers and seldom exercisers with gender-segmented samples. The error correlation estimates between the switching equation and both outcome equations are both significant at the 1% level or lower, which suggests endogeneity of regime switching. In addition these positive error correlations suggest that unobserved characteristics affect physical activity and PHQ-8 in the same direction for both males and females.

Of the 25 variables in the switching equation, 14 variables are significant at the 10% level for females. Statistical significance is more scant for the male sample—with only 6 variables significant. Recent physical health, the physical health variable is significant in the switching equation at the 1% level of significance in both male and female samples, rejecting the hypothesis of weak instrument and justifying use of the variable as an identification variable. Of the 25 variables in the regime (PHQ-8) equations, statistical significance is scant for males— with 3 variables (Summer, Hispanic, and Married) significant for seldom exercisers and 5 variables significant (Summer, Age, Age^2, Income, and Student) for regular exercisers. A lot more variables are significant in the regime equations for female, with over one half of the variables significant at the 10% level or lower. The estimates also differ greatly among male and female regular exercisers, in terms of signs, magnitudes, and statistical significance. To further exploit effects of physical activity and explanatory variables on the level of depressive symptoms, treatment effects and marginal effects of explanatory variables are discussed below.

Treatment effects of physical activity on depression

The primary purpose of estimating the switching probability model is to investigate the effect of physical activity on depressive symptoms. Average treatment effects (ATE) are calculated to describe quantitative effects of physical activity on depression between seldom and regular exercisers. The results, presented in Table 5, suggest that physical activity (regular exercise) decreases the probabilities of some levels of depressive symptoms among males and females. According to the results of ATEs, for a randomly selected male individual, a regular exerciser has 0.87% and 0.83% lower probabilities of moderate and moderately severe depressive symptoms than a seldom exerciser and for a randomly selected female individual, the regular exerciser has 2.34% and 1.00% lower probabilities of mild and moderate depressive symptoms than seldom exerciser. In terms of the levels of depressive symptoms, physical activity is most beneficial for mild and moderate depressed individuals and the effect of regular activity is most notable on mild depressed females.

Marginal effects of explanatory variables on depression

Average marginal effects of explanatory variables conditional on the probabilities of seldom and regular exercisers allow further exploration for the effects of explanatory variables on depression. Results are presented in Table 6 for males and Table 7 for females.

Marginal effects for males

Age is a key determinant of depression, and it has a negative effect on all depression categories for both seldom and regular exercisers among males. Conditional on seldom exerciser, a 10-year increase in age is associated with a 0.24% (0.11%) decrease in the probability of

moderate (moderately severe) depressive symptoms and 0.07% decrease in the probability of severe depressive symptoms, but conditional on regular exerciser, a 10-year increase in age is associated with 1.09%, 0.37%, 0.18% and 0.24% decreases in the probabilities of mild, moderate, moderately severe, and severe depressive symptoms.

As expected, income plays a role in affecting the level of depressive symptoms for both seldom and regular exercisers among males. The marginal effects of income on the probabilities of all depression categories are negative, which suggest that higher income decreases the probabilities of depressive symptoms; thus, poor males are more likely to have depressive symptoms than rich ones. Specifically, for seldom exerciser, a one-category increase in income level decreases the probabilities of mild, moderate, moderately severe, and severe depressive symptoms by 0.50%, 0.20%, 0.09%, and 0.06%, and decrease the those corresponding probabilities by 0.75%, 0.22%, 0.10%, and 0.14% conditional on regular exerciser.[5]

Season affects the depressive symptoms of seldom exercisers but not regular exercisers among males. The probabilities of being mild, moderate, moderately severe, and severe depressed are 1.70%, 0.68%, 0.31%, and 0.20% higher in the spring season than in the fall.

Supporting our hypothesis in the role of endogenous self-reported mental health condition in previous conceptual framework, recent mental health condition (number of self-reported bad mental health days in the past 30 days) has a positive impact on the level of depressive symptoms for both seldom and regular exercisers among males. A one-day increase in recent number of bad mental health days decreases the probabilities of mild, moderate, moderately severe, and severe depressive symptoms by 0.97% (0.93%), 0.38% (0.27%), 0.17% (0.12%), and 0.11% (0.17%) conditional on seldom (regular) exerciser.

---

[5] Income in this study in divided into categories from 1 to 8.

Race plays a role in affecting some categories of depressive symptoms among male regular exercisers, but it does not have significant effects on depressive symptoms of seldom exercisers. Comparing to male regular exercisers of other race, black males have 4.92% (0.88%) lower probability of being mild (severe) depressed.

Education only affects seldom exercisers among males, and compared with males who only have high school diplomas, those who have bachelor's degrees or above are less likely to be depressed. Seldom exercisers who have bachelor's degrees or above are 2.67%, 1.06%, 0.43%, and 0.29% less likely to have mild, moderate, moderately severe, and severe depressive symptoms.

Employment status affects depressive symptoms of male regular exercisers but not seldom exercisers. Compared with male homemakers, a student has 7.12%, 2.11%, 1.10%, and 1.28% lower probabilities of being mild, moderate, moderately severe, and severe depressed conditional on regular exerciser.

Marital status affects both seldom and regular exercisers among males. Specifically, married males who rarely exercise have 2.23%, 0.82%, 0.40%, and 0.24% lower probabilities of mild, moderate, moderately severe, and severe depressive symptoms than single males, and married males who exercise regularly have 2.34% and 0.46% higher probabilities of mild and severe depressive symptoms than single ones.

Marginal effects for females

Similar to results for males, age affects the depression category probabilities of females negatively. Conditional on seldom (regular) exerciser, a 10-year increase in age is associated with 0.89%, 0.27%, 0.11%, and 0.09% (1.52%, 0.53%, 0.33% and 0.30%) decreases in the

probabilities of mild, moderate, moderately severe, and severe depressive symptoms. Even though the results for both males and females are consistent with some of the previous studies (Wade and Cairney 1997; Schieman et al. 2002; Streiner et al. 2006), the effect of age on depressive symptoms is not prominent.

Income affects the depressive symptoms of females as well, and higher income decreases the probabilities of depressive symptoms of both seldom and regular exercisers. Specifically, a one-category increase in income level decreases the probabilities of mild, moderate, moderately severe, and severe depressive symptoms by 0.44%, 0.16%, 0.08%, and 0.07% among seldom exercisers, while the corresponding decreases are 0.73%, 0.25%, 0.15%, and 0.14% for regular exercisers. Our negative effects of income on depression for both males and females are similar to those reported by Zimmerman and Katon (2005).

Consistent with the effect for males, recent mental health condition has a positive impact on all depression category probabilities among females. A one-day increase in the recent number of bad mental health days decreases the probabilities of mild, moderate, moderately severe, and severe depressive symptoms by 1.22%, 0.44%, 0.21%, and 0.18% (0.92%, 0.31%, 0.19% and 0.17%) for seldom (regular) exercisers.

The effect of race on females is quite different from males, and race has more significant effects on regular exercisers. Among female regular exercisers, a white (Hispanic) female has 6.53%, 2.34%, 1.49% and 1.35% (9.36%, 3.53%, 2.09% and 1.93%) higher probabilities of mild, moderate, moderately severe, and severe depressive symptoms than females of other races. Our findings suggest that Hispanic females are more likely to be depressed than white females and this is consistent with the findings of Bromberger et al. (2004) who indicate that Hispanic women have higher odds of being depressed than white women.

Education affects both seldom and regular exercisers among females, and compared with females who only have high school diplomas, those who have bachelor's degrees or above are less likely to be depressed. Seldom (regular) exercisers who have bachelor's degrees or above are 1.85%, 0.73%, 0.27%, and 0.26% (1.99%, 0.63%, 0.39% and 0.34%) less likely to have mild, moderate, moderately severe and severe depressive symptoms. Relating these results to those of males, we find more educated people have lower risks of being depressed which coincide with previous research of Craig and Natta (1979).

Unlike the effects on males, employment status plays important roles in both seldom and regular exercisers among females. Conditional on seldom exerciser, unemployed females have 3.18%, 1.42%, 0.59%, and 0.51% higher probabilities of mild, moderate, moderately severe, and severe depressive symptoms than homemakers. Among female regular exercisers, a student has 4.83%, 1.54%, 0.96%, and 0.81% lower probabilities of mild, moderate, moderately severe, and severe depressive symptoms. In addition, unable females are more likely to be depressed than homemakers, and unable seldom (regular) exercisers have 7.12%, 3.31%, 1.46%, and 1.18% (4.49%, 1.65%, 0.99% and 0.85%) higher probabilities of being mild, moderate, moderately severe, and severe depressed.

Concluding remarks

This paper examines the effects of physical activity and socio-demographic factors on level of depressive symptoms, using data from a large national sample of the general population. PHQ-8 scores are used to measure the level of depressive symptoms, and an endogenous switching ordered probability model is developed to address the ordinal depression outcome and binary endogenous physical activity.

Our primary finding is that regular physical activity is negatively associated with depressive symptoms and doing physical activity regularly decreases the probabilities of moderate and moderately severe depressive symptoms for males, and decreases the probabilities of mild and moderate depressive symptoms for females. This finding also suggests mild and moderate depressed females will benefit more from regular physical activity.

This study is the first to evaluate the implication of physical activity and depression across major socio-demographic factors and seldom and regular exercisers. By comparing marginal effects of socio-demographic variables between seldom and regular exercisers between genders, we find some differences in the mechanism of depression among seldom and regular exercisers. For males, season and education play significant roles in affecting depression of seldom exercisers while being black and being student influence depression of regular exercisers. For females, race plays a prominent role in affecting depression of regular exercisers, and we find that being white or Hispanic increases the probabilities of all depression categories significantly.

The findings of this study can inform policy makers and doctors who are concerned about depression issues. We find that the probabilities of being depressed are higher among low income, less educated, unemployed and unable individuals and those who report bad mental health days recently, and policy makers should pay more attention to those individuals with poor living status. In clinical treatment for depression, doctors can recommend mild or moderate depressed patients to take part in physical activity regularly, which is an effective means to reduce mild and moderate depressive symptoms.

While this paper represents one of the first attempts to investigate the role of physical activity in ordinal depression, further studies might consider the use of panel data and

investigation of the depression issues among various sub-population, such as teenagers, minorities, and the disabled. Further, physical activity and other socio-demographic factors are likely to be important for general health besides depression, and interesting insights may emerge with a similar study for general health.

References

Addis, M.E. 2008. Gender and Depression in Men. *Clinical Psychology: Science and Practice* 15(3):153–168.

Babyak, M., J.A. Blumenthal, S. Herman, P. Khatri, M.Doraiswamy, K. Moore, W.E. Craighead, and T.T. Baldewicz. 2000. Exercise Treatment for Major Depression: Maintenance of Therapeutic Benefit at 10 Months. *Psychosomatic Medicine* 62:633–638.

Barnow, B.S., C.G. Cain, and A.S. Goldberger. 1980. Issues in the Analysis of Selectivity Bias. *Evaluation Studies Review Annual* 5: 43–59.

Bromberger, J.T., S. Harlow, N. Avis, H.M. Kravitz, and A. Cordal. 2004. Racial/Ethnic Differences in the Prevalence of Depressive Symptoms among Middle-Aged Women: The Study of Women's Health across the Nation (SWAN). *American Journal of Public Health* 94(8):1378–1385.

Camacho, T., R. E. Roberts, N. B. Lazarus, G. A. Kaplan, and R.D. Cohen. 1991. Physical Activity and Depression: Evidence from the Alameda County Study. *American Journal of Epidemiology* 134:220–231.

Centers for Disease Control and Prevention (CDC). 2010. Current Depression among Adults- United States 2006 and 2008. *Morbidity and Mortality Weekly Report* 59(38):1229–1235.

Chalder, M., N.J. Wiles, J. Campbell, S.P. Hollinghurst, A.M. Hase, A.H. Taylor, K.R. Fox, C. Costelloe, A. Searle, H. Baxter, R. Winder, C. Wright, K.M. Turner, M. Calnan, D.A. Lawlor, T.J. Peters, D.J. Sharp, A.A. Montgomery, and G. Lewis. 2012. Facilitated Physical Activity as A Treatment for Depressed Adults: Randomized Controlled Trial. *British Medical Journal* 344:2758–2771.

Chapman, D.P., G. S. Perry, and T.W. Strine. 2005. The Vital Link between Chronic Disease and Depressive Disorders. *Preventing Chronic Disease: Public Health Research, Practice and Policy* 2(1):1–10.

Craft, L.L. 2005. Exercise and Clinical Depression: Examining Two Psychological Mechanisms. *Psychology of Sport and Exercise* 6(2):151–171.

Craft, L.L., and F.M. Perna. 2004. The Benefits of Exercise for the Clinically Depressed. *The Primary Care Companion to the Journal of Clinical Psychiatry* 6:104–111.

Craig, T.J., and P.A.V. Natta. 1979. Influence of Demographic Characteristics on Two Measures of Depressive Symptoms. *Archives of General Psychiatry* 35:149–154.

De-Moor, M.H.M., A.L. Beem, J.H. Stubbe, D.I. Boomsma, and E.J.C.D. Geus. 2006. Regular Exercise, Anxiety, Depression and Personality: A Population–Based Study. *Preventive Medicine* 42:273–279.

De-Velde, S.V.D., P. Bracke, and K. Levecque. 2010. Gender Difference in Depression in 23 European Countries. Cross-National Variation in the Gender Gap in Depression. *Social Science & Medicine* 71:305–313.

Dooley, D., R. Catalano, and G. Wilson. 1994. Depression and Unemployment: Panel Findings from the Epidemiologic Catchment Area Study. *American Journal of Community Psychology* 22(6):745–765.

DSM–IV. 1994. *Diagnostic and Statistical Manual of Mental Disorders*. 4th ed. Washington DC: American Psychiatric Association.

Farmer, M.E., B. Z. Locke, E. K. Moscicki, A. L. Dannenberg, D. B. Larson, and L. S. Radloff. 1988. Physical Activity and Depressive Symptoms: The NHANES I Epidemiologic Follow-up Study. *American Journal of Epidemiology* 128:1340–1351.

Foley, L.S., H. Prapavessis, E.A. Osuch, J.A.D. Pace, B.A. Murphy, and N.J. Podolinsky. 2008. An Examination of Potential Mechanisms of Exercise as a Treatment for Depression: A Pilot Study. *Mental Health and Physical Activity* 7(1):69–73.

Goetzel, R.Z., K. Hawkins, R.J. Ozminkowski, and S. Wang. 2003. The Health and Productivity Cost Burden of the Top 10 Physical and Mental Health Conditions Affecting Six large U.S. Employers in 1999. *Journal of Occupational and Environmental Medicine* 45:5–14.

Goodwin, R.D. 2003. Association between Physical Activity and Mental Disorders among Adults in the United States. *Preventive Medicine* 36:698–703.

Goodwin, R.D., and I.H. Gotlib. 2004. Gender Difference in Depression: The Role of Personality Factors. *Psychiatry Research* 126:135–142.

Hamer, M., E. Stamatakis, and A. Steptoe. 2008. Dose-Response Relationship between Physical Activity and Mental Health: The Scottish Health Survey. *British Journal of Sports Medicine* 43:1111–1114.

Hansson, M., J. Chotai, and O. Bodlund. 2009. Patients' Beliefs about the Cause of Their Depression. *Journal of Affective Disorders* 124:54–59.

Harmatz, M.G., A.D. Well, C.E. Overtree, K.Y. Kawamura, M. Rosal, and I.S. Ockene. 2000. Seasonal Variation of Depression and Other Moods: A Longitudinal Approach. *Journal of Biological Rhythms* 15:344–350.

Hoffman, B.M., M.A. Babyak, W.E. Craighead, A. Sherwood, P.M. Doraiswamy, M.J. Coons, and J.A. Blumenthal. 2011. Exercise and Pharmacotherapy in Patients with Major Depression: One-Year Follow-Up of the Smile Study. *Psychosomatic Medicine* 73:127–133.

Jang, S.N., I. Kawachi, J. Chang, K. Boo, H.G. Shin, H. Lee, and S. Cho. 2009. Marital Status, Gender, and Depression: Analysis of the Baseline Survey of the Korean Longitudinal Study of Aging (KLoSA). *Social Science & Medicine* 69:1608–1615.

Jorm, A.F., A.E. Korten, P.A. Jacomb, H. Christensen, B. Rodgers, and P. Pollitt. 1997. Public Beliefs about Causes and Risk Factors for Depression and Schizophrenia. *Social Psychiatry and Psychiatric Epidemiology* 32:143–148.

Kessler, R.C., C. Foster, P.S. Webster, and J.S. House. 1992. The Relationship between Age and Depressive Symptoms in Two National Surveys. *Psychology and Aging* 7:119–126.

Kessler, R.C., K.A. McGonagle, M. Swartz, D.G. Blazer, and C.B. Nelson. 1993. Sex and Depression in the National Comorbidity Survey I: Lifetime Prevalence, Chronicity and Recurrence. *Journal of Affective Disorders* 29:85–96.

Kessler, R.C., and T.B. Ustun. 2008. *The WHO World Mental Health Surveys: Global Perspectives on the Epidemiology of Mental Disorders*. New York: Cambridge University Press.

Kroenke, K., T.W. Strine, R.L. Spitzer, J.B.W. Williams, J.T. Berry, and A.H. Mokdad. 2009. The PHQ-8 as a Measure of Current Depression in the General Population. *Journal of Affective Disorders* 114:163–173.

Li, M., and J.L. Tobias. 2005. Bayesian Analysis of Treatment Effects in an Ordered Potential Outcomes Models. Fourth Annual Advances in Econometrics Conference, Southern Methodist University.

Milette, K., M. Hudson, M. Baron, B.D. Thombs, and Canadian Scleroderma Research Group. 2010. Comparison of the PHQ-9 and CES-D Depression Scales in Systemic Sclerosis:

Internal Consistency Reliability, Convergent Validity and Clinical Correlates. *Rheumatology (Oxford)* 49:789–796.

Mirowsky, J., and C.E. Ross. 1992. Age and Depression. *Journal of Health and Social Behavior* 33(3):187–205.

Mota–Pereira, J., J. Silverio, S. Carvalho, J.C. Ribeiro, D. Fonte, and J. Ramos. 2011. Moderate Exercise Improves Depression Parameters in Treatment-resistant Patients with Major Depression Disorder. *Journal of Psychiatric Research* 45:1005–1011.

Papke, L.E, and J.M. Wooldrige. 2005. A Computational Trick for Delta-method Standard Errors. *Economics Letters* 86: 413–417.

Radloff, L.S. 1977. The CES–D Scale: A Self-Report Depression Scale for Research in the General Population. *Applied Psychological Measurement* 1(3):385–401.

Ranga, K., and R. Krishnan. 2002. Biological Risk Factors in Late Life Depression. *Biological Psychiatry* 52:185–192.

Riso, L.P., R.L. Miyatake and M.E. Thase. 2002. The Search for Determinants of Chronic Depression: a Review of Six Factors. *Journal of Affective Disorders* 70:103–115.

Robertson, R., A. Robertson, R. Jepson, and M. Maxwell. 2012. Walking for depression or depressive symptoms: A systematic review and meta-analysis. *Mental Health and Physical Activity* 5(1):66–75.

Rosenthal, N.E., D.A. Sack, J.C. Gillin, A.J. Lewy, F.K. Goodwin, Y. Davenport, P.S. Mueller, D.A. Newsome, and T.A. Wehr. 1984. Seasonal Affective Disorder, A Description of the Syndrome and Preliminary Findings with Light Therapy. *Archives of General Psychiatry* 41(1):72–80.

Roy, A.D. 1951. Some Thought on the Distribution of Earnings. *Oxford Economic Papers* 3:135–146.

Salmon, P. 2001. Effects of Physical Exercise on Anxiety, Depression, and Sensitivity to Stress: A Unifying Theory. *Clinical Psychology Review* 21(1):33–61.

Schieman, S., K.V. Gundy, and J. Taylor. 2002. The Relationship between Age and Depressive Symptoms: A Test of Competing Explanatory and Suppression Influences. *Journal of Aging and Health* 14(2):260–285.

Sieverdes, J.C., B.M. Ray, X. Sui, D.C. Lee, G.A. Hand, M. Baruth, and S.N. Blair. 2012. Association between Leisure Time Physical Activity and Depressive Symptoms in Men. *Medicine & Science in Sports & Exercise* 44(2):260–265.

Somervell, P.D., P.J. Leaf, M.M. Weissman, D.G. Blazer, and M.L. Bruce. 1989. The Prevalence of Major Depression in Black and White Adults in Five United States Communities. *American Journal of Epidemiology* 130(4):725–735.

Streiner, D.L., C. Psych, J. Cairney, and S. Veldhuizen. 2006. The Epidemiology of Psychological Problems in the Elderly. *Canadian Journal of Psychiatry* 52:185–191.

Wade, T.J., and J. Cairney. 1997. Age and Depression in a Nationally Representative Sample of Canadians: A Preliminary Look at the National Population Health Survey. *Canadian Journal of Public Health* 88(5):297–302.

Whooley, M.A., C.I. Kiefe, M.A. Chesney, J.H. Markovitz, K. Matthews, and S.B. Hulley. 2002. Depressive Symptoms, Unemployment, and Loss of Income. *Archives of Internal Medicine* 162:2614–2620.

World Health Organization (WHO). 2012. Media Center: Depression. Available at: http://www.who.int/mediacentre/ factsheets/fs369/en/index.html (Accessed 21 February 2013).

World Health Organization (WHO). 2001. Mental Health: A Call for Action by World Health Ministers. Geneva, Switzerland.

World Health Organization (WHO). 2008. The Global Burden of Disease: 2004 Update. Geneva, Switzerland: WHO Press.

Yen, S.T., D.J. Bruce, and L. Jahns. 2012. Supplemental Nutrition Assistance Program Participation and Health: Evidence from Low-income Individuals in Tennessee. *Contemporary Economic Policy* 30(1):1–12.

Yen, S.T., W.D. Shaw, and Y. Yuan. 2010. Cigarette Smoking and Self-reported Health in China. *China Economic Review* 21:532–543.

Zimmerman, F.J., and W. Katon. 2005. Socioeconomic Status, Depression Disparities, and Financial Strain: What Lies behind the Income-Depression Relationship? *Health Economics* 14:1197–1215.

Table 1

Frequency distribution of physical activity and depression categories

| Physical activity | Depressive symptoms (PHQ-8) | | | | | Total |
	No	Mild	Moderate	Mod. severe	Severe	
			Male			
Seldom	2306	395	148	75	27	2951
Regular	1528	202	59	30	28	1847
Total	3834	597	207	105	55	
			Female			
Seldom	2901	746	273	156	75	4151
Regular	2067	359	98	54	33	2611
Total	4968	1105	371	201	108	

Table 2

Definitions and sample statistics of variables in pooled, male and female samples[a]

Variable	Definitions	Male	Female
Endogenous variables			
PHQ-8	Indicator of depression level ranging from 0-4	0.32	0.43
		(0.76)	(0.86)
Physical activity	Did physical activities more than 10 times during past	0.38	0.39
	30 days (yes = 1, no = 0)		
Continuous explanatory variables			
Mental health	Days during past 30 days when physical health not good	2.93	3.84
		(7.28)	(8.04)
Age	Age in years	54.30	54.45
		(16.40)	(16.15)
Income	Annual household income level from 1 to 8	5.84	5.44
		(2.09)	(2.18)
Children18	Number of children in household age < 18	0.53	0.58
		(1.03)	(1.04)
Physical health	Days during past 30 days when physical health not good	4.13	4.43
		(8.87)	(8.76)
Binary explanatory variables (yes = 1, no = 0)			
Fall	Fall (reference)	0.25	0.26
Winter	Winter	0.21	0.21
Spring	Spring	0.27	0.26
Summer	Summer	0.27	0.27
White	Race is White	0.71	0.72
Black	Race is Black	0.01	0.01
Hispanic	Race is Hispanic	0.22	0.21
Other race	Other race (Reference)	0.06	0.06
Base	Do not have high school diploma	0.08	0.08
High school	Has a high school diploma or GED (reference)	0.28	0.26
Some college	Has some college but not a Bachelor's degree	0.24	0.28
Graduate	Has a Bachelor's degree or above	0.40	0.38
Employed	Employed	0.57	0.51
Unemployed	Unemployed	0.07	0.05
Retired	Retired	0.27	0.24
Student	Student	0.02	0.03
Unable	Unable to work	0.06	0.07
Homemaker	Homemaker (reference)	0.00	0.11
Home owner	Home owner	0.79	0.78
Married	Married	0.60	0.51

Divorced	Divorced	0.14	0.18
Widowed	Widowed	0.06	0.14
Separated	Separated	0.02	0.02
Single	Single (reference)	0.19	0.15
Sample size		4792	6762

[a] Standard deviations are in parentheses. Income is the annual household income reported as categories from 1 to 8: 1 = less than $10,000, 2= $10,000 to $15,000, 3= $15,000 to $20,000, 4= $20,000 to $25,000, 5= $25,000 to $35,000, 6= $35,000 to $50,000, 7= $50,000 to $75,000, and 8= $75,000 or more.

Table 3

Likelihood ratio and Wald tests for switching probability model against treatment effect model[a]

| Sample | df | Test statistics | |
		Likelihood ratio	Wald
Pooled sample	29	88.75 (< 0.001)	69.01 (< 0.001)
Male sample	28	59.96 (0.004)	43.30 (0.033)
Female sample	28	70.48 (< 0.001)	53.65 (0.003)

[a] *p*-values are in parentheses.

Table 4

Maximum-likelihood estimation of ordinal PHQ-8 equation with binary endogenous switching[a]

	Male			Female		
Variable	Switching: Physical Activity	PHQ-8: Seldom Exerciser	PHQ-8: Regular Exerciser	Switching: Physical Activity	PHQ-8: Seldom Exerciser	PHQ-8: Regular Exerciser
Constant	0.256 (0.412)	0.601 (0.486)	0.240 (0.551)	−0.323 (0.185)*	−0.094 (0.199)	−0.928 (0.306)***
Winter	−0.074 (0.055)	−0.020 (0.060)	−0.031 (0.105)	−0.236 (0.047)***	−0.147 (0.050)***	−0.132 (0.083)
Spring	0.018 (0.051)	0.091 (0.057)	0.034 (0.096)	−0.070 (0.043)	−0.045 (0.047)	−0.057 (0.073)
Summer	0.154 (0.051)***	0.157 (0.058)***	0.152 (0.091)*	0.105 (0.042)**	0.077 (0.047)	0.012 (0.069)
Age / 10	0.026 (0.073)	−0.026 (0.079)	−0.365 (0.125)***	0.115 (0.065)*	0.177 (0.070)**	−0.126 (0.109)
Age^2/1000	−0.007 (0.068)	0.009 (0.076)	0.273 (0.120)**	−0.117 (0.061)*	−0.205 (0.066)***	0.017 (0.103)
Income	0.005 (0.012)	−0.021 (0.013)	−0.055 (0.020)***	0.012 (0.010)	−0.021 (0.011)*	−0.045 (0.017)***
Children18	−0.025 (0.021)	−0.006 (0.023)	0.016 (0.037)	−0.038 (0.018)**	−0.025 (0.019)	−0.000 (0.029)
White	−0.196 (0.078)**	−0.135 (0.086)	−0.110 (0.125)	−0.021 (0.066)	−0.042 (0.069)	0.503 (0.128)***
Black	−0.060 (0.193)	−0.008 (0.207)	−0.500 (0.386)	−0.056 (0.165)	−0.059 (0.183)	−0.179 (0.360)
Hispanic	−0.195 (0.084)**	−0.203 (0.092)**	−0.178 (0.136)	−0.126 (0.071)*	−0.106 (0.074)	0.526 (0.135)***
Base	−0.070 (0.079)	−0.111 (0.082)	0.095 (0.129)	0.018 (0.067)	0.052 (0.066)	0.156 (0.107)
Some college	0.002 (0.052)	−0.046 (0.055)	0.067 (0.093)	0.166 (0.043)***	0.120 (0.045)***	0.073 (0.073)
Graduate	0.188 (0.049)***	−0.006 (0.056)	0.024 (0.092)	0.263 (0.043)***	0.074 (0.048)	−0.029 (0.076)
Employed	−0.592 (0.360)	−0.458 (0.434)	−0.711 (0.444)	−0.215 (0.055)***	−0.186 (0.060)***	−0.216 (0.090)**
Unemployed	−0.363 (0.366)	−0.190 (0.439)	−0.393 (0.453)	−0.184 (0.086)**	0.040 (0.089)	0.106 (0.131)
Retired	−0.349 (0.363)	−0.279 (0.438)	−0.583 (0.452)	−0.117 (0.064)*	−0.129 (0.072)*	−0.082 (0.108)
Student	−0.460 (0.384)	−0.387 (0.456)	−1.010 (0.500)**	−0.057 (0.113)	0.025 (0.120)	−0.394 (0.185)**
Unable	−0.252 (0.368)	0.039 (0.440)	−0.084 (0.456)	0.036 (0.084)	0.327 (0.083)***	0.295 (0.126)**
Home owner	0.038 (0.052)	−0.043 (0.055)	−0.016 (0.091)	−0.140 (0.043)***	−0.155 (0.045)***	−0.136 (0.067)**
Married	−0.130 (0.060)**	−0.185 (0.065)***	0.132 (0.111)	0.003 (0.052)	−0.016 (0.055)	−0.088 (0.086)
Divorced	0.048 (0.070)	0.095 (0.075)	0.194 (0.122)	0.013 (0.058)	0.043 (0.061)	0.089 (0.096)
Widowed	−0.112 (0.101)	−0.046 (0.109)	0.209 (0.182)	0.077 (0.067)	0.041 (0.073)	0.087 (0.112)
Separated	0.045 (0.145)	0.102 (0.153)	0.119 (0.239)	0.061 (0.111)	0.058 (0.113)	0.152 (0.167)
Physical health	−0.018 (0.002)***			−0.022 (0.002)***		
Mental health		0.046 (0.003)***	0.070 (0.006)***		0.054 (0.002)***	0.062 (0.004)***
μ2 , 2		0.529 (0.036)***	0.825 (0.068)***		0.665 (0.033)***	0.853 (0.052)***
μ3 , 3		0.981 (0.065)***	1.409 (0.110)***		1.177 (0.055)***	1.405 (0.082)***
μ4 , 4		1.613 (0.107)***	1.917 (0.150)***		1.791 (0.082)***	2.024 (0.122)***
0 , 1		0.901 (0.020)***	0.615 (0.106)***		0.876 (0.020)***	0.618 (0.079)***
Log likelihood	−5741.389			−8798.296		

[a] Asymptotic standard errors are in parentheses. Asterisks indicate levels of significance: *** = 1%, ** = 5%, *= 10%.

Table 5

Average treatment effects of physical activity on probabilities of PHQ-8[a]

Depressive symptoms (PHQ-8)	Average treatment effects (ATE)	
	Male	Female
No depressive symptoms (PHQ-8=0)	2.58 (0.99)***	4.10 (0.93)***
Mild depressive symptoms (PHQ-8=1)	–1.41 (0.94)	–2.34 (0.89)***
Moderate symptoms (PHQ-8=2)	–0.87 (0.52)*	–1.00 (0.50)**
Moderately severe symptoms (PHQ-8=3)	–0.83 (0.37)**	–0.56 (0.38)
Severe depressive symptoms (PHQ-8=4)	0.54 (0.30)*	– 0.20 (0.30)

[a] All effects of probability are multiplied by 100. Asymptotic standard errors are in parentheses. Asterisks indicate the level of significance: *** = 1%, ** = 5%, * =10%.

Table 6

Average marginal effects of explanatory variables on the probability of PHQ-8 categories by physical activity of male sample[a]

	Conditioned on physical activity = 0 (seldom exerciser)					Conditional on physical activity = 1 (regular exerciser)				
Variable	PHQ-8 = 0	PHQ-8 = 1	PHQ-8 = 2	PHQ-8 = 3	PHQ-8 = 4	PHQ-8 = 0	PHQ-8 = 1	PHQ-8 = 2	PHQ-8 = 3	PHQ-8 = 4
Continuous explanatory variables										
Age / 10	0.99 (0.59)*	–0.58 (0.36)	–0.24 (0.14)*	–0.11 (0.06)*	–0.07 (0.04)*	1.88 (0.74)**	–1.09 (0.46)**	–0.37 (0.15)**	–0.18 (0.07)**	–0.24 (0.09)***
Income	0.84 (0.37)**	–0.50 (0.22)**	–0.20 (0.09)**	–0.09 (0.04)**	–0.06 (0.03)**	1.21 (0.42)***	–0.75 (0.26)***	–0.22 (0.08)***	–0.10 (0.04)**	–0.14 (0.05)***
Children18	–0.35 (0.65)	0.21 (0.38)	0.09 (0.15)	0.03 (0.07)	0.02 (0.04)	–0.56 (0.76)	0.35 (0.47)	0.10 (0.14)	0.05 (0.06)	0.06 (0.09)
Physical health	–0.40 (0.05)***	0.24 (0.03)***	0.10 (0.01)***	0.04 (0.01)***	0.03 (0.00)***	–0.15 (0.04)***	0.09 (0.02)***	0.03 (0.01)***	0.01 (0.00)***	0.02 (0.00)***
Mental health	–1.63 (0.08)***	0.97 (0.06)***	0.38 (0.03)***	0.17 (0.02)***	0.11 (0.01)***	–1.49 (0.08)***	0.93 (0.07)***	0.27 (0.03)***	0.12 (0.02)***	0.17 (0.02)***
Binary explanatory variables										
Winter	–0.95 (1.75)	0.55 (1.03)	0.25 (0.42)	0.09 (0.19)	0.06 (0.12)	0.04 (2.17)	–0.03 (1.35)	–0.01 (0.39)	–0.00 (0.18)	–0.00 (0.25)
Spring	–2.89 (1.69)*	1.70 (0.99)*	0.68 (0.41)*	0.31 (0.19)*	0.20 (0.12)*	–0.59 (2.03)	0.37 (1.26)	0.11 (0.37)	0.05 (0.17)	0.07 (0.23)
Summer	–2.12 (1.73)	1.28 (1.02)	0.45 (0.41)	0.25 (0.19)	0.15 (0.12)	–2.02 (1.98)	1.24 (1.23)	0.38 (0.37)	0.17 (0.17)	0.23 (0.23)
White	0.35 (2.47)	–0.23 (1.48)	–0.03 (0.57)	–0.06 (0.26)	–0.03 (0.17)	0.75 (2.61)	–0.46 (1.62)	–0.15 (0.48)	–0.06 (0.22)	–0.08 (0.30)
Black	–1.05 (6.01)	0.61 (3.49)	0.27 (1.45)	0.11 (0.65)	0.07 (0.41)	7.89 (4.69)*	–4.92 (2.89)*	–1.39 (0.85)	–0.70 (0.47)	–0.88 (0.52)*
Hispanic	2.76 (2.44)	–1.66 (1.45)	–0.59 (0.56)	–0.32 (0.26)	–0.19 (0.17)	2.08 (2.64)	–1.29 (1.64)	–0.39 (0.47)	–0.17 (0.23)	–0.24 (0.31)
Base	2.26 (2.16)	–1.36 (1.30)	–0.50 (0.49)	–0.25 (0.23)	–0.15 (0.15)	–2.76 (2.96)	1.71 (1.83)	0.50 (0.56)	0.23 (0.25)	0.32 (0.34)
Some college	1.64 (1.51)	–0.98 (0.90)	–0.38 (0.35)	–0.17 (0.16)	–0.11 (0.10)	–1.43 (1.99)	0.88 (1.23)	0.26 (0.36)	0.12 (0.17)	0.16 (0.23)
Graduate	4.45 (1.52)***	–2.67 (0.93)***	–1.06 (0.35)***	–0.43 (0.17)***	–0.29 (0.10)***	1.01 (1.87)	–0.64 (1.17)	–0.17 (0.34)	–0.09 (0.16)	–0.12 (0.21)
Employed	2.90 (12.80)	–1.74 (7.80)	–0.52 (2.93)	–0.41 (1.33)	–0.23 (0.75)	11.06 (10.47)	–6.64 (5.97)	–2.24 (2.33)	–1.01 (1.10)	–1.17 (1.12)
Unemployed	–1.00 (12.75)	0.49 (7.22)	0.37 (3.21)	0.08 (1.47)	0.07 (0.86)	4.67 (7.41)	–2.88 (4.55)	–0.84 (1.30)	–0.40 (0.67)	–0.54 (0.90)
Retired	2.20 (11.95)	–1.34 (7.13)	–0.42 (2.76)	–0.28 (1.28)	–0.15 (0.78)	8.77 (8.01)	–5.45 (4.87)	–1.62 (1.57)	–0.76 (0.76)	–0.95 (0.85)
Student	3.67 (11.10)	–2.24 (6.60)	–0.73 (2.57)	–0.46 (1.23)	–0.25 (0.71)	11.60 (4.23)***	–7.12 (2.45)***	–2.11 (0.91)**	–1.10 (0.52)**	–1.28 (0.47)***
Unable	–7.40 (15.16)	3.95 (7.73)	2.08 (4.43)	0.87 (2.03)	0.50 (1.00)	–0.39 (9.62)	0.26 (6.01)	0.05 (1.73)	0.03 (0.80)	0.04 (1.09)
Home owner	2.45 (1.64)	–1.44 (0.96)	–0.59 (0.40)	–0.25 (0.18)	–0.16 (0.11)	0.66 (1.92)	–0.41 (1.19)	–0.12 (0.35)	–0.06 (0.16)	–0.08 (0.22)
Married	3.69 (1.89)*	–2.23 (1.14)*	–0.82 (0.44)*	–0.40 (0.20)**	–0.24 (0.12)**	–3.80 (2.22)*	2.34 (1.35)*	0.68 (0.42)	0.32 (0.20)	0.46 (0.28)*
Divorced	–2.35 (2.25)	1.39 (1.32)	0.54 (0.54)	0.26 (0.24)	0.16 (0.15)	–3.99 (2.88)	2.45 (1.76)	0.75 (0.56)	0.34 (0.25)	0.46 (0.33)
Widowed	–0.83 (3.16)	0.47 (1.84)	0.23 (0.77)	0.08 (0.35)	0.05 (0.21)	–6.06 (4.71)	3.72 (2.83)	1.13 (0.94)	0.50 (0.41)	0.71 (0.56)
Separated	–2.73 (4.69)	1.60 (2.69)	0.64 (1.16)	0.30 (0.52)	0.19 (0.32)	–2.26 (5.45)	1.40 (3.34)	0.42 (1.02)	0.19 (0.46)	0.26 (0.63)

[a] All effects on probabilities are multiplied by 100. Asymptotic standard errors are in parentheses. Asterisks indicate level of significance: *** = 1%, ** = 5%, * = 10%.

Table 7

Average marginal effects of explanatory variables on the probability of PHQ-8 categories by physical activity of female sample[a]

Variable	Conditioned on physical activity = 0 (seldom exerciser)					Conditional on physical activity = 1 (regular exerciser)				
	PHQ-8 = 0	PHQ-8 = 1	PHQ-8 = 2	PHQ-8 = 3	PHQ-8 = 4	PHQ-8 = 0	PHQ-8 = 1	PHQ-8 = 2	PHQ-8 = 3	PHQ-8 = 4
Continuous explanatory variables										
Age/10	1.37 (0.52)***	–0.89 (0.32)***	–0.27 (0.11)**	–0.11 (0.05)**	–0.09 (0.05)**	2.69 (0.67)***	–1.52 (0.40)***	–0.53 (0.14)***	–0.33 (0.09)***	–0.30 (0.08)***
Income	1.04 (0.33)***	–0.62 (0.20)***	–0.23 (0.07)***	–0.10 (0.03)***	–0.09 (0.03)***	1.26 (0.41)***	–0.73 (0.24)***	–0.25 (0.08)***	–0.15 (0.05)***	–0.14 (0.05)***
Children18	0.10 (0.58)	–0.06 (0.35)	–0.01 (0.13)	–0.01 (0.06)	–0.01 (0.05)	–0.37 (0.72)	0.22 (0.42)	0.07 (0.14)	0.04 (0.09)	0.04 (0.08)
Physical health	–0.49 (0.04)***	0.29 (0.03)***	0.11 (0.01)***	0.05 (0.01)***	0.04 (0.00)***	–0.23 (0.04)***	0.13 (0.02)***	0.04 (0.01)***	0.03 (0.01)***	0.02 (0.01)***
Mental health	–2.06 (0.06)***	1.22 (0.06)***	0.44 (0.03)***	0.21 (0.02)***	0.18 (0.01)***	–1.60 (0.07)***	0.92 (0.06)***	0.31 (0.03)***	0.19 (0.02)***	0.17 (0.02)***
Binary explanatory variables										
Winter	0.41 (1.53)	–0.27 (0.90)	–0.01 (0.34)	–0.07 (0.16)	–0.05 (0.14)	0.88 (2.02)	–0.50 (1.18)	–0.19 (0.39)	–0.11 (0.24)	–0.09 (0.21)
Spring	0.16 (1.47)	–0.10 (0.87)	–0.01 (0.32)	–0.02 (0.15)	–0.02 (0.13)	0.73 (1.80)	–0.42 (1.04)	–0.15 (0.35)	–0.09 (0.21)	–0.08 (0.19)
Summer	–0.57 (1.49)	0.35 (0.89)	0.09 (0.32)	0.07 (0.15)	0.06 (0.13)	0.76 (1.69)	–0.45 (0.98)	–0.14 (0.33)	–0.09 (0.20)	–0.08 (0.18)
White	1.12 (2.20)	–0.67 (1.31)	–0.24 (0.48)	–0.11 (0.22)	–0.10 (0.20)	–11.71 (2.51)***	6.53 (1.31)***	2.34 (0.57)***	1.49 (0.40)***	1.35 (0.36)***
Black	0.99 (5.56)	–0.59 (3.33)	–0.20 (1.18)	–0.11 (0.56)	–0.09 (0.49)	3.76 (7.94)	–2.21 (4.73)	–0.72 (1.49)	–0.44 (0.93)	–0.39 (0.79)
Hispanic	1.21 (2.27)	–0.73 (1.35)	–0.22 (0.49)	–0.14 (0.23)	–0.12 (0.20)	–16.91 (4.20)***	9.36 (2.13)***	3.53 (1.02)***	2.09 (0.64)***	1.93 (0.57)***
Base	–1.57 (2.15)	0.93 (1.26)	0.34 (0.48)	0.16 (0.22)	0.14 (0.19)	–4.01 (2.91)	2.29 (1.65)	0.80 (0.60)	0.48 (0.36)	0.44 (0.32)
Some college	–0.85 (1.43)	0.52 (0.85)	0.14 (0.31)	0.11 (0.14)	0.09 (0.13)	–0.18 (1.80)	0.09 (1.04)	0.05 (0.35)	0.02 (0.21)	0.02 (0.19)
Graduate	3.12 (1.46)**	–1.85 (0.89)**	–0.73 (0.31)**	–0.27 (0.14)*	–0.26 (0.13)**	3.36 (1.83)*	–1.99 (1.08)*	–0.63 (0.36)*	–0.39 (0.22)*	–0.34 (0.19)*
Employed	2.27 (1.90)	–1.36 (1.14)	–0.43 (0.41)	–0.26 (0.19)	–0.22 (0.17)	3.35 (2.26)	–1.92 (1.31)	–0.67 (0.45)	–0.40 (0.28)	–0.35 (0.24)
Unemployed	–5.70 (3.02)*	3.18 (1.64)*	1.42 (0.77)*	0.59 (0.34)*	0.51 (0.29)*	–4.99 (3.69)	2.87 (2.08)	0.98 (0.76)	0.59 (0.45)	0.55 (0.42)
Retired	2.27 (2.18)	–1.37 (1.31)	–0.45 (0.46)	–0.24 (0.22)	–0.21 (0.19)	0.87 (2.67)	–0.50 (1.56)	–0.18 (0.52)	–0.10 (0.31)	–0.09 (0.28)
Student	–2.24 (3.85)	1.30 (2.22)	0.52 (0.88)	0.22 (0.40)	0.20 (0.35)	8.14 (3.33)**	–4.83 (2.02)**	–1.54 (0.64)**	–0.96 (0.41)**	–0.81 (0.32)**
Unable	–13.06 (3.19)***	7.12 (1.55)***	3.31 (0.96)***	1.46 (0.43)***	1.18 (0.30)***	–7.98 (3.81)**	4.49 (2.08)**	1.65 (0.84)*	0.99 (0.51)*	0.85 (0.42)**
Home owner	2.81 (1.49)*	–1.68 (0.88)*	–0.57 (0.33)*	–0.30 (0.15)**	–0.26 (0.13)**	2.11 (1.74)	–1.21 (1.01)	–0.43 (0.35)	–0.25 (0.21)	–0.22 (0.18)
Married	0.68 (1.73)	–0.40 (1.03)	–0.15 (0.37)	–0.07 (0.17)	–0.06 (0.15)	2.30 (2.16)	–1.34 (1.26)	–0.45 (0.42)	–0.27 (0.26)	–0.24 (0.22)
Divorced	–1.37 (1.94)	0.81 (1.14)	0.30 (0.43)	0.14 (0.20)	0.12 (0.17)	–2.20 (2.49)	1.27 (1.43)	0.43 (0.50)	0.26 (0.30)	0.24 (0.27)
Widowed	0.17 (2.28)	–0.10 (1.36)	–0.06 (0.48)	–0.01 (0.23)	–0.01 (0.20)	–1.45 (2.87)	0.83 (1.64)	0.29 (0.57)	0.17 (0.34)	0.16 (0.32)
Separated	–0.84 (3.65)	0.50 (2.16)	0.16 (0.80)	0.09 (0.36)	0.08 (0.33)	–3.43 (4.53)	1.95 (2.55)	0.69 (0.92)	0.41 (0.55)	0.38 (0.51)

[a] All effects on probabilities are multiplied by 100. Asymptotic standard errors are in parentheses. Asterisks indicate level of significance: *** = 1%, ** = 5%, *=10%.

Figure 1

Frequency distribution of physical activity and depression categories