UNNC School of Economics Working Paper Series

Series H Health Economics

健康經濟學研究中心
Centre for Health Economics

Exploring the Intergenerational Correlation of Mental Health in China

Xuezheng Qin, Sharon Chen, Chee-Ruey Hsieh

School Economics Working Paper – Health Series 2019.2H

Exploring the Intergenerational Correlation of Mental Health in China

Xuezheng Qin

School of Economics, Peking University, Beijing, China, 100871

Telephone: +86-10-6275-7237

Fax: +86-10-6275-4237

Email: xqin@pku.edu.cn

Sharon Chen

Departments of Biological Engineering, School of Engineering, Massachusetts Institute of Technology

77 Massachusetts Avenue, Cambridge, Massachusetts, USA 02139

Telephone: +1-7048044893

E-mail: schen20@mit.edu

Chee-Ruey Hsieh*

School of Economics, the University of Nottingham Ningbo China

199 Taikang East Road, Ningbo, 315100, China

Telephone: +86-574-8818-0000 (ext. 8610)

E-mail: Hsieh.Chee-Ruey@nottingham.edu.cn

Acknowledgement: We are grateful to the National Natural Science Foundation of China (No. 71573003), Beijing Social Science Foundation Research Base Grant (No. 16JDLJB001) for their financial support. We thank Siyuan Liu and Yumeng Ao for their excellent research assistance. The authors are responsible for all remaining errors.

Abstract

Background

The rising global prevalence of mental illnesses such as depression, in combination with the stylized fact of low treatment rate for such illnesses, raises a serious public and academic concern on whether the untreated mental illness will further increase the disease prevalence over the next generation through the channel of intergenerational transmission.

Objective

This paper addresses the intergenerational correlation of mental health in China, by using a national representative survey dataset obtained from China Family Panel Studies.

Methods

We use a full 20-question version of the CES-D (Center for Epidemiologic Studies Depression) questionnaire to measure an individual's mental health status. Based on a parent-child matched sample, we use a stepwise regression approach to estimate the impacts of parents' CES-D scores on their children's CES-D scores.

Results

We find that both father's and mother's mental health have significant impact on their children's mental health. In addition, we find that parental mental health appears to have a greater impact on the mental health in the next generation than maternal mental health does, which is contrary to what most literature portrays.

Conclusion

Our results suggest that both maternal and paternal mental health have significant impacts on the mental health of their offspring, although paternal mental health appears to have a greater effect. Additionally, intergenerational correlation of mental health is present in all studied areas and subpopulations of China.

I. Introduction

Family is the fundamental social unit that links parents and children from generation to generation, which in turn shapes the intergenerational correlations in almost every dimension of human daily life, including education, income, health, wealth, occupational choices and health behaviors. In recent years, economic literature has accumulated significant evidence to quantify the intergenerational correlations of health and other outcomes in addition to investigating the underlying pathways of such intergenerational transmissions. A consensus of the existing studies indicates that there is a strong correlation between parents' and children's outcomes, with the correlation coefficient ranging from 0.2 to 0.6 (Black and Devereux, 2011; Johnston et al., 2013). The empirical results on the underlying pathways behind the intergenerational correlation, however, are mixed across different study contexts. The evidence provided by some studies highlights the importance of genetic factors, while other studies emphasize the relative importance of the "nurture" mechanism, including behavioral and environmental factors (Eriksson et al. 2014).

This paper revisits this classical topic with a new focus on mental health. The key issues to be addressed are threefold. First, we explore the extent to which mental health problems, such as depression, are correlated between parents and children, using a nationally representative survey dataset obtained from China. Second, we examine whether the intergeneration correlation is stronger through the maternal or paternal side. Third, we explore potential mechanisms in shaping the intergenerational persistence of mental health in China.

The prevalence of mental illnesses, including depression, has been increasing around the world in the past decades. Depression is recognized as a major contributor to mental illnesses, and a leading cause for disabilities and suicidal mortality. According to Fan, Pei, Hou, et al. (2013), over 100 million individuals experience mental disorders in a year in China, accounting for over 20% of the country's total disease burden. Zhou (2010) also suggests that mental health problems are severe in China, and that the prevalence of mental diseases is increasing.

Recent studies such as Qin, Wang and Hsieh (2018) show that mental health illnesses are highly prevalent and costly in China, and also distributed unevenly across different geographical areas with socioeconomic gradients. Four subpopulations - the elderly, rural, poor and low educated populations - were identified as being vulnerable to mental health problems. In addition,

it is known that women are more susceptible to depression and mental illnesses than men. The costs of medical expenditures due to depressive mental illnesses occupy 14.7% of China's total medical spending, with depression and depressive symptoms accounting for 6.9% and 7.8% respectively (Hsieh and Qin 2018). This number is only expected to rise as more individuals are diagnosed with mental disorders over time. With the increasing prevalence and advancing knowledge of mental depression, the research focus is being turned to analyzing reasons for the occurrence of depression and developing better treatment options. Thus, studying the intergenerational correlation of mental health conditions and its contributing factors is essential.

Since the turn of the century, China has been implementing policies to increase the diagnosis of mental disorders, increase access to treatment, and reduce stigma. Depression first started to be diagnosed in the early 1990s, during the post-Mao era. During the Mao era, it was believed that those who were depressed were traitors to the communist regime, and thus depressive symptoms were either repressed or ignored. Despite the notion that mental health disorders were starting to be recognized, individuals were still targeted and treated inferiorly for having these conditions. This belief leads individuals to avoid seeking diagnoses and treatment. Both then and now, those with mental health conditions must be treated in psychiatric hospitals as there are no psychiatric wards or professional psychiatrists at most general hospitals in China. Additionally, amongst medical professionals, psychiatry has been viewed as an inferior branch of medicine. As a result, few medical students desire to pursue psychiatry as a career path. Across the country, there is an overall shortage of mental healthcare resources, and many prospective psychiatrists do not receive adequate training to be able to properly treat their future patients.

Not only is the quality of psychiatric care lackluster in China, but the treatment availability is low and mal-distributed. Of the psychiatric centers available in China, a majority of them are located in urban areas. Until recent years, mental health treatment was not covered by most insurance policies in China, and individuals were responsible for their own mental health medical costs. Both biological and psychological studies show that parental mental health and behaviors have a large influence on the mental health and behavior of their children (Connell and Goodman, 2002). It is widely understood in the sociology and psychology fields that children learn from their parents, and attempt to emulate the behaviors and emotions demonstrated by them. Parents who are nearby but not necessarily emotionally invested or

engaged tend to raise children who are less engaged in activities and more distressed. The connection between parents' investment and children's competence suggests that parental emotional involvement matters, and highly affects children's emotional competence and regulation (Volling, 2002). Additionally, it has been shown that depression in parents affects the emotional development of children. Constant exposure to negative emotions at a young age causes children to be more prone to the development of negative emotions themselves and difficulty in regulating emotions. A study of depressed mothers, in particular, showed that mothers who are depressed have maladaptive attitudes, thoughts, and behaviors. These behaviors, as well as exposing children to a similarly stressful environment to the mother's shows that children are at risk of developing their own emotional and mental health issues (Sroufe, 2001).

Around the world, increasing numbers of studies have paid attention to the intergenerational correlation in mental health. Recently, Johnston et al. (2013) studied the intergenerational correlation of mental health in Britain, using data from the 1970 British Cohort Study. They identified that the intergenerational persistence of mental health is relevant economically, with maternal mental health associated with lasting effects on children's future. Additionally, poor parental mental health was identified to be related to high long-run economic costs. Persson and Rossin-Slater (2018) address the intergenerational transmission of mental health through the channel of fetal stress exposure. Based on Swedish administrative data, they find evidence to support the casual link between antenatal exposure to maternal stress and mental health during childhood and adulthood. Specifically, they find that *in utero* exposure to maternal stress, as measured by the death of a family member during pregnancy, has a significant impact on the offspring's mental health in later life. This is reflected by the increased consumption of prescription drugs treating mental health conditions, such as attention deficit hyperactivity disorder during childhood and anti-anxiety and depression medications in adulthood. In addition to the above studies that address the intergenerational transmission of mental health from the maternal line, some other studies also explore this issue from the paternal line. For example, based on data obtained from both Ireland and UK, Lewis et al. (2017) find evidence of a positive association between paternal and adolescent depressive symptoms, and the magnitude of this association is similar to the mother-child association of depressive symptoms.

In summary, the existing studies find consistent evidence on the strong intergenerational correlation of mental health. However, all these studies are based on a small group of European countries, which in turn indicates the importance of further intergenerational analyses by using data in varying countries around the world. As a result, our paper contributes to the growing body of literature in this line through the following three aspects. First, our study uses data obtained from China, which offers the potential to shed new insight and provide policy implications for developing or low- and middle-income countries. Second, our data on the parental information include both mothers and fathers, which in turn provide the opportunity to test the hypothesis on whether maternal or paternal impact plays a more important role in accounting for the intergenerational correlation of mental health. Third, our survey data include rich information on the personal and socioeconomic background that enables us to control for the influence of behavioral and environmental factors.

II. Methods

2.1 Data Source and Descriptive Analysis

CFPS (China Family Panel Studies) is a nationally representative longitudinal survey designed and implemented by the Institute of Social Science Surveys (ISSS) of Peking University. It was conducted in 25 Chinese provinces (these provinces jointly cover 95% of the Chinese population) in five years (2008, 2009, 2010, 2011, 2012). In each wave, the CFPS survey samples about 15,000 households nationwide using the multi-stage probability proportional to size (PPS) sampling method, and it interviews all members of the sampled households. The questionnaire gathers individual and household level information on the demographic, socioeconomic, and health-related characteristics of the respondents. While the community-level survey collects information on the infrastructure, demographic profiles, social services and economic conditions of the rural villages or urban communities that the respondents live in.

In the 2012 CFPS survey, a full 20- question version of the CES-D (Center for Epidemiologic Studies Depression) questionnaire (Radloff, 1977) was administered to assess the respondents' mental health status. CES-D questionnaire is one of the most widely used self-evaluation tools on the respondent's mental health condition. Out of the 20 questions in the CES-

D, 16 measure negative feelings and 4 measure positive feelings. Respondents are asked to rate how often they experienced the specified emotions in the past week, with the options varying from 0 to 3 for each question (0 = rarely, 1 = little, 2 = occasionally, 3 = often). The CES-D score can thus be calculated based on the responses as follows:

$$CES-D = \sum_i Score_{i,somatic} + \sum_j Score_{j,interpersonal} + \sum_k Score_{k,depressed} + \sum_l (4 - Score_{l,positive}) \quad (1)$$

where $Score_{i,somatic}$, $Score_{j,interpersonal}$, $Score_{k,depressed}$ and $Score_{l,positive}$ represent the score for the i-th question on the somatic-retarded activity, the j-th question on interpersonal relations, the k-th question on the depressed affect and the l-th question on the positive affect, respectively. Thus, the overall CES-D score ranges from 0 to 60, with a higher score indicating more frequent occurrence of depressive symptoms and higher likelihood of depression. We use the CES-D score calculated from equation (1) to measure an individual's mental health status.

For the purpose of our study, we include the families for which both the respondent and his/her parents' CES-D information is available, and we drop the respondents younger than 16 and older than 99 as well as those with missing information on key variables such as gender, age, education and income. The final study sample consists of 5,074 observations, for whom the average CES-D score is 11.23, their fathers' average CES-D is 11.70, and their mothers' average score is 13.93.

In our econometric analysis, the control variables are added into the regressions in a step-wise fashion to reflect the various dimensions of impact on an individual's mental health in addition to the parental CES-D scores[1]. The basic control variables include the respondent's demographic and socioeconomic characteristics, such as age in years (*age*), residential status (*urban*), gender (*female*), marital condition (*married*), education attainment ("primary school or below", "middle school", "high school" and "college or above"), employment types (*unemployed*) and annual personal income (*income*).

[1] Following the convention in the prior literature such as Cutler and Lleras-Muney (2010), the control variables are uniformly adjusted to reflect their detrimental impacts on the individual's mental health. For example, positive factors such as *employed* and *married* are multiplied by -1 before adding into the regressions. In addition, income variables (such as *income*) and expenditure variables (such as *medical_cost*) also are transformed with logarithm, e.g. *Adj_Income* = ln(*Income*)*(-1).

In addition to the basic demographic and socioeconomic variables, we control for the respondents' health-related behaviors, such as whether the individual smoked last month (*smoking*), whether the individual engaged in drinking alcohol last month (*drinking*), whether the individual consumed unhealthy food (such as high calorie or pickled foods) last week (*unhealthy_food*), and how often the individual partakes in physical exercise (*exercise*, ranging from 1 being "very often" to 5 being "never").

To control for the impact of healthcare accessibility on the respondent's mental health status, we use the following variables: *time_hospital* is the shortest travel time from the individual's residence to the nearest hospital, *satis_hospital* describes respondents' level of satisfaction with their local hospital system, *no_healthins* is a dummy variable indicating a person's health insurance status (equals 1 if the respondent is not covered by health insurance), *medical_cost* is the calculated out-of-pocket medical cost last year for the respondent's family, and *no_fitspending* is a dummy variable that equals 1 if the household spent no money on fitness-keeping or health preservation. Furthermore, an individual's physical health condition can also influence his/her mental health status. To control for this, we define *chronic* as a dummy variable that equals 1 if the respondent has chronic diseases (such as hypertension and diabetes), which is a proxy measure of the respondent's physical health status.

With respect to the environmental factors, we control for the following dummy variables: *no_flushtoilet* indicates the household hygiene status and means the respondent does not have flush toilet in the residential place; *no_electric* means the respondent does not have stable electricity at home; *no_cleanwater* means the household does not have treated water for cooking usage; *poor_garbage* indicates poor sanitation of household garbage disposal (e.g. garbage is directly dumped on the riverbank in the open air around the household for convenience); *crowded_housing* means the respondent lives in an over-crowded housing environment (e.g. three or more generations of family members live in one bedroom, or living room is used as a bedroom).

Table 1 provides the sample summary statistics. As indicated, 35% of the respondents in our sample are female, and 42% are married. The high representation of male population indicates that parents prefer to live with their married sons in China. The average age of our sample is 25.7, a relatively young age caused by the sample selection criteria that requires both

parents' mental health information to be available. Around 29% of the respondents live in the urban sector, which is lower than the general urbanization rate in China. This may indicate that rural residents are more likely to live with their parents compared to urban residents. In terms of the respondents' socioeconomic status (SES), 29% of the sample are currently unemployed (including retired individuals, student populations, and those who are not working for payment). The average annual personal income is 14,066 CNY. More than 37% of the sample received education beyond the middle school (9-year compulsory education) level, and those with primary school or less education (illiterate or semi-illiterate) account for 23.6% of the whole sample.

For health and healthcare related variables, Table 1 shows that 31% of our sample are currently smoking, 12% are currently engaged in alcoholic drinking, and 64% of the sample consumed unhealthy food (such as high calorie or pickled foods) within the previous week. The average frequency level for partaking in physical exercise is 3.4 on a scale of 1 to 5, indicating relative physical inactivity among the sample respondents. For the healthcare accessibility measures, 18% of the respondents have no health insurance, and the shortest time from home to the nearest hospital averages 10.78 minutes. As for the health-related variables, around 4% of the respondents have chronic diseases and the average out-of-pocket medical spending in the previous year is 3,750 CNY. More than 90% of the sample spent no money on fitness measures or health preservation, indicating a relatively low tendency towards preventive health-related consumption.

Concerning living conditions, Table 1 indicates that 16% lived in an overcrowded place and 30% had no access to clean water for cooking. 49% reported not having stable access to electricity, and the average monthly personal electricity consumption of a household is 26.3 kWh. 25% of those surveyed responded saying they had poor sanitation in garbage disposal and 55% did not have flushable toilets available. Overall, these statistics show that for those surveyed, living conditions are relatively poor, which is affected largely by the high representation of rural populations.

For the parents in the surveyed families, the average CES-D score for mothers is higher at 13.9 when compared to the CES-D score for fathers, at 11.7. The average age for mothers and fathers, respectively, is 51.1 and 52.3, respectively. In general, fathers have higher education

levels than mothers, with more men completing all levels of secondary and higher education than women.

2.2 Empirical Specifications

Following literature convention (e.g. Cutler and Lleras-Muney, 2010), we use a stepwise regression approach to estimate the impacts of parents' CES-D scores on their children's CES-D scores. Our baseline regression model takes on the following form, and the estimation is based on the conventional Ordinary Least Squares (OLS):

$$score = \beta_0 + \beta_1 score\_f + \beta_2 score\_m + X_i\alpha + u_i \quad (1)$$

where *score*, *score_f* and *score_m* denote the CES-D scores of the child respondent, the father and the mother, respectively. *X* is a set of baseline control variables including the child's basic demographic characteristics (gender, age, quadratic term of age, urban/rural status, and province fixed effects). The coefficients β_1 and β_2 in equation (1) thus represent the marginal impacts of parental mental health on children's mental health status, one of the main foci of this study.

To further explore the potential channels that explain the intergenerational correlation of mental health, we add other mediating/control variables in a stepwise fashion. These mediating variables can be categorized into five dimensions, namely parents' background (parents' education levels and ages), children's socioeconomic status (income, employment status, education levels), children's health behaviors (*exercise, smoking, drinking, unhealthy_ food)*, children's healthcare condition (*chronic, satis_hosipital, time_hospital, medical_cost, no_fitspending, no_healthins*), and the quality of living environment *(crowded_housing, no_flushtoilet, poor_garbage, no_cleanwater, no_electric, month_electric*). To facilitate the comparison among different channels, we value each variable monotonically so that an increase in these variables have a hypothesized positive effect on the child's CES-D score[2]. For example, we define the frequency of physical exercise on a scale of 1 to 5 with 1 = very often and 5 = never.

[2] Some variables in the model are monotonically transformed from their raw values. For example, adjusted income is defined as ln(Income)*(-1) and adjusted electricity consumption is defined as ln(Elec_Per+1)*(-1).

Following an approach similar to Cutler and Lleras-Muney (2010), we add the above mediating variables into equation (1) in a stepwise fashion to form the following extended regression model:

$$score = \beta_0^j + \beta_1^j \cdot score\_f + \beta_2^j \cdot score\_m + X_i^j \cdot \alpha + u_i \qquad (2)$$

where j denotes the number of mediating variable blocks being added into the model. For example, β_1^1 refers to the impact of father's CES-D when only the parents' background variables are controlled for in the regression; and β_1^2 represents such impacts when children's SES is also added into the regression. Based on the OLS estimates of equation (2) and equation (1), we can calculate the gradient of each channel, which is defined as:

$$Gradient\_f = 1 - \beta_1^j / \beta_1 \text{ and } Gradient\_m = 1 - \beta_2^j / \beta_2 \qquad (3)$$

Thus, the gradient of the intergenerational correlation channel X^j is reflected in the percentage decline in the impacts of parents' mental health compared with the baseline estimates, when such channel/mediator is controlled for (Cutler and Lleras-Muney, 2010).

III. **Estimation Results**

Table 2 presents the regression results, including the estimated marginal effects and the robust clustered standard errors for each independent variable. The baseline model containing only the parental CES-D scores as regressors is shown in column (1). Additional factors of consideration were added into analysis several at a time as to be able to analyze their individual effects on children's mental health and to gauge their mediating effects on the intergenerational transmission of depression. Children's basic information is added in column (2), parents' information in column (3), children's socioeconomic status (SES) in column (4), children's health behaviors in column (5), children's healthcare condition in column (6), and children's living environment in column (7). In addition to the variables presented in Table 2, additional baseline controls on children's demographic characteristics and parental background are held constant throughout the analysis in column (2) – (7).

We first examine the results of the baseline model. The estimations show that there is a strong and significant intergenerational correlation on parents and children's mental health

outcomes. In addition, we find that fathers' CES-D score has a greater impact on the mental health status of the child compared to mothers' CES-D score. An increase in father's CES-D score by 10 tends to increase the child's CES-D score by 1.98, compared to 1.82 for a 10-point increase in mother's CES-D score. As additional variables were added into the regression in column (2) – (7), the estimated effect of an increase in parental CES-D score on child's CES-D score decreased, suggesting the mediating effects of these factors. When accounting for all additional variables, it stands that an increase in the father's CES-D score has a slightly greater impact than an increase in the mother's CES-D score, with the impact on child's CES-D score being 1.66 and 1.56 respectively for a 10-point change in parental CES-D. Greater impacts of paternal CES-D scores on children's CES-D scores are also reflected in the gradient values, which are calculated based on the percentage difference in the key coefficient estimates between the extended and baseline models. When additional variables are added in, the gradient increases, indicating greater decline in the magnitude of parental CES-D impacts on children's mental health, i.e. the behavioral, social and environmental channels have significant compounding effects on the parental impacts, thus they serve as important channels of the intergenerational transmission of mental depression. In addition, the gradient percentage for father's CES-D score is consistently higher than that for mother's CES-D score, indicating that the channels under consideration tend to play a larger role for fathers than mothers.

With regard to other control variables, the results also yield several interesting findings. First, we find that bad health behaviors, as measured by lack of exercise, smoking and drinking, are associated with a higher CES-D score, indicating that health behaviors are important predictors for the occurrence of mental health problems. This finding is also robust to different specifications. Second, the presence of chronic diseases has a significantly positive impact on the CES-D score, indicating a tendency of co-morbidity between physical and mental health. Third, among a set of environment variables, we find that the lack of electricity has a significant deteriorating effect on the status of mental health, indicating that the lack of public investment on infrastructure may also contribute to the poor mental health.

Table 3 reports the subsample regression results. Based on the demographic characteristics, the respondents in our original sample is divided into male vs. female, rural vs. urban, and different age groups of children. Looking at the subsamples, it is notable that the

observed counts for males is significantly higher than that seen for females, and also significantly higher for rural areas than urban areas. This disparity in observed counts may be due to the oversampling of rural respondents in the CFPS surveys, and the male-female gap may be exacerbated by the urban-rural disparity. An increased number of rural households can also lead to increased male counts due to the preference for male children in China's rural communities (Murphy et al., 2011).

The regression results in Table 3 indicate that the intergeneration correlation (particularly the paternal influences) tend to be higher for rural areas. For example, father's CES-D score is shown to have consistently larger impacts in rural populations than urban populations, especially in the baseline model when we are taking into account fewer channeling variables. After controlling for all other variables, an increase by 10 points in fathers' CES-D score is seen to cause a 1.87-point increase in children's CES-D score for the rural areas, compared to a 1.21 point increase for the urban areas. However, the impact of mothers' CES-D score appears to be more consistent across rural and urban populations. The impacts of both mothers' CES-D score and fathers' CES-D score are observed to be significant for all regression models.

Comparing subgroups by gender, the regression coefficients are noticed to follow different trends. For male children, the effect of fathers' CES-D score and mothers' CES-D score on the children appear to be relatively similar. For the baseline model, an increase of 10 points in either fathers' or mothers' CES-D score appears to increase the children's CES-D score by 1.59 and 1.63 points, respectively. Similarly, the apparent increase in children's CES-D score with all channeling variables incorporated into the regression is 1.52 versus 1.58, respectively. However, for female children, fathers' CES-D score appears to have a much greater impact on children's CES-D than mothers' impact. For the baseline model, an increase of 10 points in either fathers' or mothers' CES-D score appears to increase the children's CES-D by 2.02 and 1.60 points, respectively. When all channeling variables are incorporated into the regression, the increase in children's CES-D score is 1.90 versus 1.56, respectively. It suggests a noticeable gap between the male and female children in terms of the impacts of their parents' CES-D scores. A comparison between fathers' and mothers' impact shows that mothers' CES-D scores have a relatively similar effect on children's mental health status, regardless of gender. On the other

hand, an increase in fathers' CES-D score has a larger impact on female children than male children.

Regressions on different age groups indicate that the effect of fathers' CES-D score on children's CES-D is more pronounced for two subgroups: 1) children aged 16 to 18 and 2) middle aged and older adults (aged 40 and above). With all the mediating variables incorporated, these two subgroups had regression coefficients of .213 and .217, respectively. These are much higher than the coefficient estimates for young adults (age 19 to 25) and prime age adults (age 26 to 39), which were .156 and .159, respectively. The impact of mothers' CES-D score appears to be more stable amongst various age groups, with the only notable exception being that such impact is quantitatively smaller for middle aged and older adults.

It is also worth noting that all subsample analyses produced significant regression coefficients for the relationship between fathers' and children's CES-D scores as well as between mothers' and children's CES-D scores. These strong relationships prove significant, regardless of the variables taken into account.

Furthermore, our results indicate that the gradient values are not homogenous across subsamples. Specifically, we find that the effect of father's mental health on children's mental health decreases by about 10 percent for the urban subsample and the subsample of middle aged and older adults. This indicates that children's living environment plays an important role in account for their mental health if the respondent live in urban areas as well as they are getting old.

IV. Conclusions

In this paper, we explore the intergenerational correlation of mental health as observed through mother-child and father-child relationships. This information is also presented and broken down by subgroups. Using the nationally representative CFPS 2012 dataset with CES-D metrics, our findings indicate that there is a strong intergenerational correlation of mental health. A 10-point increase in the CES-D score for mothers appears to increase children's CES-D score by 1.56 points, while a 10-point increase in the CES-D score of fathers appears to increase children's CES-D score by 1.65 points. A stronger correlation of intergenerational mental health is thus observed between father-child relationships than mother-child relationships, contrary to what has been demonstrated in most studies based on the Western countries. This correlation can

be interpreted as worrisome, as it is shown that there is significant correlation for mental health between generations, and that maternal and paternal depression are likely to transfer to the offspring. This correlation allows us to presume that the prevalence rate of mental illness will increase over time, as it develops both naturally with globalization and urbanization as well as through the transmission from parents to children. Furthermore, our stepwise regression results suggest that this intergenerational correlation of mental health occurs through several channels such as the behavioral, environmental and socioeconomic channels.

Our results have three important policy implications. Firstly, the significant and identifiable correlation of mental health between generations indicates that the current categorization of mental illness as a non-communicable disease is potentially misleading. Although individuals with mental illness is not communicable to other people at one point in time, it could be communicable over time as the younger generation can be *infected* by untreated mental illnesses through family connection. Thus there is negative externality to the society if the majority of people with mental illnesses remain undiagnosed and untreated, which in turn suggests that mental illness needs to be identified early on and treated in order to avoid the external costs to the society. Given that individuals often lack awareness to internalize such costs in their decision-making, it is important to call for the government interventions to treat the prevention of mental illnesses as public goods in a similar way as the vaccination of communicable diseases.

Secondly, there exists substantial treatment gap in mental healthcare in China due to policy- and history- induced barriers. Medical policies in China today do not have extensive or comprehensive mental health benefits, making it difficult for individuals to seek treatment that may have potentially high costs and large comorbidities. Additionally, social stigma of mental illnesses may lead to a decrease in the number of individuals willing to seek treatment and care for the fear of being ostracized by the society and community. The above arguments point to the necessity of further government and policy interventions to remove the access barriers for treating and preventing depression and reduce the stigma against mental disorders. Mental healthcare management systems must be established and maintained throughout the country, in addition to being supported by other local facilities, the government, and other health care institutions. Additionally, when comparing China to other nations, it is notable that China's

progress in mental health acceptance is slow, and that mental health benefits are lacking. Our analysis suggests that increasing investment in the mental health sector in China will yield a substantial high return to the society if we take the intergenerational correlation of mental health into account.

Thirdly, the evidence on the intergenerational correlation of mental health points to a "family focus" in policy making within the mental health sector. As mental health of both parents is important predictor for the mental health of their children, the involvement of fathers as well as mothers in early intervention is crucial for reducing the prevalence of mental illness such as depression in the younger generation.

There are several limitations to our study. First, the data collected from the 2012 CFPS represents a limited time frame, and the self-reported measures on mental health may provide inaccuracies. Additionally, sampling methodologies (an oversampling of rural populations) may skew the data reported. Thus, the strength of the relationship between mental health and certain variables may be overestimated (Lindell & Whitney, 2001), and this may also exacerbate the shown intergenerational linkage of mental health. Secondly, the CES-D questionnaire covers a relatively short period of time (the week prior to the survey), which can allow for unexpected discrepancies (caused by, for example, recent shocks such as death in the family, and frequent mood fluctuations) to appear. Studies based on other metrics, such as the PHQ-2 and PHQ-9 questionnaires, can be performed and compared to the current findings based on the CES-D questionnaire. Thirdly, other outside factors may influence one's mental health, such as air pollution, and these unobserved factors may not be captured by the current study.

References

Black, SE, Devereux, PJ. 2011. "Recent developments in intergenerational mobility." In *Handbook of Labor Economics*. Vol. 4B, edited by O. Ashenfleter and D. Card, 2011;1487-1541. Elsevier.

Connell, AM, Goodman, SH. The association between psychopathology in fathers versus mothers and children's internalizing and externalizing behavior problems: A meta-analysis. *Psychological Bulletin*, 2002;128(5): 746-773.

Cutler, DM, Lleras-Muney, A. Understanding differences in health behaviors by education. *Journal of Health Economics*, 2010;29(1): 1-28.

Eriksson, T, Pan, J., Qin, X. The intergenerational inequality of health in China. *China Economic Review*, 2014;(31): 392-409.

Fan, P, Pei, J, Hou, Z, et al. The discussion of current mental health status and mental health management strategy in China. Journal of Practical Medical Techniques, 2013;20(8): 911-912 (in Chinese).

Hsieh, CR, Qin, X. 2018. Depression hurts, depression costs: the medical spending attributable to depression and depressive symptoms in China. *Health Economics*, 2018;27(3): 525-544.

Johnston, DW, Schurer, S, Shields, MA. 2013. Exploring the intergenerational persistence of mental health: evidence from three generations. *Journal of Health Economics*, 2013;32: 1077-1089.

Lewis, G, Neary, M., Polek, E, Flouri, E, Lewis, G. The association between paternal and adolescent depressive symptoms: evidence from two population-based cohorts. *Lancet Psychiatry*, 2017; 4: 920-926.

Lindell, MK, Whitney, DJ. 2001. Accounting for common method variance in cross-sectional research designs." *Journal of Applied Psychology*, 2001;86(1): 114-121.

Murphy, R, Tao, R, Lu, X. Son preference in rural china: Patrilineal families and socioeconomic change. *Population and Development Review*, 2011;37(4): 665-690.

Persson, P, Rossin-Slater, M. "Family ruptures, stress, and the mental health of the next generation." *American Economic Review*, 2018;108(4-5): 1214-52.

Qin, X, Wang, T., Zhuang, C. Intergenerational transfer of human capital and its impact on income mobility: Evidence from China. *China Economic Review*, 2016; (38): 306-321.

Qin, X, Wang, S, Hsieh, CR. The prevalence of depression and depressive symptoms among adults in China: estimation based on a national household survey." *China Economic Review*, 2018;51, 271-282.

Radloff, LS. The CES-D scale a self-report depression scale for research in the general population. *Applied Psychological Measurement*, 1977;1(3): 385-401.

Volling, B, McElwain, N, Notaro, P, Herrera, C. Parents' emotional availability and infant emotional competence: Predictors of parent-infant attachment and emerging self-regulation." *Journal of family psychology*, 2002; 16: 447-465.

Sroufe, LA. From infant attachment to promotion of adolescent autonomy: Prospective, longitudinal data on the role of parents in development." In *Parenting and the Child's World: Influences on Academic, Intellectual, and Social-emotional Development*. Edited by J. G. Borkowski, S. L. Ramey and M. Bristol-Power. Psychology Press, 2001.

Zhou, Y. Mental health social work in America and its significance to China. *Journal of Sichuan University (Social Science Edition)*, 2010; 3: 127-132 (in Chinese).

Table 1 Sample Descriptive Statistics of Key Variables

Variable	Definition	
Children's Information		
score	CES-D score of children	11.23 (6.68)
age	Age in years	25.69 (7.73)
urban	Urban residents (1=yes)	0.290 (0.45)
female	Gender (1=female)	0.354 (0.48)
married	Married (1=yes)	0.420 (0.49)
employed	Employed (1=yes)	0.710 (0.45)
primary	Primary school or below (1=yes)	0.236 (0.42)
middle	Middle school (1=yes)	0.387 (0.49)
high	High school (1=yes)	0.212 (0.41)
college	College or above (1=yes)	0.164 (0.37)
income	Annual personal income of household (*yuan*)	14066 (18946)
exercise	Frequency of physical exercise (1=often, 5=never)	3.403 (1.62)
smoking	Smoked last month (1=yes)	0.313 (0.46)
drinking	Drank over 3 times last month (1=yes)	0.123 (0.33)

unhealthy_food	Had unhealthy food last week (1=yes)	0.642 （0.48）
chronic	Have chronic diseases (1=yes)	0.040 （0.20）
satis_hosipital	Level of satisfaction with local hospital (1=highest, 5=lowest)	2.620 （0.67）
time_hospital	Shortest travel time from home to nearest hospital (in minutes)	10.78 （14.92）
medical_cost	Family out-of-pocket medical cost last year (*yuan)*	3750 （10467）
no_fitspending	Did not spend on fitness or health preservation last year (1=yes)	0.910 (0.29)
no_healthins	Have no health insurance (1=yes)	0.183 （0.39）
crowded_housing	Live in an overcrowded housing environment (1=yes)	0.160 (0.37)
no_flushtoilet	No flush toilet (private or public) available in residential place (1=yes)	0.553 （0.50）
poor_garbage	Poor sanitation in garbage disposal (1=yes)	0.250 (0.44)
no_cleanwater	Have no clean water for cooking (1=yes)	0.300 (0.46)
no_electric	Have no stable electricity (1=yes)	0.490 (0.50)
month_electric	Monthly personal electricity consumption of household (kW ћ)	26.312 (26.34)
Father's Information		
score_f	CES-D score of father	11.696 （7.34）
primary_f	Primary school or below for father (1=yes)	0.507 （0.5）

middle_f	Middle school for father (1=yes)	0.316 (0.46)
high_f	High school for father (1=yes)	0.142 (0.35)
college_f	College or above for father (1=yes)	0.036 (0.19)
age_f	Age in years for father	52.939 (9.28)
Mother's Information		
score_m	CES-D score of mother	13.931 (8.21)
primary_m	Primary school or below for mother (1=yes)	0.677 (0.47)
middle_m	Middle school for mother (1=yes)	0.217 (0.41)
high_m	High school for mother (1=yes)	0.090 (0.29)
college_m	College or above for mother (1=yes)	0.017 (0.13)
age_m	Age in years for mother	51.105 (8.84)
Observation	Sample Size	5,074

Note: Data Resource: China Family Panel Studies (2012). The reported statistics are the sample mean with standard deviation in parentheses. All respondents aged 16 or above. We only keep respondents whose father and mother's information is both available in the dataset.

Table 2 Regression Results on the Intergeneration Transmission of Depression – Full Sample

Variable	Baseline model	Adding children's basic info	Adding parents' info	Adding children's SES	Adding children's health behavior	Adding children's healthcare condition	Adding children's living environment
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
score_f	0.198***	0.173***	0.173***	0.173***	0.172***	0.167***	0.165***
	(0.0131)	(0.0131)	(0.0132)	(0.0133)	(0.0133)	(0.0133)	(0.0133)
score_m	0.182***	0.162***	0.162***	0.159***	0.159***	0.157***	0.156***
	(0.0117)	(0.0118)	(0.0118)	(0.0119)	(0.0119)	(0.0119)	(0.0119)
Gradient_f	-	-	0.00%	0.00%	0.58%	3.47%	4.62%
Gradient_m	-	-	0.00%	1.85%	1.85%	3.09%	3.70%
Children's basic condition		Yes	Yes	Yes	Yes	Yes	Yes
Parents' information			Yes	Yes	Yes	Yes	Yes
Children's education				Yes	Yes	Yes	Yes
Adj_Income				-0.0568	-0.0538	-0.0196	-0.0174
				(0.0835)	(0.0837)	(0.0840)	(0.0845)
Children's unemployment				Yes	Yes	Yes	Yes
exercise					0.164***	0.149**	0.157***
					(0.0590)	(0.0589)	(0.0589)
smoking					0.580**	0.556**	0.558**
					(0.231)	(0.230)	(0.230)
drinking					0.615**	0.619**	0.621**
					(0.282)	(0.282)	(0.281)
unhealthy_food					0.00311	-0.0140	-0.0254
					(0.188)	(0.187)	(0.187)
chronic						4.339***	4.283***
						(0.440)	(0.440)
satis_hosipital						0.738***	0.724***
						(0.129)	(0.130)
time_hospital						-0.00305	-0.00266
						(0.00620)	(0.00628)
medical_cost						-9.92e-06	-9.70e-06
						(8.21e-06)	(8.23e-06)
no_fitspending						0.253	0.235

						(0.309)	(0.309)
no_healthins						0.226	0.241
						(0.234)	(0.234)
crowded_housing							-0.106
							(0.246)
no_flushtoilet							0.285
							(0.250)
poor_garbage							-0.0811
							(0.225)
no_cleanwater							-0.270
							(0.219)
no_electric							0.702***
							(0.181)
Adj_month_electric							-0.0883
							(0.104)
Constant	6.393***	2.142	1.575	1.303	1.161	-0.0764	-0.552
	(0.191)	(1.554)	(1.633)	(1.931)	(1.931)	(1.952)	(1.988)
Observations	5,074	5,074	5,074	4,982	4,982	4,910	4,908
R-squared	0.138	0.175	0.175	0.185	0.189	0.212	0.214

Note: 1) The reported statistics are the coefficients of the explanatory variables with the clustered robust standard errors shown in the parentheses. *, **, *** denote statistical significance at 10%, 5%, 1% levels, respectively. 2) Children's basic information includes age, whether urban resident or not, gender, province. Parents' information includes education level and age. Children's SES (socioeconomic status) includes whether employed or not, annual income, education level. 3) Adjusted income is defined as ln(Income)*(-1); adjusted electricity consumption is defined as ln(month_electric+1)*(-1).

Table 3 Estimated Intergeneration Transmission of Depression – Subsample Results

Variable	Adding children's basic info	Adding parents' info	Adding children's SES	Adding children's health behavior	Adding children's healthcare condition	Adding children's living environment
	(1)	(2)	(3)	(4)	(5)	(6)
			All sample (obs = 5074)			
score_f	0.173***	0.173***	0.173***	0.172***	0.167***	0.165***
	(0.0131)	(0.0132)	(0.0133)	(0.0133)	(0.0133)	(0.0133)
score_m	0.162***	0.162***	0.159***	0.159***	0.157***	0.156***
	(0.0118)	(0.0118)	(0.0119)	(0.0119)	(0.0119)	(0.0119)
Gradient_f	-	0.00%	0.00%	0.58%	3.47%	4.62%
Gradient_m	-	0.00%	1.85%	1.85%	3.09%	3.70%
			Gender			
			Male (obs = 3277)			
score_f	0.159***	0.158***	0.160***	0.158***	0.154***	0.152***
	(0.0165)	(0.0166)	(0.0167)	(0.0167)	(0.0167)	(0.0168)
score_m	0.163***	0.164***	0.160***	0.160***	0.159***	0.158***
	(0.0146)	(0.0146)	(0.0148)	(0.0147)	(0.0147)	(0.0147)
Gradient_f	-	0.63%	-0.63%	0.63%	3.14%	4.40%
Gradient_m	-	-0.61%	1.84%	1.84%	2.45%	3.07%
			Female (obs = 1797)			
score_f	0.202***	0.205***	0.203***	0.202***	0.196***	0.190***
	(0.0215)	(0.0218)	(0.0220)	(0.0220)	(0.0222)	(0.0224)
score_m	0.160***	0.159***	0.157***	0.161***	0.158***	0.156***
	(0.0201)	(0.0203)	(0.0204)	(0.0204)	(0.0205)	(0.0206)
Gradient_f	-	-1.49%	-0.50%	0.00%	2.97%	5.94%
Gradient_m	-	0.62%	1.88%	-0.63%	1.25%	2.50%
			Region			
			Rural (obs = 3604)			

score_f	0.193***	0.192***	0.194***	0.192***	0.188***	0.187***
	(0.0156)	(0.0157)	(0.0159)	(0.0159)	(0.0160)	(0.0160)
score_m	0.161***	0.163***	0.158***	0.159***	0.156***	0.156***
	(0.0139)	(0.0140)	(0.0141)	(0.0141)	(0.0141)	(0.0141)
Gradient_f	-	0.52%	-0.52%	0.52%	2.59%	3.11%
Gradient_m	-	-1.24%	1.86%	1.24%	3.11%	3.11%

Urban (obs = 1470)

score_f	0.133***	0.134***	0.135***	0.137***	0.128***	0.121***
	(0.0241)	(0.0244)	(0.0244)	(0.0244)	(0.0243)	(0.0245)
score_m	0.155***	0.156***	0.155***	0.153***	0.158***	0.158***
	(0.0221)	(0.0222)	(0.0223)	(0.0222)	(0.0222)	(0.0222)
Gradient_f	-	-0.75%	-1.50%	-3.01%	3.76%	9.02%
Gradient_m	-	-0.65%	0.00%	1.29%	-1.94%	-1.94%

Age Groups

Children (age 16 to 18) (obs = 931)

score_f	0.212***	0.222***	0.215***	0.211***	0.215***	0.213***
	(0.0312)	(0.0316)	(0.0317)	(0.0315)	(0.0315)	(0.0319)
score_m	0.148***	0.153***	0.160***	0.159***	0.150***	0.151***
	(0.0285)	(0.0287)	(0.0287)	(0.0285)	(0.0286)	(0.0288)
Gradient_f	-	-4.72%	-1.42%	0.47%	-1.42%	-0.47%
Gradient_m	-	-3.38%	-8.11%	-7.43%	-1.35%	-2.03%

Young Adult (age 19 to 25) (obs = 2038)

score_f	0.163***	0.161***	0.159***	0.159***	0.158***	0.156***
	(0.0199)	(0.0201)	(0.0203)	(0.0203)	(0.0208)	(0.0209)
score_m	0.157***	0.159***	0.156***	0.157***	0.157***	0.157***
	(0.0179)	(0.0181)	(0.0183)	(0.0182)	(0.0184)	(0.0184)
Gradient_f	-	1.23%	2.45%	2.45%	3.07%	4.29%
Gradient_m	-	-1.27%	0.64%	0.00%	0.00%	0.00%

Prime Age Adult (age 26 to 39) (obs = 1758)

score_f	0.160***	0.159***	0.165***	0.167***	0.161***	0.159***
	(0.0229)	(0.0231)	(0.0231)	(0.0232)	(0.0229)	(0.0230)
score_m	0.173***	0.174***	0.173***	0.172***	0.171***	0.167***
	(0.0203)	(0.0204)	(0.0205)	(0.0205)	(0.0202)	(0.0203)
Gradient_f	-	0.62%	-3.13%	-4.38%	-0.63%	0.62%
Gradient_m	-	-0.58%	0.00%	0.58%	1.16%	3.47%

Middle Aged and Older Adult (age 40 and above) (obs = 347)

score_f	0.242***	0.245***	0.232***	0.210***	0.206***	0.217***
	(0.0575)	(0.0581)	(0.0585)	(0.0592)	(0.0590)	(0.0595)
score_m	0.158***	0.162***	0.165***	0.171***	0.140**	0.128**
	(0.0535)	(0.0547)	(0.0546)	(0.0544)	(0.0544)	(0.0551)
Gradient_f	-	-1.24%	4.13%	13.22%	14.88%	10.33%
Gradient_m	-	-2.53%	-4.43%	-8.23%	11.39%	18.99%

Note: 1) The reported statistics are the coefficients of the explanatory variables with the clustered robust standard errors shown in the parentheses. *, **, *** denote statistical significance at 10%, 5%, 1% levels, respectively. 2) Children's basic information includes age, whether urban resident or not, gender, province. Parents' information includes education level and age. Children's SES (socioeconomic status) includes whether employed or not, annual income, education level.