Network for Studies on Pensions, Aging and Retirement

NETSPAR ACADEMIC SERIES

Exposure in utero to adverse events and health late-in-life

Evidence from China

Jiyuan Wang, Rob Alessie, Viola Angelini

DP 10/2023-068

Received: 3 February 2022 | Revised: 29 August 2022 | Accepted: 26 October 2022

DOI: 10.1002/hec.4632

RESEARCH ARTICLE

Health Economics WILEY

Exposure *in utero* to adverse events and health late-in-life: Evidence from China

Jiyuan Wang[1] | **Rob Alessie**[2] | **Viola Angelini**[2]

[1]School of Insurance, China Institute for Actuarial Science, Central University of Finance and Economics, Beijing, China

[2]University of Groningen and Netspar, Groningen, The Netherlands

Correspondence
Viola Angelini, Faculty of Economics and Business, University of Groningen and Netspar, Nettelbosje 2, Groningen 9747AE, The Netherlands.
Email: v.angelini@rug.nl

Funding information
National Institute on Ageing of the National Institute of Health, Grant/Award Numbers: 1-R21-AG031372-01, 1-R01-AG037031-01, 3-R01AG037031-03S1; National Science Foundation, Grant/Award Number: SBS-1021134; Natural Science Foundation of China, Grant/Award Numbers: 70773002, 70910107022, 71130002; Peking University; World Bank, Grant/Award Numbers: 7145915, 7159234

Abstract

This paper estimates the effect of in utero exposure to adverse events on late life diabetes, cardiovascular disease risks and cognition deficiency. We merge data on the regional violence during the Cultural Revolution and the excessive death rates during the Chinese Great Famine with data from the China Health and Retirement Longitudinal Study survey. Results show that female babies who were exposed in utero to the famine have higher diabetes risks, while male babies who were exposed to the Cultural Revolution are shown to have lower cognitive abilities.

KEYWORDS
cardiovascular disease, Chinese Great Famine, cognition, cultural revolution, diabetes, early life conditions

JEL CLASSIFICATION
I10, J11, J14

1 | INTRODUCTION

A growing body of literature has shown that exposure to adverse conditions around birth can have long-term negative consequences on health. The long shadow of early-life circumstances can be explained by both biological and social mechanisms. According to the "Fetal Origin Hypothesis" by Barker (see, e.g., Almond & Currie, 2011; Barker et al., 1993; Barker, 1995), malnutrition during pregnancy might negatively affect fetal growth and program later coronary heart disease and type-2 diabetes (see, e.g., Dinkelman, 2017; Doblhammer et al., 2013; Hu et al., 2017; Kim et al., 2017; Lumey et al., 2011; Portrait et al., 2011; Xu et al., 2017; Van den Berg et al., 2015). Moreover, prenatal stress might also be responsible for developing diseases such as depression and cognitive deficiency in the longer run (see, e.g., Akbulut-Yuksel, 2017; Bramsen et al., 2007; Gade & Wenger, 2011; Grimard & Laszlo, 2014; Islam et al., 2017; Strauss et al., 2011; Teerawichitchainan & Korinek, 2012). The pathway framework (Kuh & Shlomo, 2004), on the other hand, rests less on the biological imprinting and states that the impact of early life conditions on health later in life depends also on the interaction of the individual with the environment. The idea is that adverse conditions around birth might set in motion lifetime trajectories of health-related disadvantages. For example, children born under adverse circumstances might be exposed to poor diets, smoking, worse educational and job opportunities, and all these factors in turn can have a negative impact on health.

In this paper we estimate the effect of in utero exposure to the Chinese Great Famine (1959–1961) and the Cultural Revolution (mainly the Red Guard Army movement 1966–1971) on a range of physical and mental health outcomes in later life. While the famine was responsible for severe malnutrition for pregnant mothers living in rural areas, the Cultural Revolution caused a sharp increase in violence and stress in urban areas. We draw data from the China Health and Retirement Longitudinal Study (CHARLS) survey, a nationally representative survey of the Chinese population aged 45 or more. The 2011 and 2015 waves of the survey include data from blood biomarkers, which allow us to construct an 8-year risk score for diabetes and a 10-year risk score for cardiovascular disease (CVD). Indeed some in utero effects might only manifest themselves later in life but early signals of disease might be detected at younger ages using blood-based biomarkers. For mental health, the data include standard survey scales to measure the presence of depressive symptoms and assess cognitive abilities.

Our results support the idea that adverse events around birth potentially have a long-lasting negative effect on both physical and mental health. We also find that there are gender differences in this effect: in particular exposure in utero to the famine heightens the risk of type-2 diabetes only for women. We also find that the famine increases depressive symptoms for men, although this result is less robust. On the other hand, the Cultural Revolution mainly harms the cognitive abilities of men.

This paper contributes to the existing literature in three main ways. First, we study the long-term health effects of prenatal exposure not only to the Great Famine but also to the Cultural Revolution. Using two natural experiments is interesting because the Great Famine and the Cultural Revolution affected different populations, that is, the rural and urban respectively, and through different mechanisms. While the Great Famine primarily affected the nutrition of mothers during pregnancy, the Cultural Revolution had an effect on the stress levels of pregnant women. Second, we do not rely on self-reported measures of diabetes and cardiovascular disease but construct clinical risk measures thanks to the availability of blood-based biomarkers. Third, we estimate gender-specific effects.

The rest of the paper is organized as follows. Section 2 explains the background of the Chinese Great Famine and the Cultural Revolution, and their potential health effects. Section 3 introduces the data and the construction details of the main variables. Section 4 describes the empirical model and Section 5 presents the results. Finally, Section 6 concludes the paper.

2 | THE CHINESE GREAT FAMINE, THE CULTURAL REVOLUTION AND HEALTH LATER IN LIFE

2.1 | The Chinese Great Famine

The Chinese Great Famine (CGF thereafter) is by far one of the largest famines in human history in terms of both the severity and the size of the affected population. The famine occurred mainly between 1959 and 1961, when grain production substantially dropped. Given the fact that grain was the major source of food, the decline in grain production caused 16.5 to 45 million deaths during the famine period (see, e.g., Meng et al., 2015). In the year 1958, some regions had already experienced some grain production drops, but the famine had not yet become a nationwide disaster. After the famine, grain production increased monotonically until it reached the normal level.

The famine hit primarily the rural areas. Indeed, during the famine period, China was running a centrally planned economy, and rural households were only able to eat in local communal kitchens while urban households were allowed to consume normal food products within quotas depending on the household size. Moreover, rural households were not allowed to store food privately. So when the famine arrived, rural households were immediately hit by the food shortage shock in the communal kitchens. What was worse, the grain produced in rural areas was over-procured by the central government so it exacerbated the starvation of the rural households.

The famine not only produced substantial deaths nationwide but also had a long-term impact on the health outcomes of survivors. With little food available for rural pregnant mothers, the fetuses in the famine cohort were more likely to suffer from malnutrition than any other cohort. Babies who experience malnutrition in utero are shown to have higher risks of type-2 diabetes (see, e.g., Vaiserman & Lushchak, 2019) and higher risks of CVD (see, e.g., Alessie et al., 2019). The literature finds that maternal malnutrition creates intrauterine growth restriction, which then produces fetal adipose tissue and pancreatic β-cell dysfunction for the fetuses. To be more specific, the β-cell dysfunction is essential to explain the mechanisms behind the association. The β-cell is made from the stem cells when the pancreas of the embryo is developing, and it is the basis for the islets of Langerhans (the home of hormone). The decreased β-cell in reaction to the environmental cues will lead to irreversibly reduced insulin secretion. That is why babies exposed to malnutrition will develop higher diabetes risk.

The risks of having CVD also increase because of two reasons: first, since type-2 diabetes can also cause (or caused by) CVD, the risks of developing CVD also increase. Second, the lack of amino acids leads to a reduction in β-cells, while

glucose and oxygen deficiencies will directly lead to a reprogramming of the neuro-endocrine system (see, e.g., Almond & Mazumder, 2011). Then the CVD risks will be developed along with the type-2 diabetes risk.

The annual province-level death rates are provided by Lin and Yang (2000) and are available upon request from the authors. During the famine years death rates are clearly higher than in other non-famine years. Based on this, we calculate the excessive death rates and employ them to measure the famine severity across provinces. We calculate the excessive death rates in 0.1% unit by subtracting the average annual death rates between 1956 and 1958 from those in the famine years (1959–1961). During non-famine years, the measure of severity equals zero.

2.2 | The Cultural Revolution

The Cultural Revolution (CR thereafter) was a major adverse political event that happened between 1966 and 1976 (see, e.g., Walder & Su, 2003) in China. It is often referred to as China's "lost 10 years". In this period, millions of victims suffered from various types of political movements and prosecutions, and the negative impacts persist throughout the victims' whole life course. A more detailed description of the CR can be found in Bonnin (2006). The event itself can be divided into two sub-periods: the Red Guard Army movement between 1966 and 1971, and the rustication program between 1971 and 1976. The Red Guard Army movement was accompanied by massive conflicts and associated victims, while the rustication movement produced very few conflicts and victims. The extant studies either employ the Red Guard Army movement to study the effect of violence (see, e.g., Wang, 2019) or employ the Rustication movement to study the impact of interrupted education (see, e.g., Meng & Gregory, 2002).

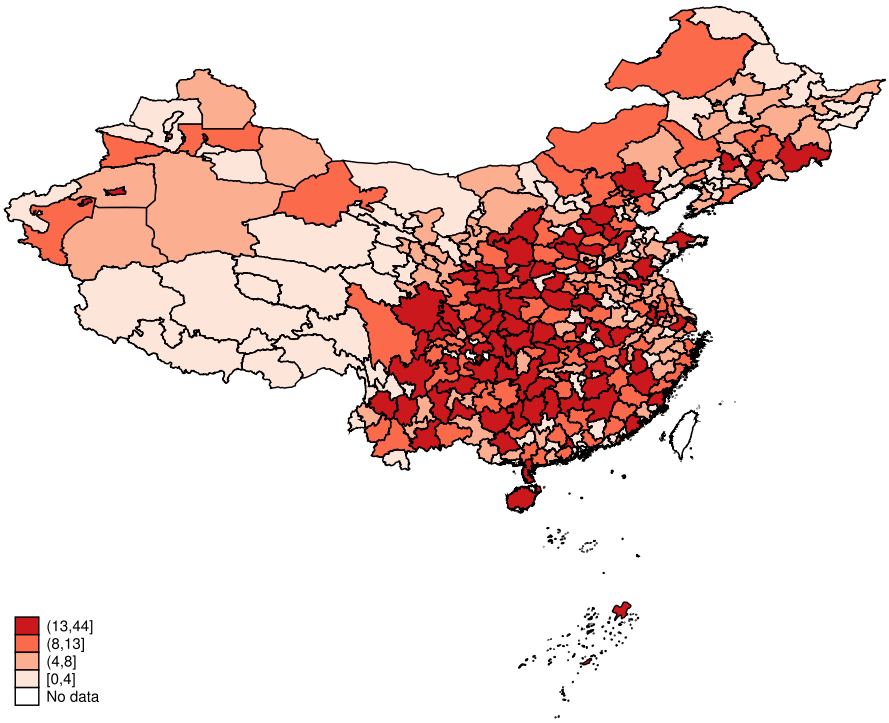
In this paper, we focus on the Red Guard Army Movement during the CR because we would like to focus on the effect of conflicts. During this period, students were encouraged by Mao to rebel against the central/local governments. Students worshiped him, followed his extreme-left ideology and created a lot of violence in this movement. In response, the governments repressed those student activities, producing a large number of victims and deaths. Those massive conflicts, victims, injuries and deaths were documented by local annals and gazetteers (see, e.g., Walder, 2014).

Pregnant mothers who experienced or witnessed the violence were more likely to suffer from prenatal stress (perhaps also malnutrition). This stress can be passed on to the fetuses, as shown by recent studies. The hypothalamic-pituitary-adrenal (HPA) axis plays a crucial role in understanding the mechanism,[1] especially during early fetal development because plasticity is very high (see, Braithwaite et al., 2014). Basically, the HPA axis is a neuroendocrine system that regulates the body's response to stress. Higher prenatal stress implies more responses or reactions from the HPA axis, which in turn affects the brain formation of babies and cognitive abilities later in life.[2]

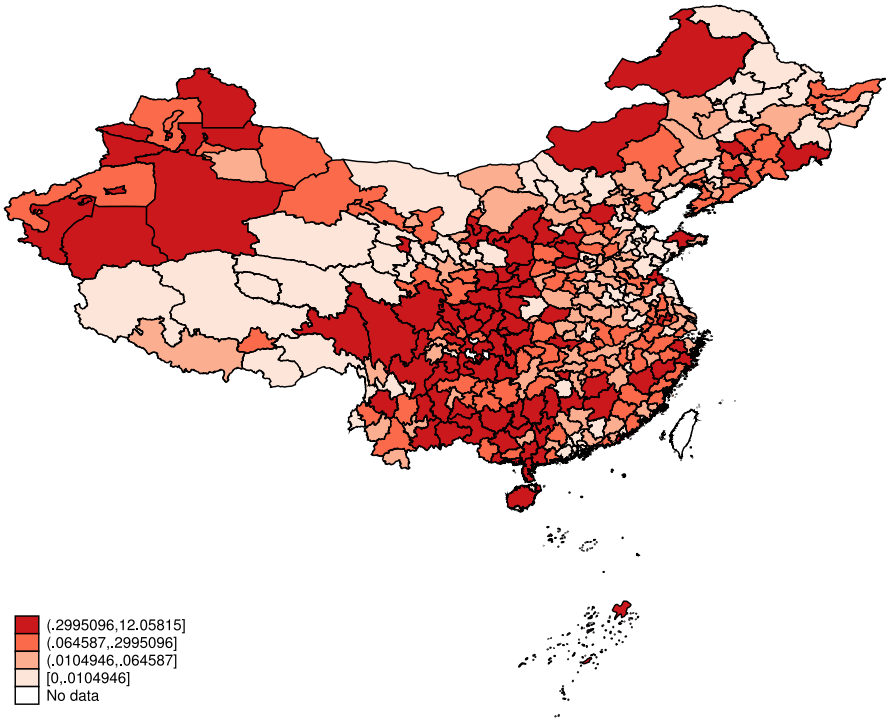
Mothers in different regions and in different years suffered from different levels of violence, and we make use of this geographical and temporal variation to identify the health effects. More specifically, we employ the number of collective conflicts as the measure of violence intensity during the CR (Red Guard Army movement 1966–1971) years. The measure equals 0 during the non-CR years. In a sensitivity analysis, we also use an alternative measure based on the number of deaths due to these collective conflicts. A similar strategy is also employed by Bai (2014) and Wang (2019). In Figure 1, we show the severity of total violence during the CR years by prefecture.

2.3 | Role of gender

In the literature, the evidence on the sex-specific health effects of prenatal exposure to adverse events is mixed (Alessie et al., 2019). For example, in terms of exposure to prenatal stress, some studies reveal that it is mostly male offspring's emotionality to be negatively affected, while others find the opposite effect, that is, that female offspring's anxiety and depressive symptoms respond significantly to the adverse environment (see, e.g., Bock et al., 2015, for a review of both animal and human studies). A strand of research finds that the gender-specific effects in response to stress can be explained by the timing of stress exposure (see, e.g., Mueller & Bale, 2008). Male babies are more responsive to prenatal stress during early gestational periods while female babies are more responsive during late gestational periods. In terms of exposure to malnutrition, there is also some evidence showing that males are affected differently than females. For example, Eriksson et al. (2010) find that boys grow faster than girls in utero, therefore boys need more nutrition from mothers. If mothers experience malnutrition during pregnancy, boys in utero will more likely be affected. Moreover, boys in utero tend to develop their brains first. If exposed to undernutrition, boys' visceral development will be sacrificed to sustain the nutrition supply for their brains. In the long run, boys' kidney development might be compromised, and some diseases such as hypertension and CVD will incur due to the underdevelopment



(a) Number of Conflicts* 1000 per capita



(b) Number of Deaths* 1000 per capita

FIGURE 1 Collective violence during the cultural revolution 1966–1971. [Colour figure can be viewed at wileyonlinelibrary.com]

of their kidney functions (see, e.g., Barker et al., 2006). In this paper, we will also empirically investigate the gender-specific health effects of exposure to the CR and the famine to see whether males and females are affected differently.

3 | DATA ON HEALTH OUTCOMES AND DESCRIPTIVE STATISTICS

We draw data from the China Health and Retirement Longitudinal Study (CHARLS) 2011 and 2015 surveys. The CHARLS survey is designed to be nationally representative, and it focuses only on the population aged 45 and above. It has one national study with four regular waves (wave 2011, 2013, 2015 and 2018) and a retrospective wave (wave 2014). In this paper, we use the 2011 and 2015 waves of the national sample because they are the only waves that include a biomarker section with blood tests. Biomarkers are powerful preclinical or pre-morbid signals for diseases such as diabetes or CVD and allow us to construct risk scores. There are two reasons why it is important to use biomarkers. First, diabetes is severely underdiagnosed and, therefore, underreported in survey data. Mao et al. (2019) report that 64%–76% of diabetes in China is undiagnosed. This is especially true for those living in rural areas and with limited access to health care and for younger individuals as diabetes is often asymptomatic at younger ages. Notice that we do assign value 1 to the risk score if the patient has been diagnosed with diabetes but using biomarkers allows us also to detect (the risk of) diabetes in the undiagnosed population. Second, looking at 8-year and 10-year risks rather than at the presence of the disease allows us also to focus on the younger population as CVD and diabetes are more likely to manifest themselves at older ages.

As the CHARLS data set does not provide suitable measures for the severity of the CGF and the CR, we merge it with two additional data sources. We use the province-level excessive death rates from Lin and Yang (2000) to measure the famine severity, and we use prefecture-level conflicts and death counts from Walder (2014) to measure the CR severity. We merge each individual in CHARLS with these two additional data sets using the information on the prefecture/province and the year at birth. In CHARLS we also have information on the *hukou* status at birth, a household registration system used in China that identifies a person as a resident in a rural or urban area.

3.1 | Sample selection

In the original data sets, there are 17,708 observations in the 2011 wave and 21,095 observations in the 2015 wave. Only 11,847 respondents in the 2011 wave and 13,420 respondents in the 2015 wave have participated in the blood-based biomarker sections.[3] When we merge the different sections of the questionnaire, we are left with 10,138 observations in wave 2011 and 13,280 observations in wave 2015. We then construct a cross-sectional data set by pooling the two waves, and we use the latest information for each individual in the pooled data set. In other words, if a respondent has been interviewed in both 2011 and 2015, we use the information in 2015, which results in a sample size of 16,674.

We restrict our sample to those who were born between 1950 and 1971, which leaves us with 10,959 observations. We choose this period for the following reasons: first, the period not only covers both the CGF (1959–1961) and the Red Guard Army movement of the CR (1966–1971) but also includes a relatively peaceful time (1950–1959, 1962–1965). Second, we do not want to include war periods such as the Second World War (1938–1945) and the civil war (1945–1949). After merging with the CR violence data and the CGF excessive death rates data, the sample size reduces to 9794. We also drop observations with missing values in *hukou* status (144 observations), cognitive tests (595 respondents), depression symptoms (687) and the blood-based biomarkers (483).

In the final sample, there are 7885 respondents distributed in 254 prefectures in 29 provinces.[4] In this sample, there are 4341 female and 3544 male respondents. Of those 7885 respondents, 7232 individuals have rural *hukou* at birth and 653 individuals have non-rural *hukou* at birth. There are 3668 individuals who were born between 1950 and 1958, 799 individuals born between 1959 and 1961, 1758 individuals born between 1961 and 1965, and 1660 individuals born between 1966 and 1971.

3.2 | Health outcomes

3.2.1 | Diabetes and CVD

Different from the extant studies which employ mostly self-reported health measures, we use biomarker data to construct risk scores for developing diabetes and CVD. For type-2 diabetes we consider the 8-year risk score, which clinically can be served as an early monitor for diabetes prevalence after 8 years. We use various sources of information to construct the risk score from the Framingham Heart Study (FHS, see, Wilson et al., 2007): age, gender, BMI, blood pressure, blood High Density Level Cholesterol (HDL Cholesterol), Triglycerides, and fasting blood glucose. The detailed calculations are available upon request from the

authors. The risk score works appropriately for undiagnosed patients but works inappropriately for those who have been diagnosed with diabetes and are taking medications. Therefore, we assign a risk score of 1 to those who were diagnosed with diabetes but have normal blood sugar levels because they are taking medications, which we can identify from the CHARLS data. The distribution of the risk score can be found in Table 1. In our sample, on average, the risk of getting diabetes in 8 years is 32.5%, and 16.2% of the respondents have already been diagnosed with diabetes. The figures are comparable with other data sources or public reports. For example, Xu et al. (2013) report that nearly half of the Chinese adult population are prediabetic, and the 2015 Report on Nutrition and Chronic Disease in Chinese Residents reveals that the prevalence rate of diabetes among Chinese adults aged 40 is 9.7% (see, Burns & Liu, 2017, chapter 6). We also disaggregate the numbers by cohort, gender and *hukou* status, and the statistics can be found in Table 2. From Panel A of Table 2, we find that normally the younger cohorts have lower diabetes risks except for the famine cohort (1959–1961). The 8-year diabetes risk for the famine cohort is 35.7%, which is higher than that of both the pre- and post-famine cohorts (34.1% and 31.9%),[5] indicating that the famine might play a role in increasing the risk of getting diabetes. Panel B of the table shows that females have lower CVD risk but higher prevalence rates.

For CVD, we follow D'agostino et al. (2008) and construct a 10-year risk score from the FHS. Other studies employ the Systematic COronary Risk Evaluation (hereinafter SCORE) (see, e.g., Alessie et al., 2019; Conroy et al., 2003) but we focus on the FHS because Selvarajah et al. (2014) shows that it performs better than SCORE for the Asian population. The details on the construction of the risk score are available upon request. As for diabetes, we assign value one to the CVD risk score for respondents with a history of CVD events such as heart failure and stroke. The descriptive statistics can be found in Table 1. On average, 8.27% of the sample has already experienced a CVD event (stroke and heart failures), and the average CVD risk in the sample is 15.4%. The figure is close to external sources. For example, the reported prevalence rate of CVD in China is about 20% (one in five adults in China has a CVD).[6]

3.2.2 | Cognition and depression

We employ various measures to describe cognitive abilities and depressive symptoms of older Chinese individuals. The first measure of cognition, *episodic memory*, uses the word recalling questions in CHARLS (see, e.g., Lei et al., 2012). Respondents were requested to read 10 Chinese words, and then they were asked to recall as many words as possible at two different

TABLE 1 Descriptive statistics CHARLS 2011/15: Born during 1950–1971

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variables	**Mean**	**S.D.**	**Min**	**Max**	**Quartile 1**	**Median**	**Quartile 3**
Sample size 7885							
Female (0–1)	0.551	0.497	0	1	0	1	1
Rural hukou (0–1)	0.917	0.276	0	1	1	1	1
Cognition and depression							
Episodic memory (word recall) (0–10)	3.699	1.765	0	10	2.500	4	5
Mental intactness (TICS) (0–10)	6.815	2.754	0	10	5	7	9
Graphical cognition (0–1)	0.694	0.461	0	1	0	1	1
CESD-10 (0–30)	7.792	6.295	0	30	3	6	11
CVD and diabetes							
Diabetes	0.162	0.368	0	1	0	0	0
8-year diabetes risk	0.325	0.341	3.51e–05	1	0.0478	0.166	0.551
CVD	0.0827	0.275	0	1	0	0	0
10-year CVD risk	0.154	0.266	0.00220	1	0.0249	0.0572	0.128

Note: Sample size 7885 include all the respondents in wave 2011 and 2015, and we keep the latest information if the repondents appear twice. Rural *hukou* equals one if the repondents have the rural *hukou* at birth, zero otherwise. Episodic memory is constructed by asking the respondents recall the 10 words listed on cards, ranging from 0 (no words recalled) to 10 (all words recalled). Mental intactness is measured through questions which ask the respondent to subtract 7 from 100 and keep subtracting 7 from the result for a maximum of five times and to name the correct date (day of the week, month, day, year, and season). The maximum score is 10. Graphical cognitive ability is measured by a dummy which is equal to 1 if the respondent is able to draw a picture showed by the interviewer. CESD-10 measures depressive symptoms, ranging from 0 (no depression at all) to 30 (the most serious depression). Diabetes is a dummy which equals one if the respondents have been diagnosed with diabetes and/or taking medications to control the blood sugar level. The 8-year risk score for diabetes is constructed by following Wilson et al. (2007). CVD is a dummy for actual cardiovascular disease and equals one if the respondents have been diagnosed with stroke or heart failures. The 10-year CVD risk measure is from Framingham Heart Study (D'agostino et al. (2008)).

TABLE 2 Descriptive statistics of health outcomes by groups in CHARLS 2011 and 2015[a]

Panel A: By cohort														
	Pre-famine		**Famine**				**Pre-CR**				**CR cohort**			
	1950–1959		**1959–1961**				**1962–1965**				**1966–1971**			
	(*N* = 3668)		**(*N* = 799)**				**(*N* = 1758)**				**(*N* = 1660)**			
Health variables	**Mean**	**SD**	**Mean**	**SD**	**Diff.**[b]	***p*-value**[c]	**Mean**	**SD**	**Diff.**	***p*-value**	**Mean**	**SD**	**Diff.**	***p*-value**
Female	0.524	0.499	0.546	0.498	−0.022	0.2587	0.572	0.495	−0.026	0.2193	0.589	0.492	−0.017	0.3143
Rural	0.910	0.286	0.900	0.300	0.010	0.3748	0.922	0.269	−0.022*	0.0647	0.937	0.243	−0.015*	0.0878
Cognition and depression														
Episodic memory (word recall)	3.334	1.743	3.683	1.684	−0.349***	0.0001	4.011	1.715	−0.328***	0.0001	4.183	1.728	−0.172***	0.0035
Mental intactness (TICS)	6.417	2.880	6.780	2.662	−0.363***	0.0011	7.011	2.592	−0.231**	0.0384	7.502	2.514	−0.491***	0.0001
Graphical cognition	0.622	0.485	0.701	0.458	−0.079***	0.0001	0.763	0.425	−0.062***	0.0009	0.779	0.415	−0.016	0.2659
CESD-10	8.282	6.470	7.458	6.073	0.824***	0.0015	7.659	6.190	−0.201	0.4440	7.013	6.019	0.646***	0.0020
CVD and diabetes														
Diabetes	0.187	0.390	0.184	0.388	0.000	0.8437	0.141	0.348	0.043***	0.0053	0.117	0.321	0.024**	0.0365
8-year diabetes risk	0.341	0.346	0.357	0.357	−0.016	0.2390	0.319	0.332	0.038***	0.0089	0.282	0.324	0.037***	0.0010
CVD	0.110	0.313	0.100	0.300	0.010	0.4098	0.0666	0.249	0.033***	0.0033	0.0301	0.171	0.037***	0.0001
10-year CVD risk	0.207	0.294	0.170	0.286	0.037***	0.0012	0.119	0.241	−0.05***	0.0033	0.0674	0.169	0.0516***	0.0001

Panel B: By hukou and gender												
	Male		**Female**				**Rural**		**Non-rural**			
	(*N* = 3544)		**(*N* = 4341)**				**(*N* = 7232)**		**(*N* = 653)**			
Health variables	**Mean**	**SD**	**Mean**	**SD**	**Diff. (Male-female)**	***p*-value**	**Mean**	**SD**	**Mean**	**SD**	**Diff (Rur-NonRur).**	***p*-value**
Female	-	-	-	-	-	-	0.550	0.497	0.551	0.498	−0.001	0.9607
Rural	0.917	0.275	0.917	0.276	0.000	1.0000	-	-	-	-	-	-
Cognition and depression												
Episodic memory (word recall)	3.735	1.694	3.670	1.821	0.065	0.1038	3.608	1.750	4.715	1.611	−1.107***	0.0001
Mental intactness (TICS)	7.477	2.417	6.274	2.891	1.203***	0.0001	6.706	2.777	8.015	2.149	−1.309***	0.0001
Graphical cognition	0.793	0.405	0.613	0.487	0.180***	0.0001	0.680	0.466	0.850	0.357	−0.170***	0.0001
CESD-10	6.620	5.750	8.750	6.554	−2.130***	0.0001	7.924	6.336	6.332	5.618	1.5920***	0.0001

(Continues)

TABLE 2 (Continued)

Panel B: By hukou and gender												
	Male		Female				Rural		Non-rural			
	(*N* = 3544)		(*N* = 4341)				(*N* = 7232)		(*N* = 653)			
Health variables	**Mean**	**SD**	**Mean**	**SD**	**Diff. (Male-female)**	***p*-value**	**Mean**	**SD**	**Mean**	**SD**	**Diff (Rur-NonRur).**	***p*-value**
CVD and diabetes												
Diabetes	0.166	0.373	0.158	0.364	0.008	0.3371	0.157	0.364	0.213	0.410	−0.056***	0.0002
8-year diabetes risk	0.321	0.347	0.328	0.335	−0.007	0.3638	0.319	0.339	0.395	0.356	−0.076***	0.0001
CVD	0.0643	0.245	0.0977	0.297	−0.033***	0.0001	0.0784	0.269	0.130	0.337	−0.052***	0.0001
10-year CVD risk	0.182	0.233	0.132	0.288	0.050***	0.0001	0.150	0.260	0.206	0.319	−0.056***	0.0001

[a]In this table we disaggregate the statistics from Table 1 by cohort, gender and *hukou* status. Panel A exhibits the statistics by cohort, and Panel B exhibits the ones by gender and *hukou*. Total sample size: 7885.

[b]We compare the mean differences between two consecutive cohorts.

[c]The *p*-value are calculated based on a two-tailed test.

*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

moments in time. The first is just after reading all the words, and the second is after answering several other questions. We measure the number of correct words by taking the average of the two counts. The second measure of cognition is *mental intactness*, which is generated from 10 questions of the Telephone Interview of Cognitive Status (TICS hereinafter), (see, e.g., Huang & Zhou, 2013; Lei et al., 2012). This measure is constructed based on whether the respondent is able to subtract 7 from 100 and keep subtracting 7 from the result for a maximum of five times, and to name the correct date (day of the week, month, day, year, and season). The maximum score is 10. The third measure considers graphical cognitive ability, which is a dummy indicating whether the respondent is able to draw a picture shown by the interviewer.

For depressive symptoms, we employ the well-known CESD-10 measure (see, e.g., Björgvinsson et al., 2013). Respondents were asked to report how often they had experienced the following situations: being bothered by small things, having difficulty in keeping their mind on what they were doing, feeling depressed, being tired of doing things, feeling hopeful, feeling fearful, feeling restless, feeling happy, feeling lonely, and feeling hard to get going. The answers were recorded on a four-point scale from 0 to 3, corresponding to rarely, some days, occasionally, and most of the time for the negative questions, and reversed for the positive ones. The total score ranges between 0 and 30, where a cutoff score of 10 or higher indicates clinically relevant depression.

In Panel A of Table 2, we observe that older cohorts have on average lower cognitive abilities (episodic memory, mental intactness and graphical cognition), whereas depressive symptoms are more severe among younger respondents. Panel B shows that females have lower cognitive performance than males, and they have higher levels of depressive symptoms. Rural-born residents have more depressive symptoms and lower cognitive abilities than their urban counterparts.

4 | EMPIRICAL FRAMEWORK

The model to be estimated is of the following form:

$$y_{igpt} = \alpha + \beta_1 \cdot EDR_{pt} \cdot Rural_{igp} + \beta_2 \cdot Vio_{gt} \cdot Urban_{igp} + X_{igp}\gamma + \lambda_t + \theta_g + \epsilon_{igpt}, \quad (1)$$

where y_{igpt} is one of the health outcomes we have mentioned in Section 3. The subscript i refers to the individual, g denotes prefecture (254), p means province (29), and t stands for the year of birth. The variable excess death rate EDR_{pt} is the excess mortality rate in province p in the year in which the respondent was born t for the CGF years (1959–1961). It is computed by subtracting the average death rate between 1956 and 1958 from the annual death rate during the famine years. During the non-famine years, EDR_{pt} equals zero. As within a province there are both rural and urban areas but the CFG only affected rural areas, we interact the provincial excess death rate with a dummy $Rural_{igp}$, which equals one if the individual was born in a rural area (having a rural *hukou* at birth) and zero if the individual was born in an urban area (having an urban *hukou* at birth). As we only consider the *hukou* status at birth, this dummy is time-invariant. The variable violence Vio_{gt} is measured at the level of the prefecture g in year t during the CR and is the number of conflicts per thousand individuals during the CR years. It is equal to zero in the non-CR years. Since also within prefectures there are both rural and urban areas but the Cultural Revolution mainly hit urban areas, we interact Vio_{gt} with the dummy $Urban_{igp}$.

The covariate vector X_{igp} includes variables like gender and *hukou* status. The cross-cohort correlations are captured by the cohort fixed effect λ_t, and we also introduce the prefecture fixed effects θ_g to account for unobserved cross-regional heterogeneity. The rest of the unobserved shocks are included in the idiosyncratic error term ϵ_{igpt}. To account for the correlation within each prefecture and each cohort, the standard errors are clustered at both the prefecture and the year of birth level.

5 | RESULTS

5.1 | Main results

We estimate model (1) for the full sample, which consists of all individuals who were born between 1950 and 1971, and the results can be found in Table 3. Column (1) exhibits the regression estimates using the predicted 8-year type-2 diabetes risk as the dependent variable. Column (2) shows the estimates using the predicted 10-year CVD risk. Column (3) displays the estimates using depression as the dependent variable which is measured by the CESD 10 scale. Columns (4–6) are the estimates for three cognition measures.

The results show that mothers' exposure to a 1% (the unit of the excess death rates in the raw data is 0.1%) more severe famine in rural areas during pregnancy would increase the risk of type-2 diabetes of the offspring by 0.03% points (Column [1]) on average, while the effect on CVD risk is insignificant. These results suggest that exposure to famine might have triggered the

TABLE 3 Main results: CHARLS 2011 and 2015

	(1)	(2)	(3)	(4)	(5)	(6)
	Diabetes	**CVD**	**Depression**	**Cognition**		
	Predicted 8-year risk	**Predicted 10-year risk**	**CESD-10**	**Graphical cognition**	**Mental intactness (TICS)**	**Episodic memory (Word recall)**
Variables	%	%	**Score 0–30(Worst)**	**0–1(Best)**	**0–10(Best)**	**0–10(Best)**
EDR*Rural	0.00354**	0.000783	0.0597**	−0.000878	−0.0112	−0.00961
	(0.00170)	(0.00122)	(0.0291)	(0.00220)	(0.0123)	(0.00805)
Conflicts*Urban	−0.0108	−0.0358	−0.0410	−0.0644	0.369	0.0276
	(0.0372)	(0.0253)	(0.677)	(0.0437)	(0.255)	(0.210)
Rural hukou	−0.0325	−0.0185	1.292***	−0.138***	−0.815***	−0.862***
	(0.0247)	(0.0157)	(0.353)	(0.0269)	(0.129)	(0.122)
Female	0.00872	−0.0443***	2.204***	−0.193***	−1.274***	−0.134**
	(0.0120)	(0.00781)	(0.158)	(0.0162)	(0.111)	(0.0541)
Constant	0.484**	0.216	0.808	1.053***	10.45***	5.689***
	(0.193)	(0.145)	(3.487)	(0.252)	(1.468)	(0.964)
Observations	7885	7885	7885	7885	7885	7885
R-squared	0.090	0.149	0.127	0.152	0.191	0.153
Interview year dummy	YES	YES	YES	YES	YES	YES
Year of birth FE	YES	YES	YES	YES	YES	YES
Prefecture FE	YES	YES	YES	YES	YES	YES

Note: Standard errors are clustered at the prefecture and year of birth levels and reported in parentheses. In Columns (1–2) we examine the impacts of violence (measured by log of number of collective conflicts) and famine (measured by excessive death rates) on the predicted 8-year diabetes risk and 10-year CVD risk. Column (3) shows the results using CESD-10 as the measure of depression, and columns (4–6) use three measures of cognition: (i) ability to draw a required picture, (ii) mental intactness, and (iii) memory ability using word recall questions (see the paper for details). *EDR* is the province-level excessive death rates in 1959, 1960 and 1961, which is calculated by subtracting the average death rates between year 1956–1958 from the annual death rates in the three famine years. The variable *Conflicts* measures the level of violence in each of the years between 1966 and 1971. It is measured at the prefecture level per thousand population. The variable *Rural* is the indicator for rural *hukou* at birth.

*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

"thrifty genotype" mechanism which induces insulin secretion during adult life. The magnitude of the effect is non-negligible because it is around 10.9% of the sample mean and around 10.4% of a standard deviation. This increase in the risk of developing diabetes might generate substantial healthcare costs and welfare losses for society. A 1% increase in famine severity leads to a 0.6 points increase in the depression scale, while there is no significant effect on cognitive abilities. Exposure to conflicts in utero has no significant long-run effects on any of the health outcomes.

5.2 | Heterogeneous effects by gender

Some studies, such as Eriksson et al. (2010), suggest that boys might be affected differently than girls (see Section 2.3). To account for potential heterogeneous effects between males and females, we interact the famine severity and regional conflicts with gender. The results can be found in Table 4.

The results show that the effect of the famine on diabetes risk matters only for females, and a 1% increase in excessive death rate would imply a 0.06% points higher risk of getting type-2 diabetes. The effect is around 17.5% of the sample mean and 16.7% of a standard deviation, which is statistically and economically relevant. The effects for males are statistically insignificant. The results might also be driven by selective mortality, namely males were more likely to die earlier. In terms of the impact on CVD risk, we also find insignificant results. The pattern for depression is different: a 1% increase in famine-induced deaths increases males' depression score by 0.8, suggesting that boys' emotionality is more affected by the adverse prenatal environment, especially prenatal stress. The results also show that male babies who were exposed to the Cultural Revolution have lower (graphical) cognitive abilities. For female babies, we find a positive effect on mental intactness, which is not in line with our expectations and might point to a resilience effect.

TABLE 4 Gender specific: CHARLS 2011 and 2015

	(1)	(2)	(3)	(4)	(5)	(6)
	Diabetes	CVD	Depression	Cognition		
	Predicted 8-year risk	Predicted 10-year risk	CESD-10	Graphical cognition	Mental intactness	Episodic memory
			Score		(TICS)	(Word recall)
Variables	%	%	0–30(Worst)	0–1(Best)	0–10(Best)	0–10(Best)
Male*EDR*Rural	0.000457	0.00138	0.0803**	0.00162	−0.0138	−0.00825
	(0.00222)	(0.00193)	(0.0402)	(0.00303)	(0.0172)	(0.0115)
Female*EDR*Rural	0.00566***	0.000376	0.0454	−0.00258	−0.00937	−0.0105
	(0.00210)	(0.00170)	(0.0348)	(0.00260)	(0.0147)	(0.00962)
Male*Conflicts*Urban	−0.00619	−0.0622	0.869	−0.167**	−0.421	−0.444
	(0.0531)	(0.0391)	(0.937)	(0.0764)	(0.395)	(0.283)
Female*Conflicts*Urban	−0.0138	−0.0189	−0.621	0.00115	0.871***	0.329
	(0.0438)	(0.0316)	(0.799)	(0.0547)	(0.319)	(0.272)
Rural hukou	−0.0324	−0.0188	1.302***	−0.139***	−0.823***	−0.867***
	(0.0248)	(0.0158)	(0.356)	(0.0271)	(0.129)	(0.121)
Female	0.00638	−0.0442***	2.231***	−0.192***	−1.286***	−0.139**
	(0.0118)	(0.00810)	(0.163)	(0.0161)	(0.112)	(0.0542)
Constant	0.486**	0.216	0.783	1.052***	10.46***	5.692***
	(0.193)	(0.145)	(3.487)	(0.251)	(1.468)	(0.963)
Observations	7885	7885	7885	7885	7885	7885
R-squared	0.091	0.149	0.127	0.153	0.192	0.153
Interview year dummy	YES	YES	YES	YES	YES	YES
Year of birth FE	YES	YES	YES	YES	YES	YES
Prefecture FE	YES	YES	YES	YES	YES	YES

Note: Standard errors are clustered at the prefecture and year of birth levels and reported in parentheses. This table show the gender-specific estimates with the full sample. In Columns (1–2) we examine the impacts of violence (measured by log of number of collective conflicts) and famine (measured by excessive death rates) on the predicted 8-year diabetes risk and 10-year CVD risk. Column (3) shows the results using CESD-10 as the measure of depression, and columns (4–6) use three measures of cognition: (i) ability to draw a required picture, (ii) mental intactness, and (iii) memory ability using word recall questions (see the paper for details). *EDR* is the province-level excessive death rates in 1959, 1960 and 1961, which is calculated by subtracting the average death rates between year 1956–1958 from the annual death rates in the three famine years. The variable *Conflicts* measures the level of violence in each of the years between 1966 and 1971. It is measured at the prefecture level per thousand population. The variable *Rural* is the indicator for rural *hukou* at birth.

*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

5.3 | Robustness checks

We conduct a number of robustness checks to ensure that the effects are not spurious. First, as we have different outcome variables, we apply a correction for multiple hypothesis testing. For comparability and to be conservative, we apply this correction also to all other robustness checks reported in this section. In particular, we divide the dependent variables into two groups: physical health (diabetes and CVD risk) and mental health (depression and the cognition measures). Within each group, we compute *p*-values that adjust for multiple hypothesis testing using the Romano-Wolf correction (Romano and Wolf (2005)).[7] As we have not only multiple outcomes but also multiple treatments, this procedure should be taken with some caution. The results are reported in Table 5 and show that the effects of the famine on diabetes and of the Cultural Revolution on cognitive abilities remain significant, while the results on depression are less robust.

Second, since the number of conflicts might not capture the full damage during the CR, we use an alternative violence measure: the (log) number of deaths due to the collective conflicts. Third, since the results might be sensitive to the choice of the control group, in a sensitivity analysis we restrict the control group to those who were born after 1955 (the original sample includes all respondents who were born after 1949).

TABLE 5 Gender specific: CHARLS 2011 and 2015

	(1)	(2)	(3)	(4)	(5)	(6)
	Diabetes	**CVD**	**Depression**	**Cognition**		
	Predicted 8-year risk	**Predicted 10-year risk**	**CESD-10**	**Graphical cognition**	**Mental intactness**	**Episodic memory**
			Score		**(TICS)**	**(Word recall)**
Variables	**%**	**%**	**0–30(Worst)**	**0–1(Best)**	**0–10(Best)**	**0–10(Best)**
Male*EDR*Rural	0.000457	0.00135	0.0803	0.00162	−0.0138	−0.00825
	(0.00222)	(0.00193)	(0.0402)	(0.00303)	(0.0172)	(0.0115)
	[0.8492]	[0.6484]	[0.3497]	[0.1928]	[0.3497]	[0.5045]
Female*EDR*Rural	0.00566***	0.000393	0.0454	−0.00258	−0.00937	−0.0105
	(0.00210)	(0.00170)	(0.0348)	(0.00260)	(0.0147)	(0.00962)
	[0.0030]	[0.7662]	[0.7502]	[0.3437]	[0.3437]	[0.2627]
Male*Conflicts*Urban	−0.00619	−0.0623**	0.869	−0.167*	−0.421	−0.444
	(0.0531)	(0.0391)	(0.937)	(0.0764)	(0.395)	(0.283)
	[0.8352]	[0.0270]	[0.7403]	[0.0819]	[0.7592]	[0.7592]
Female*Conflicts*Urban	−0.0138	−0.0188	−0.621	0.00115	0.871***	0.329**
	(0.0438)	(0.0316)	(0.799)	(0.0547)	(0.319)	(0.272)
	[0.9181]	[0.4146]	[0.4406]	[0.4406]	[0.0090]	[0.0350]
Rural hukou	−0.0324***	−0.0189***	1.302***	−0.139***	−0.823***	−0.867***
	(0.0248)	(0.0157)	(0.356)	(0.0271)	(0.129)	(0.121)
	[0.0010]	[0.0010]	[0.0010]	[0.0010]	[0.0010]	[0.0010]
Female	0.00638	−0.0445***	2.231***	−0.192***	−1.286***	−0.139*
	(0.0118)	(0.00812)	(0.163)	(0.0161)	(0.112)	(0.0542)
	[0.6024]	[0.0010]	[0.0010]	[0.0010]	[0.0010]	[0.0939]
Constant	0.486**	0.253*	0.783	1.052***	10.46***	5.692***
	(0.193)	(0.145)	(3.487)	(0.251)	(1.468)	(0.963)
Observations	7885	7885	7885	7885	7885	7885
R-squared	0.091	0.149	0.127	0.153	0.192	0.153
Interview year dummy	YES	YES	YES	YES	YES	YES
Year of birth FE	YES	YES	YES	YES	YES	YES
Prefecture FE	YES	YES	YES	YES	YES	YES

Note: Romano-Wolf Correction for Multiple Hypothesis Testing. The *p*-values in brackets underlying the significance levels are Romano-Wolf stepdown adjusted *p*-values which consider multiple hypothesis testing across different health outcomes. Standard errors are clustered at the prefecture and year of birth levels and reported in parentheses. This table show the gender-specific estimates with the full sample. In Columns (1–2) we examine the impacts of violence (measured by log of number of collective conflicts) and famine (measured by excessive death rates) on the predicted 8-year diabetes risk and 10-year CVD risk. Column (3) shows the results using CESD-10 as the measure of depression, and columns (4–6) use three measures of cognition: (i) ability to draw a required picture, (ii) mental intactness, and (iii) memory ability using word recall questions (see the paper for details). *EDR* is the province-level excessive death rates in 1959, 1960 and 1961, which is calculated by subtracting the average death rates between year 1956–1958 from the annual death rates in the three famine years. The variable *Conflicts* measures the level of violence in each of the years between 1966 and 1971. It is measured at the prefecture level per thousand population. The variable *Rural* is the indicator for rural *hukou* at birth.

*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

In Table 6 we report the results when we replace the measure of CR violence with the log of the number of deaths. Using an alternative measure does not substantially alter the effect of the famine on diabetes and CVD risk scores. The estimated effects of violence on cognition are less robust, although both results in Column 4 of Tables 3 and 5 show that (graphical) cognition of males is negatively affected by violence exposure.

The results for the third robustness check are available from request from the authors. When we restrict the sample to those who were born after 1955, the estimates do not qualitatively change from the ones in Table 4. For instance, the estimated coefficient in Column (1) for the female-EDR-rural interaction term changes from 0.00566 to 0.00558. The effect of the

TABLE 6 Gender specific: CHARLS 2011 and 2015

	(1)	(2)	(3)	(4)	(5)	(6)
	Diabetes	CVD	Depression	Cognition		
	Predicted 8-year risk	Predicted 10-year risk	CESD-10	Graphical cognition	Mental intactness (TICS)	Episodic memory (Word recall)
Variables	%	%	Score 0–30(Worst)	0–1(Best)	0–10(Best)	0–10(Best)
Male*EDR*Rural	0.000443	0.00130	0.0807	0.00163	−0.0135	−0.00790
	(0.00222)	(0.00193)	(0.0402)	(0.00303)	(0.0172)	(0.0115)
	[0.8591]	[0.6633]	[0.3297]	[0.1928]	[0.3297]	[0.4855]
Female*EDR*Rural	0.00566***	0.000344	0.0458	−0.00262	−0.00934	−0.0104
	(0.00209)	(0.00170)	(0.0348)	(0.00260)	(0.0147)	(0.00962)
	[0.0020]	[0.7662]	[0.7463]	[0.3646]	[0.3646]	[0.2887]
Male*Deaths*Urban	−0.891	−0.691*	15.00	−2.562	−7.533	−7.455
	(0.883)	(0.667)	(15.98)	(1.753)	(6.729)	(4.416)
	[0.1598]	[0.0849]	[0.6114]	[0.3027]	[0.6114]	[0.4316]
Female*Deaths*Urban	0.390	0.882	−30.04	−0.318	16.24**	−1.765
	(1.106)	(0.737)	(16.78)	(1.210)	(7.067)	(5.082)
	[0.8072]	[0.8072]	[0.3926]	[0.9700]	[0.0360]	[0.9700]
Rural hukou	−0.0321***	−0.0143***	1.266***	−0.136***	−0.837***	−0.881***
	(0.0247)	(0.0157)	(0.352)	(0.0269)	(0.127)	(0.115)
	[0.0010]	[0.0010]	[0.0010]	[0.0010]	[0.0010]	[0.0010]
Female	0.00601	−0.0445***	2.230***	−0.191***	−1.282***	−0.134
	(0.0119)	(0.00810)	(0.163)	(0.0162)	(0.112)	(0.0544)
	[0.6244]	[0.0010]	[0.0010]	[0.0010]	[0.0010]	[0.1269]
Constant	0.486**	0.253*	0.790	1.051***	10.45***	5.686***
	(0.193)	(0.145)	(3.486)	(0.251)	(1.468)	(0.963)
Observations	7885	7885	7885	7885	7885	7885
R-squared	0.091	0.149	0.128	0.153	0.192	0.153
Interview year dummy	YES	YES	YES	YES	YES	YES
Year of birth FE	YES	YES	YES	YES	YES	YES
Prefecture FE	YES	YES	YES	YES	YES	YES

Note: Alternative Violence Measure during the CR. The *p*-values in brackets underlying the significance levels are Romano-Wolf stepdown adjusted *p*-values which consider multiple hypothesis testing across different health outcomes. Standard errors are clustered at the prefecture and year of birth levels and reported in parentheses. This table shows the results using the log of number of deaths due to Cultural Revolution as the measure of violence. In Columns (1–2) we examine the impacts of violence (measured by log of number of deaths due to collective conflicts) and famine (measured by excessive death rates) on the predicted 8-year diabetes risk and 10-year CVD risk. Column (3) shows the results using CESD-10 as the measure of depression, and columns (4–6) use three measures of cognition: (i) ability to draw a required picture, (ii) mental intactness, and (iii) memory ability using word recall questions (see the paper for details). *EDR* is the province-level excessive death rates in 1959, 1960 and 1961, which is calculated by subtracting the average death rates between year 1956–1958 from the annual death rates in the three famine years. The variable *Deaths* measures the number of deaths due to CR violence in each year between 1966 and 1971. It is measured at the prefecture level per thousand population. The variable *Rural* is the indicator for rural *hukou* at birth.

*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

male-EDR-rural interaction term in Column (3) does not change at all. The coefficient on the male-conflict-urban interaction term in Column (4) changes from −0.167 to −0.128. Overall, the results are robust to changes in the size of the control group.

Finally, we replace our measure of famine severity with an alternative measure based on cohort size shrinkage rather than on (excessive) death rates, using census data in 1990. This alternative measure was used in other studies (see, e.g., Xu et al. (2017)) and has the advantage that it can be computed at the prefecture level. However, it also has two important disadvantages. First, as it is based on cohort size shrinkage, it is confounded by (reduced) fertility and, most importantly, by migration to other prefectures

TABLE 7 Gender specific: CHARLS 2011 and 2015

	(1)	(2)	(3)	(4)	(5)	(6)
	Diabetes	CVD	Depression	Cognition		
	Predicted 8-year risk	Predicted 10-year risk	CESD-10	Graphical cognition	Mental intactness	Episodic memory
			Score		(TICS)	(Word recall)
Variables	%	%	0–30(Worst)	0–1(Best)	0–10(Best)	0–10(Best)
Male*CSSI*Rural	0.0630	0.0463	2.946	0.0877	0.311	−0.327
	(0.0850)	(0.0736)	(2.897)	(0.107)	(0.631)	(0.412)
	[0.3277]	[0.9540]	[0.8282]	[0.4096]	[0.8282]	[0.7193]
Female*CSSI*Rural	0.0928**	0.102	2.365	−0.0492	0.378	−0.315
	(0.112)	(0.0747)	(1.570)	(0.102)	(0.597)	(0.392)
	[0.0470]	[0.1489]	[0.6873]	[0.2308]	[0.6873]	[0.4126]
Male*Conflicts*Urban	−0.00609	−0.0650**	0.825	−0.167*	−0.437	−0.441
	(0.0532)	(0.0391)	(0.938)	(0.0764)	(0.395)	(0.282)
	[0.8452]	[0.0280]	[0.7473]	[0.0839]	[0.7752]	[0.7752]
Female*Conflicts*Urban	−0.0145	−0.0199	−0.662	−0.000128	0.857**	0.333**
	(0.0435)	(0.0316)	(0.798)	(0.0547)	(0.320)	(0.272)
	[0.9151]	[0.4446]	[0.4685]	[0.4685]	[0.0100]	[0.0390]
Rural hukou	−0.0336***	−0.0218***	1.221***	−0.141***	−0.846***	−0.859***
	(0.0253)	(0.0156)	(0.348)	(0.0289)	(0.131)	(0.126)
	[0.0010]	[0.0010]	[0.0010]	[0.0010]	[0.0010]	[0.0010]
Female	0.00775	−0.0470***	2.237***	−0.189***	−1.287***	−0.140*
	(0.0123)	(0.00839)	(0.168)	(0.0170)	(0.117)	(0.0562)
	[0.5415]	[0.0010]	[0.0010]	[0.0010]	[0.0010]	[0.0949]
Constant	0.492**	0.265*	1.057	1.051***	10.51***	5.661***
	(0.193)	(0.146)	(3.491)	(0.252)	(1.470)	(0.964)
Observations	7885	7885	7885	7885	7885	7885
R-squared	0.091	0.149	0.127	0.153	0.192	0.153
Interview year dummy	YES	YES	YES	YES	YES	YES
Year of birth FE	YES	YES	YES	YES	YES	YES
Prefecture FE	YES	YES	YES	YES	YES	YES

Note: Alternative Famine Measure. The *p*-values in brackets underlying the significance levels are Romano-Wolf stepdown adjusted *p*-values which consider multiple hypothesis testing across different health outcomes. Standard errors are clustered at the prefecture and year of birth levels and reported in parentheses. This table show the gender-specific estimates with the full sample. In Columns (1–2) we examine the impacts of violence (measured by log of number of collective conflicts) and famine (measured by cohort size shrinkage indices) on the predicted 8-year diabetes risk and 10-year CVD risk. Column (3) shows the results using CESD-10 as the measure of depression, and columns (4–6) use three measures of cognition: (i) ability to draw a required picture, (ii) mental intactness, and (iii) memory ability using word recall questions (see the paper for details). *CSSI* (*Cohort Size Shrinkage Indices*) is constructed from the 1% sample of the 1990 China Population Census. This measure is obtained by subtracting the famine cohort (1959–1961) size from the average non-famine cohort (1956–1958 and 1962–1964) size during famine years and divided by the average non-famine cohort size.

*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

that happened in the 30 years between the famine period and 1990. Second and related to the previous point, the measure is based on the prefecture of residence in 1990 and not on the prefecture of birth. Table 7 shows that our results are qualitatively confirmed.

6 | CONCLUSION

We estimate the long-run impact of exposure in utero to the Chinese Great Famine and the Cultural Revolution on physical and mental health later in life. We merge data on the regional violence during the Cultural Revolution and the excessive death rates during the Chinese Great Famine with data from the China Health and Retirement Longitudinal Study (CHARLS) survey.

While the Chinese Great Famine was responsible for severe malnutrition and a stressful environment for pregnant mothers in rural areas, the Cultural Revolution mainly affected pregnant mothers in urban areas.

The results show that *in utero* exposure to the famine increases the risk of getting type-2 diabetes later in life, and that the effect is more pronounced for females. A 1% increase in excessive death rates increases diabetes risk by around 0.06% points for girls who were exposed to famine. Boys exposed to the famine are shown to have a higher tendency to develop depressive symptoms later in life, while exposure to the Cultural Revolution has a negative effect on their cognitive abilities. However, the results on mental health seem to be less robust than those on physical health. As in our data we only observe those who survived into adulthood, our results are likely to underestimate the negative effect of adverse events at birth on health later in life.

Overall, our results suggest that early life intervention is crucial. Future research should focus on disentangling the mechanisms through which adverse conditions around birth translate into poor health outcomes later in life, which is crucial to identify potential targets for prevention and intervention strategies.

ACKNOWLEDGMENTS

We are grateful to the two data providers: the China Health and Retirement Longitudinal Study (CHARLS) team at Peking University and Professor Andrew G. Walder at Stanford University. The CHARLS was supported by the behavioral and Social Research division of the National Institute on Aging of the National Institute of Health (Grant Nos. 1-R21-AG031372-01, 1-R01-AG037031-01, and 3-R01AG037031-03S1), the Natural Science Foundation of China (Grant Nos. Seventy,773,002, 70,910,107,022, and 71,130,002), the World Bank (Contract Nos. Seventy one45915 and 7,159,234), and Peking University. Collection work for the second data set "China Political Events Dataset, 1966–1971" was directed by Professor Andrew G. WAlder, with support from National Science Foundation (Grant SBS-1021134) for his project, "Political Movements in an Authoritarian Hierarchy" (2010–2013).

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support this study are available at: https://charls.charlsdata.com/pages/data/111/en.html.

ORCID

Jiyuan Wang https://orcid.org/0000-0003-4054-1058
Viola Angelini https://orcid.org/0000-0002-5177-3447

ENDNOTES

[1] There are also other mediation pathways to explain the association, but the one with HPA axis is commonly accepted.

[2] Various biological literature tries to find biomarkers to quantify the activity of the HPA axis and the prenatal stress, such as hair cortisone levels (see, e.g., Molenaar et al., 2019) and dehydroepiandrosterone (DHEA) and dehydroepiandrosteronesulfate (DHEAS) (see, e.g., Schmelter et al., 2019).

[3] Not all individuals participated in the biomarker sections. In the non-blood biomarker section, the response rate is around 78.9%, while in the blood biomarker section, the response rate is around 65%.

[4] Provinces of Xizang, Chongqing (province-level municipality) and Hainan are not included in CHARLS.

[5] The difference between famine and post-famine cohorts is statistically significant.

[6] The statistics are provided by World Heart Federation: https://www.world-heart-federation.org.

[7] In our calculations we made use of the user written Stata command rwolf2 developed by Clarke et al. (2020).

REFERENCES

Akbulut-Yuksel, M. (2017). War during childhood: The long run effects of warfare on health. *Journal of Health Economics*, *53*, 117–130. https://doi.org/10.1016/j.jhealeco.2017.02.005

Alessie, R. J., Angelini, V., van den Berg, G. J., Mierau, J. O., & Viluma, L. (2019). Economic conditions at birth and cardiovascular disease risk in adulthood: Evidence from post-1950 cohorts. *Social Science & Medicine*, *224*, 77–84. https://doi.org/10.1016/j.socscimed.2019.02.006

Almond, D., & Currie, J. (2011). Killing me softly: The fetal origins hypothesis. *The Journal of Economic Perspectives*, *25*(3), 153–172. https://doi.org/10.1257/jep.25.3.153

Almond, D., & Mazumder, B. (2011). Health capital and the prenatal environment: The effect of ramadan observance during pregnancy. *American Economic Journal: Applied Economics*, *3*(4), 56–85. https://doi.org/10.1257/app.3.4.56

Bai, L. (2014). Economic legacies of the cultural revolution. *Job Market Paper*. https://scholar.google.com/citations?view_op=view_citation&hl=en&user=x8V3jLoAAAAJ&citation_for_view=x8V3jLoAAAAJ:Y0pCki6q_DkC

Barker, D. J. (1995). Fetal origins of coronary heart disease. *BMJ*, *311*(6998), 171–174. https://doi.org/10.1136/bmj.311.6998.171

Barker, D. J., Bagby, S. P., & Hanson, M. A. (2006). Mechanisms of disease: In utero programming in the pathogenesis of hypertension. *Nature Clinical Practice Nephrology*, *2*(12), 700–707. https://doi.org/10.1038/ncpneph0344

Barker, D. J., Godfrey, K. M., Gluckman, P. D., Harding, J. E., Owens, J. A., & Robinson, J. S. (1993). Fetal nutrition and cardiovascular disease in adult life. *The Lancet*, *341*(8850), 938–941. https://doi.org/10.1016/0140-6736(93)91224-a

Björgvinsson, T., Kertz, S. J., Bigda-Peyton, J. S., McCoy, K. L., & Aderka, I. M. (2013). Psychometric properties of the CES-D-10 in a psychiatric sample. *Assessment*, *20*(4), 429–436. https://doi.org/10.1177/1073191113481998

Bock, J., Wainstock, T., Braun, K., & Segal, M. (2015). Stress in utero: Prenatal programming of brain plasticity and cognition. *Biological Psychiatry*, *78*(5), 315–326. https://doi.org/10.1016/j.biopsych.2015.02.036

Bonnin, M. (2006). The "lost generation": Its definition and its role in today's Chinese elite politics. *Social Research: International Quarterly*, *73*(1), 245–274. https://doi.org/10.1353/sor.2006.0017

Braithwaite, E. C., Murphy, S. E., & Ramchandani, P. G. (2014). Prenatal risk factors for depression: A critical review of the evidence and potential mechanisms. *Journal of Developmental Origins of Health and Disease*, *5*(5), 339–350. https://doi.org/10.1017/s2040174414000324

Bramsen, I., Deeg, D. J., van der Ploeg, E., & Fransman, S. (2007). Wartime stressors and mental health symptoms as predictors of late-life mortality in World War II survivors. *Journal of Affective Disorders*, *103*(1), 121–129. https://doi.org/10.1016/j.jad.2007.01.014

Burns, L. R., & Liu, G. G. (2017). *China's healthcare system and reform*. Cambridge University Press.

Clarke, D., Romano, J. P., & Wolf, M. (2020). The romano–wolf multiple-hypothesis correction in stata. *STATA Journal*, *20*(4), 812–843. https://doi.org/10.1177/1536867x20976314

Conroy, R., Pyörälä, K., Fitzgerald, A. e., Sans, S., Menotti, A., De Backer, G., De Bacquer, D., Ducimetiere, P., Jousilahti, P., Keil, U., et al. (2003). Estimation of ten-year risk of fatal cardiovascular disease in Europe: The SCORE project. *European Heart Journal*, *24*(11), 987–1003. https://doi.org/10.1016/s0195-668x(03)00114-3

D'agostino, R. B., Vasan, R. S., Pencina, M. J., Wolf, P. A., Cobain, M., Massaro, J. M., & Kannel, W. B. (2008). General cardiovascular risk profile for use in primary care. *Circulation*, *117*(6), 743–753. https://doi.org/10.1161/circulationaha.107.699579

Dinkelman, T. (2017). Long-run health repercussions of drought shocks: Evidence from South African homelands. *The Economic Journal*, *127*(604), 1906–1939. https://doi.org/10.1111/ecoj.12361

Doblhammer, G., van den Berg, G. J., & Lumey, L. H. (2013). A re-analysis of the long-term effects on life expectancy of the great Finnish famine of 1866–68. *Population Studies*, *67*(3), 309–322. https://doi.org/10.1080/00324728.2013.809140

Eriksson, J. G., Kajantie, E., Osmond, C., Thornburg, K., & Barker, D. J. (2010). Boys live dangerously in the womb. *American Journal of Human Biology*, *22*(3), 330–335. https://doi.org/10.1002/ajhb.20995

Gade, D. M., & Wenger, J. B. (2011). Combat exposure and mental health: The long-term effects among us vietnam and gulf war veterans. *Health Economics*, *20*(4), 401–416. https://doi.org/10.1002/hec.1594

Grimard, F., & Laszlo, S. (2014). Long-term effects of civil conflict on women's health outcomes in Peru. *World Development*, *54*, 139–155. https://doi.org/10.1016/j.worlddev.2013.08.004

Hu, X. F., Liu, G. G., & Fan, M. (2017). Long-term effects of famine on chronic diseases: Evidence from China's great leap forward famine. *Health Economics*, *26*(7), 922–936. https://doi.org/10.1002/hec.3371

Huang, W., & Zhou, Y. (2013). Effects of education on cognition at older ages: Evidence from China's great famine. *Social Science & Medicine*, *98*, 54–62. https://doi.org/10.1016/j.socscimed.2013.08.021

Islam, A., Raschky, P., & Smyth, R. (2017). The long-term health effects of mass political violence: Evidence from China's cultural revolution. *Social Indicators Research*, *132*(1), 257–272. https://doi.org/10.1007/s11205-015-1030-6

Kim, S., Fleisher, B., & Sun, J. Y. (2017). The long-term health effects of fetal malnutrition: Evidence from the 1959–1961 China great leap forward famine. *Health Economics*, *26*(10), 1264–1277. https://doi.org/10.1002/hec.3397

Kuh, D., & Shlomo, Y. B. (2004). *A life course approach to chronic disease epidemiology*. Oxford University Press. number 2.

Lei, X., Hu, Y., McArdle, J. J., Smith, J. P., & Zhao, Y. (2012). Gender differences in cognition among older adults in China. *Journal of Human Resources*, *47*(4), 951–971. https://doi.org/10.1353/jhr.2012.0033

Lin, J. Y., & Yang, D. T. (2000). Food availability, entitlements and the Chinese famine of 1959–61. *The Economic Journal*, *110*(460), 136–158. https://doi.org/10.1111/1468-0297.00494

Lumey, L. H., Stein, A. D., & Susser, E. (2011). Prenatal famine and adult health. *Annual Review of Public Health*, *32*(1), 237–262. https://doi.org/10.1146/annurev-publhealth-031210-101230

Mao, W., Yip, C.-M. W., & Chen, W. (2019). Complications of diabetes in China: Health system and economic implications. *BMC Public Health*, *19*(1), 1–11. https://doi.org/10.1186/s12889-019-6569-8

Meng, X., & Gregory, R. G. (2002). The impact of interrupted education on subsequent educational attainment: A cost of the Chinese cultural revolution. *Economic Development and Cultural Change*, *50*(4), 935–959. https://doi.org/10.1086/342761

Meng, X., Qian, N., & Yared, P. (2015). The institutional causes of China's great famine, 1959–1961. *The Review of Economic Studies*, *82*(4), 1568–1611. https://doi.org/10.1093/restud/rdv016

Molenaar, N. M., Tiemeier, H., van Rossum, E. F., Hillegers, M., Bockting, C., Hoogendijk, W., van den Akker, E., van den Lambregtse-Berg, M., & El Marroun, H. (2019). Prenatal maternal psychopathology and stress and offspring hpa axis function at 6 years. *Psychoneuroendocrinology*, *99*, 120–127. https://doi.org/10.1016/j.psyneuen.2018.09.003

Mueller, B. R., & Bale, T. L. (2008). Sex-specific programming of offspring emotionality after stress early in pregnancy. *Journal of Neuroscience*, *28*(36), 9055–9065. https://doi.org/10.1523/jneurosci.1424-08.2008

Portrait, F., Teeuwiszen, E., & Deeg, D. (2011). Early life undernutrition and chronic diseases at older ages: The effects of the Dutch famine on cardiovascular diseases and diabetes. *Social Science & Medicine*, *73*(5), 711–718. https://doi.org/10.1016/j.socscimed.2011.04.005

Romano, J. P., & Wolf, M. (2005). Stepwise multiple testing as formalized data snooping. *Econometrica*, *73*(4), 1237–1282. https://doi.org/10.1111/j.1468-0262.2005.00615.x

Schmelter, V., Antonaci, L., Fröhlich, S., Herebian, D., Verde, P., Heil, M., & Hoehn, T. (2019). Quantification of fetal steroids in nails of neonates to quantify prenatal stress and growth restriction. *Biological Psychology*, *140*, 81–85. https://doi.org/10.1016/j.biopsycho.2018.12.001

Selvarajah, S., Kaur, G., Haniff, J., Cheong, K. C., Hiong, T. G., van der Graaf, Y., & Bots, M. L. (2014). Comparison of the Framingham Risk Score, SCORE and WHO/ISH cardiovascular risk prediction models in an Asian population. *International Journal of Cardiology*, *176*(1), 211–218. https://doi.org/10.1016/j.ijcard.2014.07.066

Strauss, K., Dapp, U., Anders, J., von Renteln-Kruse, W., & Schmidt, S. (2011). Range and specificity of war-related trauma to posttraumatic stress; depression and general health perception: Displaced former World War II children in late life. *Journal of Affective Disorders*, *128*(3), 267–276. https://doi.org/10.1016/j.jad.2010.07.009

Teerawichitchainan, B., & Korinek, K. (2012). The long-term impact of war on health and wellbeing in northern Vietnam: Some glimpses from a recent survey. *Social Science & Medicine*, *74*(12), 1995–2004. https://doi.org/10.1016/j.socscimed.2012.01.040

Vaiserman, A., & Lushchak, O. (2019). Prenatal malnutrition-induced epigenetic dysregulation as a risk factor for type 2 diabetes. *International Journal of Genomics*, *2019*, 1–11. https://doi.org/10.1155/2019/3821409

Van den Berg, G. J., Pinger, P. R., & Schoch, J. (2015). Instrumental variable estimation of the causal effect of hunger early in life on health later in life. *The Economic Journal*, *126*(591), 465–506. https://doi.org/10.1111/ecoj.12250

Walder, A. G. (2014). Rebellion and repression in China, 1966–1971. *Social Science History*, *38*(3–4), 513–539. https://doi.org/10.1017/ssh.2015.23

Walder, A. G., & Su, Y. (2003). The cultural revolution in the countryside: Scope, timing and human impact. *The China Quarterly*, *173*, 74–99. https://doi.org/10.1017/s0009443903000068

Wang, J. (2019). Impoverishing or improving? The impact of the Chinese cultural revolution on lifetime income. *Job Market Paper*. https://jiyuanwang.weebly.com/uploads/1/2/0/9/120933858/jmp-wang-nov.pdf

Wilson, P. W., Meigs, J. B., Sullivan, L., Fox, C. S., Nathan, D. M., & D'Agostino, R. B. (2007). Prediction of incident diabetes mellitus in middle-aged adults: The framingham offspring study. *Archives of Internal Medicine*, *167*(10), 1068–1074. https://doi.org/10.1001/archinte.167.10.1068

Xu, H., Zhang, Z., Li, L., & Liu, J. (2017). Early life exposure to China's 1959–61 famine and midlife cognition. *International Journal of Epidemiology*, *47*(1), 109–120. https://doi.org/10.1093/ije/dyx222

Xu, Y., Wang, L., He, J., Bi, Y., Li, M., Wang, T., Wang, L., Jiang, Y., Dai, M., Lu, J., Min, X., Yichong, L., Nan, H., Jianhong, L., Shengquan, M., Chung-Shiuan, C., Guangwei, L., Yiming, M., Jiajun, Z., …, & for the 2010 China Noncommunicable Disease Surveillance Group. (2013). Prevalence and control of diabetes in Chinese adults. *JAMA*, *310*(9), 948–959. https://doi.org/10.1001/jama.2013.168118

How to cite this article: Wang, J., Alessie, R., & Angelini, V. (2023). Exposure *in utero* to adverse events and health late-in-life: Evidence from China. *Health Economics*, *32*(3), 541–557. https://doi.org/10.1002/hec.4632