Economy and health

Viluma, Laura

CHAPTER 3

3. ECONOMIC CONDITIONS AT BIRTH AND CARDIOVASCULAR DISEASE RISK IN ADULTHOOD: EVIDENCE FROM POST-1950 COHORTS[3]

3.1. INTRODUCTION

An expanding body of literature has documented negative effects of adverse economic conditions around the time of birth on a range of late-life mortality and morbidity outcomes (see, for instance, Doblhammer *et al.*, 2013; van den Berg *et al.*, 2006, 2009, 2011, 2013; Lindeboom *et al*., 2010). Insights from various fields have led to the formulation of fetal programming hypothesis – alternatively referred to as the critical period, biological imprint, biological embedding or developmental origins hypothesis – which suggests that certain exposures early in life permanently and irreversibly alter the structure and/or function of organs, tissues and systems (Rasmussen et al., 2001; Kuh and Hardy, 2002; Kuh and Ben-Shlomo, 2004; Case *et al.*, 2005; Wadhwa *et al.*, 2009; Barker, 1995). According to this literature, in-utero deprivation is an important source of, amongst others, cardiovascular disease (CVD) (Barker, 1995). While in his seminal work, Barker (1995) suggests maternal malnutrition as the main mechanism linking economic conditions at birth and health outcomes, other authors also highlight the importance of changes in other health-related consumption and time use (Deheija and Lleras-Muney, 2004), and stress

[3] This chapter is based on Alessie, R., Angelini, V., Mierau, J.O., van den Berg, G., and Viluma, L., 2019. Economic conditions at birth and cardiovascular disease risk in adulthood: Evidence from post-1950 cohorts. *Social Science and Medicine,* (forthcoming).

This study benefitted from comments received at CINCH Academy, 2017, RES Junior Symposium, 2017, SOM PhD conference, 2017, Healthwise congress, 2016, EuHEA conference, 2016, IBEO workshop, 2016, and NIDI/RUG conference, 2015. I thank Tom Wansbeek for his helpful suggestions.

(Alessie *et al*., 2018). In addition, changes in cohort composition may play a role. If low socioeconomic status (SES) families postpone fertility during periods economic hardship more than high SES families, the average health of the individuals that are born might even improve (Deheija and Lleras-Muney, 2004, Alessie *et al*., 2018).

This paper adds to the existing literature by studying the effect of economic conditions at the time of birth on cardiovascular health in adulthood in post-1950 cohorts. To avoid endogeneity problems with the use of measures of economic conditions in the individual household, many studies, including ours, exploit exogenous variation in contextual conditions, such as the state of the business cycle at birth, as an indicator of the economic conditions that the households face (see, for instance, van den Berg *et al.*, 2006, 2009, 2011, 2013, Angelini and Mierau, 2014, Cutler *et al.*, 2016, and the overviews in Almond and Currie, 2011 and van den Berg and Lindeboom, 2014). This literature focuses predominantly on over-all or cause-specific mortality as outcome measures, because mortality data is objective, reliable and readily available. However, since mortality is an end-state phenomenon, these analyses necessarily examine cohorts born before World War II. There are several reasons why the findings from this time frame might not be applicable to the current cohorts.

First, it is possible that the extent to which an economic downturn leads to adverse economic conditions within households may be smaller for cohorts born after World War II. Western European countries have established social safety nets that include relatively generous unemployment benefits and welfare payments for those without work. This could mean that the business cycle has become a weaker indicator for economic shocks in the household.

Second, the nature of the effects of economic conditions might have changed. Since the social safety nets protect households from sudden income losses and, hence, deprivation and malnutrition, other mechanisms become important instead. For instance, an economic downturn may increase stress, fear of job loss, or lower opportunity costs of pregnancy. The consequences of these exposures might be different from the effects of business cycle fluctuations without the social safety nets.

Third, compared to pre-war birth cohorts, in modern cohorts female labor force participation in adulthood tends to be higher. This may alter the effect of business cycle

fluctuations in more recent birth cohorts, through increased stress exposure of the mothers. In addition, the long-run effects on cohorts of daughters may depend on their own exposure to labor market fluctuations later in their life. All this may lead to changes in the size of average long-run effects of business cycles across eras. Moreover, the increase in the female labor participation suggest possible gender differences in these changes in the relationship between the business cycle at birth and adult health in recent cohorts. However, the size and direction of the changes remains an empirical question that we address in this paper.

Another reason to expect different effects among men and women is the well-documented gender differences in effects of the conditions early in life. Biological evidence documents gender-differences in fetal sensitivity (Catalano *et al.*, 2005) which may be driven by different intra-uterine growth strategies between male and female fetuses (see, Eriksson *et al.*, 2010). Focusing on the effects of the business cycle, studies with cohorts born a long time ago tend to find smaller over-all effects among women than among men, at least in terms of mortality outcomes (see the above references; some studies focus mainly or solely on men). However, Yeung *et al.* (2014) report that effects on CVD mortality are stronger among women than among men.

In sum, the current paper contributes to the literature by studying the gender-specific effects of the state of the business cycle at birth on cardiovascular disease (CVD) risk in adulthood in post-World War II birth cohorts. Since CVD is the leading cause of death in Europe and around the world (Nichols *et al.,* 2014), it is essential to understand the various causes of CVD over the human lifecycle. Moreover, the early-life conditions literature since its inception has shown that early-life conditions are an important determinant of CVD (see the above references). Indeed, the fetal programming hypothesis (Barker, 1995) was formulated specifically in terms of CVD. As discussed above, since our focus is on cohorts born after World War II, the ordinarily used mortality outcomes are not suitable for our analysis. Even actual CVD is not very common among individuals from these cohorts, as they are still relatively young. To proceed, we operationalize CVD *risk,* using biomarker data. Biomarkers enable the observation of a quantitative outcome variable at ages when mortality and morbidity cannot yet be observed. The fact that reliable and consistent biomarkers exist for predicting CVD is an additional advantage of

using CVD risk as an outcome. Notice, however, that the literature on the developmental origins of late-life health puts CVD in the same range of outcomes as type-2 diabetes, mental health and cognitive impairments (see again the above references and, e.g., Lumey *et al.*, 2011), suggesting that CVD biomarkers are also informative on the risks of those health outcomes.

The use of biomarker data has an additional advantage. Policy interventions aimed at preventing high-age CVD have to rely on predictors of CVD such as the biomarkers we examine. If a relation with economic conditions early in life exists then the collection of the relevant biomarker data among adults can be targeted to those born under adverse conditions.

The biomarker data we use are collected as part of the Lifelines – a large-scale study covering over 75,000 individuals born between 1950 and 1992 (see Scholtens *et al.*, 2015). The Lifelines baseline interviews and biomarker data collection, on which we base our analysis, was performed in the northern part of the Netherlands between 2006 and 2013. The raw variables (such as blood cholesterol levels and systotic blood pressure) are used to compute the so-called Systematic COronary Risk Evaluation (SCORE) index. The latter was developed and validated by the European Society of Cardiology to estimate gender- and age-specific 10-year absolute risks of a fatal CVD event (see, Conroy *et al.,* 2003, for a detailed description).

We use spatial and temporal fluctuations in the unemployment rate (taken from Statistics Netherlands) as exogenous indicators of economic conditions early in life. One advantage of this measure, as opposed to GDP, is that it is available at the level of a province for our full observation window. At the time, the Netherlands had 11 provinces, and most individuals in the sample were born in the three that constitute the area from which the original sample was drawn in 2013. Combining the biomarker-based SCORE index with provincial level unemployment data enables us to analyse the relationship between adverse conditions at birth and CVD risk later in life.

The remainder of the paper is set up as follows. Section 2 describes the data and section 3 outlines our empirical strategy. Section 4 presents and discusses our results and the final section concludes.

3.2. DATA

For the purpose of our analysis we combine individual health data from Lifelines with data on regional unemployment from Statistics Netherlands. In what follows we describe each in turn and explain the construction of our key variables of interest.

3.2.1. LIFELINES

Lifelines is a large population-based cohort study and biobank carried out in the northern part of the Netherlands that was established as a resource for research on complex interactions between environment, phenotypic and genomic factors in the development of chronic diseases and healthy ageing. The study adopted three recruitment strategies: recruitment of an index population, aged 25 to 49, via general practitioners (GP), subsequent inclusion of their family members, and online self-registration. Patients who were unable to read Dutch or who had limited life expectancy due to severe illness were excluded by the GP and not invited for participation. The participants visited one of the Lifelines research sites for a physical examination. Also, blood and urine samples were collected, and participants completed extensive questionnaires. The baseline data were collected for 167,729 participants, aged from 6 months to 93 years. Controlling for differences in the demographic composition, the Lifelines adult study population is broadly representative for the adult population of the northern Netherlands (see Klijs *et al.*, 2015, and Scholtens *et al.*, 2015, for a detailed description of the study).

Furthermore, since we are working with retrospective data to discover the effect of the conditions at the time of birth on the current health of the participants, we must consider two potential sources of bias. First, we are concerned with survivorship bias because we only observe the individuals who have survived till the moment of data collection and are in good enough health to participate. The biological programming framework predicts that the effect of adverse early life conditions on current health is negative. If those who are the most affected, do not survive or are unable to participate, our estimates may be biased towards zero. However, the mortality effects of adverse early life conditions have been shown to present only late in life – after the age of 65 (van den Berg et al., 2011). Since our sample contains only participants younger than 65, survivorship bias seem unlikely. Another concern is the possibility of selective abortion – the individuals

that have been affected the most by adverse circumstances might not even be born. Such effects, if present, could also cause bias towards zero in our estimates. However, in the context of the Netherlands after 1950, spontaneous abortions due to fluctuations in the business cycle seem implausible. To sum up, survivorship bias and selective abortion are unlikely issues in our dataset, but in case they are present, our estimates could be considered as conservative or as the lower bound of the true effect.

For the purposes of this paper, Lifelines supplies us with a sample of 95,422 individuals born between 1950 and 1992.[4] For our analysis, we select only individuals born in the Netherlands, which reduces the sample size to 80,820. Further, we exclude any observations that do not contain all of the information necessary to calculate our main outcome variable (the CVD risk SCORE) such as smoking status, age, gender, total cholesterol and blood pressure. Our final sample consists of 76,566 individuals.

By linking the Lifelines data to birth certificate data from the Municipal Personal Records Database (in Dutch: *Gemeentelijk Basisadministratie*), we obtain information on

Table 3.1: Sample size per province

Province	Frequency	Percent
Friesland	27,864	36.39
Groningen	22,112	28.88
Drenthe	13,757	17.97
Zuid-Holland	2,955	3.86
Noord-Holland	2,788	3.64
Overijssel	2,684	3.51
Gelderland	1,775	2.32
Utrecht	1,045	1.36
Noord-Brabant	899	1.17
Limburg	350	0.46
Zeeland	182	0.24
Flevoland	155	0.2
Total	76,566	100

[4] Lifelines also contains respondents born before 1950 and after 1992. For the former, however, we do not have regional unemployment data, while the latter were administered a different survey due to their age. We are using the Lifelines Baseline sample for ages 18 to 63 (release 201303, made available in 2014).

the province of birth of each sample member. Owing to the high quality of both data sources, no observations are lost in the matching process. While Lifelines contains individuals born all over the Netherlands, births from the three northern provinces are naturally overrepresented. Table 3.1 presents the sample sizes per province.

For our purpose, the most important feature of the Lifelines cohort study is that it includes biomarkers concerning cardiovascular disease (CVD) risk, which we can use to construct the Systematic COronary Risk Evaluation (SCORE) index. The SCORE index was developed and validated by the European Society of Cardiology to estimate the 10-year risk of a fatal CVD event. The European Society of Cardiology recommends using SCORE system for assessing individuals' CVD risk in clinical practice in European countries (Pieopoli et al., 2016). In contrast to an actual fatal CVD event (e.g., a major heart attack), the advantage of a surrogate endpoint such as the SCORE index is that we are able to consider relatively younger individuals who are more representative of current cohorts.

The SCORE index is constructed according to an algorithm (see the Appendix and Conroy *et al.*, 2003) that uses age, gender, smoking status, blood cholesterol levels and blood pressure as inputs to estimate the 10-year absolute risk of a fatal CVD event. We can draw all the constituent parts of the SCORE index from the Lifelines data and are, therefore, able to associate the CVD risk to each individual in our sample. A detailed overview of the CVD risk and in its distribution is contained in Table 3.2 as well as Figure 3.1. Both the table and the figure highlight that CVD risk exhibits a strong gender specific pattern. Moreover, we note that CVD risk is strongly skewed, with most individuals displaying limited CVD risk – an issue that we address in the analysis of our results.

In addition, to tentatively assess *selection* effects we exploit the fact that Lifelines contains some socioeconomic indicators of the family into which the individual was born, albeit the information is limited. In this respect we use the age of the respondent's mother when he/she was born and whether or not she was smoking during the pregnancy. Both indicators have been shown to have a strong relationship with a family's socioeconomic status[5] – with families from lower socioeconomic groups tending to have children earlier and being more likely to smoke (Cutler and Glaeser, 2005; Cutler and Lleras-Muney, 2006).

[5] Naturally, the impact of these indicators need not only run through the socioeconomic status of the parents but may also have direct effects in their own right. Smoking during pregnancy in particular has been associated to a plethora of adverse health outcomes for babies.

The summary statistics of these and our other variables of interest are provided in Table 3.3.

Table 3.2: Detailed descriptive statistics for CVD risk score (%)

	All	Males	Females
Percentiles			
1%	$1.07*10^{-4}$	0.003	$7.15*10^{-5}$
25%	0.035	0.376	0.015
50%	0.174	0.478	0.085
75%	0.582	1.075	0.268
99%	3.604	4.765	1.908
Observations	76566	31411	45155
Mean	0.470	0.809	0.235
Std. Dev.	0.778	1.025	0.399
Skewness	3.995	3.073	3.758
Kurtosis	30.29	19.13	24.70

Figure 3.1: Histograms of absolute 10-year risk of fatal cardiovascular disease event for men and women.

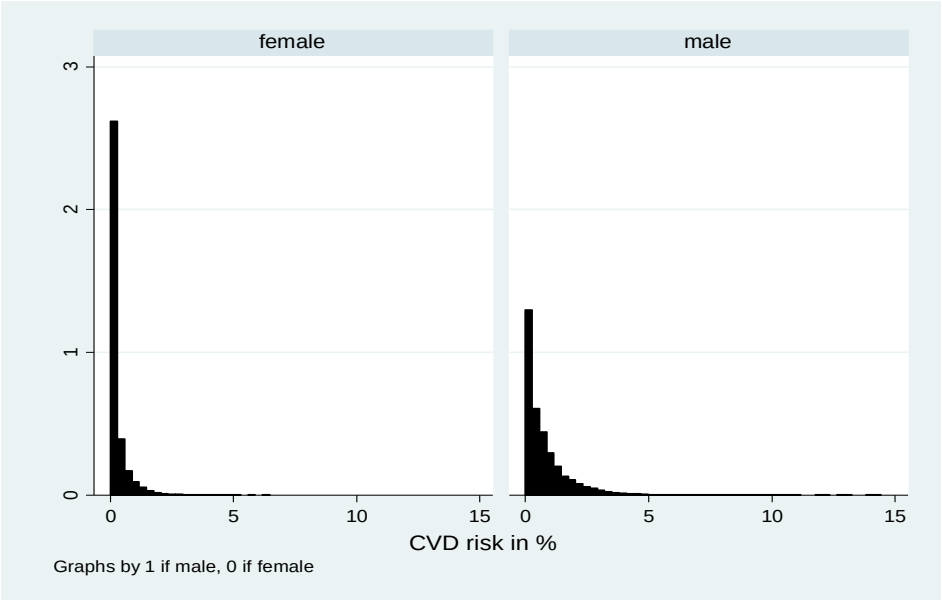


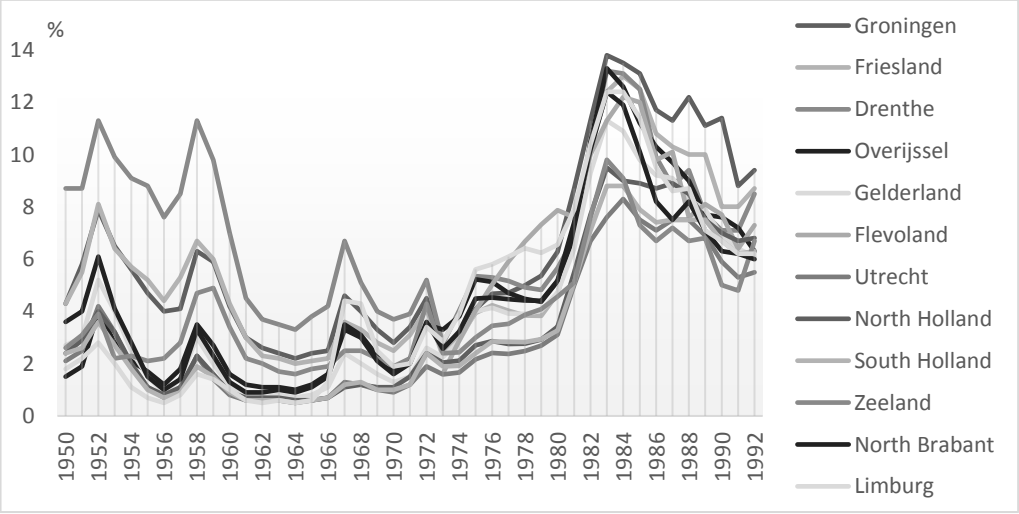
Table 3.3: Descriptive statistics

Variable	Observations	Mean	Std. Dev.	Min	Max
Provincial unemployment rate (%)	76566	4.620	3.031	0.500	13.80
Male	76566	0.410	0.492	0	1
Age at the first visit	76566	42.42	9.526	20.00	63.04
Birth year	76566	1968	9.562	1950	1992
CVD risk score (%)	76566	0.470	0.778	0	14.12

3.2.2. UNEMPLOYMENT DATA

We proxy the early life conditions by using provincial unemployment data from Statistics Netherlands. The unemployment rate provides us with a contextual variable that serves as a proxy of the socioeconomic conditions under which the individual was born without suffering from the endogeneity of individual level socioeconomic indicators. Provincial level unemployment data are available from 1950 onward, which creates the lower bound for the birth year in the data. We display the development of unemployment between 1950 and 1992 in Figure 3.2.

Figure 3.2: Provincial unemployment level in the Netherlands, 1950-1992



Source: Statistics Netherlands (www.cbs.nl)

During our sample period, the Netherlands went through all phases of the business cycle multiple times. After World War II, the Netherlands enjoyed a period of substantial economic growth with low associated unemployment rates. At the end of the 1970s and for much of the early 1980s, the Netherlands were hit by a strong recession due to the second oil crisis. This recession was particularly strong in the northern Netherlands where unemployment peaked at well over 10% at the height of the recession. In the early 1990s alongside the world-wide economic boom, unemployment rates dropped significantly all over the country. While the data display a clear common trend among the provinces, we also observe substantial differences in both levels and trends of unemployment between the provinces. Implying that province-level unemployment data provides us with additional variation from which to identify our relationship of interest.

3.3. METHODS

Our main interest lies in the relationship between unemployment in the birth year and CVD risk later in life. To this end, we start with a simple linear specification with CVD risk as the outcome variable and the provincial unemployment level in the birth year as main explanatory variable. To account for the structure of the data, we also include a birth year and a province fixed effect. We allow the effect of unemployment at birth to be gender specific to account for potential gender differences in the relationship between early-life conditions and health later in life.

Beyond a direct impact on the SCORE index, age can act as a risk factor or as modifier of the other risk factors – the levels of cholesterol and blood pressure increase with age and increase earlier in life in men than in women. Since we already account for birth year fixed effects in the model, there is a risk of multicollinearity if we include also age in the model. However, since the Lifelines data were gathered over a 7-year period (2006-2013), sample members with identical ages can have a variety of different birth years.

Since it is well known in the econometrics that the effects of age on health are highly non-linear, it is important to allow for this non-linearity in the model. Commonly it is done by including a polynomial of age or dummy variables for age categories. However, the coefficients of a polynomial are difficult to interpret while categorizing age omits a lot of available information. An approach that does not have these problems is using a series of

linear splines, joined at "knot points", to model the effects of age. We control for age in the model by including linear splines with knots at 30, 40, 50 and 60 years. This means that we allow different linear slopes for age from 20 to 30 years, 30 to 40, 40 to 50, 50 to 60 and beyond 60 years of age.

To sum up, we estimate the following specification:

$$CVD_{ipt} = \alpha + \beta_1 u_{pt} * m_i + \beta_2 u_{pt} * f_i + \beta_3 m_i + \sum_{k=1}^{K} \beta_4 s_{ipt} + \theta_t + \rho_p + \varepsilon_{ipt} \qquad (3.1)$$

where CVD_{ipt} denotes ten-year fatal CVD event risk for individual i born in province p in year t; u_{pt} is the unemployment rate in province p and birth year t, m_i is a dummy variable taking value 1 if male and 0 if female, while f_i takes the opposite value; s_{ipt} stands for a linear spline with K knots in age of the individual i born in province p and year t; ρ_p is a province fixed effect, and θ_t is a year-of-birth fixed effect. We estimate the specification in (1) by ordinary least squares (OLS).

Since we are interested in the gender specific effects of unemployment at birth, we have included two interaction terms in the specification – one that interacts unemployment level with a dummy that takes value 1 if the individual is male and 0 if female and one that in interacts unemployment level with a dummy that takes value 1 if the individual is female and 0 if male. Accordingly, to avoid multicollinearity, we do not include the "main" effect of unemployment in the model. This approach allows us to interpret our main parameters of interest, β_1 and β_2, directly. β_1 represents the marginal effect of unemployment at birth in males while β_2 represents the effect in females.

Since our main explanatory variable – the unemployment level – is measured at the province level, the error terms are likely to be correlated within provinces. Therefore, cluster-robust standard errors are required for statistical inference. The province fixed effects partially control for the within-province correlation, but perhaps not completely. In addition, since there are only 12 provinces in the Netherlands, the number of clusters is small (due to the small sample sizes, the three provinces with smallest number of observations – Limburg, Flevoland and Zeeland – were grouped together so effectively we have 10 clusters) which means that the estimated variance matrix of the OLS estimator is likely to be downwards biased (Cameron and Miller, 2015). Therefore, we apply the bias-correction of Bell and McCaffrey (2002), which was named CR2VE by Cameron and Miller (2015, p.342 and p. 346), to the standard cluster-robust variance estimates. CR2VE

correction implies scaling the province specific vector of residuals $\hat{u}_p$ so that $\tilde{u}_p = \left(I_{Np} - H_{pp}\right)^{-0.5} \hat{u}_p$, where $H_{pp} = X_p(X'X)^{-1}X'_p$ with X_p being an $N_p \times K$ matrix, where K is the number of variables included in the model, and N_p is the size of the sample in province p. Stacking X_p over P provinces yields X. In addition, since the number of observations varies considerably across provinces, the effective number of clusters is reduced even further (Imbens and Kolesar, 2016). To address the unbalanced cluster sizes, we base the Wald tests on a $t(\nu^*)$-distribution where the degrees of freedom $\nu^*$are determined by the data as proposed by Imbens and Kolesar (2016). According to a Monte Carlo study performed by Cameron and Miller (2015), the null hypothesis is rejected too often if we use the "standard formula" for cluster robust standard errors when the number of clusters is small. However, the use of CR2VE residual and of the data-determined degrees of freedom leads to a considerable improvement in inference. That is, the actual size of the t-test (the probability of Type I error given the sample size) is close to the nominal size of the test (the desired significance level α).

3.4. RESULTS

We present our main estimation results in Table 3.4. Column 1 contains the OLS estimate of the model specified in Equation (3.1) with the cluster robust standard errors determined as outlined above. The results suggest that while the impact of unemployment on women's CVD risk is significant at the 1% level – even after taking into account the CR2VE standard errors – the impact on men does not differ significantly from zero.[6] More precisely, for women a 1 percentage point increase in the provincial unemployment level leads to a 0.02 percentage point increase in the risk of experiencing a fatal CVD event in the coming 10 years. While this effect may seem relatively small, comparing it to the effect of ageing indicates that, for instance, for a 45-year-old woman born when the unemployment rate was elevated by 1 percentage point, the CVD risk is equivalent to that of an identical woman who is 6 months older but born when the unemployment rate was not elevated.

[6] Regardless of the early-life conditions, men do have a higher CVD risk than women, as is reflected by the substantial point estimate of the gender dummy.

Table 3.4: CVD risk and unemployment level at birth by gender.

	(1) SCORE CVD death risk %	(2) SCORE CVD death risk %
female x unemp	0.021***	0.021***
	0.002	*0.002*
male x unemp	-0.006	-0.006
	0.013	*0.013*
Male	0.679***	0.681***
	0.058	*0.057*
Age of mother at birth		$-3.5*10^{-4}$
		0.001
Mother smoked during pregnancy		0.034***
		0.006
Linear spline in age	*YES*	*YES*
Province fixed effects	*YES*	*YES*
Birth year fixed effects	*YES*	*YES*
Observations	*76,566*	*76,566*
IK degrees of freedom:		
female x unemp	6.8	6.8
male x unemp	6.9	6.9
Test male*unemp - female*unemp	***	***

Note: OLS regression results. CR2VE standard errors clustered at the province level are reported in italics under the coefficients. The Imbens-Kolesar degrees of freedom used in the t-tests for the key variables are reported at the bottom of the table (*** signifies p<0.01, ** signifies p<0.05, and * signifies p<0.1).

In addition to the biological programming, the impact of early-life conditions on later life outcomes may be due to selection effects. While Lifelines does not include information about the individual socioeconomic conditions into which a child was born, we do have knowledge of the age of the respondent's mother when he/she was born and whether or not she was smoking. Both indicators have been shown to have a strong relationship with a family's socioeconomic status (SES) – with family's from lower socioeconomic groups tending to have children earlier and being more likely to smoke. Even though these variables cannot control for all the variation in parental SES, if parental

SES explains the relationship between unemployment level at birth and CVD risk, we will observe a change in the coefficients of unemployment when we include these available parental characteristics in the model. Therefore, we add the two indicators to our main specification and report the estimation results in column (2) of Table 4. The results indicate that while age of the mother is not associated with elevated CVD risk, being born to a mother who smoked is strongly associated to heightened CVD risk later in life. However, after accounting for these two variables, the magnitude of the impact of unemployment at birth on CVD risk later in life is essentially unaffected. This result leads us to believe that selection effects do not rationalize our results. This is also in line with other literature on long-term health effects of conditions at birth that finds that the composition of new-borns does not vary systematically over the business cycle (see e.g. Kåreholt, 2001; van den Berg *et al*. 2009 and 2011, and an overview in Van den Berg and Modin, 2013).

We have performed a range of sensitivity analyses to verify whether the results are robust with respect to aspects of the model specification (results available upon request). In general, the key results are not qualitatively affected by modifications of the specification; that is, sign and significance remain preserved. For example, the results are robust to adding other biomarkers from blood samples to the SCORE index (such as glucose levels and inflammation markers). Results by subintervals of the birth-year window suggest that the effects for females are somewhat stronger for more recent birth cohorts within our observation window. This points to the importance of female labor force participation which started to increase at the national level in the Netherlands only after 1970. Whether this is connected to increased stress exposure is a topic for further research.

One exception to the general robustness is found when birth-year fixed effects in the model specification are replaced by a second-degree polynomial in the year of birth. Apparently, the polynomial is not able to capture the effects of the major restructuring of the economy in the early 1980s, or the combination of a low-degree polynomial and an age spline is not sufficiently flexible to fit secular time and age patterns in the data. Also, low-degree polynomials may be less suitable than birth-year fixed effects if the operationalization of the definition of unemployment changes over calendar time or if there are institutional changes in the ease with which transitions into out-of-the-labor-force states such as disability and early retirement can be made. As a final sensitivity

analysis, we estimated models for actual CVD occurrence. As expected, due to the low occurrence of CVD, none of the estimated effects is significantly different from zero. The signs of the effects are in line with our results.

3.5. CONCLUSIONS

We find that women exposed to unfavorable business-cycle conditions at birth are at an increased risk for fatal CVD events in adult life, among relatively young cohorts of women born after World War II. We interpret this as evidence that unfavorable conditions in the household at birth cause an elevated CVD risk in adult women. The fact that studies using data from much earlier birth cohorts did not unambiguously find strong evidence for such an effect among women may reflect a gradual increase in the size of this causal effect over the past century. This is supported by the finding that the effect is stronger among females born later in our observation window.

For men, CVD risk, on average, is unaffected by early-life exposure to recessions. As explained in the paper, this does not necessarily entail that causal effects of adverse economic conditions at the individual level are absent. Instead, the business cycle might not be capturing fluctuations in economic conditions well, due to improving social safety nets over the 20th century.

There are several possible explanations for the gradual increase in the causal effect in women, and the gender differences in general. First, this gradual increase in women coincides with the increase in female labor force participation, so it is easy to perceive that these two trends might be related and that the effects of exposure in utero might be intensified by the stress of labor participation later in life. Second, compared to pre-war cohorts, smoking and obesity has increased among women from post-war cohorts. There is a consensus in the medical literature that at younger ages (under 50) smoking and type 2 diabetes increases CVD risk in women significantly more than in men (Maas and Appelman, 2010). These risk factors might also magnify the effects of the early life conditions leading to larger and increasing effects in women under the age of 64. However, whether any of these explanations are true remains a topic for further research.

To conclude, relatively recent birth cohorts are potentially more representative of current and future cohorts. And the usage of biomarkers allows us to detect elevated health

risks well before health events occur. Taken together, this means that the results point at increased risks of actual CVD in the near future for women born when unemployment was high.

APPENDIX: CALCULATING 10-YEAR RISK ESTIMATES FOR FATAL CARDIOVASCULAR DISEASE (SCORE)

(Conroy et al., 2003)

Step 1: Calculate the underlying risks for coronary heart disease and for non-coronary cardiovascular disease separately for the person's age now and for their age in ten years' time, using the values for α and p shown in Table A1. The underlying survival probability, S_0, is given by:

$S_0(age) = exp\ \{ - (exp(\alpha))(age-20)^p\}$

$S_0(age+10) = exp\ \{ - (exp(\alpha))(age-10)^p\}$

Table A1: Coefficients for Step 1

		CHD		NON-CHD CVD	
		α	p	α	p
LOW RISK	Men	−22.1	4.71	−26.7	5.64
	Women	−29.8	6.36	−31.0	6.62
HIGH RISK	Men	−21.0	4.62	−25.7	5.47
	Women	−28.7	6.23	−30.0	6.42

Step 2: Using the coefficients in Table A2, calculate the weighted sum, w, of the risk factors cholesterol, smoking and systolic blood pressure. Two weighted sums will have to be calculated, one for coronary heart disease and one for non-coronary cardiovascular disease. Smoking is coded as 1 for current and 0 for non-smoker, so no value for smoking has to be entered if the person is a non-smoker. Cholesterol is measured in mmol/L and SBP is measured in mmHg. The weighting for each risk factor is denoted by β.

$w = \beta_{chol}(cholesterol-6) + \beta_{SBP}(SBP-120) + \beta_{smoker}(current)$

Table A2: Coefficients for Step 2

	CHD	NON-CHD CVD
CURRENT SMOKER	0.71	0.63
CHOLESTEROL (MMOL/L)	0.24	0.02
SYSTOLIC BP (MMHG)	0.018	018 0.022

Step 3: Combine the underlying risks for coronary heart disease and for non-coronary cardiovascular disease, at the person's age and at their age ten years from now (four calculations) which were calculated in step 1 with the weighted sum of a person's risk factors from step 2 for the two end-points, coronary heart disease and non-coronary cardiovascular disease, to get the probability of survival at each age for each cause.

$S(age) = \{S_0(age)\}^{exp(w)}$

$S(age+10) = \{S_0(age+10)\}^{exp(w)}$

Step 4: For each cause, calculate the 10-year survival probability based on the survival probability for the person's current age and their age in 10 years' time:

$S_{10}(age) = S(age+10) / S(age)$

Step 5: Calculate the 10 year risk for each end-point as

$Risk_{10} = 1 - S_{10}(age)$

Step 6: Combine the risks for coronary heart disease and non-coronary cardiovascular disease by adding them:

$CVDRisk_{10}(age) = [CHDRisk(age)] + [Non\text{-}CHDRisk(age)]$