

NBER WORKING PAPER SERIES

HEALTH AND ECONOMIC ACTIVITY OVER THE LIFECYCLE:
LITERATURE REVIEW

Daniel Prinz
Michael Chernew
David Cutler
Austin Frakt

Working Paper 24865
<http://www.nber.org/papers/w24865>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2018

The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at <http://www.nber.org/papers/w24865.ack>

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2018 by Daniel Prinz, Michael Chernew, David Cutler, and Austin Frakt. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Health and Economic Activity Over the Lifecycle: Literature Review
Daniel Prinz, Michael Chernew, David Cutler, and Austin Frakt
NBER Working Paper No. 24865
July 2018
JEL No. I1,I15

ABSTRACT

We systematically review the literature linking health to economic activity, particularly education and labor market outcomes, over the lifecycle. In the first part, we review studies that link childhood health to later-life outcomes. The main themes we focus on are in-utero exposures, birthweight, physical health and nutrition, mental health, and the environment. In the second part, we review studies of the impact of health on labor market success for adults. The main themes we focus on are the environment, disability, physical health shocks, within-household spillovers, cancer, and mental health.

Daniel Prinz
National Bureau of Economic Research
1050 Massachusetts Avenue
Cambridge, MA 02138
dprinz@g.harvard.edu

Michael Chernew
Harvard Medical School
Dept. of Health Care Policy
180 Longwood Avenue
Boston, MA 02115
and NBER
chernew@hcp.med.harvard.edu

David Cutler
Department of Economics
Harvard University
1875 Cambridge Street
Cambridge, MA 02138
and NBER
dcutler@harvard.edu

Austin Frakt
VA Boston
afrakt@gmail.com

Contents

1	Introduction	5
2	Childhood health	7
2.1	Introduction	7
2.2	Fetal origins: In-utero exposure	8
2.2.1	Introduction	8
2.2.2	Summary of methods	13
2.2.3	Studies: The 1918 influenza pandemic	16
2.2.4	Summary: The 1918 influenza pandemic	20
2.2.5	Studies: Other in-utero exposures	21
2.2.6	Summary: Other in-utero exposures	29
2.3	Fetal origins: Birthweight	29
2.3.1	Introduction	29
2.3.2	Summary of methods	33
2.3.3	Studies	34
2.3.4	Summary	38
2.4	Child health and nutrition	39
2.4.1	Introduction	39
2.4.2	Summary of methods	40
2.4.3	Studies: Historical data	43
2.4.4	Summary: Historical data	45
2.4.5	Studies: Modern data	45
2.4.6	Summary: Modern data	48
2.5	Child mental health	49
2.5.1	Introduction	49
2.5.2	Summary of methods	50

2.5.3	Studies	54
2.5.4	Summary	59
2.6	Environmental health	60
2.6.1	Introduction	60
2.6.2	Summary of methods	61
2.6.3	Studies	63
2.6.4	Summary	66
2.7	Summary	67
3	Adult health	68
3.1	Introduction	68
3.2	Environmental health and productivity	70
3.2.1	Introduction	70
3.2.2	Summary of methods	71
3.2.3	Studies	73
3.2.4	Summary	75
3.3	Disability	75
3.3.1	Introduction	75
3.3.2	Summary of methods	76
3.3.3	Studies	79
3.3.4	Summary	80
3.4	Health shocks	80
3.4.1	Introduction	80
3.4.2	Summary of methods	81
3.4.3	Studies	85
3.4.4	Summary	89
3.5	Within-household spillovers	90
3.5.1	Introduction	90

3.5.2	Summary of methods	91
3.5.3	Studies	93
3.5.4	Summary	96
3.6	Cancer	96
3.6.1	Introduction	96
3.6.2	Summary of methods	97
3.6.3	Studies	99
3.6.4	Summary	101
3.7	Mental health	102
3.7.1	Introduction	102
3.7.2	Summary of methods	102
3.7.3	Studies	105
3.7.4	Summary	108
3.8	Summary	109
	References	110

List of Tables

1	Fetal origins: In-utero exposure - Summary	10
2	Fetal origins: In-utero exposure - Summary	11
3	Fetal origins: In-utero exposure - Summary	12
4	Fetal origins: Birthweight - Summary	31
5	Fetal origins: Birthweight - Summary	32
6	Child health and nutrition - Summary	41
7	Child health and nutrition - Summary	42
8	Child mental health - Summary	52
9	Child mental health - Summary	53
10	Environmental health - Summary	62
11	Environmental health and productivity - Summary	72
12	Disability - Summary	78
13	Health shocks - Summary	83
14	Health shocks - Summary	84
15	Within-household spillovers - Summary	92
16	Cancer - Summary	98
17	Mental health - Summary	104

1 Introduction

A voluminous literature studies the relationship between health at various points in life and economic outcomes. The goal of this literature review is to summarize what is known about the impact of health on economic activity throughout the life course. We consider the life course in two parts. During the first part of life, childhood, health capital and general human capital are built through investment, possibly affected by health shocks. During the second part of life, adulthood, people enter the labor market. Their productivity is influenced by the health capital and general human capital built during childhood, depreciates over time, and is also subject to shocks.

In the childhood section of this review, we examine how children's health influences educational outcomes, which in turn likely influence labor market outcomes in adulthood. We also examine how children's health may be connected with labor market outcomes through other channels. Labor market outcomes, including employment and earnings, are natural endpoints to be interested in, however, the majority of studies is only able to examine educational outcomes, because the examination of labor market outcomes requires long-term follow up and more data. In the adult section, we examine how health is connected with labor market outcomes, including employment and earnings.

To understand the scope of this review, it may be useful to consider the types of studies that we are not including. We do not include studies that consider how economic circumstances impact health. We also omit studies that have health as their outcome variable, including studies of how in utero and childhood health shocks influence health in childhood or in adulthood. Instead we focus only on studies with economic outcomes, like education and earnings. Our focus is the United States and other developed countries, many of which have better data access for high-quality empirical studies, thus we omit studies of developing countries.

Each subsection of the children and adults sections presents a systematic review of a particular literature. The systematic review was done by searching the economics literature

using Google Scholar, specifying appropriate search terms for each subsection (e.g., “in utero” OR “fetal” OR “prenatal” for Section 2.2 on the impact of in utero health), excluding studies that were not relevant. In addition to the Google Scholar search, we added a small number of studies that were referenced by the studies we found through the systematic search but did not themselves come up in the search (usually because they did not contain the keywords in their title). While we reference a few classic papers in each section, we limit our review in most sections to 2005-2017, and indicate review articles that summarize earlier work where available. Mindful of the slow publication process and the fast pace of some of the subliterations covered in our review, we also included relevant NBER Working Papers from 2015-2017 from the Children, Health, and Labor Studies programs. More specific details on the systematic review are given at the beginning of each subsection.

We structure each of the two sections, on children and adults, similarly. We give a short overview of the broad topics of childhood and adult health and discuss the most recent comprehensive reviews. Then we review the subliterations on different types of health. In the childhood section, these are in-utero exposures, birthweight, physical health and nutrition, mental health, and the environment. In the adult section, these are the environment, disability, physical health shocks, within-household spillovers, cancer, and mental health. Each of the subsections on different types of health is structured similarly. After a brief introduction, which discusses previous important reviews where available and the search terms used in the systematic review, we first summarize the methods that are common among the studies in the subsection (but typically quite different across subsections), and then go through the studies. We also provide summary tables for each subsection and conclude each section with a summary discussion, which includes directions for future work.

2 Childhood health

2.1 Introduction

This section focuses on the impact of health on the educational attainment of children and the long-term consequences of childhood health. While a voluminous literature shows that later-life health is strongly related to childhood health, we prioritize studies that also establish the importance of childhood health on outcomes other than later-life health. These “non-health endpoints” are cognitive and noncognitive ability, educational attainment, and longer-term labor market outcomes. We also limit our attention to studies that were done in the United States and other developed economies, some of which have much better access to administrative data, allowing researchers to connect data on health to data on measures of educational attainment and economic outcomes.

Two relatively recent general reviews on the topic are [Currie \(2009\)](#) and [Almond and Currie \(2011a\)](#). These reviews are updated in [Almond, Currie and Duque \(2017\)](#). [Currie \(2009\)](#) summarizes the literature that connects parental socioeconomic status (SES) to children’s health and children’s health to education and future income. In this review, we are interested in the latter of these connections, the relationship between childhood health, education, and future income, but understanding the former, the relationship between parental SES and childhood health is helpful to conceptualize the intergenerational transmission of economic status. The conclusion of [Currie \(2009\)](#) is that child health has a measurable impact on future educational and labor market outcomes, although the exact magnitude of this connection has not been clearly established by the literature. [Almond and Currie \(2011a\)](#) is a more general review of the literature on human capital development in early childhood. The authors discuss in detail the literatures on the long-term consequences of both the prenatal environment and the early childhood environment. They conclude that the literature establishes that prenatal and early childhood health matter for longer-term outcomes at the population level and that in most, though not all cases, prenatal impacts are

larger than postnatal impacts. They single out mental health conditions and non-cognitive skills as examples health conditions that become important in early childhood after birth. [Almond, Currie and Duque \(2017\)](#) summarize the developments of the fetal origins and early childhood literatures since their earlier reviews. They argue that we can identify some interventions during the in utero and early life periods that can significantly improve long-term outcomes, but there is still much to learn about heterogeneous effects and the magnitude of the effect of various shocks and interventions. Moreover, they emphasize that little is known about what they term the “missing middle years”, that is, early and middle childhood proxies for longer-term effects.

The section on children considers six main themes in the literature: (1) the fetal origins literature that studies various in-utero exposures, (2) the related fetal origins literature that studies the impact of low birth weight on longer-term outcomes, (3) the literature that studies the longer-term impact of child physical health and nutrition, (4) the literature on the longer-term impact of child mental health, and finally (5) the literature on environmental health. Each of the thematic subsections starts with a discussion of the common methodologies that have been used in the particular literature and then proceeds to discuss the most important studies.

2.2 Fetal origins: In-utero exposure

2.2.1 Introduction

While David Barker was not the first one to work on this topic, the fetal origins hypothesis, the idea that fetal conditions are persistent, latent, and may act through “programming”, is sometimes called Barker’s hypothesis. [Barker and Osmond \(1986\)](#), [Barker \(1990\)](#), and [Barker \(1995\)](#) present empirical work on the association between fetal conditions and later life health. These studies were followed by a large literature in economics that established the longer-term health impact of fetal origins. Here we focus on contributions that examine non-health endpoints, such as education and earnings. An optimistic review of this literature is

presented in [Almond and Currie \(2011b\)](#) and its update, [Almond, Currie and Duque \(2017\)](#). Given the large number of studies that have been written on a single historical pandemic, the Spanish Flu of 1918, after discussing the common methodologies in the in-utero exposure literature, we first discuss papers studying the 1918 influenza pandemic, and then separately discuss papers studying other exposures.

This subsection is based on a systematic review which started with searching the economics literature on Google Scholar using the search terms "in utero" OR "fetal" OR "prenatal" for years 2005-2017. Our search returned an initial 164 results, however we only consider 25 studies that are relevant to our review. These need to consider the impact of health on economic outcomes, rather than the other way around, examine non-health endpoints (education or economic outcomes), and be done in the US or other developed countries. In addition to the Google Scholar search, we added a small number of studies that were referenced by the studies we found through the systematic search but did not themselves come up in the search (usually because they did not contain the keywords in their title). Mindful of the slow publication process and the fast pace of some of the literature covered here, we also included relevant NBER Working Papers from 2015-2017 from the Children, Health, and Labor Studies programs.

Table 1: Fetal origins: In-utero exposure - Summary

Study	Exposure	Area	Time period	Outcomes	Data	Approach	Findings
Almond and Mazumder (2005)	Influenza	US	1918	Health	SIPP	Cohort	Individuals who were in utero at the peak of the influenza pandemic have worse health; e.g., 16 percent more likely to rate their health as fair or poor
Almond (2006)	Influenza	US	1918	Education SES	IPUMS	Cohort Geographic	Individuals who were in utero at the peak of the influenza pandemic have worse socioeconomic outcomes; e.g., 4-5% less likely to complete high school, earn about \$1,000 less per year, 8% more likely to be disabled, 8% more likely to be on welfare
Bengtsson and Helgertz (2015)	Influenza	Sweden	1918	Health SES	Administrative	Cohort Geographic	Large negative effects for health; Weak positive effects for earnings
Karlsson, Nilsson and Pichler (2014)	Influenza	Sweden	1918	Health SES	Administrative	Geographic	No detectable earnings effect
Neelsen and Stratmann (2012)	Influenza	Switzerland	1918	Education SES	Census	Cohort Geographic	Some negative impact on education and socioeconomic status
Lin and Liu (2014)	Influenza	Taiwan	1918	Education	Census	Cohort Geographic	Large negative impact on education
Richter and Robling (2016)	Influenza	Sweden	1918	Education	Administrative	Cohort Geographic	for women, 3-4 months less schooling and 3-5 percentage points lower college attendance if mother in utero during pandemic; for men 4-7 months lower schooling and 7-11 percentage points lower college attendance if father in utero during pandemic; no mother-son and father-daughter transmission
Richter and Mazumder (2016)	Influenza	Sweden	1918	Education	Administrative	Cohort Geographic	suggestive evidence that second generation effect through direct health channel positive and effect through indirect household environment channel negative
Kelly (2011)	Influenza	UK	1957	Health Education	NCDS	Geographic	negative effect on physical health of children born to less healthy mothers; negative effects on test scores across the SES distribution; birthweight does not capture impact
Almond, Edlund and Palme (2009)	Chernobyl	Sweden	1986	Health Education	Administrative	Cohort Geographic Family FE	substantial negative impact on test scores and probability of high school graduation (correspond to up to 3% income loss); not mediated through birthweight or health outcomes

Table 2: Fetal origins: In-utero exposure - Summary

Study	Exposure	Area	Time period	Outcomes	Data	Approach	Findings
Black, Bütikofer, Devereux and Salvanes (2013)	Nuclear testing	Norway	1956-1966	Health Education SES	Administrative	Cohort Geographic Family FE	Large impacts on IQ at age 18, years of education, high school completion, and earnings
Sanders (2012)	Air pollution	Texas	1980s	Education	School test scores	Cohort Geographic	TSP in the year of birth is negatively correlated with high school test performance
Bharadwaj, Gibson, Zivin and Neilson (2016)	Air pollution	Chile	1992-2001	Education	Administrative	Cohort Geographic Family FE	Significant negative correlation between in utero carbon monoxide exposure and fourth grade math and language test results
Persico, Figlio and Roth (2016)	Environmental toxicants	Florida	1994-2002	Education	Administrative	Cohort Geographic Family FE	children conceived before the cleanup within 2 miles of Superfund sites were 7.4 percentage points (40%) more likely to repeat a grade compared; 6.6 percentage points more likely to be suspended from school; 10 percentage points more likely to be diagnosed with a cognitive disability.
Almond and Mazumder (2011)	Ramadan	Michigan	1989-2006	Birthweight	Birth certificates	Cohort	20 gram negative effect on birthweight; reduction in the number of male births
Almond, Mazumder and van Ewijk (2015)	Ramadan	England	1998-2007	Education	Administrative	Cohort	overlap of Ramadan with pregnancy is associated with a 0.05-0.08 standard deviation reduction in test scores at age 7
Greve, Schultz-Nielsen and Tekin (2015)	Ramadan	Denmark	1985-1995	Education	Administrative	Cohort	Large negative effects on the test scores of low-SES students
Neelsen and Stratmann (2011)	Famine	Greece	1941-1942	Education SES	Census	Cohort Geographic Timing	Educational attainment significantly reduced, by as much as 3 months for those exposed during infancy, corresponding to 3.3 percentage point drop in upper secondary school completion
Lavy, Schlosser and Shany (2016)	Micronutrient deficiency	Israel	1991	Health Education	Administrative	Timing	Students already in Israel in the first trimester 12 percentage points (50%) more likely to obtain a high school diploma; 12 percentage points less likely to repeat a grade; 7 percentage points less likely to drop out of high school; No differences in birthweight.
von Hinke Kessler Scholder, Wehby, Lewis and Zuccolo (2014)	Alcohol	England	1991-1992	Education	ALSPAC		Strong negative association between alcohol consumption and schooling outcomes

Table 3: Fetal origins: In-utero exposure - Summary

Study	Exposure	Area	Time period	Outcomes	Data	Approach	Findings
Nilsson (2016)	Alcohol	Sweden	1964-1972	SES	Administrative	Cohort Geographic	Exposed cohorts earn 24% less; 8.3 percentage point increase in share of individuals with no labor income; Schooling reduced on by 0.32 years; share of individuals in the bottom third of the cognitive ability score distribution increases by 27% (16% for non-cognitive ability)
Aizer, Stroud and Buka (2015)	Maternal stress	Massachusetts Rhode Island	1959-1965	Education	NCPP	Family FE	Only very high levels of cortisol are predictive of lower IQ and fewer years of education; no significant impact on birth outcomes
Black, Devereux and Salvanes (2016)	Maternal stress	Norway	1960-2010	Education SES	Administrative	Timing	Negative but small impact on birth outcomes; no evidence of persistent impacts on education completed, employment, earnings, or cognitive ability
Persson and Rossin-Slater (2016)	Maternal stress	Sweden	1973-2011	Mental health	Administrative	Timing	Large impacts on birth outcomes and later-life takeup of ADHD, anti-anxiety, and depression medications
Lee (2014)	War	South Korea	1950-1953	Education SES	Census	Cohort Geographic	Significantly lower educational attainment; Worse employment outcomes

2.2.2 Summary of methods

The main identification challenge in the in utero exposure literature is that in general it is not plausible to assume that children who experience a worse in utero environment are otherwise similar to children who experience a better in utero health environment. For example, if we wanted to estimate the impact of maternal smoking during pregnancy and simply regressed children's outcomes on a measure of maternal smoking behavior, omitted variable bias would likely make the results impossible to interpret as causal estimates. The reason for this is that we know that smoking is correlated with a variety of other maternal characteristics, for example socioeconomic status that are in turn correlated with children's outcomes. Omitted variable bias will be a concern even if we control for observable maternal characteristics because unobservable characteristics correlated with smoking behavior can play a large role in determining children's outcomes as well. Given this basic identification challenge, the literature on in utero exposures has primarily relied on four research designs to arrive at causal estimates: cohort discontinuities in exposure, geographic variation in exposure, variation in the timing of exposure, and where possible family fixed effect specifications. In this section we discuss in broad terms identification under each of these research designs, including the underlying assumptions and possible violations.

Cohort discontinuities in exposure arise when an exposure can be relatively precisely timed, allowing researchers to compare people in utero during the exposure to people in utero before and after. Because many outcomes (e.g, educational attainment) can have strong secular time trends, it is important to have data on cohorts born both before and after the exposure in question. A prime example of the use of a cohort discontinuity design is [Almond \(2006\)](#) which estimates the long-term impact of in utero exposure to the 1918 influenza pandemic. Using data on people who were in utero in each year from 1912 to 1922, the author is able to compare the 1919 cohort (in utero in 1918) to the previous and later cohorts. In this context, secular time trends are important, because for example each successive cohort got more years of education independent of the exposure under study.

Therefore estimation is done relative to these time trends. The main assumption required for the identification of causal effects is that different cohorts are in fact comparable, in particular, that the exposure does not coincide with changes in observable and unobservable characteristics correlated with the outcomes. For example as discussed below, the potential contemporaneous impact of World War I on the socioeconomic characteristics of families giving birth in 1919 is a challenge to [Almond \(2006\)](#).

Some studies rely on geographic variation in exposure either in their primary or in their secondary analysis. Instead of comparing people born at different times, they compare people born at the same time but in different areas of a country which had different levels of exposure. An example of a strong study that relies on geographic (as well as temporal) variation in exposure is [Black, Bütikofer, Devereux and Salvanes \(2013\)](#) which exploits variation across municipalities and months in low-level radioactive exposure in Norway. The main assumption required for the identification of causal effects is that people born in different regions are in fact comparable, at least conditional on observables. This would for example be violated if people knew about the potential exposure and moved around the country accordingly, and the people who chose to relocate were different, for example, because they care more about their children in unobservable ways, than the people who chose to stay and be exposed. The assumption is most plausible if the exposure was sudden and unexpected (as was arguably the case with nuclear fallout from Chernobyl reaching Sweden in 1986, studied in [Almond, Edlund and Palme, 2009](#)) or not known to people (as was arguably the case with nuclear fallout from Soviet testing reaching Norway in the 1950s and 1960s, studied in [Black, Bütikofer, Devereux and Salvanes, 2013](#)).

Some of the studies below contrast the outcomes of people who experienced an exposure at one point in their life with the outcomes of people who experienced the exposure at a different point in their life. In [Lavy, Schlosser and Shany \(2016\)](#), different trimesters of pregnancies are contrasted, in [Black, Devereux and Salvanes \(2016\)](#) and [Persson and Rossin-Slater \(2016\)](#) bereavement events in the mothers family during versus before or after

the pregnancy are contrasted. The idea behind this strategy is that by using only people who experienced the exposure it is able to eliminate concerns about selection into the exposure, while isolating the effect of the timing of the exposure. For example, in the context of the in utero micronutrient deficiency among Ethiopian Jews immigrating to Israel during a military operation studied by [Lavy, Schlosser and Shany \(2016\)](#), all individuals in the data experience the deficiency but in different trimesters, so the identifying assumption is that the trimester of the pregnancy is random relative to the date of the military operation studied. Similarly, in the two maternal stress studies ([Black, Devereux and Salvanes, 2016](#); [Persson and Rossin-Slater, 2016](#)) all mothers experience the bereavement event. In this context, the identifying assumption is whether the child is in utero or has already been born is random relative to the bereavement event causing the variation in stress.

Some more recent studies that make use of comprehensive administrative data are able to connect parents to children and siblings to each other to control for family fixed effects. The idea is that family background has a large impact on birth and later-life outcomes and is likely correlated with various exposures. Moreover, family background can be thought to subsume a variety of unobserved characteristics which are hard to control for. In a family fixed effects specification, the variation exploited is within families across pregnancies, increasing the credibility of estimates. For example, [Black, Bütikofer, Devereux and Salvanes \(2013\)](#) only relies on siblings born at different times to the same mother and therefore subject to the same family inputs but different levels of exposure. Besides the large data requirements, the need to connect parents, children, and siblings and throw away most of the data where children don't have siblings, the main challenge is the possibility of change over time in family inputs related to the exposure. If in fact family inputs change contemporaneously with and in a way related to the exposure, the identifying assumption of constant family background is no longer valid, moreover omitted variable bias can be worse because the impact of any violation is increased by reliance on within-family variation. Such problems might for example plague [Aizer, Stroud and Buka \(2015\)](#) which relies on variation in cortisol

levels across pregnancies of a woman to measure the impact of maternal stress on children's outcomes. Since the variation in cortisol levels is connected to changes in socioeconomic status and therefore likely to changes in related unobservable characteristics, it is unclear whether it is in fact the impact of stress that is identified.

2.2.3 Studies: The 1918 influenza pandemic

A large number of studies investigate the long-term impact of in utero exposure to the 1918 influenza pandemic, often called the Spanish Flu. The pandemic infected 500 million people all over the world, on all continents. Below, we discuss the papers that originated this wave of literature by studying long-term impacts in the US and then turn to critiques of the original studies as well as research attempting to extend the findings to European and Asian countries.

[Almond and Mazumder \(2005\)](#) used data from the Survey of Income and Program Participation (SIPP) to show that cohorts in utero during the 1918 influenza pandemic had worse health outcomes decades later. Relying on SIPP data from 1984 to 1996, this study examines the long-term outcomes of approximately 25,000 individuals of ages 60 to 82 at the time of the survey, born in the U.S. in the 1915-1923 period. The authors estimate linear regressions of health measures on indicators of quarter (or in other specifications month) of birth (4th quarter of 1918, 1st quarter of 1919, 2nd quarter of 1919, etc.) and control for survey year, age, and season of birth effects. The key result of the study is that individuals born in the second quarter of 1919, who were in utero at the peak of the influenza pandemic, fare significantly worse for a number of health measures than individuals born just before or just after. For example, they are 16 percent more likely to rate their health as fair or poor (as opposed to excellent, very good, or good).

In perhaps the most famous article in this literature, [Almond \(2006\)](#) presents complementary evidence using census data, establishing the negative consequences of exposure to the 1918 influenza pandemic for educational attainment, disability, income, and socioeconomic

status. Making use of the Integrated Public Use Microdata Series (IPUMS), a 1% (for 1960), 3% (for 1970), and 5% (for 1980) sample of microdata from the U.S. census, the author studies the outcomes in 1960, 1970, and 1980 of individuals born in 1919, and therefore exposed to the pandemic, relative to trends for individuals born between 1912 and 1922. Most results are reported for white men. The key regression framework in the paper relates educational and economic outcomes to trends over time (parametrized as a linear and quadratic year of birth effect) allowing for a deviation for those born in 1919. He finds that the 1919 birth cohort received on average 1.5 months less education and was 4-5% less likely to complete high school relative to the trend. Because of course not all mothers contracted influenza during the pandemic, these average effects are larger, an estimated 13-15%, for those who did contract the disease. The incomes of the 1919 birth cohort were also lower according to the census data. They made \$850-\$1,100 (updating the author's figures to 2016 dollars) less per year than the surrounding cohorts. The estimate for the wage effect is \$3,100 for sons of infected mothers. Men are 6% more likely to have work-limiting and 8% more likely to have work-preventing disability (17% and 25%, respectively, for those born to infected mothers). They are also 8% more likely to be on welfare. Though not discussed at length, most effects that the author estimates are smaller for women and non-white individuals than for white men who comprise his main analytic sample. There is also some discussion and use of a secondary method that relies on geographic variation in exposure. A key limitation to be noted in translating the "reduced-form" estimates of the impact of exposure to the pandemic into the "treated" estimates of the impact of being born to a mother who actually contracted influenza is that the author relies on an estimated one-third infection rate rather than microdata connecting sons with their mothers. Some of the recent studies we discuss below that use administrative data from Nordic countries were able to solve this issue, actually connecting parents with their children.

[Almond \(2006\)](#) has been influential in spurring a large literature that has aimed to establish the impact of the pandemic in other countries. It has also drawn some criticism

that questions the validity of the identifying assumptions in the original study.

In a recent critique, [Brown and Thomas \(2016\)](#) argue that the families whose children were in utero during the 1918 influenza pandemic were systematically lower SES and that controlling for SES there is little evidence that in utero exposure has an impact on longer-term outcomes. The key insight of this critique is that the estimates in [Almond \(2006\)](#) are unbiased causal estimates of in utero exposure to the pandemic if the 1919 cohort is “statistically exchangeable” with the surrounding cohorts. In other words, the worse educational and labor market outcomes that are found by the original study can be attributed to the impact of the pandemic if there are no characteristics that are either unobserved or observed but not controlled for in the original regressions and are correlated with both the outcomes (educational attainment, labor market success) and the explanatory variable of interest (exposure to the pandemic). The authors argue that this assumption does not hold and the original results suffer from substantial omitted variable bias or lack of exchangeability. In particular, they show, using data from the 1920 and 1930 censuses that the fathers of men born in 1919 were much less likely to be literate and in general had worse socioeconomic backgrounds. They first show that individuals born in 1919 and found by the original study to be poorer in adulthood had poorer fathers. In a second exercise they also demonstrate that controlling for parental background, there is no evidence that these individuals are in fact poorer as adults and they might in fact be better off. In their discussion, the authors cite World War I as a possible reason for difference between the distribution of paternal characteristics of the 1919 cohort relative to the surrounding cohorts. The two particular reasons related to the war that they mention is that wealthier families might have differentially postponed childbearing because of the war and wealthier men were also more likely to be deployed during the war. In fact they show that that WWI veteran status is a “sufficient statistic” for paternal SES in the regressions: when an indicator for father’s veteran status is included in the regressions, the original results, that in utero exposure to the pandemic had a negative effect on adult outcomes, either disappear or are reversed.

Examining rich administrative data from Sweden, [Bengtsson and Helgertz \(2015\)](#) conclude that while effects on later-life health are large and negative, income is weakly and positively affected by exposure to the pandemic. The authors are able to study the universe of individuals living in Sweden at any time between 1968 and 2012 using data that includes longitudinal information on hospitalizations, educational attainment, and income. They are also able to exploit month-by-county level variation in influenza exposure and rule out changes in paternal SES around the time of the pandemic, the essence of the [Brown and Thomas \(2016\)](#) critique. They both argue that Sweden did not actively participate in the war and show using aggregate data that paternal SES is not related to birth quarter around the pandemic. While the authors find significant morbidity effects (the affected cohorts are hospitalized at an approximately 3% higher rate), long-term mortality effects were small in Sweden. The earnings of cohorts exposed in utero to influenza were not lower, and there is some indication that they were higher, than those of the surrounding cohorts. The affected cohorts earn about 2% more and may be more likely to be in high SES occupations. Providing complementary evidence using geographic variation in exposure rates, [Karlsson, Nilsson and Pichler \(2014\)](#) find that there was no detectable effect on earnings in Sweden. [Neelsen and Stratmann \(2012\)](#) rely on 1970 census data from Switzerland and as a secondary analysis exploit not only variation across time but regional variation in influenza severity. They uncover negative impacts on education and socioeconomic status. [Lin and Liu \(2014\)](#) rely on 1980 census data from Taiwan for educational outcomes and also explore regional variation to some extent. They find large negative effects on education.

In a recent pair of studies, [Richter and Robling \(2016\)](#) and [Richter and Mazumder \(2016\)](#) investigate whether the 1918-1919 influenza pandemic had multigenerational impacts on socioeconomic outcomes. That is, they look at what happens to individuals whose parents were in utero during the pandemic. Using administrative data from Sweden [Richter and Robling \(2016\)](#) establishes that for women, schooling was 3-4 months lower and college attendance was reduced by 3-5 percentage points if their mothers were in utero during the

pandemic, while for men schooling was 4-7 months lower and college attendance was reduced by 7-11 percentage points if their fathers were in utero during the pandemic. Adjusting for the proportion of people who were actually infected during the pandemic, these estimates may be as high as 12-16 months for women and 16-28 months for men. Interestingly, they find that mothers don't transmit the negative impact to their sons and fathers don't transmit the negative impact to their daughters. In [Richter and Mazumder \(2016\)](#), the authors look at a sample of adoptees in the same Swedish administrative data in order to disentangle direct (health) and indirect (household environment) mechanisms in the intergenerational transmission of the impacts of the influenza pandemic. While they are underpowered to draw firm conclusions, they find suggestive evidence that the second generation effect through the direct health channel may be positive and the effect through the indirect household environment channel may be negative.

2.2.4 Summary: The 1918 influenza pandemic

Even for a particular exposure, the 1918 Spanish Flu, estimates of long-term impact can vary a lot across different contexts. The literature has presented numbers from large, 4-5% estimates of the impact of exposure to the pandemic on educational outcomes in [Almond \(2006\)](#) to no impact in studies using administrative data in Sweden ([Bengtsson and Helgertz, 2015](#); [Karlsson, Nilsson and Pichler, 2014](#)). It unclear is whether the later studies deliver estimates so much smaller than the early studies because they rely on more detailed administrative data or because the context is different, and European welfare states were better able to compensate for the similar underlying negative impacts of the influenza pandemic. It also remains an open question whether we can generalize about the importance of in-utero health from one historic influenza pandemic to current circumstances. The following section which examines other in-utero exposures suggests that this may be the case.

2.2.5 Studies: Other in-utero exposures

While the 1918 influenza pandemic has received the most attention in the literature, there are studies that were written using similar methods but focused on in-utero exposure to other events. These exposures include a 1957 influenza pandemic, the 1986 Chernobyl accident, nuclear weapon testing, fasting during Ramadan, and most recently maternal stress and alcohol consumption during pregnancy.

Using outcome data from the National Child Development Study (NCDS), [Kelly \(2011\)](#) shows that in utero exposure to the 1957 Asian influenza pandemic in the UK had negative effects on the cognitive development of children. The NCDS is an ongoing cohort study that follows all, approximately 17,000, British children born between March 3 and March 9, 1958. This means that the study is able to rely on very detailed data on the socioeconomic background and development of the children included. It also means that all children included in the data were potentially exposed to the Asian flu, which hit the UK between September and November 1957, in utero, making the cohort discontinuity design that is presented as the primary analysis in most of this literature impossible to use in this case. Instead, the author relies on geographic variation across local authorities in the incidence of influenza, a research design that is presented as the secondary analysis in much of the preceding literature. The analysis essentially regresses the outcomes of interest (birthweight, test scores at ages 7 and 11) on the number of pneumonia notifications per 100,000 in a local authority (a measure of influenza exposure) and local authority-level and child-level control variables. The author concludes that exposure to influenza had a negative effects on the physical health of children born to less healthy mothers and had negative effects on test scores across the SES distribution. The paper also discusses the fact that birthweight does not capture the impact of influenza exposure on cognitive development which suggests that the birthweight literature discussed below should be treated separately from the in-utero literature that relies on various exposures.

[Almond, Edlund and Palme \(2009\)](#) study the negative impact of in-utero exposure to

fallout from the 1986 Chernobyl nuclear accident on school outcomes in Sweden. The authors rely on administrative data for the universe of Swedish people born between 1983 and 1988. They study birth, hospitalization, and schooling outcomes only as in 2009 it was too early to study labor market outcomes for individuals born in 1986. Because they rely on high-quality administrative data that allows them to connect parents to children and siblings with each other, the authors are able to control for family fixed effects. In essence, they compare children within the same family, one of whom was in utero during the nuclear accident, while the others were not. This may be a major improvement relative to the previous literature that relied on cohort discontinuities and/or geographical variation only and that were possibly confounded by unobserved socioeconomic differences across families having births at different times. The primary finding is a substantial negative impact on test scores and the probability of high school graduation (projecting forward these correspond to up to 3% income loss) which is not mediated through birthweight or health outcomes (measured in hospitalization data) which are not affected by in utero exposure to the fallout. The authors argue that the impact then must come through latent cognitive ability. Similarly to other studies in the in-utero literature, their secondary analysis considers regional variation in the fallout and finds evidence consistent with the primary findings: the magnitude of test score drops is larger in more exposed areas. Considering heterogeneity in effects, the authors find that the negative impact on human capital is concentrated among low-SES families.

Also studying exposure to nuclear fallout, [Black, Bütikofer, Devereux and Salvanes \(2013\)](#) use variation across time and municipalities in Norway, created by the location and timing of Soviet nuclear weapon testing in the 1950s and 1960s. They find declines in cognitive ability, educational attainment, and earnings associated with exposure to low-dose nuclear radiation. Similarly to other studies from Nordic countries, the authors are able to examine rich administrative data, which means that they are able to control for a variety of parental characteristics and connect parents with their children and children with their siblings. Given that nuclear testing was done for many years, rather than being a discreet event

like Chernobyl or the 1918 influenza pandemic, the authors cannot rely on a simple cohort discontinuity design. However, the substantial heterogeneity in exposure across municipalities and across time allows them to exploit geographic heterogeneity interacted with time (measured at the month level). The ability to connect families in the data means that they can include family fixed effects, comparing siblings born at various times, corresponding to different exposure levels. As discussed previously, this can be important in alleviating concerns that families with different observed and unobserved characteristics have children at different times and places in a way that is correlated with exposure and later-life outcomes. This paper uncovers large impacts on IQ at age 18, years of education, high school completion, and earnings. Because of the richness of the data, the authors were able to match another generation of people in addition to those directly exposed in utero and their parents. Remarkably, the effects on IQ at age 18 persist, with a 60% transmission rate, in a second generation, that is among the children of people who were in utero in the 1950s and 1960s.

[Sanders \(2012\)](#) studies the impact of in utero exposure to air pollution. He exploits geographic and temporal variation in air pollution generated by a wave of industrial recession that hit Texas in the early 1980s. The author implements an instrumental variable (IV) strategy, using annual employment changes at the county level in manufacturing employment as the instrument for total suspended particulate matter (TSPs). In the first stage, he shows the strong relationship between the instrument and pollution and then examines the impact of pollution on educational outcomes, in particular, high school test performance. The key is to link the TSP in a student's year and county of birth to her test scores. (Since the test score data does not contain information on students' place of birth, the author assumes that they are taking the test in the county where they were born.) The main finding of this study is that TSP in the year of birth is negatively correlated with high school test performance.

[Bharadwaj, Gibson, Zivin and Neilson \(2016\)](#) use similar methods but studies the impact of in utero exposure to air pollution in Chile and have better identification because they are able to control for family fixed effects. They connect data on national standardized test

scores at the individual level to a dataset on births that allow the authors to match siblings through a unique mother identifier. Knowing the birth date of each student, they match on environmental quality at the municipality level for the time when the student was in utero. They uncover a significant negative correlation between in utero carbon monoxide exposure and fourth grade math and language test results.

Using detailed data on siblings born before, during, and after the cleanup in areas near toxic waste sites in Florida, [Persico, Figlio and Roth \(2016\)](#) study the impact of in utero exposure to toxic waste. Between 1993 and 2002, the period under study in this paper, the Environmental Protection Agency (EPA) cleaned up a large number of so-called Superfund sites that released neurotoxicants into the air, water, and ground in Florida. The authors exploit within-family variation in the timing of the births of sibling pairs. Because they have data on every child born in Florida between 1994 and 2002 and educated in Florida in public schools between 1996-2012, they are able to trim their data a great deal. In particular, of the 1.6 million children in the data, they only study the 13,000 families that live within five miles of the sites and who are exposed to a change in environmental toxicity in the sense that they have one child before and one child after the cleanup. While findings on birth outcomes are mixed, the authors find a strong relationship between environmental toxicity and cognitive development, as measured by test scores, grade repetition, behavioral incidents, and cognitive disability. Their main estimates suggest that children conceived before the cleanup within 2 miles of the Superfund sites were 7.4 percentage points (40%) more likely to repeat a grade compared to their own siblings born after the cleanup. They are 6.6 percentage points more likely to be suspended from school and 10 percentage points more likely to be diagnosed with a cognitive disability.

[Almond and Mazumder \(2011\)](#), [Almond, Mazumder and van Ewijk \(2015\)](#) and [Greve, Schultz-Nielsen and Tekin \(2015\)](#) describe the negative impact of fasting during Ramadan while pregnant, in the United States, England, and Denmark, respectively. Both papers exploit the fact that Ramadan, the ninth month of the Islamic (Hijri) calendar moves by 10

to 12 days each year in the Gregorian calendar. For example, in 2016 Ramadan started on June 6, but in 2017 it will start on or around May 27, and in 2018 it will start on or around May 16. In the years 2006-2008, it started in September, in the years 1996-1997, it started in January. [Almond and Mazumder \(2011\)](#) use approximately 50,000 birth records for mothers of Arab ancestry living in Michigan for their period 1989-2006. The authors also account for the fact that a quarter of Michigan Arabs are Christians. Perhaps more importantly, they account for the variation in hours of exposure to fasting generated by the fact that Muslims are expected to fast during daylight hours, the length of which varies (from 9 to 15) across the year. They also control for a variety of parental characteristics. They uncover an approximately 20 gram negative effect on birthweights and a reduction in the number of male births. Turning to educational outcomes, [Almond, Mazumder and van Ewijk \(2015\)](#) use data from a register of students of Bangladeshi and Pakistani origin, and find that an overlap of Ramadan with pregnancy is associated with a 0.05-0.08 reduction in test scores at age 7. Similarly, [Greve, Schultz-Nielsen and Tekin \(2015\)](#) find large negative effects on the test scores of low-SES Muslim students in Denmark, using administrative data for the 1985-1995 period.

In a study of the impact of experiencing famine in utero, in infancy, or during early childhood, [Neelsen and Stratmann \(2011\)](#) examine the long-term educational and SES outcomes of individuals who were born in 1940-1942 and were therefore exposed to the 1941-1942 Greek famine. Using individual-level census data from 1971, 1981, 1991, and 2001, they compare the outcomes of the 1940-1942 birth cohorts to the overall trends in the 1936-1946 cohorts. They find that in urban areas, which were more exposed to the famine than rural areas, educational attainment is significantly reduced, by as much as 3 months for those exposed during infancy, corresponding to a 3.3 percentage point drop in upper secondary school completion.

Studying a more recent period, [Lavy, Schlosser and Shany \(2016\)](#) use the variation in in utero micronutrient (iodine, iron, and folic acid) availability created by the immigration of

Ethiopian Jews to Israel to find large impacts on educational achievement. In May 1991, the government of Israel airlifted 14,000 Jews from Ethiopia to Israel. The authors study this event, often referred to as “Operation Solomon” as a discontinuous change in the micronutrient and healthcare environment for infants in utero at the time of the airlift. Because the airlift was covert action by the Israel Defense Forces and kept secret by military censorship, selection concerns are arguably not too important in this context. To isolate the impact of the change on in utero development, they compare pregnancies in the first and second trimester with pregnancies in the third trimester at the time. The authors follow all 600 children born after the airlift in Israel but incepted in Ethiopia in healthcare and educational data. They find that exposure to better environmental conditions earlier in a pregnancy lead to significantly better educational outcomes, higher high school graduation rates and better test scores, especially among girls. Students who were already in Israel in the first trimester are 12 percentage points (50%) more likely to obtain a high school Baccalaureate diploma than students who arrived to Israel in the third trimester, and they are also 12 percentage points less likely to repeat a grade and 7 percentage points less likely to drop out of high school. Similarly to other studies, they find no differences in birthweight.

Using a genetic instrument, [von Hinke Kessler Scholder, Wehby, Lewis and Zuccolo \(2014\)](#) studies the educational consequences of in utero exposure to alcohol. Exploiting variation in the presence of two different genotypes of an alcohol metabolism gene, the author finds in the Avon Longitudinal Study of Parents and Children (ALSPAC) that among a large sample of children born in the Avon region of England, there was a strong negative association between alcohol consumption and schooling outcomes.

Using administrative data from Sweden, [Nilsson \(2016\)](#) finds that prenatal exposure to alcohol consumption damages cognitive and non-cognitive ability, educational attainment, and labor market outcomes; effects are mostly through prenatal health. Using administrative data to follow all children born between 1964 and 1972 in Sweden, the author exploits a policy that increased the availability of strong beer products both in general and to under

21 people in particular in two regions between November 1967 and July 1968. The author first confirms that the policy change did indeed lead to large increases in alcohol sales in the affected regions and then goes on to estimate its impact using a triple differences (difference-in-difference-in-differences) strategy, comparing children of under/above 21 mothers, regions affected/not affected by the policy change, before/after the policy change. He finds in 1990-2004 administrative data that adults exposed in utero to maternal alcohol consumption had substantially worse labor market outcomes, mediated by cognitive and non-cognitive ability, and education. Effects are stronger for men than for women. On average, exposed cohorts earn 24% less, and there is an 8.3 percentage point increase in the share of individuals with no labor income. Schooling is reduced on average by 0.32 years, and the share of individuals in the bottom third of the cognitive ability score distribution increases by 27% (16% for non-cognitive ability).

In recent contributions, [Aizer, Stroud and Buka \(2015\)](#), [Black, Devereux and Salvanes \(2016\)](#), and [Persson and Rossin-Slater \(2016\)](#) study the impact of maternal stress during pregnancy, in the United States, Norway, and Sweden, respectively. The evidence appears mixed: using variation in maternal cortisol levels, [Aizer, Stroud and Buka \(2015\)](#) find negative effects on cognition, health, and educational attainment; exploiting bereavement events in the mother's family, [Black, Devereux and Salvanes \(2016\)](#) find small negative effects on birth outcomes but no adverse effects on education and labor market outcomes in Norway, while [Persson and Rossin-Slater \(2016\)](#) uncover large impacts on later-life mental health in Sweden.

The data in [Aizer, Stroud and Buka \(2015\)](#) come from the National Collaborative Perinatal Project (NCP) which collected data for pregnancies between 1959 and 1965. They focus on the 1,100 pregnancies in Boston and Providence for which children were followed into adulthood. To measure stress, they use data on cortisol levels in mothers which has been linked to stress. Given obvious concerns about cross-sectional endogeneity of cortisol levels to observed economic status and important unobserved characteristics, they exploit

variation across pregnancies for siblings born to the same mother. They first establish that this variation in cortisol levels follows changes in the economic status of the mother. They then show that cortisol levels are associated with children's outcomes in a highly nonlinear way. Essentially only very high levels of cortisol are predictive of lower IQ and fewer years of education. They find no significant impact on birth outcomes.

In two closely related studies [Black, Devereux and Salvanes \(2016\)](#) and [Persson and Rossin-Slater \(2016\)](#) use very detailed administrative data matched across generations from Norway and Sweden, respectively. Both studies exploit a death in the mother's family as a shock to maternal stress levels and present various empirical strategies to account for the possibly nonrandom nature of exposure to death during a pregnancy. Importantly, they compare children who are exposed to a death in the mother's family while in utero to children who are exposed to the same shock shortly after birth. [Black, Devereux and Salvanes \(2016\)](#) use the universe of Norwegian birth and death records for the last 50 years which they connect to comprehensive administrative data on education, families, labor market outcomes, taxes, and earnings. IQ data come from military records. There are negative but small impacts on birth outcomes. However, there is no evidence of persistent impacts on important long-term outcomes, such as education completed, employment, earnings, or cognitive ability. [Persson and Rossin-Slater \(2016\)](#) use the universe of Swedish birth and death records for the last 40 years which they connect to comprehensive administrative data on inpatient and prescription drug claims. They find large impacts on birth outcomes and later-life takeup of ADHD, anti-anxiety, and depression medications.

In recent work on in utero exposure to war, which is a mixture of displacement, stress, and hunger, [Lee \(2014\)](#) finds that the educational attainment and socioeconomic status of South Korean cohorts that were in utero during the worst part of the Korean War (1951 birth cohort) were negatively affected relative to general trends in these outcomes. The 1951 birth cohort has significantly lower educational attainment, is less likely to have professional employment, and these effects are larger for individuals who were in utero in areas more

affected by the war. There is also some evidence of heterogeneity by gender and pregnancy stage.

2.2.6 Summary: Other in-utero exposures

The studies discussed in this section suggest that a range of different in-utero exposures (e.g, nuclear fallout, deficient nutrition, maternal stress) can have long lasting effects on individuals. These effects are often not mediated through birthweight or even early-life health but perhaps through latent cognitive ability. The wide variety of exposures that have been found to have substantial impacts on later-life outcomes suggests that the in-utero environment is very important both for early development and later-life economic outcomes. A key open question is how much improving the in-utero environment through various interventions can affect later-life outcomes in the US or other developed countries. In other words, it is not yet clear whether we can generalize from the large shocks that are typically studied in this literature to more conventional variation in the in-utero environment.

2.3 Fetal origins: Birthweight

2.3.1 Introduction

A literature related to fetal origins studies that examine the longer-term impact of exposure to various negative health shocks considers birthweight as a proxy for the gestational environment and the resulting health endowment. We choose to consider this literature separately for three reasons: (1) there are many studies interested in the impact of birthweight as the explanatory health measure rather than in the impact of in utero shocks, (2) several of the studies focused on in utero exposures show that birthweight can be relatively unaffected by exposures that impact later life outcomes, and (3) the near exclusive reliance of birthweight studies on twin fixed effects represents a methodological approach different from in utero exposure studies that rely primarily on variation across cohorts or across geographic areas.

This subsection is based on a systematic review which started with searching the economics literature on Google Scholar using the search terms "birthweight" OR "birth weight" for years 2005-2017. Our search returned an initial 128 results, however we only consider 15 studies that are relevant to our review. These need to consider the impact of birthweight on economic outcomes, rather than the other way around, examine non-health endpoints (education or economic outcomes), and be done in the US or other developed countries. In addition to the Google Scholar search, we added a small number of studies that were referenced by the studies we found through the systematic search but did not themselves come up in the search (usually because they did not contain the keywords in their title). Mindful of the slow publication process and the fast pace of some of the literature covered here, we also included relevant NBER Working Papers from 2015-2017 from the Children, Health, and Labor Studies programs.

Table 4: Fetal origins: Birthweight - Summary

Study	Area	Time period	Outcomes	Data	Approach	Findings
Currie and Hyson (1999)	UK	1958	Health Education SES	NCDS	Cross-sectional	large and significant negative association between LBW and the number of O-Level examinations passed at age 16, self-reported health, and employment outcomes measured at ages 23 and 33
Behrman and Rosenzweig (2004)	Minnesota	1936-1955	Education SES	MTR Authors' survey	Family + Twin FE	Large effects on education and wages
Miller, Mulvey and Martin (2005)	Australia	1964-1971	Education SES	Survey	Twin FE	100 gram increase in birthweight would be associated with a 1.4% increase in earnings
Currie and Moretti (2007)	California	1989-2001	Education SES	Birth records	Family FE	Intergenerational persistence of LBW is substantial, effects on SES and educational attainment
Black, Devereux and Salvanes (2007)	Norway	1967-1997	Health Education SES	Administrative	Family + Twin FE	10% increase in birthweight is associated with a 1% higher probability of high school graduation and with 1% higher earnings
Lin, Liu and Chou (2007)	Taiwan	1978-1982	Education	Administrative	Cross-sectional	An LBW child with low-education parents is 4.5% less likely to attend university, an LBW child with high-education parents is 2.8% less likely to attend university
Oreopoulos, Stabile, Walld and Roos (2008)	Manitoba	1978-1985	Health Education SES	Administrative	Family + Twin FE	Infant health, measured by birthweight and some other indicators, is found to be strongly associated with educational and labor force outcomes

Table 5: Fetal origins: Birthweight - Summary

Study	Area	Time period	Outcomes	Data	Approach	Findings
Royer (2009)	California	1960-1982	Health Education	Birth records ECLS-B	Family + Twin FE	200 gram increase in birthweight would be associated with a 0.03-0.04 year increase in schooling
Cesur and Kelly (2010)	1979	US	Cognitive skills	NLSY ECLS-K	Cross-sectional, IV	Very high birthweight has a negative effect on cognitive test scores
Nakamuro, Uzuki and Inui (2013)	Japan	2012	Education SES	Online survey	Twin FE	lower birthweight is associated with worse academic achievement at age 15, but has no statistically significant impact on further educational attainment and earnings
Figlio, Guryan, Karbownik and Roth (2014)	Florida	1992-2002	Health Education	Birth records School records	Family + Twin FE	10% increase in birthweight is associated with a one-twentieth of one standard deviation increase in test scores; 100 gram increase in birthweight would be associated with a 0.03-0.04 year increase in schooling
Chatterji, Kim and Lahiri (2014)	US	1997-2007	Education	PSID	Family FE	10% increase in birthweight being associated with a 0.036 standard deviation increase in childhood math test score
Xie, Chou and Liu (2016)	Taiwan	1978-2006	Education	Administrative	Family + Twin FE	a 200 gram increase in birthweight would be associated with a 0.03-0.04 year increase in schooling
Bharadwaj, Eberhard and Neilson (2017)	Chile	1992-2002	Education	Administrative	Family + Twin FE	a 10% increase in birthweight is associated with a one-twentieth of one standard deviation increase in test scores
Bharadwaj, Lundborg and Rooth (2017)	Sweden	1926-1958	Education SES	Administrative	Twin FE	10% increase in birthweight is associated with a 7-10 percentage point increase in the probability of high school completion and a 9% increase in income between ages 25-33

2.3.2 Summary of methods

The empirical economics literature that connects birthweight, which is thought to be associated with in utero nutritional environment to health, educational, and economic outcomes relies almost exclusively on twin fixed effects estimation strategies. The identification problem that this strategy is supposed to solve is that birthweight is expected to be strongly correlated with family background, and other observed and unobserved characteristics that are also correlated with children's outcomes. The idea is that once these observed and unobserved family background characteristics are controlled for, we can consider variation in birthweight within twin pairs to be reflective of randomly arising variation in in utero nutrition and environment.

Twins have been used to study a variety of causal effects in economics, for example, the returns to schooling. In most other contexts, the credibility of the twin research design is questionable because even within twin pairs there will be omitted variable bias that can be blown up by the fixed effect specification, making estimates more, not less biased. For example, in returns to schooling estimates, comparing a twin with more and a twin with less schooling, it is likely that the twin getting more schooling is different in a number of ways, most importantly potential earnings, from the twin getting less schooling.

However, in the context of studies of birthweight, twin studies might be valid, considering that birthweight is measured at the very beginning of life. The main concern is that birthweight is correlated with or proxies for unmeasured endowments at the beginning of life, and it is the effect of those endowments rather than birthweight that the studies attribute to birthweight. Another potential concern is that parental actions, for example in the form of educational investments, may reinforce or counteract birth endowments. Variation in birthweight within twin pairs turns out to be quite substantial. For example, in the sample of all US twins used in [Almond, Chay and Lee \(2005\)](#) the median difference is 220 grams, in the sample of all Norwegian twins used in [Black, Devereux and Salvanes \(2007\)](#) the difference between the medians is 310 grams. To put these in context, the mean birthweight is around

2,500 grams and the standard deviation is around 600 in each of the studies.

2.3.3 Studies

Using data from the UK's National Child Development Survey (NCDS), [Currie and Hyson \(1999\)](#) find that low birth weight (LBW) is associated in the long term with education, self-reported health, employment, and wages for both low-SES and high-SES children.¹ The authors find large and significant negative associations between LBW and the number of O-Level examinations passed at age 16, self-reported health, and employment outcomes measured at ages 23 and 33, controlling for various measures of family background. The authors examine the interaction between SES and LBW, but find that for the most part these interactions are not statistically significant. They conclude that the effect exists for children of low-SES and high-SES backgrounds as well. This study is important in the development of the literature because it represents an early attempt to study the long-term consequences of in utero conditions. However, causality is hard to evaluate as the approach is essentially cross-sectional, comparing LBW and non-LBW children, only controlling for SES. There are possibly many unobserved factors that bias the estimates if we want to interpret them as causal.

[Behrman and Rosenzweig \(2004\)](#) collect data on monozygotic twins in Minnesota and find that increasing birthweight would increase educational attainment at most birthweights and would have positive labor market payoffs for LBW infants. The authors identified same-sex twins in the Minnesota Twin Registry (MTR) which contains birth records for all twins born in the state between 1936 and 1955. They sent out questionnaires to about 6,000 same-sex twin pairs but estimation relies on female same-sex twin pairs with enough information provided on returned questionnaires, so their twin sample size is only 800. Their estimates are derived from variation in birthweight within twin pairs. In addition to twin fixed effects, they also include some other control variables. They uncover large effects on education and

¹For more details on the NCDS, see Section [2.2.5](#).

wages.

Analyzing a sample of Australian identical twins in survey data, [Miller, Mulvey and Martin \(2005\)](#) find that birthweight has no impact on schooling levels and only a small impact of earnings. They estimate that a 100 gram increase in birthweight would be associated with a 1.4% increase in earnings.

Using data on births in California, [Currie and Moretti \(2007\)](#) conclude that birthweight is intergenerationally transmitted, the transmission is stronger for mothers in higher-poverty areas, and contributes to the transmission of SES. Using birth records for all infants born in California between 1989 and 2001, the authors attempt to link mothers' birth records from the 1970-1974 period, which is only possible for mothers also born in California. Their sample contains about 640,000 birth records. To study intergenerational transmission of birthweight, they regress the birthweight of children on the birthweight of their mothers, adjusting for some SES control variables. They are also able to include "grandmother fixed effects", or in other words, rely on estimates within families. The intergenerational persistence of LBW is substantial, women whose mothers were LBW are 50% more likely to be LBW, controlling for SES. The transmission is found to be stronger for poor women. They also find effects on which zipcodes women end up living in (a proxy for SES) and their educational attainment.

[Black, Devereux and Salvanes \(2007\)](#) use within-twin methods and comprehensive data from Norway to establish that LBW has negative long-term consequences for IQ, education, and earnings. The researchers rely on rich administrative data from Norway that includes all births between 1967 and 1997 connected with education, family, tax, and earnings registers for the years 1986 to 2002. Military records give information on the IQ of all men. In order to study twins, the authors connect a survey of twins that contains information on zygosity. Birthweight is found to significantly effect adult IQ, education, earnings, and the birthweight of the next generation. For example, a 10% increase in birthweight is associated with a 1% higher probability of high school graduation and with 1% higher earnings.

[Oreopoulos, Stabile, Walld and Roos \(2008\)](#) use data from Manitoba and twin and sibling

methods to show that poor infant health, as measured by birthweight and other measures, is associated with elevated mortality up to age 17, and also with lower educational attainment and worse labor force outcomes. The authors match a dataset that contains birth records for all infants born in Manitoba in the 1978-1985 period with administrative records on education, later-life outpatient care, and social program uptake. Infant health, measured by birthweight and some other indicators, is found to be strongly associated with educational and labor force outcomes.

In another twin study, [Royer \(2009\)](#) finds that LBW is related to educational attainment and intergenerational transmission of LBW. The author relies on a dataset that contains birth records for twins born in California between 1960 and 1982 and a separate dataset, the Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), a nationally representative survey that follows infants born in 2001. Her estimation strategy relies on variation in birth weight within pairs of twins, and finds statistically significant but small long-term associations between birthweight and education. There is also evidence of intergenerational transmission of LBW. However, the impact of birthweight on educational outcomes is small: a 200 gram increase in birthweight would be associated with a 0.03-0.04 year increase in schooling.

[Nakamuro, Uzuki and Inui \(2013\)](#) study Japanese twins using a web-based survey. They find that lower birthweight is associated with worse academic achievement at age 15, but has no statistically significant impact on further educational attainment and earnings.

Using merged data on the birthweight and school records of all children born between 1992 and 2002 in Florida, [Figlio, Guryan, Karbownik and Roth \(2014\)](#) document the effects of poor neonatal health in the form LBW on educational outcomes. Similarly to the majority of other studies in the literature the authors use twin and sibling fixed effects to account for unobserved family-specific characteristics that could bias the results. They find that a 10% increase in birthweight is associated with a one-twentieth of one standard deviation increase in test scores. Using data from a longitudinal follow-up of the sample used in [Royer \(2009\)](#)

to account for the relationship between test scores and years of schooling, they estimate the association between birthweight and years of schooling to be almost exactly the same as in [Royer \(2009\)](#): a 200 gram increase in birthweight would be associated with a 0.03-0.04 year increase in schooling.

[Chatterji, Kim and Lahiri \(2014\)](#) use data for years 1997-2007 from the PSID and combine sibling fixed effects with an attempt to instrument for birthweight with fetal growth rate. They find relatively large effects, a 10% increase in birthweight being associated with a 0.036 standard deviation increase in childhood math test scores. They find that this effect is concentrated among those with LBW (below 2.5kg).

Using all birth records for years 1978-2006 from Taiwan, [Xie, Chou and Liu \(2016\)](#) report sibling and twin fixed effects estimates of the impact of birthweight on short-run child health outcomes, medium-run academic outcomes and adult health outcomes, and long-run schooling outcomes. In the short run, a 10% increase in birthweight is associated with an 11 deaths per 1,000 birth reduction in mortality for siblings (8.7 for male twins, 6.8 for female twins). While in the medium run, there is a positive association between birthweight and adult health, there is no significant relationship between graduating from a top high school or entering private university. A 10% increase in birthweight is associated with a 0.5% increase in the chance of entering private university. A 10% increase in birthweight is associated with a one percentage point increase in the probability of completing college for sibling pairs, but not for male or female twins. Finally, there is a positive association between birthweight and completed years of schooling that matches the [Royer \(2009\)](#) estimate: a 200 gram increase in birthweight would be associated with a 0.03-0.04 year increase in schooling. In a related study [Lin, Liu and Chou \(2007\)](#) find substantial heterogeneity: more educated parents are able to offset about 40% of the negative effect of low birthweight in Taiwan.

Using universal data on Chilean births between 1992 and 2002 and universal data on school test scores up to high school graduation, [Bharadwaj, Eberhard and Neilson \(2017\)](#) follow the evolution of the association between children's achievement and birthweight. Their

findings are seemingly contradictory: negative estimates of the relationship between birthweight and test scores are persistent over the course of education if twin fixed effects are included but are declining in magnitude if sibling fixed effects are included. The authors argue that this reflects parental investments and that in particular parents invest more in a less healthy child but do not differentiate across twins. Quantitatively, their results are about the same as the results in [Figlio, Guryan, Karbownik and Roth \(2014\)](#), a 10% increase in birthweight is associated with a one-twentieth of one standard deviation increase in test scores.

[Bharadwaj, Lundborg and Rooth \(2017\)](#) use universal administrative data on Swedish twins born 1926-1958, to study the association between birthweight and education, permanent income, and other outcomes. They find that a 10% increase in birthweight is associated with a 7-10 percentage point increase in the probability of high school completion and a 9% increase in income between ages 25-33. Interestingly, they find fading effects on income for males at later ages (above 50) but not for females.

In a rare study of *high* birthweight, [Cesur and Kelly \(2010\)](#) find that we may be concerned about birthweight being too high rather than too low for some individuals. Defining high birthweight as above 4,500 grams and using data from the NLSY and the ECLS-K, the authors find that having high birthweight is associated with lower test scores.

2.3.4 Summary

Recent studies on the impact of birthweight on educational outcomes find remarkably similar quantitative estimates. A 10% or 200 gram increase in birthweight could be associated with an approximately one-twentieth of one standard deviation increase in test scores, a 0.03-0.04 year increase in schooling and a 1% increase in high school graduation rates. To the extent that birthweight proxies for the in-utero environment, this suggests that in-utero interventions that can impact birthweight could have substantial returns. The three main caveats are that (1) relatively large increases in birthweight are needed for these impacts

in educational outcomes, (2) birthweight may be correlated with the in-utero environment but not be the actual cause of educational outcomes, which would suggest that increasing birthweight itself would not be sufficient to improve educational outcomes, and (3) these estimates typically rely on twin fixed effects designs which may be problematic if there is large non-random variation in birthweight within twin pairs.

2.4 Child health and nutrition

2.4.1 Introduction

This section discusses studies linking postnatal early life physical health and nutrition to later life outcomes. Following a methodological discussion, we cover the studies in this literature in two parts. We first discuss research that uses geographic variation and policy changes to study the historical disease environment of the US. Then we turn to studies that focus on modern data and policies.

This subsection is based on a systematic review which started with searching the economics literature on Google Scholar using the search terms "child health" OR "childhood health" OR "early-life health" OR "early life health" for years 2005-2017. Our search returned an initial 121 results, however we only consider 13 studies that are relevant to our review. These need to consider the impact of child health and nutrition on economic outcomes, rather than the other way around, examine non-health endpoints (education or economic outcomes), and be done in the US or other developed countries. In addition to the Google Scholar search, we added a small number of studies that were referenced by the studies we found through the systematic search but did not themselves come up in the search (usually because they did not contain the keywords in their title). Mindful of the slow publication process and the fast pace of some of the literature covered here, we also included relevant NBER Working Papers from 2015-2017 from the Children, Health, and Labor Studies programs.

2.4.2 Summary of methods

Compared to the in utero exposure literature that relies on geographic and time variation generated by exposure events (e.g., pandemics, nuclear events) and the birthweight literature that relies on twin fixed effects strategies, identification in the child health literature is much more difficult for a number of reasons. Measurement of child health after birth is challenging: while the date and place of birth (essential for in utero exposure studies) and birthweight (essential for birthweight studies) are well-documented in birth records and administrative data, child health is not universally documented in the United States. Hence US studies need to rely on health measures reported in survey data (e.g., NHIS, HRS, NLSY79) for the most part. Moreover, these self-reported measures are typically adults' recalls of their childhood health in surveys with the exception of studies that rely on height which the authors of these studies take as an objective measure of childhood health. Aside from studies using historical data, quasiexperimental variation in child health is hard to come by. Therefore the studies in this literature are best thought of as establishing statistical associations rather than causal relationships. The exceptions are studies of the historical disease environment in the Southern United States and the studies of the consequences of the improvement of healthcare for African Americans, but these have their own limitations because of data availability. A common methodology used in the historical studies relies on two sources of variation: geographic variation in pre-intervention rates of the condition and temporal variation created by the intervention in question. The key assumptions are that pre-intervention rates are not correlated with unobserved characteristics of areas and that the timing of the intervention is not correlated with other changes.

Table 6: Child health and nutrition - Summary

Study	Area	Time	Condition	Outcomes	Data	Approach	Findings
Bleakley (2007)	Southern US	1910s	Hookworm	Education SES	Surveys IPUMS Reports	Cohort + Geographic variation	A one standard deviation difference in hookworm infection rates across areas is associated with an 11% wage difference. The impact of the eradication of hookworm across cohorts could have increased adult wages 43%. Hookworm exposure was associated with an around 50% drop in the returns to schooling. 22% of the North-South income gap in attributable to hookworm infection in the South.
Bleakley (2010)	Southern US	1920s	Malaria	SES	IPUMS NAPP	Cohort + Geographic variation	Childhood malaria infection could reduce adult income by approximately 50%
Barreca (2010)	Southern US	1900- 1936	Malaria	Education SES	Census	Cohort + Geographic variation	Exposure to an environment with 10 additional malaria deaths per 100,000 is associated with 13% lower income, 0.4 fewer years of education, and a 5.1% reduction in the probability of graduating from high school. Malaria exposure can account for 25% of the difference in educational attainment across exposed and non-exposed states
Feyrer, Politi and Weil (2016)	Southern US	1920s	Iodine deficiency	IQ	Enlistment data	Cohort + Geographic variation	Men more affected by the introduction of iodized salt experience 15 differential increase in their IQ in affected regions, implying a 3.5 point nationwide differential
Niemesch (2015)	US	1940s	Iron deficiency	Education SES	IPUMS	Cohort + Geographic variation	Iron fortification program increased incomes at the individual level by 5.25% for those affected; 1.82% increase in the number of weeks worked (labor supply); 3.43% increase in hourly wage (productivity). One standard deviation difference in iron consumption associated with a 1 percentage point difference in school enrollment. Children affected experienced a 2% income gain on average.
Case, Fertig and Paxson (2005)	UK	1958	Childhood health	Health Education SES	NCDS	Cross-sectional	Relative to an average of 2 O-Levels passed in the sample, a chronic condition at age 7 is associated with 0.3 fewer O-Levels passed, an if the chronic condition is still present at age 16, it is associated with an additional 0.2 fewer O-Levels passed. Children who are born with low birthweight pass 0.5 fewer O-Levels on average. An additional chronic condition at age 16 is associated with a 4-5% lower probability of employment as an adult.

Table 7: Child health and nutrition - Summary

Study	Area	Time	Condition	Outcomes	Data	Approach	Findings
Case and Paxson (2008a)	US	1920-1950	Childhood health (height)	Health SES	HRS	Cross-sectional	Height is associated with a variety of outcomes
Case and Paxson (2008b)	US, UK	1958 1970 1986-2004 1998-2000	Childhood health (height)	Education SES	NCDS BCS NLSY79 FFCWS	Cross-sectional + Family FE	Lot of the association between height and labor markets outcomes is explained by the association between height and cognitive test scores
Case and Paxson (2009)	US	1920-1950	Childhood health (mortality)	Cognitive function	HRS	Cross-sectional	Cognitive function old age significantly associated with childhood disease environment
Case and Paxson (2010)	US	1986-2006	Childhood health (height)	Health Education SES	NLSY79	Cross-sectional + Family FE	Childhood health associated with later life SES
Smith (2009)	US	1968-1999	Childhood health	Education SES	PSID	Cross-sectional + Family FE	Poor childhood health has a quantitatively large effect on family income, household wealth, individual earnings, and labor supply, but not on education
Fletcher and Lehrer (2011)	US	1994-2002	Adolescent health (Genetic information)	Education	Add Health	Cross-sectional + Family FE	Family fixed effects cannot fully account for endogeneity if health, genetic markers can be helpful instrumental variables
Lundborg, Nilsson and Rooth (2014)	Sweden	1969-1997	Adolescent health	Education SES	Administrative	Cross-sectional + Family FE	Large labor market effects of health. For example, using a global health measure and controlling for a variety of observable characteristics, such as ability, moving someone from perfect health to health just above the level required for military service is associated with a 10-15% earnings penalty

2.4.3 Studies: Historical data

In the early 20th century, diseases that today we see in developing countries were endemic in the American South and nutritional deficiencies were a problem in the United States as well. [Bleakley \(2007\)](#), [Bleakley \(2010\)](#), and [Barreca \(2010\)](#) study this historic disease environment, focusing on hookworm disease and malaria. [Feyrer, Politi and Weil \(2016\)](#) and [Niimesh \(2015\)](#) study nutritional deficiencies, focusing on iodine and iron deficiency.

In [Bleakley \(2007\)](#), the long-term impact of the eradication of hookworm disease in the American South around 1910 is found to have increased school enrollment, attendance, and literacy, leading to long-term income gains, possibly through an increase to the returns to schooling. This study exploits variation across geographic areas in the levels of hookworm infection prior to the eradication campaign, and shows that areas with higher levels of hookworm infection, that is, areas that experienced a larger change during the eradication campaign, saw greater increases in school enrollment, attendance, and literacy. Following cohorts who were in school during the eradication efforts, the author finds substantially higher earnings for those more affected by the eradication. To put the paper's quantitative estimates in context, the author calculates that comparing a cohort with no childhood exposure to a cohort with full childhood exposure to hookworm, a one standard deviation difference in hookworm infection rates across areas is associated with an 11% wage difference. The impact of the eradication of hookworm across cohorts could have increased adult wages by as much as 43%. Importantly, much of this wage effect came through increases in returns to schooling: hookworm exposure was associated with an around 50% drop in the returns to schooling. The author attributes 22% of the North-South income gap in the period to hookworm infection in the South.

[Bleakley \(2010\)](#) finds that in utero and postnatal exposure to malaria in the United States was associated with lower levels of educational attainment and higher levels of poverty among those exposed to malaria. Using a research design very similar to the hookworm eradication study, relying on pre-eradication variation across US states in malaria prevalence and the

beginning of eradication campaigns around 1920, the paper shows that individuals born in states and cohorts affected by malaria experienced greater gains in socioeconomic status. To summarize his results, the author calculates that childhood malaria infection could reduce adult income by approximately 50%.

[Barreca \(2010\)](#) also studies the impact of exposure to malaria in the year of birth. In addition to using malaria eradication as a source of variation, the author introduces an instrumental variable strategy which uses temperature, a factor in how hospitable the environment is to malaria, as an instrument for malaria. He uncovers large negative effects of early life malaria exposure on educational attainment and earnings. Exposure to an environment with 10 additional malaria deaths per 100,000 is associated with 13% lower income (not statistically significant), 0.4 fewer years of education, and a 5.1% reduction in the probability of graduating from high school. The author estimates that malaria exposure can account for 25% of the difference in educational attainment across exposed and non-exposed states.

In historical work using data from the US, [Feyrer, Politi and Weil \(2016\)](#) and [Niimesh \(2015\)](#) study the impact of decreasing iodine and iron deficiency, respectively. [Feyrer, Politi and Weil \(2016\)](#) follow a strategy similar to the hookworm and malaria studies in exploiting two sources of variation: geographic variation in pre-intervention rates of iodine deficiency and the 1924 introduction of iodized salt which creates a cohort discontinuity. To establish the pre-intervention rates of iodine deficiency in different areas of the US, they use data from the World War I draft on simple goiter, a condition caused by iodine deficiency. To study impacts on cognitive ability, they use data from the World War II draft on test scores. They uncover large — 15 point for affected regions, implying 3.5 point nationwide — differential increases in IQ for men more affected by the introduction of iodized salt.

[Niimesh \(2015\)](#) uses a similar strategy, but focuses on iron deficiency. The intervention in his case is the 1943 mandate to fortify bread with iron to reduce iron deficiency. Areas that had lower levels of iron consumption before the intervention and were therefore more impacted by the intervention experienced faster increases in school enrollment and income.

By 1970, adults who were more affected by the introduction of iron fortification experienced larger income gains. The author estimates that the inexpensive iron fortification program increased incomes at the individual level by 5.25% for those who were affected because they previously suffered from iron deficiency, which came through a 1.82% increase in the number of weeks worked (labor supply) and an implied 3.43% increase in hourly wage (productivity). Part of this came through schooling: a one standard deviation difference in iron consumption in this sample is associated with a 1 percentage point difference in school enrollment. In addition to these contemporaneous impacts, the author estimates the individual long-term impact of reducing iron deficiency for children through exposure to the program for all of childhood. Children who were affected because they previously suffered from iron deficiency, could have experienced a 2% income gain on average.

2.4.4 Summary: Historical data

The impact of different diseases and deficiencies is hard to compare with each other. The studies discussed in this section suggest that historically, now eradicated diseases and nutritional deficiencies had large economic impacts in the US. If similar diseases and nutritional deficiencies were identified today, government policies to combat them could have large returns. However, these seem less relevant today than focusing on overall child health discussed in the next section.

2.4.5 Studies: Modern data

[Case, Fertig and Paxson \(2005\)](#) analyze data from the NCDS and find that childhood health has a negative impact on educational attainment and SES in adulthood.² The authors find large and significant negative associations between childhood health and the number of O-Level examinations passed at age 16, self-reported health, and employment outcomes measured at ages 23, 33, and 42 controlling for various measures of family background.

²For more details on the NCDS, see Section [2.2.5](#).

Relative to an average of 2 O-Levels passed in the sample, a chronic condition at age 7 is associated with 0.3 fewer O-Levels passed, and if the chronic condition is still present at age 16, it is associated with an additional 0.2 fewer O-Levels passed. Children who are born with low birthweight pass 0.5 fewer O-Levels on average. An additional chronic condition at age 16 is associated with a 4-5% lower probability of employment as an adult. These are cross-sectional results in the sense that no exogenous shocks to childhood health are exploited. Therefore these results are perhaps best interpreted as establishing associations rather than causality.

In a series of studies, Anne Case and Christina Paxson have documented the association between height, taken as a proxy for early-life health, and later-life outcomes. [Case and Paxson \(2008a\)](#) provide evidence that height is associated with a variety of outcomes measuring adults' height and wellbeing in the Health and Retirement Study (HRS). This finding is again cross-sectional, the essence is a correlation between height on the one hand and educational attainment and occupational choice on the other hand, controlling for childhood SES. Late-life cognitive ability is also found to be strongly correlated with height, controlling for early-life SES. In its reliance on the HRS, a survey of older Americans, this study also demonstrates that it is hard to come by reliable measures of childhood health connected with reliable data on SES in the US. Extending these findings, [Case and Paxson \(2008b\)](#) uses data from the UK and the US and finds that a lot of the association between height and labor markets outcomes is explained by the association between height and cognitive test scores, evidence of the association between children's health and human capital accumulation. The data come from the NCDS discussed above in Section 2.2.5, as well as the British Cohort Study (BCS), and the 1979 National Longitudinal Study of Youth (NLSY79). The BCS has followed all individuals living in Great Britain born in the week of April 5, 1970. The NLSY79 is a well-known American survey that has followed 13,000 Americans who were originally entered into the survey as youth in 1979. In another extension, [Case and Paxson \(2010\)](#), using data from the NLSY79 survey, presents further evidence using

height data of the relationship between childhood health and later-life SES. This study is also able to control for mother fixed effects, alleviating some of the concerns about unobserved family background biasing estimates. A simple quantitative summary of the Case and Paxson findings on height and earnings that is repeated throughout these papers is that moving the height of men from the 25th to the 75th percentile (an increase by 4 inches) is associated with a 9-10% increase in earnings. In related work, [Case and Paxson \(2009\)](#) analyze data reported in the HRS on cognitive function in old age connected with childhood disease environment and find significant association.

[Smith \(2009\)](#) finds that poor childhood health has a quantitatively large effect on family income, household wealth, individual earnings, and labor supply, but not on education. Relying on a representative sample of 5,000 families in the Panel Study of Income Dynamics (PSID), the author is able to control for family fixed effects, alleviating the concern that instead of childhood health per se it is unmeasured family background and SES factors that drive much of the results in other studies. The estimation exploits only variation in childhood health within sibling pairs, but unfortunately still relies on retrospective recall of childhood health by adults which can be problematic both because of measurement error and because the causality between poor reported childhood health and adult economic status might be reversed: it might be that poorer individuals recall being less healthy as children because they are poorer, rather than them being poorer now because they were actually less healthy.

[Fletcher and Lehrer \(2011\)](#) argue that genetic variation within families can be useful to study the impact of mental health outcomes on educational outcomes. In a departure from a literature that attempted to measure childhood health from adults' recall in surveys, they rely on the National Longitudinal Study of Adolescent (Add Health) which collected genetic information on participants in addition to data on their health and educational outcomes. Estimation is done within a family fixed effects framework using genetic variation as an instrumental variable for health. The key observation is that even within families, variation across siblings in health can be correlated with unobserved factors, biasing estimates. The

authors argue that using genetic variation as an instrument is helpful to arrive at causal estimates of adolescent health on educational outcomes given this possible within-family endogeneity.

In a study of adolescent health, [Lundborg, Nilsson and Rooth \(2014\)](#) show from universal Swedish registers that adolescent health is also strongly associated with adult labor market outcomes. Several major challenges to the earlier literature relying on representative surveys of the US population are solved by the universal Swedish register data exploited by this study. The size of the data is much larger, allowing for improved precision. The quality of the data is much better, not relying on self-reported measures and measures based on recall, reducing measurement error and concerns about reverse causality (along the lines that poorer people recall having worse health as children). The data is much broader, which means that the authors don't need to rely on one-dimensional measures like height to proxy for childhood health. The authors are also able to include sibling and twin fixed effects and link the adolescent health data with later life economic outcomes. They find large labor market effects and also find that mental health conditions are the most important. For example, using a global health measure and controlling for a variety of observable characteristics, such as ability, moving someone from perfect health to health just above the level required for military service is associated with a 10-15% earnings penalty.

2.4.6 Summary: Modern data

We take the literature on the relationship between child physical health and longer-term outcomes using modern data to suggest that there is a strong relationship between childhood health and education and economic outcomes, but note that it is not obvious what interventions would be beneficial. We interpret the first group of studies that we discussed in this section—[Case, Fertig and Paxson \(2005\)](#), [Case and Paxson \(2008a\)](#), [Case and Paxson \(2008b\)](#), [Case and Paxson \(2009\)](#), [Case and Paxson \(2010\)](#)—as providing evidence of a strong association between childhood health and longer-term outcomes, rather than describing a

causal pathway. They clearly show that individuals who were less healthy as children or are less tall get less education and earn less. However, these studies don't rely on any identifying variation in health or height, which means that omitted factors could be important. The other main limitation of this group of studies is that they rely on surveys, so measurement error could be large. Nevertheless these studies are important in the literature connecting childhood health to longer-term outcomes because they establish an important association without relying on administrative data that is not available in the US. [Chay, Guryan and Mazumder \(2009\)](#) and [Chay, Guryan and Mazumder \(2014\)](#) stand apart from the rest of the studies because of their focus on the Black-White gap and hospital desegregation. An important contribution of their studies is showing that actual government policies can impact childhood health and longer-term outcomes in the US. It is plausible that improving hospital care for low-SES children could still yield gains in longer-term education and labor market outcomes.

2.5 Child mental health

2.5.1 Introduction

A more recent literature has started examining whether in addition to physical health and nutrition, childhood mental health can also be linked to educational achievement and adult outcomes. The measurement of childhood mental health is even more difficult than the measurement of childhood physical health and is likely to be more affected by variation and potential endogeneity in diagnostic criteria and diagnosis rates. Nevertheless policy-makers and researchers have become more concerned with mental health conditions, so in this section we examine what has been learned about the long-term impact of childhood mental health.

This subsection is based on a systematic review which started with searching the economics literature on Google Scholar using the search terms "mental health" OR "psychological" OR "depression" OR "ADHD" OR "behavioral") AND ("child" OR "childhood" OR "youth" OR "adolescent") for years 2005-2017. Our search returned an initial 55 results, however

we only consider 16 studies that are relevant to our review. These need to consider the impact of children’s mental health on economic outcomes, rather than the other way around, examine non-health endpoints (education or economic outcomes), and be done in the US or other developed countries. In addition to the Google Scholar search, we added a small number of studies that were referenced by the studies we found through the systematic search but did not themselves come up in the search (usually because they did not contain the keywords in their title). Mindful of the slow publication process and the fast pace of some of the literature covered here, we also included relevant NBER Working Papers from 2015-2017 from the Children, Health, and Labor Studies programs.

2.5.2 Summary of methods

The majority of the studies in the children’s mental health literature relies on surveys, most commonly the National Longitudinal Survey of Youth (NLSY79), the National Longitudinal Survey of Children and Youth (NLSCY) from Canada, and the National Longitudinal Study of Adolescent Health (Add Health). Most of the studies try to exploit the panel structures of these surveys and most also rely on family fixed effects, in order to try to account for unobserved family-level confounders by only comparing children in the same family. As already discussed in other sections, while the inclusion of family fixed effects can be important in order to account for some the obvious confounders, it is unable to account for individual-level confounders that will likely vary even within families. Some articles in this literature use administrative data from outside the US, from Canada ([Currie, Stabile, Manivong and Roos, 2010](#)) and Denmark ([Dalsgaard, Nielsen and Simonsen, 2014](#)). Administrative data can allow for much bigger samples and more accurate measurement of both mental health and the outcomes of interest, such as education or crime. In addition to panel data combined with family fixed effects, a few studies use other identification strategies, such as variation across physicians in ADHD treatment ([Dalsgaard, Nielsen and Simonsen, 2014](#)) and discontinuity in treatment assignment ([Cuellar and Dave, 2015](#)). Another idea, used

by [Ding, Lehrer, Rosenquist and Audrain-McGovern \(2009\)](#), is to rely on genetic markers that are taken to be predetermined and as a valid instrument for the actual manifestation of mental health conditions. In our view, for the most part, the studies of the impact of mental health on children's development should be seen as describing associations rather than causal relationships.

Table 8: Child mental health - Summary

Study	Area	Time	Condition	Outcomes	Data	Approach	Findings
Currie and Stabile (2006)	US Canada	1990s	ADHD	Education	NLSY79 NLSCY	Panel + Family FE	ADHD has a substantial negative impact on test scores and educational attainment; mental health can be more important than physical health; having a score at the 90th percentile of the hyperactivity score distribution increases the probability of grade repetition by 6-7%, while having a chronic condition like asthma does not change the probability of grade repetition
Currie and Stabile (2009)	US Canada	1990s	ADHD Anxiety/ depression Conduct disorders	Education	NLSY79 NLSCY	Panel + Family FE	Conduct disorders and anxiety/depression have negative impacts too, though ADHD has the largest and most consistent effect
Fletcher and Wolfe (2008)	US	1990s	ADHD	Education	Add Health	Panel + Family FE	The impact of ADHD is not robust to the inclusion of sibling fixed effects
Fletcher (2014)	US	1990s	ADHD	SES	Add Health	Panel + Family FE	Childhood ADHD is associated with a 10-14% reduction in adult employment and 33% reduction in earnings
Fletcher (2008)	US	1990s	Depression	Education	Add Health	Panel	Negative correlation between adolescent depression and various measures of educational attainment (high school completion, college enrollment, quality of college), but only for females
Fletcher (2010)	US	1990s	Depression	Education	Add Health	Panel + Family FE	Large negative effects of adolescent depression on measures of educational attainment
Currie, Stabile and Jones (2014)	Quebec	1990s	ADHD	Education	NLSCY	Policy change over time and across provinces	The use of ADHD medications increased as they became more accessible; academic outcomes (test scores, grade repetition) did not improve and possibly even worsened
Dalsgaard, Nielsen and Simonsen (2014)	Denmark	1990s	ADHD	Crime	Administrative	Physicians across hospitals	Decreased criminal behavior as a result of medication use in ADHD
Ding, Lehrer, Rosenquist and Audrain-McGovern (2009)	Virginia	1998- 2003	ADHD Depression Obesity	Education	GATOR	IV: Genetic markers	Depression and obesity have substantial negative effects on academic achievement, but ADHD does not seem to be related to academic achievement

Table 9: Child mental health - Summary

Study	Area	Time	Condition	Outcomes	Data	Approach	Findings
Aizer (2008)	US	1998-2006	ADD	Education	ECLS-K	Panel	Increasing the number of children with undiagnosed ADD in a classroom by one standard deviation lowers the average test scores in the classroom by 10 to 20% of a standard deviation
Frank and Meara (2009)	US	1979-2004	Depression	Education	NLSY79	Panel	Don't find any effect of maternal depression on cognitive development of children.
Currie, Stabile, Manivong and Roos (2010)	Manitoba	1979-1987	ADHD Conduct disorder	Education SES	Administrative	Panel + Family FE	Early physical health can be linked to young adult outcomes through young adult health, while early mental health has predictive power for young adult outcomes even conditional on future health and health at birth.
Smith and Smith (2010)	US	1968-2007	Depression Alcohol, drugs Other psychological problems	Education SES	PSID	Panel + Family FE	Children affected by mental health problems go onto earn 20% (or \$10,4000) less and have \$18,000 less in household assets; the lifetime costs of childhood mental health conditions amount to \$300,000 in lost family income
Salm and Schunk (2012)	Osnabrück	2002-2005	Hyperactivity Emotional symptoms Peer problems Conduct problems	Cognitive development	Administrative	Panel + Family FE	Physical health condition sare not strongly associated with cognitive development, but mental health conditions, in particular hyperactivity, can explain a lot of the variation in cognitive development at the time children enter school.
Webbink, Vujić, Koning and Martin (2012)	Australia	1989-1990 1996-2000	Conduct disorder	Education Crime	ATR	Panel + Family FE	Conduct disorder is associated with highly significant 15 percentage point (13 percentage point for depression) increase in the probability of attacking others and 7 percentage point (no effect for depression) increase in the probability of being arrested.
Cuellar and Dave (2015)	Washington	2004-2009	Mental health treatment	Education	Administrative	Discontinuity in treatment assignment	Behavioral interventions can have an effect on school dropout and school completion for youth in the justice system.

2.5.3 Studies

In a pair of studies, [Currie and Stabile \(2006\)](#) and [Currie and Stabile \(2009\)](#) use data from longitudinal surveys from the US and Canada to argue that childhood mental problems can have substantial negative impacts on test scores and educational attainment, possibly a larger impact than physical health. Their American data comes from the National Longitudinal Survey of Youth (NLSY79) which has followed 12,000 youth from 1979 onwards and the children of female respondents in the original survey sample from 1986 onwards. They relate hyperactivity scores measured between 1990 and 1994 to children's outcomes in the later waves of the survey. Their Canadian data comes from the National Longitudinal Survey of Children and Youth (NLSCY) which began in 1994. They relate mental health screener scores measured in 1994 to children's outcomes in the later waves of the survey. Both studies are able to control for family fixed effects using siblings. [Currie and Stabile \(2006\)](#) find that Attention Deficit Hyperactivity Disorder (ADHD) has a substantial negative impact on test scores and educational attainment; they claim that mental health can be more important than physical health. For example, while having a score at the 90th percentile of the hyperactivity score distribution increases the probability of grade repetition by 6-7%, having a chronic condition like asthma does not change the probability of grade repetition. [Currie and Stabile \(2009\)](#) extend the findings to some other mental problems, such as conduct disorders and anxiety/depression, though ADHD has the largest and most consistent effect.

[Fletcher and Wolfe \(2008\)](#) attempt to extend the findings in [Currie and Stabile \(2006\)](#) to older children and broader measures of human capital. They rely on the National Longitudinal Study of Adolescent Health (Add Health) which followed adolescents in the 1990s. They rely on retrospective self-reports of young adults about whether they had ADHD as children. They find that even though OLS estimates of the impact of ADHD are large and negative, they are not robust to the inclusion of sibling fixed effects. On the other hand, in follow up work, [Fletcher \(2014\)](#) connects ADHD with labor markets outcomes and finds that childhood ADHD is associated with large reductions in employment and earnings. This

study also relies on Add Health data and retrospective self-reports of ADHD diagnosis, but for labor market outcomes the large and negative coefficients survive the inclusion of family fixed effects. The author claims that childhood ADHD is associated with a staggering 10-14% reduction in adult employment and 33% reduction in earnings.

In related work, using the same data from Add Health but studying adolescent depression, [Fletcher \(2008\)](#) shows that there is negative correlation between adolescent depression and various measures of educational attainment (high school completion, college enrollment, quality of college), but only for females. Using the same data but now controlling for sibling fixed effects, [Fletcher \(2010\)](#) also finds large negative effects of adolescent depression on measures of educational attainment.

Given that ADHD seems to be related to later-life outcomes, a possible policy prescription would be increasing access to treatment. However, examining the impact of improved access to ADHD medications through expanded insurance coverage for prescription drugs in Quebec, [Currie, Stabile and Jones \(2014\)](#) find little improvement in children's medium-term and long-term academic outcomes. As with their previous studies of ADHD in Canada, the data come from the NLSCY and the authors exploit the panel structure of the survey combined with the variation generated by the policy change. Since the policy that increased access to ADHD medications was only introduced in Quebec and only starting in 1996, the authors are able to arrive at estimates of the impact of increased access on use of these medications and on academic outcomes, comparing the post-1996 period in Quebec to earlier years and to children living in other provinces. While the use of ADHD medications is found to have increased as they became more accessible, academic outcomes (test scores, grade repetition) did not improve and possibly even worsened. Looking at the same issue in Denmark, but exploiting variation in physicians' propensity to prescribe, [Dalsgaard, Nielsen and Simonsen \(2014\)](#) document decreased criminal behavior as a result of medication use in ADHD. This study is able to use hospital and crime register data for the universe of Danish children born in the 1990s. The obvious missing piece is an examination of educational outcomes: the

reduced form relationship between ADHD medication use and crime is hard to understand without considering the role of education.

[Ding, Lehrer, Rosenquist and Audrain-McGovern \(2009\)](#) use variation in genetic markers to identify the impact of ADHD, depression, and obesity on academic achievement. Their data comes from the Georgetown Adolescent Tobacco Research (GATOR) study, which was a small survey of students in five Northern Virginia high schools between 1999-2003. In addition to questionnaires, the original study collected DNA samples which allowed the authors to connect some genetic markers to ADHD, depression, and obesity. They use genetic markers to instrument for health conditions. Depression and obesity are found to have substantial negative effects on academic achievement, but ADHD does not seem to be related to academic achievement in their data.

[Aizer \(2008\)](#) studies the spillovers of ADD on peers using data from the Early Childhood Longitudinal Survey – Kindergarten Cohort (ECLS-K). This dataset follows 20,000 children who started kindergarten in 1998. Crucially, not only the parents of children were surveyed, but their teachers too, which gives the researcher data on the participating child’s classroom. From the parent surveys, the author is able to estimate the number of children with diagnosed and undiagnosed ADD in the participating child’s classroom. A child is taken to have undiagnosed ADD if she is diagnosed with ADD later in the data. Increasing the number of children with undiagnosed ADD in a classroom by one standard deviation lowers the average test scores in the classroom by 10 to 20% of a standard deviation. Exploiting the panel nature of the data, the author controls for individual fixed effects and looks at what happens to a child academic achievement when a peer is diagnosed with and treated for ADD. The main finding is that when children are diagnosed and treated, their behavior improves but their own academic achievement does not. However, the academic achievement of their peers improves.

[Frank and Meara \(2009\)](#) follow children from the NLSY79 survey to find that maternal depression affects the social and emotional development of children. However, they fail to

uncover any meaningful effect on cognitive development (reading and mathematics). Their sample consists of 1,600 children of female respondents to the NLSY for whom all the data is available. There are obvious empirical challenges such as reverse causation, meaning that children's problems could cause maternal depression and the usual concerns about omitted variable bias caused by factors related to both maternal mental health and children's development. The authors try to get around these problems by exploiting the panel structure of the survey. In particular, they rely on maternal mental health problems that precede the time at which children's outcomes are measured. This of course doesn't solve the problem that it is possible that unmeasured, latent problems impact maternal mental health which are then taken to cause the children's problems once those problems are measured. The researchers also control for a wide set of variables and use propensity score models as well.

Using sibling data from Manitoba, [Currie, Stabile, Manivong and Roos \(2010\)](#) find that early physical health can be linked to young adult outcomes through young adult health, while early mental health has predictive power for young adult outcomes even conditional on future health and health at birth. This study relies on universal administrative health insurance records for all, approximately 50,000, children born in the Canadian province of Manitoba between 1979-1987. Moreover, it only relies on siblings for identification which allows the authors to account for unobserved family-specific background measures. On the outcome side, the researchers are able to link the healthcare data to universal administrative data on education and on welfare uptake.

Following siblings for 40 years in the Panel Study of Income Dynamics (PSID), [Smith and Smith \(2010\)](#) find that childhood psychological problems are associated with worse educational attainment and substantially lower adult incomes. The strength of this study is the long panel and the ability to control for sibling fixed effects. The weakness is that childhood health is measured retrospectively, through a child health history addition to the 2007 PSID. As discussed earlier, the problem with retrospective measurement is that it is possible that people who have worse SES recall being sicker in their childhood, leading to

reverse causation. This study also clearly discusses the impacts of mental health problems in dollar terms. The authors report that children affected by mental health problems go on to earn 20% (or \$10,4000) less and have \$18,000 less in household assets. They estimate that the lifetime costs of childhood mental health conditions amount to \$300,000 in lost family income.

[Salm and Schunk \(2012\)](#), analyzing data from a German town, find that mental health conditions, in particular hyperactivity, rather than physical health conditions, can explain the transmission of human capital from parents' education to children's development. Their data come from Osnabrück, a German town, and include the initial medical examination records of all, approximately 4,000, children entering school between 2002 and 2005. Their sibling sample include approximately 900 children. Their measures of cognitive development also come from the same medical examinations. They find that physical health conditions are not strongly associated with cognitive development, but mental health conditions, in particular hyperactivity, can explain a lot of the variation in cognitive development at the time children enter school.

[Webbink, Vujić, Koning and Martin \(2012\)](#) find that childhood conduct disorder reduces human capital accumulation and increases violent and criminal behavior. The data come from the Australian Twin Registry (ATR). The approximately 4,300 twin pairs in the register were born in the 1964-1971 period and were entered in the 1980-1982 period. They were then surveyed in 1989-1992 and 1996-2000. The authors estimate the impact of conduct disorder and depression using variation within twin pairs. There is some evidence that educational attainment is reduced by both conditions. In the most stringent specification that relies on identical twin pairs only, conduct disorder is associated with highly significant 15 percentage point (13 percentage point for depression) increase in the probability of attacking others and 7 percentage point (no effect for depression) increase in the probability of being arrested. Of course, one should be careful about interpreting these findings as causal effects.

[Cuellar and Dave \(2015\)](#) examine the impact of mental health interventions for youth

in the justice system and find lower dropout rates, increased high school completion, but negative or non-significant effects on grades. The study relies on administrative data from the state-wide juvenile justice system and the state-wide school system in Washington. The authors assess the impact of being assigned to various mental health treatments (Multi-systemic Therapy (MST), Functional Family Therapy (FFT), and Aggression Replacement Training (ART) under a state policy. While there is possibly strong selection in the uptake of the various therapies, the researchers also exploit a discontinuity in the eligibility of youth for mental health treatment. They also use assignment to a treatment as an instrument for receiving that treatment. They find that behavioral interventions can have an effect on school dropout and school completion.

2.5.4 Summary

Much like the in the case of the studies of childhood physical health, we understand the studies presented in this section on childhood mental health as describing associations rather than causal pathways from mental health to economic outcomes. The evidence suggests that mental health is strongly correlated with longer-term outcomes, the correlation may even be stronger than for physical health, likely because (latent) cognitive ability is more correlated with mental than with physical health. As we already discussed mental health is one of the most difficult areas to study because of reporting and measurement issues. We discussed [Cuellar and Dave \(2015\)](#) which studies a mental health intervention for youth in the justice system as an example of studies that look at actual interventions, however, there is ample room for studying interventions that can improve mental wellbeing and educational and economic outcomes.

2.6 Environmental health

2.6.1 Introduction

This section discusses recent studies that have connected environmental exposures, in particular exposure to lead and to air pollution, that have a negative effect on health to longer-term outcomes. Recent surveys of the economics literature that connect environmental health to human capital and labor market behavior are presented in [Graff Zivin and Neidell \(2013\)](#) and [Currie, Zivin, Mullins and Neidell \(2014\)](#). [Graff Zivin and Neidell \(2013\)](#) argue that this literature has not only provided quasi-experimental evidence on the impact of air pollution on health and economic outcomes but also shows that pollution can have negative long-term impacts even if its present in quantities below regulatory standards. [Currie, Zivin, Mullins and Neidell \(2014\)](#) conclude that there is strong evidence on the impact of pollution on health, though we can still learn more about longer-term economic impacts. Both surveys emphasize the contributions of economics to this particular field of research through the introduction of conceptual models and credible identification strategies that allow researchers to estimate causal effects.

This subsection is based on a systematic review which started with searching the economics literature on Google Scholar using the search terms ("environment" OR "environmental" OR "pollution" OR "exposure" OR "lead" OR "pollen") AND ("child" OR "childhood" OR "children" OR "long-term") for years 2005-2017. Our search returned an initial 64 results, however we only consider 9 studies that are relevant to our review. These need to consider the impact of environmental exposures on economic outcomes, rather than the other way around, examine non-health endpoints (education or economic outcomes), and be done in the US or other developed countries. In addition to the Google Scholar search, we added a small number of studies that were referenced by the studies we found through the systematic search but did not themselves come up in the search (usually because they did not contain the keywords in their title). Mindful of the slow publication process and the fast

pace of some of the literature covered here, we also included relevant NBER Working Papers from 2015-2017 from the Children, Health, and Labor Studies programs.

2.6.2 Summary of methods

The literature studying environmental health, most importantly air pollution and exposure to lead has relied on stronger identification strategies than most of the literature studying the long-term impact of childhood mental and physical health. This is possible because many policies have affected the quality of the environment in geographically heterogeneous ways, which presents useful identifying variation. In addition, in many cases, the evolution of environmental quality is also well-documented at fine geographical levels, typically through data from individual monitoring stations. Moreover, administrative data on test scores in the education system have benefitted many studies, and in recent work administrative earnings records have also been exploited.

Table 10: Environmental health - Summary

Study	Area	Time	Exposure	Outcomes	Data	Approach	Findings
Rau, Reyes and Urza (2013)	Chile	1984-1989	Lead	Education SES	Administrative	Geographic variation in exposure	Large reductions in test scores and a significant impact on earnings
Reyes (2015)	US	1980s	Lead	Crime	NLSY79 NLSY97 NHANES	Phase-out of leaded gasoline, geographic variation	Significant effect of lead on child behavioral problems, teen aggression, and young adult criminal behavior
Grönqvist, Nilsson and Robling (2014)	Sweden	1970s 1980s	Lead	Crime	Administrative	Phase-out of leaded gasoline, geographic variation	7 to 14% drop in crime; lead levels only affect crime above a threshold; impact is larger among low-SES families
Aizer, Currie, Simon and Vivier (2016)	Rhode Island	1997-2004	Lead	Education	Administrative	Policy requiring lead-safe certificates for rental homes, geographic variation in pre-policy exposure	The reduction of lead exposure in Rhode Island can alone explain 37-76% of the decline in racial disparities in test scores because African American children were previously disproportionately exposed to lead
Currie et al. (2009)	Texas	1996-2001	Air pollution	Education	Administrative	Geographic and temporal variation in pollution levels	The level of CO is significantly and positively correlated with school absences
Lavy, Ebenstein and Roth (2014)	Israel	2000-2002	Air pollution	Education	Administrative	Student fixed effects with temporal variation in pollution levels	Large and negative association between pollution and test performance
Marcotte (2015)	US	2003-2012	Pollen	Education	Aggregate test scores	State-year panel	Pollen levels in the weeks leading up to tests are negatively related to performance, but not as strongly as pollen levels on the day of the test
Bensnes (2016)	Norway	2008-2011	Pollen Air pollution	Education	Administrative	Student fixed effects with temporal variation in pollen levels	Pollen levels are significantly associated with performance on high-stakes tests
Isen, Rossin-Slater and Walker (2016)	US	1969-1974	Air pollution	SES	Administrative	Variation in air quality across birth counties	As air pollution levels were lowered, children born in these years went on to earn more

2.6.3 Studies

A recent literature shows that lead exposure is connected to worse academic performance and lower earnings and that when lead exposure is reduced by government policies, such as the phasing out of leaded gasoline, outcomes improve.

Using data from a city in Chile, [Rau, Reyes and Urza \(2013\)](#), connect lead exposure to academic performance and earnings. The variation that they exploit is the dumping of more than 20,000 tons of toxic waste in a city between 1984 and 1989 which interacted with the city's expansion, and led to lead exposure in a geographically heterogeneous way. They use data in which they can identify the distance of families to the waste and information on lead exposure. They connect these data on exposure to administrative data on test scores and adult earnings. They find large reductions in test scores and a significant impact on earnings.

Using data from the NLSY79 and the NLSY97, [Reyes \(2015\)](#) connects lead exposure to risky and antisocial behavior in adolescence. The identification in this paper is based on the phaseout of leaded gasoline in the US between 1975 and 1990. Over successive cohorts, blood lead levels were reduced over time but there was also some variation across states in the speed of the reduction. The data on blood levels comes from the National Health and Nutrition Examination Survey (NHANES II) and is combined with government data on leaded gasoline to establish the connection between the state-level geographic distribution of leaded gasoline and blood lead levels. The author then estimates the impact of blood lead levels on child behavioral problems, teen aggression, and young adult criminal behavior, finding significant effects.

In Sweden, [Grönqvist, Nilsson and Robling \(2014\)](#) show reductions in criminal behavior as response to the phasing out of leaded gasoline. The identification strategy is similar to [Reyes \(2015\)](#): leaded gasoline was phased out in Sweden in the 1970s and 1980s, but there was geographic variation in lead levels. However, the data used in this study is much better than in [Reyes \(2015\)](#). First, to study criminal behavior, the authors obtain data on

every criminal conviction in Sweden since 1985. Second, they obtain comprehensive data on every child born in three birth cohorts (1972-1974, 1977-1979, 1982-1984). Third, they obtain data on the municipality-level moss lead level for 1,000 survey locations for three years (1975, 1980, 1985). Connecting these datasets allows for a nuanced and precise analysis of the effects of lead exposure on crime. The authors find a large, 7 to 14% drop associated with the drop in blood levels caused by the phaseout and also describe substantial heterogeneity and nonlinearity. They find that lead levels only affect crime above a threshold and that the impact is larger among low-SES families.

[Aizer, Currie, Simon and Vivier \(2016\)](#), using data from Rhode Island, show improvements in test scores and argues that reduction in lead exposure is a major reason for the decrease in the black-white test score gap. The authors connect data on 57,000 children born in Rhode Island with data from blood lead level tests (available for approximately 80% of children in the state) and data on third grade test scores. The key variation that they exploit is generated by a 1997 policy that required the owners of rental properties to obtain “lead-safe” certificates. Bringing in neighborhood-level data on the share of old (pre-war) buildings that have much more leaded paint, the researchers exploit geographic variation in the change of blood levels among children living in different neighborhoods. They find that there is a strong negative association between elevated blood levels and test performance. They argue that the reduction of lead exposure in Rhode Island can alone explain 37-76% of the decline in racial disparities in test scores because African American children were previously disproportionately exposed to lead.

[Currie et al. \(2009\)](#) connect air quality data from Texas with school absences, finding a strong relationship. They adopt a difference-in-difference-in-differences strategy that only relies on variation in air pollution within schools and years, across attendance periods, six-week periods within a school year. They connect data on absences by each public school student in grades 1 to 8 studying at a school within 10 miles of a pollution monitor for the years 1996-2001. Their key outcome variable is a school-level average absentee rate for a six-

week attendance period, which they relate to a local measure of pollution (carbon monoxide (CO), ozone, and other small particulate matter (PM10)), namely for each pollutant, the percent of days in the attendance period the level of the pollutant exceeded various percentages (25%, 50%, 75%, 100%) of the EPA threshold. They find that the level of CO is significantly and positively correlated with school absences.

Using data Israel, the US, and Norway, respectively, [Lavy, Ebenstein and Roth \(2014\)](#), [Marcotte \(2015\)](#), and [Bensnes \(2016\)](#) find that air pollution and pollen have large negative impacts on high-stakes high school exam scores.

[Lavy, Ebenstein and Roth \(2014\)](#) uses data on the test scores of all students (about half of each cohort) who at the end of high school take the Bagrut exam, which is required to enter universities, for the years 2000-2002. They are able to connect these test scores at the individual student level to demographic information and to the fine particulate matter (PM2.5) and CO measure from the monitoring station nearest to the student's school. There are 139 monitoring stations, so the geographic variation is observed at a relatively fine level. Further strengthening identification, each student takes exams after each grade, the authors are able to include student fixed effects, using only within-student, across-exam variation. They find a large and negative association between pollution and test performance. Since these are high stakes exams that are important in university admissions, this means that the allocation of students to higher education is also affected.

In work very similar to [Lavy, Ebenstein and Roth \(2014\)](#), [Bensnes \(2016\)](#) uses universal data on high school exams for the 2008-2011 period from Norway, connected to pollen spread predicted from data collected at twelve stations across the country. Thus the variation that the author exploits comes from the variation of pollen levels across time and across municipalities. Additionally, he is able to include student fixed effects because each student takes exams on multiple days. Therefore identification comes only from within-student, across-exam date variation in pollen levels. He finds significant effects and argues that because only a portion of students are actually sensitive to pollen, for these students the

effects are even larger. Because the exams examined in the study are high stakes high school exams, they also influence the way students sort into higher education.

[Marcotte \(2015\)](#) is a closely related study which has less clean identification but uses data from the US. The author obtained data on daily pollen counts for 16 monitoring stations across America for the 2003-2012 period and connected it to third grade state test data from 48 school districts that are near the stations. He then estimates the relationship between pollen counts on the day the state test was taken and the proportion of students scoring proficient or advanced on the tests. He finds that pollen levels in the weeks leading up to tests are negatively related to performance, but not as strongly as pollen levels on the day of the test.

Using variation created by the 1970 Clean Air Act Amendments, [Isen, Rossin-Slater and Walker \(2016\)](#) find lower labor force participation and earnings among those exposed to air pollution in their year of birth. The amendments lead to the designation of each US county as an attainment or a nonattainment based on the level of pollution in 1970. Nonattainment counties were forced to lower pollution levels and experienced large improvements in air quality. The authors use data from 148 counties in 24 states, connecting data on the change in air pollution in these counties to the earnings records of 90% of the US population using a longitudinal dataset from the Census Bureau. They show that as air pollution levels were lowered, children born in these years went on to earn more.

2.6.4 Summary

The first group of studies discussed in this section on environmental health investigated the impacts of exposure to lead on various longer-term outcomes. It appears that exposure to lead, even below quantities below regulatory levels leads to worse academic outcomes, lower earnings, and more criminal behavior. The studies here rely on quasi-experimental variation, usually connected to the introduction of some policy (e.g., phase out of leaded gasoline) over time. This is useful both to better identify causal impacts rather than just associations, and

also because they show that feasible government policies can have substantial impacts in this area. The second group of studies concentrates on air pollution, in particular, pollution below regulatory levels. These are identified off of variation in weather which causes variation in pollution. Air pollution clearly impacts student performance, though if we take this literature together with the literature on air pollution and adult productivity discussed in Section 3.2.3, it appears that at these levels of pollution, at least in these studies, it primarily impacts one-time test scores rather than longer-term cognitive ability.

2.7 Summary

In this section we considered various aspects of childhood health and summarized what has been found in the literature on the association of these aspects of childhood health and longer-term outcomes. Clearly, childhood health, whether physical or mental health, in utero, early childhood, or adolescent health, matters for education and labor market market success. If one had to answer the question "what matters most?", ranking different areas of health, the difficulty would be that some effects are large but less well-identified, while others are cleanly identified but smaller. Studies of child mental health recover very large effects: children with certain mental conditions go on to earn 10-20-30% less, but the mental health studies are probably better understood as uncovering associations rather than causal effects. Historic studies of nutrition and public health interventions also recover large effects, and they have a more plausible claim at identification, but data limitation are of course important. The non-historic studies of nutrition are suggestive of associations, rather than causal effects. On the other hand, in utero exposure to negative shocks is associated with earnings declines more on the order of 10% or even less, but these effects are generally cleanly identified in a large number of studies, making use of both cross-cohort and cross-geography variation. Studies of environmental health are perhaps somewhat less identified, but still relatively clean, and usually find important though not huge effects. Birthweight studies uncover the smallest effects. The twin research design used in these studies has been challenged, though

birthweight studies are probably one of the most appropriate uses for this design. However, there are still important open issues. Below we highlight two of these.

First, the magnitude of the effect of various aspects of health on outcomes is often hard to establish. Many studies do not report the magnitude of their estimates clearly, they appear satisfied with finding that there is an effect but do not report it in easily interpretable and comparable units or put it into context. To evaluate which aspects of health are most important for longer-term outcomes and just as importantly in which areas we could have cost-effective interventions we would need to have clear, interpretable, and comparable estimates of magnitudes of effect sizes.

Second, the majority of studies that we discussed do not study interventions, but either rely on naturally arising variation or study exogenous shocks. This means that we do not have a direct handle on what sorts of interventions would be effective. Therefore one of the important directions for future research is the study, either through natural experiments relying on different times/places of introduction or through randomized controlled trials, of interventions that affect health and longer-term outcomes.

3 Adult health

3.1 Introduction

This section focuses on the impact of adult health on labor market outcomes, including labor force participation, employment, hours, and earnings. There are several ways in which health can impact labor market outcomes. In generic terms, it can affect both the costs and benefits of work. People in worse health might find it harder to work or be less productive at work. People with high out-of-pocket medical costs might also need to attain higher income. The generosity and other features of available social insurance programs will have a large impact on the association of health and labor market outcomes. For example, in the United States, job lock, the phenomenon that some people need to keep working at an

employer to keep their health insurance is another distinct channel through which health can affect employment. [Currie and Madrian \(1999\)](#) present a comprehensive review of the literature connecting adult health and labor markets generated between about 1980 and 1999. The authors conclude that poor health is associated with lower earnings and labor force participation, but the magnitude of the association is sensitive to the dimension of health measured. In this section, we describe more recent contributions linking health and labor market outcomes.

One of the reasons that it is difficult to make progress in this literature is that researchers have not been allowed to connect US administrative data on health with administrative data on labor market outcomes, such as wages and labor force participation. Much of the literature using US data is based on surveys, such as the Panel Study of Income Dynamics (PSID) and the Health and Retirement Study (HRS). Surveys have some important limitations: health and labor market outcomes are typically self-reported which means measures are potentially imprecise and endogenous and sample sizes are small.

There is a literature that relies on US Social Security Disability Insurance (SSDI) data, though much of this literature is concerned with the impact of receiving disability insurance (DI) payments, rather than with the effects of health on labor market outcomes, such as transition to disability. There are also studies that are able to rely on high-quality administrative data from Western and Northern European countries. Another recent strand of the literature has studied the dynamic relationship between health and labor market outcomes: instead of investigating whether people in poor health fare worse in the labor market, researchers have looked at what happens to the labor market outcomes of an individual when a negative health shock hits. An advantage of studying the response to health shocks over studying health status is that estimates are now within individual. Essentially, the question asked is holding a person's other characteristics fixed, what happens when her health "suddenly" worsens. The dynamics of health can also be endogenous, though some of the factors causing endogeneity might be observable in administrative panel data.

We organize this section around six main themes: (1) environmental health and productivity, (2) disability, (3) the consequences of physical health shocks, (4) within-household spillovers from health, (5) cancer, (6) mental health.

3.2 Environmental health and productivity

3.2.1 Introduction

A large literature, discussed in Section 2.6 establishes the consequences of the in utero and childhood environment on adult outcomes. This section focuses on recent work that estimates the impact of environmental health conditions on adults' contemporaneous productivity and wages. The most important environmental factor that has been linked to productivity is air pollution. The impact of different types of air pollution can be large even for workers working inside and need not be limited to developing countries. The studies that we discuss in this section have demonstrated the impact of air pollution on the productivity of agricultural workers in California, call-center workers in China, and the general working population in Mexico City.

This subsection is based on a systematic review which started with searching the economics literature on Google Scholar using the search terms ("environment" OR "environmental" OR "pollution" OR "exposure" OR "air") AND ("productivity" OR "labor") for years 2005-2017. Our search returned an initial 132 results, however we only consider 5 studies that are relevant to our review. These need to consider the impact of environmental exposures on economic outcomes, rather than the other way around, examine non-health endpoints (economic outcomes), and be done in the US or other developed countries. In addition to the Google Scholar search, we added a small number of studies that were referenced by the studies we found through the systematic search but did not themselves come up in the search (usually because they did not contain the keywords in their title). Mindful of the slow publication process and the fast pace of some of the literature covered here, we also included relevant NBER Working Papers from 2015-2017 from the Health and Labor Studies

programs.

3.2.2 Summary of methods

From a methodological perspective, the studies that are discussed in this section can be put into two groups. The first group includes [Graff Zivin and Neidell \(2012\)](#), [Chang, Graff Zivin, Gross and Neidell \(2016a\)](#), and [Chang, Graff Zivin, Gross and Neidell \(2016b\)](#). These studies rely on detailed productivity and wage data from a single company combined with day-to-day variation in pollution. The measurement of productivity is aided by the fact that workers are paid based some kind of piece rate (e.g., number of berries picked, phone calls made), which means that the companies keep precise measures of output and workers are incentivized to maximize output. Moreover, hours are also logged, allowing the researchers to distinguish extensive margin (work or not), intensive margin (hours worked), and productivity (output per hour worked) effects. Using panel data that follows the same worker across days, the authors are also able to control for worker fixed effects, further strengthening the identification.

The second group of studies includes [Lehmann and Wadsworth \(2011\)](#) and [Hanna and Oliva \(2015\)](#). These studies rely on more general variation in environmental conditions across cohorts, the exposure to nuclear fallout from Chernobyl and exposure to air pollution from an oil refinery, respectively. In these studies, the measurement of productivity is less precise, since they rely on a proxy for productivity, namely wages. Moreover, it is possible that other things change over time in addition to the change in the environment. Since large populations are affected, these studies are also measuring more general equilibrium type effects.

Table 11: Environmental health and productivity - Summary

Study	Area	Time	Exposure	Data	Identifying variation	Findings
Lehmann and Wadsworth (2011)	Ukraine	2003-2007	Chernobyl	ULMS	Residence in 1986	Adults who resided relatively close to Chernobyl are 13 percentage points more likely to report poor health 20 years later which translates into worse labor market performance
Graff Zivin and Neidell (2012)	California	2009-2010	Ozone	Company IT system	Day-to-day variation in pollution	A 10 ppb (one standard deviation / the decrease in federal ozone standard proposed by the EPA) increase in average ozone exposure decreases worker productivity by 5.5 percent
Chang, Graff Zivin, Gross and Neidell (2016a)	China	2010-2012	Air pollution	Company IT system	Day-to-day variation in pollution	A 10-unit increase in their air pollution increase (relative to a mean of 66 units) is associated with a 0.35 percent drop in the number of calls per workday
Chang, Graff Zivin, Gross and Neidell (2016b)	California	2001-2003	Air pollution	Company payroll records	Day-to-day variation in pollution	A 6 percent decrease in worker productivity is associated with a 10-unit (one standard deviation) change in pollution
Hanna and Oliva (2015)	Mexico	1989-1993	Air pollution	ENEU	Distance	The closure of a refinery had a significant negative impact on the level of sulfur-dioxide and the improvement in air quality had a positive effect on labor supply, people on average worked an extra 1.3 hours (3.5%) more.

3.2.3 Studies

In Section 2.2.5 we discussed studies of the in utero impact of nuclear fallout, including the fallout from Chernobyl. Turning to adults, [Lehmann and Wadsworth \(2011\)](#) study the impact of Chernobyl on adult health and labor market outcomes. Their data comes from the 2003, 2004, and 2007 waves of the Ukrainian Longitudinal Monitor Survey (ULMS) which surveyed about 9,000 Ukrainians on their health and labor market outcomes, but also asked them about their place of residence in 1986. The authors use 1986 place of residence as a source of variation in exposure to fallout from the catastrophe. They find that adults who resided relatively close to Chernobyl are 13 percentage points more likely to report poor health 20 years later. This negative impact on health translates into worse labor market performance.

[Graff Zivin and Neidell \(2012\)](#) document the impact of ozone pollution on productivity. The authors exploit variation across work days in ambient ozone concentration at a large farm in the Central Valley in California connected with daily measures of agricultural worker productivity. Worker productivity is typically difficult to measure, but in this context workers are paid through piece rate contracts, depending on the number of blueberries and grapes that they pick. This contract design allows the researchers to measure worker productivity precisely, as the number of pieces harvested, and to study the impact of environmental conditions on productivity. Moreover, they are able to observe the decision to work and the number of hours worked, which turn out not to be affected by ozone concentration. This is crucial for isolating the effect on productivity. The main finding is that higher ozone concentration has a substantial negative effect on worker productivity: a 10 ppb (approximately the cross-day standard deviation of ozone pollution and also the decrease in federal ozone standard proposed by the EPA) increase in average ozone exposure decreases worker productivity by 5.5 percent.

In the same vein, [Chang, Graff Zivin, Gross and Neidell \(2016a\)](#) connect data on Chinese call-center workers' productivity to daily measures of pollution and weather. Employees

receive performance pay and productivity can be measured precisely as the number of phone calls made, the length of phone calls, and the amount of time that a worker is available to take calls. Pollution is measured as an air pollution index, based on particulate matter in the air on a given day. The main finding is that productivity decreases substantially with higher pollution levels and the productivity decrease happens because workers take longer breaks when pollution is higher. The main estimate presented by the authors suggests that a 10-unit increase in their air pollution increase (relative to a mean of 66 units) is associated with a 0.35 percent drop in the number of calls per workday.

In another related study, [Chang, Graff Zivin, Gross and Neidell \(2016b\)](#) study the impact of air pollution on pear packers working indoors. Fine particulate matter has a significant negative impact on their productivity, while pollutants that do not go indoors do not impact output. As in the previous studies, workers are paid based on output (number of boxes packed), allowing the researchers to measure productivity precisely. In line with previous findings, indoor air pollution also does not affect labor supply but affects the productivity of labor. The main estimate points to a 6 percent decrease in worker productivity associated with a 10-unit (roughly one standard deviation) change in pollution.

Environmental health can potentially affect labor supply in addition to productivity. [Hanna and Oliva \(2015\)](#) use the closure of a large oil refinery in Mexico City as a plausibly exogenous shock to air pollution in the city. The authors first show that the closure of the refinery had a significant negative impact on the level of sulfur-dioxide, a pollutant with known negative health effects, in the air. Using labor market data from the National Employment Panel Survey of Mexico (ENEU), the authors show that the improvement in air quality in turn had a positive effect on labor supply, people on average worked an extra 1.3 hours (3.5%) more.

3.2.4 Summary

The studies discussed in this section demonstrate that environmental conditions or environmental health impact the productivity of labor significantly. The magnitudes are hard to compare across various pollutants but we consider the micro evidence presented well-identified. Beyond improving observable health, interventions that can improve environmental conditions are likely to improve the productivity of labor and perhaps even increase labor supply.

3.3 Disability

3.3.1 Introduction

Disability may be thought of as a reduction in one's ability to work or as very bad health. Receiving disability insurance (DI) payments can be regarded as a labor market outcome in itself, possibly as a consequence of health shocks. Several recent studies have analyzed the impact of becoming disabled on worker's earnings but also workers' transitions to DI from working following health shocks.

This subsection is based on a systematic review which started with searching the economics literature on Google Scholar using the search terms ("disability" OR "disabled") AND ("labor" OR "earnings" OR "income") for years 2005-2017. Our search returned an initial 41 results, however we only consider 5 studies that are relevant to our review. These need to consider the impact of disability (and not disability insurance) on economic outcomes, rather than the other way around, examine non-health endpoints (economic outcomes), and be done in the US or other developed countries. In addition to the Google Scholar search, we added a small number of studies that were referenced by the studies we found through the systematic search but did not themselves come up in the search (usually because they did not contain the keywords in their title). Mindful of the slow publication process and the fast pace of some of the literature covered here, we also included relevant NBER Working

Papers from 2015-2017 from the Health and Labor Studies programs.

3.3.2 Summary of methods

Typically in this literature studies rely on panel data, looking at what happens to a person over time and mostly use data from surveys such as the Panel Study of Income Dynamics (PSID) and the Health and Retirement Study (HRS). One of the key challenges in this literature has been the endogeneity of reported disability status to labor market outcomes. This concern is exacerbated in studies of disability rather than general physical health. While general physical health can also be endogeneous to labor market outcomes (i.e., a person's health can worsen when she loses her job, rather than the person losing her job because her health worsened), there are at least some relatively objective measures of health (e.g., did the person have a stroke?) and some studies are able to measure physical health from administrative data, such as hospital claims. On the other hand, disability status is not an objective condition. First, in surveys, it's quite probable that people are more likely to report themselves as being disabled if they are not working. (This likely happens with general physical health as well to some extent.) Second, we know that people are more likely to claim disability when their labor market prospects worsen or more generally as a substitute for other social programs (see e.g., [Black, Daniel and Sanders, 2002](#); [Autor and Duggan, 2003](#); [Borghans, Gielen and Luttmer, 2014](#)). The studies we discuss in section generally use some variant of work-limiting conditions rather than disability program participation as a measure of disability.

The general approach in this literature is to use panel data that follow individuals over time and look at what happens to individuals as they become disabled. This approach is able to account for personal characteristics that are fixed over time, but not for variables that may change in parallel with disability (e.g., labor market conditions). Some studies use matching methods to create a control group that does not become disabled and then compare the treatment and control groups rather than just relying on the within-person changes over

time. This matching method is credible to the extent that we think that individuals matched on observables are truly comparable.

Table 12: Disability - Summary

Study	Area	Time	Data	Approach	Findings
Charles (2003)	US	1968-1993	PSID	Panel with individual FE	The onset of disability is followed by a sharp drop in earnings, followed by strong recovery around two years after the onset, and then by a slower upward trend in the following years
Mok, Meyer, Charles and Achen (2008)	US	1968-1993	PSID	Panel with individual FE	Relative to Charles (2003) , earnings losses are in fact several times larger and recovery is somewhat weaker
Meyer and Mok (2013)	US	1968-2009	PSID	Panel with individual FE	In addition to findings in Mok, Meyer, Charles and Achen (2008) , large declines in earnings and consumption
Cutler, Meara and Richards-Shubik (2011)	US	1994-2008	HRS	Panel with individual FE	Among near-elderly workers, major health events make people apply for DI at high rates but there is also a lot of variation both across demographic groups and by underlying health status
Lechner and Vazquez-Alvarez (2011)	Germany	1984-2002	GSOEP	Panel with individual FE and matching	Employment is reduced by 9-13%, loss of income is small (due to the welfare state) and unemployment does not increase, that is, disabled individuals who would like to work are able to find jobs

3.3.3 Studies

[Charles \(2003\)](#) uses data from the PSID for the 1968-1993 period to study what happens to the earnings of an individual when their self-reported ability to work becomes limited because of a physical impairment which the paper takes as its definition of disability. An important limitation is disability is defined to work-limiting condition, which may introduce an upward bias in the effect of disability on earnings if people whose impairments reduce their earnings are more likely to report that they have such an impairment than people who are able to better adapt to their impairments. The paper relies on person fixed effects to eliminate bias due to person-specific factors. It also takes into account the evolution of self-reported disability after onset, a possible measure of the extent to which the disability is chronic. The main finding of the paper is that the onset of disability is followed by a sharp drop in earnings, followed by strong recovery around two years after the onset, and then by a slower upward trend in the following years. On balance, disability causes substantial earnings losses. Initial earnings losses are larger and recovery is weaker for older, nonwhite, and less educated individuals. In follow-up work, however, [Mok, Meyer, Charles and Achen \(2008\)](#) revised the original estimates in [Charles \(2003\)](#) and found that earnings losses are in fact several times larger and recovery is somewhat weaker. Using PSID data extended to 2009, [Meyer and Mok \(2013\)](#) also report large declines in earnings and consumption.

In related work, focusing on near-elderly (ages 50-64) workers, [Cutler, Meara and Richards-Shubik \(2011\)](#) study the extent to which older workers transition to DI from working when they experience a major adverse event. Using longitudinal data from the HRS, the authors follow individuals after they experience a major health shock, and analyze their job characteristics, earnings, and welfare program takeup. They find that at this age, major health events make people apply for DI at high rates but that there is also a lot of variation both across demographic groups and by underlying health status.

Using German survey data, [Lechner and Vazquez-Alvarez \(2011\)](#), study the effect of becoming disabled on employment, unemployment, labor force participation, and earnings.

The German Socio-Economic Panel (GSOEP) a large panel study of West German residents aged 40-59 for the years 1984-2002 contains detailed information of individuals' health, labor market variables, and other characteristics. Since disability is measured as a continuous rather than as a binary variable, the authors defined becoming disabled as moving from 0% disability to over 30% disability. They use matching to create a control group of individuals who don't become disabled. While employment is reduced by 9-13%, loss of income is small (due to the welfare state) and unemployment does not increase, that is, disabled individuals who would like to work are able to find jobs.

3.3.4 Summary

The studies we summarized on the impact of becoming disabled suggest that disability is associated with declines in employment. However, it is important to note that it is difficult to define disability, at least in survey data in the US, independent of transitioning from employment to DI. Clearly important research is missing on the possibility of transitioning back from DI to employment and on the employability of workers whom we currently consider disabled. Whether the decline in employment at the onset of disability is associated with a decline in earnings is dependent on the particular DI system of each country. In the US, DI payments are not particularly generous and being on DI makes it virtually impossible to work due to strict earnings rules. In European welfare states, DI payments are more generous but receiving partial DI payments and working at the same time is also possible. From a policy perspective, understanding how introducing some of the policies that are used in other countries (e.g., partial disability) would be a great avenue for further research.

3.4 Health shocks

3.4.1 Introduction

A large and growing part of the literature that connects health to economic outcomes for adults relies on health shocks for identification. The key idea behind this approach is that

while cross-sectional estimates are likely to suffer from omitted variable bias and reverse causality, estimates that rely on longitudinal data “within” a person are able to account for persistent individual-level unobservables. Moreover, the extent to which health shocks are sudden or exogenous is testable by estimating the extent to which they are associated with pre-trends on average. The impacts of health also may manifest over time, thus following individuals year-to-year before and after a health event can be especially informative.

This subsection is based on a systematic review which started with searching the economics literature on Google Scholar using the search terms "health shock" OR "health event" OR "injury" OR "injuries" OR "accident" for years 2005-2017. Our search returned an initial 92 results, however we only consider 13 studies that are relevant to our review. These need to consider the impact of health shocks on economic outcomes, rather than the other way around, examine non-health endpoints (economic outcomes), and be done in the US or other developed countries. In addition to the Google Scholar search, we added a small number of studies that were referenced by the studies we found through the systematic search but did not themselves come up in the search (usually because they did not contain the keywords in their title). Mindful of the slow publication process and the fast pace of some of the literature covered here, we also included relevant NBER Working Papers from 2015-2017 from the Health and Labor Studies programs.

3.4.2 Summary of methods

The studies we discuss in this section use longitudinal data to follow individuals before and after they suffer a health shock, like a hospitalization. The earlier papers rely on surveys like the HRS in the US, the GSOEP in Germany, and the European Community Household Panel (ECHP) in comparisons of several European countries. More recent work has utilized administrative data from Denmark, Austria, and the Netherlands. Administrative data, which allows researchers to connect health shock measures (e.g., claims from hospitalizations) with outcomes (e.g., income) at the individual level, allows for a lot more precision than

surveys. Not only does administrative data give precise measurements of the timing and size of shocks and of economic outcomes, it can also reduce selection concerns. Typically studies in this literature estimate event studies, effectively plotting outcomes over time relative to a health shock, which also allows us to assess whether there were pre-trends in outcomes before the shock happened, which could bias estimates of the impact of the shock and could also be indicative of endogeneity. In addition to event studies that include person fixed effects, many of the studies also use matching methods, comparing individuals who experienced a shock with a matched control group of individuals who did not.

Table 13: Health shocks - Summary

Study	Area	Time	Data	Approach	Findings
McClellan (1998)	US	1992-1994	HRS	Panel with individual FE	Health events are more prevalent among people with lower socioeconomic status and have varying consequences depending on the type of the event; major new health events make people much more likely to leave the labor force and to retire; new chronic illnesses have a weaker but still substantial effect on retirement; accidents on can make people not work for some time, but do not push them to retire
Bound, Schoenbaum, Stinebrickner and Waidmann (1999)	US	1992-1996	HRS	Panel with individual FE	Controlling for initial health levels, people whose health declines are more likely to leave the labor force, primarily applying for DI; evidence of adaptive behavior in the form of job changes following declines in health
Smith (2005)	US	1992-2000	HRS	Panel with individual FE	Health shocks among older workers have a negative impact on labor supply, household income, and wealth; the key channel is work behavior, rather than medical costs
Riphahn (1999)	Germany	1984-1994	GSOEP	Panel with individual FE	The probability of dropping out of the labor force triples, unemployment goes up by 90%; associated with significant drops in income, though government welfare programs allow people to recover some of the income they lose from not working
García-Gómez and López Nicolás (2006)	Spain	1994-2001	ECHP	Panel with individual FE and matching	Individuals are 5% less likely to remain in employment, and 3.5% more likely to become inactive after a health shock
García-Gómez (2011)	Europe	1994-2001	ECHP	Panel with individual FE and matching	In seven out of nine countries, the probability of employment drops significantly after a health shock, while in two countries (France and Italy) the drop is not statistically significant
Trevisan and Zantomio (2016)	Europe	2002-2013	ELSA SHARE	Panel with individual FE and matching	Labor market impacts of experiencing health shocks are large, and slightly higher for women; lower educated men are more likely to exit the labor force than higher educated men, but for women the relationship is reversed

Table 14: Health shocks - Summary

Study	Area	Time	Data	Approach	Findings
García-Gómez, van Kippersluis, O'Donnell and van Doorslaer (2013)	Netherlands	1998-2005	Administrative	Panel with individual FE and matching	Health shocks in the form of acute hospital admissions are associated with large reductions in employment and earnings with little recovery, and high rates of DI transition; effects are larger for older workers and lower-income individuals; important spousal spillovers as well, larger for male spouses
Møller Danø (2005)	Denmark	1981-2000	Administrative	Panel with individual FE and matching	Large impacts for men but little impact for women, both in terms of employment and earnings; effects are larger for lower-income workers
Crichton, Stillman and Hyslop (2011)	New Zealand	1999-2004	Administrative	Panel with individual FE and matching	Large drops in employment and earnings for injured workers relative to matched non-injured workers; effects are concentrated among workers who suffer injuries that result in a longer duration of injury compensation by insurance; impacts are larger for women, older workers, and individuals who have lower incomes and weaker labor market attachment
Halla and Zweimüller (2013)	Austria	1997-2007	Administrative	Panel with individual FE and matching	People who have an accident are more likely to be unemployed and also to transition to disability retirement; women, older workers, and blue-collar workers are more likely to be unemployed or leave the labor market after health shocks
Lindeboom, Lena-Nozal and van der Klaauw (2016)	UK	1981-2000	NCDS	Panel with individual FE	People who have accidents are much more likely to transition to disability, but if they don't transition to disability, their employment is unaffected
Dobkin, Finkelstein, Kluender and Notowidigdo (2016)	US	1992-2012 2000-2011	HRS Hospital records Credit reports	Panel with individual FE	Even when people have health insurance, non-elderly individuals face significant economic risks as they can lose their uninsured labor earnings

3.4.3 Studies

A relatively early study connecting health events to labor supply is [McClellan \(1998\)](#). Relying on data from just two waves of the HRS, the author studies what happens to individuals who experience a major health event, chronic illness, or accident relative to those who don't. The main limitation of this study relative to later studies that use health shocks is the shortness of its panel: it only contains two data points per person, which makes it impossible to evaluate pre-trends which can lead to selection. Health shocks have a large negative impact on labor supply. The author finds that health events are more prevalent among people with lower socioeconomic status, and have varying consequences depending on the type of the event. Major new health events make people much more likely to leave the labor force and to retire, new chronic illnesses have a weaker but still substantial effect on retirement. Accidents on the other hand can make people not work for some time, but do not push them to retire.

In another study using the HRS, [Bound, Schoenbaum, Stinebrickner and Waidmann \(1999\)](#) find that health changes, as opposed to health levels, are very important in explaining labor force transitions among older workers. Controlling for initial health levels, people whose health declines are more likely to leave the labor force, primarily applying for DI. There is also evidence of adaptive behavior in the form of job changes following declines in health.

In similar work, [Smith \(2005\)](#) extends this analysis to five waves of the HRS. This allows the author to present evidence on trends in economic outcomes preceding health events. There is a natural downward trend in labor force participation and earnings as people age, however, new health events are associated with large drops in the probability of working and in income. In addition to these immediate impacts, there is substantial spillover to later years, the drop in labor force participation and income continues. The intensive margin (hours worked) response is smaller than the extensive margin (probability of working) response. The immediate drop and the cumulative impact are both smaller for hours worked.

In the European context, [Riphahn \(1999\)](#) uncovers very large effects using survey data from Germany. Her data come from eleven waves of the GSOEP for the years 1984-1994.

Defining health shocks as sudden deteriorations in a person's health, she finds effects much larger than the studies in the American context: the probability of dropping out of the labor force triples, unemployment goes up by 90%. These impacts are associated with significant drops in income, though government welfare programs allow people to recover some of the income they lose from not working. A useful way to reconcile the American and European estimates might be to consider that labor force impacts are dependent on the availability of government welfare programs and the replacement rates under these programs. If the deterioration of health is associated with an increase in the disutility of working or lower wages, people will be more likely to stop working if they can rely on other sources of income.

Extending earlier findings for Spain presented in [García-Gómez and López Nicolás \(2006\)](#), [García-Gómez \(2011\)](#) uses the European Community Household Panel (ECHP), a panel survey that covered nine European countries (Belgium, Denmark, France, Greece, Ireland, Italy, the Netherlands, Portugal, and Spain), for 1994-2001, to study the effect of health shocks on employment and discusses how these effects differ across countries and welfare regimes. The author relies on propensity score matching techniques to come up with what she considers a suitable control group for individuals who suffer health shocks. A health shock is defined in two ways as: (i) the change of health status from "good" or "very good" to "fair", "bad", or "very bad" across survey waves, (ii) the report of a chronic physical or mental health problem, illness, or disability that changes across survey waves. In seven out of nine countries, the probability of employment drops significantly after a health shock, while in two countries (France and Italy) the drop is not statistically significant. Importantly, there is significant heterogeneity across countries in the size of the drop. The author argues that this heterogeneity is associated with the disability policy of each country. The key parameters of the disability programs that matter are generosity and the ability to work while drawing DI. In some countries, the generosity of unemployment insurance (UI) policies can also have an impact because after a health shock individuals can transition to either DI or UI.

In another study comparing several European countries, [Trevisan and Zantomio \(2016\)](#)

focus on particular kinds of acute health shocks (myocardial infarction, stroke, and cancer) and older workers. The set of countries included is wider and allows for a broader comparison of what the authors refer to as four different welfare models: Nordic (Sweden, Denmark, and Netherlands), the Anglo-Saxon (England), the Continental (Austria, Germany, France, Belgium, and Switzerland), the Mediterranean (Spain, Italy, and Greece), and the Eastern European (Czech Republic, Poland, Estonia, and Slovenia). The data used in this study come from the English Longitudinal Study of Ageing and the Survey of Health, Ageing and Retirement in Europe for the years 2002-2013. The authors compare the labor market outcomes of individuals who experience one of the types of health shocks and matched individuals who do not experience a shock. Labor market impacts of experiencing these kinds of health shocks are found to be large, and slightly higher for women. Interestingly, lower educated men are more likely to exit the labor force than higher educated men, but for women the relationship is reversed.

Focusing on only the Netherlands, [García-Gómez, van Kippersluis, O'Donnell and van Doorslaer \(2013\)](#) study the impact of health shocks on a variety of economic outcomes. The authors use administrative data from Dutch hospitals connected with administrative employment and earnings records, as well as DI and UI records. Health shocks are defined as acute hospitalizations, unscheduled admissions that require immediate treatment. The authors employ propensity score matching and a difference-in-differences estimator to estimate the impact of health shocks on economic outcomes. Connecting high-quality administrative records from the entire Dutch population allows the authors to derive precise estimates and analyze heterogeneity along various dimensions. They are also able to connect households and estimate spillover effects. The panel structure of the data allows them to analyze the time path of economic outcomes before and after the health shock. Their findings suggest that health shocks in the form of acute hospital admissions are associated with large reductions in employment and earnings with little recovery, and high rates of DI transition. The effects are larger for older workers and lower-income individuals. There are important

spousal spillovers as well, larger for male spouses.

[Møller Danø \(2005\)](#) conducts a similar analysis using road injuries as plausibly exogenous health shocks. The author relies on an administrative panel dataset that contains a 10% random sample of the Danish population for the years 1981-2000. She focuses on individuals ages 20-54 and connects employment and earnings data with hospital admissions for road accidents. As in other studies, a matching estimator is used to compare people who have an accident with people who don't have an accident but are otherwise similar. There are large impacts for men but little impact for women, both in terms of employment and earnings. The author suggests that this might be because women are much more likely to be employed in the public sector and therefore enjoy higher job security. Effects are also larger for lower-income workers.

Using matched employer-employee data from New Zealand, [Crichton, Stillman and Hyslop \(2011\)](#) find large drops in employment and earnings for injured workers relative to matched non-injured workers. These effects are concentrated among workers who suffer injuries that result in a longer duration of injury compensation by insurance, which we might think of as more serious injuries. The heterogeneity of these estimates is consistent with findings from other studies: impacts are larger for women, older workers, and individuals who have lower incomes and weaker labor market attachment.

In another study of unexpected accidents, [Halla and Zweimüller \(2013\)](#) study what happens when workers are injured while commuting to work. Using administrative workers' compensation data connected to social security data covering all Austrians, they follow individuals who have an accident in the years 2000-2002 and a matched sample of workers who do not have an accident. They study labor market outcomes in a difference-in-differences framework and find that people who have an accident are more likely to be unemployed and also to transition to disability retirement. There is some suggestive evidence that women, older workers, and blue-collar workers are more likely to be unemployed or leave the labor market after health shocks.

Using the National Child Development Survey (NCDS) study sample from the UK, [Lindeboom, Llena-Nozal and van der Klaauw \(2016\)](#), study the impact of health shocks on labor market outcomes and disability transitions. We discussed several studies in the section on children that use this dataset.³ For this study, the authors rely on individuals in the panel survey who are included after they turn 23 in 1981. Defining health shocks as hospitalizations for accidents, the authors find that people who have such accidents are much more likely to transition to disability, but if they don't transition to disability, their employment is unaffected.

A recent study by [Dobkin, Finkelstein, Kluender and Notowidigdo \(2016\)](#) uses hospital admissions data from California linked with consumer credit reports, as well as HRS survey data to study the financial consequences of hospital admissions among the elderly and among non-elderly with and without health insurance. The authors of this study take an event study approach, essentially describing the time path, following individuals using panel data, of the financial outcomes under study relative to a hospitalization event. The consequences that they consider are out-of-pocket medical expenses, income, and credit report outcomes, such as unpaid bills and bankruptcy. They find that even when they have health insurance, non-elderly individuals face significant economic risks as they can lose their uninsured labor earnings.

3.4.4 Summary

The first group of studies that we considered in this section demonstrate that in the US context, new health shocks are an important driver of labor market transitions in the form of reduced labor supply, retirement, and DI. The second group of studies shows a similar phenomenon in European countries. How much health shocks affect labor market behavior is very much dependent on the institutional environment in each country, particularly the generosity of DI and UI programs. Finally, in a recent US-based study [Dobkin, Finkelstein,](#)

³For more details on the NCDS, see Section [2.2.5](#).

[Kluender and Notowidigdo \(2016\)](#) demonstrated that in the US there is significant financial risk associated with health shocks. Taken together, the studies in this section suggest that studying the consequences of health shocks in adulthood is important for understanding labor market behavior and understanding what policies can be implemented to reduce their adverse consequences is an important area for future research.

3.5 Within-household spillovers

3.5.1 Introduction

Economic theory and empirical research has addressed within-household economic decision-making and intra-household dependencies in outcomes such as female labor supply starting with the seminal work of Gary Becker and Jacob Mincer in the 1960s. It is natural to also analyze the impact of health on labor market outcomes in the household context, studying outcomes such as the impact of health shocks on spousal labor supply. Like the impact of own health on labor supply, a priori the impact of spousal health on labor supply also has an ambiguous sign. It depends on the extent to which the labor supplies of couples are complements (one may want to spend more time with or care for a sick spouse) or substitutes (if someone is unable to work, their spouse might need to work more to recover income for the household). The structure and generosity of social insurance will also impact labor supply decisions in response to spousal health, much like it does for own health. This section summarize a relatively novel set of empirical studies of within-household spillovers of health on labor market outcomes.

This subsection is based on a systematic review which started with searching the economics literature on Google Scholar using the search terms "health" AND ("couples" OR "spouses" OR "spousal" OR "household") for years 2005-2017. Our search returned an initial 83 results, however we only consider 8 studies that are relevant to our review. These need to consider the spillover effects of health, examine non-health endpoints (economic outcomes), and be done in the US or other developed countries. In addition to the Google

Scholar search, we added a small number of studies that were referenced by the studies we found through the systematic search but did not themselves come up in the search (usually because they did not contain the keywords in their title). Mindful of the slow publication process and the fast pace of some of the literature covered here, we also included relevant NBER Working Papers from 2015-2017 from the Health and Labor Studies programs.

3.5.2 Summary of methods

Most of the studies discussed in this section rely on surveys because in the US connecting administrative health data to administrative records of labor market outcomes and connecting administrative records pertaining to spouses has not been possible. On the other hand, surveys like the HRS often explicitly incorporate households, interviewing all members' of a participant's household. In a novel study using Danish data, [Fadlon and Nielsen \(2015\)](#) connect administrative records on the health of spouses and their labor market behavior. The predominant empirical approach in these studies is an event study strategy, studying what happens to individuals before and after the health of their spouse declines.

Table 15: Within-household spillovers - Summary

Study	Area	Time	Data	Approach	Findings
Coile (2004)	US	1992-2002	HRS	Panel with individual FE	Large labor supply effects when a worker experiences a health shock, the effects are larger for men; when one's spouse experiences a shock, no response/decrease by women, small increase by men
Van Houtven and Coe (2010)	US	1992-2006	HRS (linked to SSA)	Panel with individual FE	Men are responsive to their wives' health shocks in terms of employment and hours worked, but do not retire earlier; women are not response to their husbands' health shocks in terms of employment and hours worked, but retire more quickly when their husband has a new chronic condition
Gupta, Kleinjans and Larsen (2015)	Denmark US	1993-2001 1994-2002	Administrative HRS	Panel with individual FE	Little difference in the probability of remaining employed in the two countries; in both countries, the diagnosis of a new cancer makes people 5-10 percentage points less likely to work and a stroke or heart attack makes them around 20 percentage points less likely to work, with the exception of Danish men for whom the effect is only 8 percentage points
Charles (1999)	US	1992-1994	HRS	Panel with individual FE	Husbands reduce their labor supply when their wives are in poor health; wives increase their labor supply when their husbands are in poor health
Wu (2003)	US	1992-1994	HRS	Panel with individual FE	The impact of a health shock experienced by a wife is much larger than the impact of a health shock experienced by a husband; the earnings impact of health shocks comes through retirement decisions; some medical expenditure impact
Au, Crossley and Schellhorn (2005)	Canada	1994-2001	NPHS	Panel with individual FE	For both genders, a one standard deviation increase in health is associated with a 15-20 percentage point increase in the probability of being employed
Disney, Emmerson and Wakefield (2006)	UK	1991-1998	BHPS	Panel with individual FE	Adverse health shocks have a strong predictive effect for retirement behavior
Fadlon and Nielsen (2015)	Denmark	1980-2011	Administrative	Panel with individual FE	Widows increase their labor force participation by 11% on average when their husband dies; widowers decrease their labor supply when their wife dies; there is no economically significant labor supply response by either gender to a severe spousal health shock

3.5.3 Studies

Using data from the HRS, [Coile \(2004\)](#) studies how older workers and their spouses change their labor supply after a health shock. Using 1992-2002 data, she defines a health shock as either an acute health event (heart attack, stroke, new cancer), the onset of new chronic illnesses (diabetes, lung disease, heart failure, arthritis), or an accidental injury or fall. There are large labor supply effects when a worker herself or himself experiences a health shock, the effects are larger for men. When one's spouse experiences a shock, in most cases we can't reject no response by women, though in some cases they actually decrease their labor supply, while for men there is a small increase.

[Van Houtven and Coe \(2010\)](#) study the impact of health shocks on spousal retirement decisions, taking into account financial incentives. The authors use the HRS data for the 1992-2006 waves linked to Social Security Administration (SSA) data. This means that they have access to health information (from the HRS) and employment and pension information (from the SSA). They use various definitions of health shocks, including heart attacks, new cancer diagnoses, new functional limitations, and inpatient hospitalizations. They conclude that financial incentives are more important for retirement decisions than health shocks in terms of effect sizes. Men are responsive to their wives' health shocks in terms of employment and hours worked, but do not retire earlier as a response to spousal health shocks. On the other hand, women are not responsive to their husbands' health shocks in terms of employment and hours worked, but do retire more quickly when their husband has a new chronic condition.

[Gupta, Kleinjans and Larsen \(2015\)](#) compare the impact of severe health shocks on the labor force participation of older (ages 56-64) workers in Denmark and in the US. In Denmark workers have universal health insurance with low out-of-pocket costs, while in the US workers typically have employer-provided health insurance and potentially face high out-of-pocket costs. Their US data comes from five waves of the HRS between 1994 and 2002, while from Denmark they use a 20% random sample of the universal population register

between 1993 and 2001. Contrary to expectations that Americans would be more likely to continue working after a health shock because of large out-of-pocket costs and insurance job lock, the authors find little difference in the probability of remaining employed in the two countries. Examining effects by country and gender, in both countries, the diagnosis of a new cancer makes people 5-10 percentage points less likely to work and a stroke or heart attack makes them around 20 percentage points less likely to work, with the exception of Danish men for whom the effect is only 8 percentage points. The authors argue that the lack of differential impact between the two countries is explained by differences in diagnostic propensities, treatments, and differential survival that counteract the work incentives created by the different health insurance regimes.

In early work examining spousal spillovers of health shocks, [Charles \(1999\)](#) used data from the first two waves (1992 and 1994) of the HRS to study the impact of one's husband's or wife's health and disability status on one's labor supply. Individuals' own disability has a large negative impact on their labor supply. But importantly, spousal health and disability has a different impact depending on gender. Husbands reduce their labor supply when their wives are in poor health, but wives increase their labor supply when their husbands are in poor health.

In a similar study, [Wu \(2003\)](#) uses the first two waves (1992 and 1994) of the HRS to study the economic consequences of health shocks for couples. Health shocks are defined as heart conditions, strokes cancers, lung diseases, and diabetes with an onset between the two waves. Economic status is defined as total household wealth. The main finding of the paper is that the impact of a health shock experienced by a wife is much larger than the impact of a health shock experienced by a husband. The earnings impact of health shocks comes through retirement decisions, and there is also some medical expenditure impact. The author argues that the asymmetry arises because illness experienced by wives results in an increase in living expenses but illness experienced by husbands does not.

[Au, Crossley and Schellhorn \(2005\)](#) use data from four waves (1994-1995, 1996-1997,

1998-1999, and 2000-2001) of the Canadian National Population Health Survey (NPHS) to study the impact of health shocks on the employment outcomes of older (ages 50-64) workers. They use both self-reported health measures and an objective health index and find that for both genders, a one standard deviation increase in health is associated with a 15-20 percentage point increase in the probability of being employed.

In a similar study, using data from the British Household Panel Survey (BHPS) for 1991-1998, [Disney, Emmerson and Wakefield \(2006\)](#) analyze the role health shocks play in older workers' (ages 50-64) retirement decisions. The BHPS is a representative longitudinal survey of British workers. The longitudinal structure of the data allows the authors to analyze the impact of changes in health on changes in labor force status and retirement. The main finding of the paper is adverse health shocks have a strong predictive effect for retirement behavior and this effect might well be symmetric.

In recent work, [Fadlon and Nielsen \(2015\)](#) use comprehensive administrative data from Denmark to study the impact of spousal health shocks and spousal death on labor supply. Their dataset contains all married and cohabiting Danish couples from 1980 and 2011, which allows the authors to minimize selection concerns and provide precise estimates of any impacts. 500,000 households experienced a spousal death, while 70,000 households experienced a severe health shock (heart attack or stroke) which did not result in death. Importantly, instead of comparing households that experience a death or severe health shock to households that do not, the authors are able to compare households who experience the shock with households who would experience the shock in the future. The main empirical finding of the paper is that widows increase their labor force participation by 11% on average when their husband dies, while widowers decrease their labor supply when their wife dies. On the other hand, there is no economically significant labor supply response by either gender to a severe spousal health shock, which is likely explained by the fact that Danish workers are well-insured against the economic consequences of health shocks.

3.5.4 Summary

The issue of within-household spillovers remains an important open area for research. With the exception of the recent [Fadlon and Nielsen \(2015\)](#), all studies rely on surveys, which means that sample sizes and precision are limited. The existing literature in this area establishes that within-household spillovers should be a first-order consideration for policy as labor supply responses to spousal health shocks, especially women's response to health shocks experienced by their husbands, are large. There is virtually no research on children's response to health shocks experienced by their parents. The most important limitation in this area appears to be access to US administrative data in which family members could be linked to each other.

3.6 Cancer

3.6.1 Introduction

A number of studies have focused specifically on the consequences of cancer, including its impact of labor supply. Cancer may be a physical health condition particularly amenable to study because its diagnosis is relatively objective and treatment guidelines are relatively uniform. We may also think that some forms of cancer can be interpreted as exogenous health shocks or at least more so than many other health problems that could be endogenous to labor market outcomes.

This subsection is based on a systematic review which started with searching the economics literature on Google Scholar using the search terms "cancer" AND ("labor" OR "work" OR "earnings" OR "income") for years 2005-2017. Our search returned 6 results. In addition to the Google Scholar search, we added a small number of studies that were referenced by the studies we found through the systematic search but did not themselves come up in the search (usually because they did not contain the keywords in their title), for a total of 9 studies. Mindful of the slow publication process and the fast pace of some of

the literature covered here, we also included relevant NBER Working Papers from 2015-2017 from the Health and Labor Studies programs.

3.6.2 Summary of methods

Four of the studies discussed in this section ([Bradley, Neumark, Bednarek and Schenk, 2005](#); [Bradley, Neumark, Luo and Bednarek, 2007](#); [Bradley, Neumark and Barkowski, 2013](#); [Moran, Short and Hollenbeak, 2011](#)) are quite unique in their use of datasets and surveys that are specifically focused on individuals with cancer and cancer survivors. The advantage of using surveys specific to cancer survivors is that very detailed information is available on these individuals and there is much more power to estimate effect sizes for them. The disadvantage of these datasets is that there is no clear control group to which cancer survivors can be compared. The authors bring in data from other surveys, the CPS and the PSID, and use matching techniques to establish a control group, but in general this is a less credible empirical approach than studies focused on one consistent population part of which experiences cancer. One study ([Bradley, Bednarek and Neumark, 2002](#)) relies on HRS data connected with social security earnings data. Three studies ([Heinesen and Kolodziejczyk, 2013](#); [Kolodziejczyk and Heinesen, 2016](#); [Heinesen et al., 2017](#)) use administrative data from Denmark and one study ([Jeon, 2017](#)) uses census and administrative data from Canada. The empirical approach in all of these studies is to follow individuals with cancer diagnoses over time and compare their labor market outcomes to the outcomes of a matched control groups. In addition, to increase credibility, [Heinesen and Kolodziejczyk \(2013\)](#) and [Kolodziejczyk and Heinesen \(2016\)](#) also present estimates based on comparing people who have been diagnosed with cancer with people who would be diagnosed in the future.

Table 16: Cancer - Summary

Study	Area	Time	Data	Approach	Findings
Bradley, Bednarek and Neumark (2002)	US	1992	HRS (linked to SSA)	Cross-sectional	Breast cancer survivors are 10 percentage points less likely to work; conditional on working, they work longer hours at higher wages, resulting in higher total earnings
Bradley, Neumark, Bednarek and Schenk (2005)	Detroit	2001-2002	MDCSS + CPS	Panel with individual FE	17 percentage point drop in employment six months after diagnosis; effect is twice as strong for African-American women
Bradley, Neumark, Luo and Bednarek (2007)	Detroit	2001-2002	MDCSS + CPS	Panel with individual FE	Breast cancer decreases the labor supply to a greater extent among women insured by their spouse's policy than among women insured through their own employment
Bradley, Neumark and Barkowski (2013)	Virginia	2007-2011	Survey + CPS	Panel with individual FE	Women whose health insurance is dependent on being employed have a weaker labor supply response to breast cancer diagnosis; difference in the probability of remaining employed is small and usually not statistically significant; reduction in hours worked is 8-11% smaller for women who continue to work and need their job for health insurance; survey results consistent with job lock
Moran, Short and Hollenbeak (2011)	Pennsylvania and Maryland	2002	PSCCS + PSID	Cross-sectional, difference-in-differences, matching	Employment rate of cancer survivors is 28-30 percentage points lower if they develop new cancers, and 3-4 percentage points lower if they remain cancer free; the reduction in hours worked is about 16 hours for the group with new cancers and 3-4 hours for survivors without new cancers
Heinesen and Kolodziejczyk (2013)	Denmark	2000-2004	Administrative	Panel with individual FE and matching	Individuals with basic education are 10 percentage points more likely to be out of the labor force; 9 percentage points more likely to be on disability pension; vocational education: 7 percentage points for labor force participation and 5.5 percentage points for disability pension; with further or higher education: 4.5 percentages points for labor force participation and 3.5 percentage points for disability pension
Kolodziejczyk and Heinesen (2016)	Denmark	2000-2006	Administrative	Panel with individual FE and matching	In the public sector the impact of breast cancer on the likelihood of being out of the labor force is 11.5 percentage points vs 3.8 percentage points for the low- and high-educated, in the private sector the corresponding estimates are 6.2 and 3.2 percentage points
Heinesen et al. (2017)	Denmark	2006-2008	Administrative	Panel	Return-to-work probability is correlated with job satisfaction but cannot explain the observed educational gradient
Jeon (2017)	Canada	1991-2003	Census + Administrative	Matching	Cancer survivors are 5 percentage points less likely to work and earn 10% less on average than individuals who didn't have cancer

3.6.3 Studies

Connecting data from the first 1992 wave of the HRS with longitudinal social security earnings data, [Bradley, Bednarek and Neumark \(2002\)](#) study what happens to women who survive breast cancer relative to women who never had breast cancer. They find that breast cancer survivors are approximately 10 percentage points less likely to work. However, conditional on working, they work longer hours at higher wages, resulting in higher total earnings. The authors think that the counterintuitive finding of higher wages, hours, and earnings is not driven by selection bias but might be consistent with survivors needing to replenish financial assets, survivors' need for a full-time job for health insurance, and other causal mechanisms.

In related work, [Bradley, Neumark, Bednarek and Schenk \(2005\)](#) use data from the Metropolitan Detroit Cancer Surveillance System (MDCSS) database to identify women diagnosed with breast cancer and analyze how their labor market attachment changes in the six months following diagnosis. They use a sample of women from the Current Population Survey (CPS) as controls for women diagnosed with cancer. The authors find a 17 percentage point drop in employment six months after diagnosis. The effect is about twice as strong for African-American women.

In a pair of studies [Bradley, Neumark, Luo and Bednarek \(2007\)](#) and [Bradley, Neumark and Barkowski \(2013\)](#) study how insurance impacts women's labor market response to breast cancer diagnosis. The authors follow 500 employed, married women after they are diagnosed with breast cancer. Initially, these women either have employment-contingent health insurance offered by their employer or are covered by their spouses' health insurance. They also construct a control group using the CPS. As expected, women whose health insurance is dependent on being employed have a weaker labor supply response to breast cancer diagnosis. The difference in the probability of remaining employed is small and usually not statistically significant. The reduction in hours worked is 8-11% smaller for women who continue to work and need their job for health insurance. The authors also present results from a survey where they ask respondents about their reasons for continuing to work, these are also consistent

with health insurance induced job lock after cancer diagnosis.

Using survey data on cancer survivors from the Penn State Cancer Survivor Survey (PSCSS), [Moran, Short and Hollenbeak \(2011\)](#) study the longer-term employment consequences of surviving cancer. The PSCSS followed a sample of 1,800 cancer survivors from four large medical centers in the late 1990s and early 2000s. The authors also create a matched control sample from the PSID. This study is focused on younger workers, ages 28-54, relative to studies that use the HRS. Relative to people who are not cancer survivors, the employment rate of cancer survivors is 28-30 percentage points lower if they develop new cancers, and 3-4 percentage points lower if they remain cancer free. On average, the reduction in hours worked is about 16 hours for the group with new cancers and 3-4 hours for survivors without new cancers.

Using administrative data from Denmark, [Heinesen and Kolodziejczyk \(2013\)](#) study what happens to people diagnosed with colorectal or breast cancer relative to individuals not diagnosed with these cancers. The richness of the comprehensive administrative data which contains all Danes diagnosed with breast cancer or colorectal cancer in the years 2000-2004 allows for several credible identification strategies. First, the authors use propensity score matching to identify a control group and then employ a difference-in-differences estimator to analyze the differences in outcomes over time between the cancer and the control group. They are also able to match people who are diagnosed with cancer with people who would be diagnosed with cancer later. They are also able to study heterogeneity by education and other factors. Three years after a cancer diagnosis, individuals with basic education are 10 percentage points more likely to be out of the labor force (9 percentage points more likely to be on disability pension). The estimates are smaller for people with vocational education (7 percentage points for labor force participation and 5.5 percentage points for disability pension) and for people with further or higher education (4.5 percentages points for labor force participation and 3.5 percentage points for disability pension). In follow-up work, [Kolodziejczyk and Heinesen \(2016\)](#) show that there are differences in the consequences of

cancer between the public and the private sector, and in particular the education gradients are different. While in the public sector the impact of breast cancer on the likelihood of being out the labor force is 11.5 percentage points vs 3.8 percentage points for the low- and high-educated (a significant difference), in the private sector the corresponding estimates are 6.2 and 3.2 percentage points (a marginally significant difference). In another follow-up article, [Heinesen et al. \(2017\)](#) bring in survey data matched with the same administrative datasets to consider the extent to which job satisfaction before cancer diagnosis can explain heterogeneity in return-to-work behavior after diagnosis. They find that return-to-work probability is correlated with job satisfaction but cannot explain the observed educational gradient.

[Jeon \(2017\)](#) links data from the Canadian Cancer Registry to administrative income records. Using matching techniques to adjust for differences between individuals who survive cancers and individuals who never have cancer, the author finds that cancer survivors are 5 percentage points less likely to work and earn 10% less on average than individuals who didn't have cancer.

3.6.4 Summary

Studies of cancer consistently find large labor supply effects, though naturally the magnitude would vary across contexts and types of cancer. In the US, our best evidence comes from a series of small-scale studies of women who survived breast cancer, while [Heinesen and Kolodziejczyk \(2013\)](#) present complementary evidence using Danish administrative data. As highlighted in other sections, important evidence is still lacking on what sorts of interventions could decrease the impact of cancer on labor market outcomes. To understand what interventions may be helpful, understanding the mechanisms through which cancer reduces employment will be important.

3.7 Mental health

3.7.1 Introduction

As we make progress on physical health, the relative importance of mental health in determining labor market outcomes may be increasing. However, research on this topic is quite difficult because even the measurement of mental health is challenging. Moreover, the endogeneity of mental health to labor market outcomes is likely quite strong and stronger than for physical health. This section summarizes the sparse literature on the impact of mental health on labor market outcomes with the caveat that most identification methods are not as credible as what we have seen for physical health.

This subsection is based on a systematic review which started with searching the economics literature on Google Scholar using the search terms ("mental" OR "psychiatric" OR "depression") AND ("labor" OR "work" OR "earnings" OR "income") for years 2005-2017. Our search returned an initial 74 results, however we only consider 11 studies that are relevant to our review. These need to consider the impact of mental health on economic outcomes, rather than the other way around, examine non-health endpoints (education or economic outcomes), and be done in the US or other developed countries. In addition to the Google Scholar search, we added a small number of studies that were referenced by the studies we found through the systematic search but did not themselves come up in the search (usually because they did not contain the keywords in their title). Mindful of the slow publication process and the fast pace of some of the literature covered here, we also included relevant NBER Working Papers from 2015-2017 from the Health and Labor Studies programs.

3.7.2 Summary of methods

Studies of the impact of mental health on labor market outcomes rely on various surveys such as the National Comorbidity Survey (NCS). Given the endogeneity of mental health to labor

market outcomes and the likely importance of important unobservables, several instrumental variables strategies have been proposed in the literature. The four most common instruments are (i) pre-18 mental health conditions, (ii) parents' mental health, (iii) religiosity, and (iv) the availability of social services and support. Studies usually report OLS estimates and 2SLS estimates using these instruments. The key methodological consideration is whether these instruments can be thought of as truly exogeneous. It is quite unlikely that in most cases they would not be associated with labor market outcomes except through their effect on mental health.

Table 17: Mental health - Summary

Study	Area	Time	Data	Instruments	Findings
Ettner, Frank and Kessler (1997)	US	1990-1992	NCS	Pre-18, Parental mental health	Recent psychiatric problems reduce employment rates by 11 percentage points; conditional on working, the reduction in hours is small, but the reduction in earnings is substantial
Hamilton, Merrigan and Dufrense (1997)	Montreal	1985-1987	Survey		Employment status affects mental health, but mental health also has a strong effect on employment
Berndt et al. (1998)	US	1996	Survey	Randomized trial	Strong evidence that when antidepressants reduce depressive symptoms, the subjective evaluation of work productivity improves
Kessler et al. (1999)	US	1990-1992 1996	NCS MIDUS		People with major depression have 1.5 to 3.2 more short-term work disability days
Dewa and Lin (2000)	Ontario	1990-1991	OHS		Mental health conditions have the largest impact on the number of partial effort and extra effort days
Alexandre and French (2001)	Miami-Dade County	1996-1997	Survey	Religiosity, Social support	Employment is 19 percentage points lower for the depressed; conditional on working, the depressed work 7 to 8 weeks less per year than the non-depressed
Marcotte and Wilcox-Gök (2003)	US	1990-1992	NCS	Family history	Mental illness has a much larger than average earnings effect at the lower tail of the earnings distribution
Chatterji, Alegria, Lu and Takeuchi (2007)	US	2002-2003	NLAAS	Past mental health, Religious activity	Mental health problems have a large negative effect on the labor market outcomes of Latino men and women; reduction in employment for Asian men is large; among Asian women, the negative labor market consequences of recent mental health problems are weaker and less consistent
Chatterji, Alegria and Takeuchi (2011)	US	2001-2003	NCS-R	Past mental health, Parental mental health	Mental disorders have a strong negative impact on the labor force participation and employment of males; association is somewhat less stable for women; impact of mental health on labor market outcomes is along the extensive: conditional on working, no impact on hours or earnings
Ojeda, Frank, McGuire and Gilmer (2010)	US	2002	NSDUH	Social support	Immigrants' labor supply is less responsive to mental health problems than the labor supply of natives
Frijters, Johnston and Shields (2014)	Australia	2002-2011	HILDA	Death of a friend	Declines in mental health reduce employment significantly and the response is stronger for older workers

3.7.3 Studies

[Ettner, Frank and Kessler \(1997\)](#) use data from the National Comorbidity Survey (NCS), a 1990-1992 interview survey of approximately 4,500 Americans of ages 15-54 to study the impact of mental health conditions on labor market outcomes, including employment, hours, and earnings. Mental health problems are measured as diagnosis of a set of conditions, including depression, anxiety disorder, panic disorder, and others. In addition to simple OLS and probit regressions of labor market outcomes on mental health diagnoses, the authors use an instrumental variables approach, instrumenting for adult mental health with psychiatric disorders experienced before age 18 and psychiatric disorders exhibited by the respondent's parents. The authors find that recent psychiatric problems reduce employment rates by around 11 percentage points. Conditional on working, the reduction in hours is small, but the reduction in earnings is substantial. The authors take this as evidence of a wage effect, which we might interpret as a negative impact of mental illness on productivity.

Using a survey of approximately 250 Montreal residents for the years 1985-1987, [Hamilton, Merrigan and Dufrense \(1997\)](#) study the impact of mental health on employment. Mental health is measured using the Psychiatric Symptom Index (PSI). In order to account for the endogeneity of mental health to employment, the authors use a maximum likelihood, simultaneous equation generalized probit model. They find that employment status affects mental health, but also mental health has a strong effect on employment.

In a rare study of a medical intervention, [Berndt et al. \(1998\)](#) study the productivity effects of providing individuals suffering from chronic depression with one of two antidepressant drugs for 12 weeks. The data come from a medical randomized clinical trial of Zoloft, a selective serotonin reuptake inhibitor (SSRI) drug which was compared to an older antidepressant drug called imipramine. 635 entered the trial and 493 completed it. Depression severity is measured using the Hamilton Depression Rating Scale (HAMD), while work performance is measured using six different indices which included questions on the subjective evaluation of work performance (Hamilton Depression Rating Scale, Cornell Dysthymia

Rating Scale, L.I.F.E., Beck Inventory Scale, Quality of Life Questionnaire, and SAS-SR). The authors find strong evidence that when antidepressants reduce depressive symptoms, the subjective evaluation of work productivity improves.

In another study of short-term productivity, [Kessler et al. \(1999\)](#) use data from the NCS and the Midlife Development in the United States (MIDUS) surveys. The authors report the correlation between 30-day major depression and short-term work disability. They find that the two are strongly related. People with major depression have 1.5 to 3.2 more short-term work disability days.

Using data from the Ontario Health Survey's Mental Health Supplement from 1990-1991, [Dewa and Lin \(2000\)](#) describe the impact of mental health problems on productivity. They find that controlling for SES variables and work conditions, mental health conditions have the largest impact on the number of partial effort and extra effort days. In other words, when people with mental health problems show up for work, they require greater effort to function at work.

[Alexandre and French \(2001\)](#) study the relationship between depression and substance use and employment in a poor area, Miami-Dade County, Florida. They use survey data collected locally in 1996 and 1997 and measure depression using the Zung Self-Rating Diagnosis Scale. They instrument for depression using religiosity and the availability of social support services. They find that depression is strongly associated with employment. In an environment where only 43 percent of the non-depressed are employed, employment is 24 percent, 19 percentage points lower for the depressed. Even conditional on working, the depressed worked 7 to 8 weeks less per year than the non-depressed.

In another analysis using the NCS, [Marcotte and Wilcox-Gök \(2003\)](#) estimate the impact of mental health at different points of the earnings distribution. They use family history of mental illness as an instrument for whether a worker suffers from mental health problems, measured using the diagnostic interview measurement from the survey. The authors find that mental illness has a much larger than average earnings effect at the lower tail of the

earnings distribution.

[Chatterji, Alegria, Lu and Takeuchi \(2007\)](#) use data from the National Latino and Asian American Study (NLAAS), a 2002-2003 survey of approximately 2,600 Latinos and 2,100 Asians. Mental health is measured by whether the individual met diagnostic criteria for at least one of fourteen mental health disorders in the past years. The authors use several methods to attempt to control for the endogeneity of mental health to employment, including controlling for a large set of demographic and other background variables, controlling for past mental health in order to isolate the impact of current mental health, and using various instruments for mental health (past mental health, religious activity). The authors find that mental health problems have a large negative effect on the labor market outcomes of Latino men and women. The reduction in employment for Asian men is large, but particularly among Asian women, the negative labor market consequences of recent mental health problems are weaker and less consistent. The authors argue that this difference between Latinos and Asians might arise because Asians are much more educated on average, and they have better resources to cope with mental health problems.

In another paper [Chatterji, Alegria and Takeuchi \(2011\)](#) use the National Comorbidity Survey-Replication (NCS-R), a nationally representative survey from 2001-2003 to study the impact of psychiatric disorders on labor market outcomes, including employment, hours worked, and earnings. The definition of mental health diagnosis is similar to their previous paper. In addition to OLS and bivariate probit estimates, the authors also try to come up with plausibly exogeneous instruments for mental disorders. They use the early onset of a disorder and parental mental health, controlling for a vector of personal characteristics as their instruments. They also present results from an alternative selection on observables approach, proxying for selection along unobserved dimensions with selection along observed dimensions. They find that mental disorders have a strong negative impact on the labor force participation and employment of males, the association is somewhat less stable for women. The impact of mental health on labor market outcomes is along the extensive

margin: conditional on working, they find no impact on hours or earnings.

Ojeda, Frank, McGuire and Gilmer (2010) study how the impact of mental health on work behavior differs between men and women and between immigrants and natives. They use data from the 2002 National Survey on Drug Use and Health and define mental illness based on the K6 scale of Mental Health and on the presence of Mania and Delusions symptoms. They attempt to instrument for mental illness using social support. They find that immigrants' labor supply is less responsive to mental health problems than the labor supply of natives.

Frijters, Johnston and Shields (2014) use panel data in order to overcome some of the identification challenges inherent in the cross-sectional studies. Their data comes from ten waves, between 2002 and 2011, of the Household, Income and Labour Dynamics in Australia (HILDA) survey. This is a fairly large (approximately 14,000 participants) and representative survey of the Australian population. They define mental health using the standard Short-Form General Health Survey (SF-36) measure. In addition to using the panel structure of the data to study the impact of changes in mental health on employment, they also employ an instrumental variable, the death of a close friend. They find that declines in mental health reduce employment significantly and the response is stronger for older workers.

3.7.4 Summary

As we have already highlighted, studying mental health is much more difficult than studying physical health for a number of reasons: diagnoses are less clear and studies tend to rely on surveys with small samples and a lot of potential measurement error. Nevertheless, studying the impact of mental health on labor market outcomes is becoming more and more important as treatments for physical health problems become more available and the health burden shifts more towards mental health. The opioid epidemic presents its own challenges not only for the healthcare system but for the labor market.

3.8 Summary

In this section we have highlighted several aspects of health that have well-identified and large impacts on labor market outcomes. Studies have produced micro evidence that environmental health impacts productivity negatively, health shocks are important in explaining labor market transitions, and cancer and mental health issues are important conditions for adult health and labor market outcomes. Here we highlight open issues and challenges in the literature.

The first issue is data access. In most of the subsections on adult health, there is typically one study from a Scandinavian country that brings well-identified evidence using administrative data on the consequences of health issues. However, currently virtually no studies in the American context are able to rely on administrative data to understand the impacts of health on labor market outcomes because data from the two contexts are not possible to link. Because the labor market impact of health is dependent on institutions (e.g., the generosity of disability insurance), extrapolating from evidence produced using Scandinavian data is difficult.

The second issue is the importance of within-household spillovers. There is some evidence establishing that health shocks impact spousal labor supply, but there is clearly more to do on this question, and there is no work on spillover on children's labor market outcomes. This is partly also a data issue as data linked across family members is required to make progress on these questions.

The third issue is the importance of mental health. Studying mental health is more difficult than studying physical health because of measurement issues. However, mental health problems are becoming increasingly important for understanding adults' labor market outcomes. For example, the recent opioid epidemic affects groups with weak labor market attachment, and so it has potentially large labor market consequences.

References

- Aizer, Anna.** 2008. “Peer Effects and Human Capital Accumulation: The Externalities of ADD.” National Bureau of Economic Research Working Paper 14354.
- Aizer, Anna, Janet Currie, Peter Simon, and Patrick Vivier.** 2016. “Do Low Levels of Blood Lead Reduce Children’s Future Test Scores?” National Bureau of Economic Research Working Paper 22558.
- Aizer, Anna, Laura Stroud, and Stephen Buka.** 2015. “Maternal Stress and Child Well-Being: Evidence from Siblings.” *Journal of Human Resources*, Forthcoming.
- Alexandre, Pierre K., and Michael T. French.** 2001. “Labor Supply of Poor Residents in Metropolitan Miami, Florida: The Role of Depression and the Co-Morbid Effects of Substance Use.” *Journal of Mental Health Policy and Economics*, 4(4): 161–173.
- Almond, Douglas.** 2006. “Is the 1918 Influenza Pandemic Over? Long-Term Effects of In Utero Influenza Exposure in the Post-1940 U.S. Population.” *Journal of Political Economy*, 114(4): 672–712.
- Almond, Douglas, and Bhashkar Mazumder.** 2005. “The 1918 Influenza Pandemic and Subsequent Health Outcomes: An Analysis of SIPP Data.” *American Economic Review Papers & Proceedings*, 95(2): 258–262.
- Almond, Douglas, and Bhashkar Mazumder.** 2011. “Health Capital and the Prenatal Environment: The Effect of Ramadan Observance during Pregnancy.” *American Economic Journal: Applied Economics*, 3(4): 56–85.
- Almond, Douglas, and Janet Currie.** 2011*a*. “Human Capital Development Before Age Five.” In *Handbook of Labor Economics*. Vol. 4b, , ed. Orley Ashenfelter and David Card, Chapter 15, 1315–1486. Elsevier.

- Almond, Douglas, and Janet Currie.** 2011*b*. “Killing Me Softly: The Fetal Origins Hypothesis.” *Journal of Economic Perspectives*, 25(3): 153–172.
- Almond, Douglas, Bhaskhar Mazumder, and Reyn van Ewijk.** 2015. “In Utero Ramadan Exposure and Children’s Academic Performance.” *Economic Journal*, 125(8): 1501–1533.
- Almond, Douglas, Janet Currie, and Valentina Duque.** 2017. “Childhood Circumstances and Adult Outcomes: Act II.” National Bureau of Economic Research Working Paper 23017.
- Almond, Douglas, Kenneth Y. Chay, and David S. Lee.** 2005. “The Costs of Low Birth Weight.” *Quarterly Journal of Economics*, 120(3): 1031–1083.
- Almond, Douglas, Lena Edlund, and Mårten Palme.** 2009. “Chernobyl’s Subclinical Legacy: Prenatal Exposure to Radioactive Fallout and School Outcomes in Sweden.” *Quarterly Journal of Economics*, 124(4): 1729–1772.
- Au, Doreen Wing Han, Thomas F. Crossley, and Martin Schellhorn.** 2005. “The Effect of Health Changes and Long-term Health on the Work Activity of Older Canadians.” *Health Economics*, 14(10): 999–1018.
- Autor, David H., and Mark G. Duggan.** 2003. “The Rise in the Disability Rolls and the Decline in Unemployment.” *Quarterly Journal of Economics*, 118(1): 157–206.
- Barker, David.** 1990. “The Fetal and Infant Origins of Adult Disease.” *BMJ*, 301(6761): 1111.
- Barker, David.** 1995. “Fetal Origins of Coronary Heart Disease.” *BMJ*, 311(6998): 171–174.
- Barker, David, and Clive Osmond.** 1986. “Infant Mortality, Childhood Nutrition, and Ischaemic Heart Disease in England and Wales.” *Lancet*, 327(8489): 1077–1081.

- Barreca, Alan I.** 2010. “The Long-term Economic Impact of In Utero and Postnatal Exposure to Malaria.” *Journal of Human Resources*, 45(4): 865–892.
- Behrman, Jere R., and Mark R. Rosenzweig.** 2004. “Returns to Birthweight.” *Review of Economics and Statistics*, 86(2): 586–601.
- Bengtsson, Tommy, and Jonas Helgertz.** 2015. “The Long Lasting Influenza: The Impact of Fetal Stress during the 1918 Influenza Pandemic on Socioeconomic Attainment and Health in Sweden 1968-2012.” IZA Discussion Paper 9327.
- Bensnes, Simon Søbstad.** 2016. “You Sneeze, You Lose: The Impact of Pollen Exposure on Cognitive Performance During High-stakes High School Exams.” *Journal of Health Economics*, 49(1): 1–13.
- Berndt, Ernst R., Stan N. Finkelstein, Paul E. Greenberg, Robert H. Howland, Alison Keith, A. John Rush, James Russell, and Martin B. Keller.** 1998. “Workplace Performance Effects from Chronic Depression and its Treatment.” *Journal of Health Economics*, 17(5): 511–535.
- Bharadwaj, Prashant, Juan Eberhard, and Christopher A. Neilson.** 2017. “Health at Birth, Parental Investments and Academic Outcomes.” *Journal of Labor Economics*, Forthcoming.
- Bharadwaj, Prashant, Matthew Gibson, Joshua Graff Zivin, and Christopher A. Neilson.** 2016. “Gray Matters: Fetal Pollution Exposure and Human Capital Formation.” *Journal of the Association of Environmental and Resource Economists*, Forthcoming.
- Bharadwaj, Prashant, Petter Lundborg, and Dan-Olof Rooth.** 2017. “Birth Weight in the Long Run.” *Journal of Human Resources*, Forthcoming.

- Black, Dan, Kermit Daniel, and Seth Sanders.** 2002. “The Impact of Economic Conditions on Participation in Disability Programs: Evidence from the Coal Boom and Bust.” *American Economic Review*, 92(1): 27–50.
- Black, Sandra E., Aline Bütikofer, Paul J. Devereux, and Kjell G. Salvanes.** 2013. “This Is Only a Test? Long-Run Impacts of Prenatal Exposure to Radioactive Fallout.” National Bureau of Economic Research Working Paper 18987.
- Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes.** 2007. “From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes.” *Quarterly Journal of Economics*, 122(1): 409–439.
- Black, Sandra, Paul J. Devereux, and Kjell Salvanes.** 2016. “Does Grief Transfer Across Generations? In-utero Deaths and Child Outcomes.” *American Economic Journal: Applied Economics*, 8(1): 193–223.
- Bleakley, Hoyt.** 2007. “Disease and Development: Evidence from Hookworm Eradication in the American South.” *Quarterly Journal of Economics*, 122(1): 73–117.
- Bleakley, Hoyt.** 2010. “Malaria Eradication in the Americas: A Retrospective Analysis of Childhood Exposure.” *American Economic Journal: Applied Economics*, 2(2): 1–45.
- Borghans, Lex, Anne C. Gielen, and Erzo F. P. Luttmer.** 2014. “Social Support Substitution and the Earnings Rebound: Evidence from a Regression Discontinuity in Disability Insurance Reform.” *American Economic Journal: Economic Policy*, 6(4): 34–70.
- Bound, John, Michael Schoenbaum, Todd R. Stinebrickner, and Timothy Waidmann.** 1999. “The Dynamic Effects of Health on the Labor Force Transitions of Older Workers.” *Labour Economics*, 6(2): 179–202.

- Bradley, Cathy J., David Neumark, and Scott Barkowski.** 2013. “Does Employer-provided Health Insurance Constrain Labor Supply Adjustments to Health Shocks? New Evidence on Women Diagnosed with Breast Cancer.” *Journal of Health Economics*, 32(5): 833–849.
- Bradley, Cathy J., David Neumark, Heather L. Bednarek, and Maryjean Schenk.** 2005. “Short-term Effects of Breast Cancer on Labor Market Attachment: Results from a Longitudinal Study.” *Journal of Health Economics*, 24(1): 137–160.
- Bradley, Cathy J., David Neumark, Zhehui Luo, and Heather L. Bednarek.** 2007. “Employment-contingent Health Insurance, Illness, and Labor Supply of Women: Evidence from Married Women with Breast Cancer.” *Health Economics*, 16(7): 719–737.
- Bradley, Cathy J., Heather L. Bednarek, and David Neumark.** 2002. “Breast Cancer Survival, Work, and Earnings.” *Journal of Health Economics*, 21(5): 757–779.
- Brown, Ryan, and Duncan Thomas.** 2016. “On the Long Term Effects of the 1918 U.S. Influenza Pandemic.” *Mimeo*.
- Case, Anne, and Christina Paxson.** 2008*a*. “Height, Health, and Cognitive Function at Older Ages.” *American Economic Review Papers & Proceedings*, 98(2): 463–467.
- Case, Anne, and Christina Paxson.** 2008*b*. “Stature and Status: Height, Ability, and Labor Market Outcomes.” *Journal of Political Economy*, 116(3): 499–532.
- Case, Anne, and Christina Paxson.** 2009. “Early Life Health and Cognitive Function in Old Age.” *American Economic Review Papers & Proceedings*, 99(2): 104–109.
- Case, Anne, and Christina Paxson.** 2010. “Causes and Consequences of Early-life Health.” *Demography*, 47(1): S65–S85.
- Case, Anne, Angela Fertig, and Christina Paxson.** 2005. “The Lasting Impact of Childhood Health and Circumstance.” *Journal of Health Economics*, 24(2): 365–389.

- Cesur, Resul, and Inas Rashad Kelly.** 2010. “From Cradle to Classroom: High Birth Weight and Cognitive Outcomes.” *Forum for Health Economics and Policy*, 13(2).
- Chang, Tom, Joshua Graff Zivin, Tal Gross, and Matthew Neidell.** 2016*a*. “The Effect of Pollution on Worker Productivity: Evidence from Call-Center Workers in China.” National Bureau of Economic Research Working Paper 22328.
- Chang, Tom, Joshua Graff Zivin, Tal Gross, and Matthew Neidell.** 2016*b*. “Particulate Pollution and the Productivity of Pear Packers.” *American Economic Journal: Economic Policy*, 8(3): 141–169.
- Charles, Kerwin Kofi.** 1999. “Sickness in the Family: Health Shocks and Spousal Labor Supply.” University of Michigan Gerald R. Ford School of Public Policy Working Paper 00-011.
- Charles, Kerwin Kofi.** 2003. “The Longitudinal Structure of Earnings Losses among Work-Limited Disabled Workers.” *Journal of Human Resources*, 38(3): 618–646.
- Chatterji, Pinka, Dohyung Kim, and Kajal Lahiri.** 2014. “Birth Weight and Academic Achievement in Childhood.” *Health Economics*, 23(9): 1013–1035.
- Chatterji, Pinka, Margarita Alegria, and David Takeuchi.** 2011. “Psychiatric Disorders and Labor Market Outcomes: Evidence from the National Comorbidity Survey-Replication.” *Journal of Health Economics*, 30(5): 858–868.
- Chatterji, Pinka, Margarita Alegria, Mingshan Lu, and David Takeuchi.** 2007. “Psychiatric Disorders and Labor Market Outcomes: Evidence from the National Latino and Asian American Study.” *Health Economics*, 16(10): 1069–1090.
- Chay, Kenneth Y., Jonathan Guryan, and Bhashkar Mazumder.** 2009. “Birth Cohort and the Black-White Achievement Gap: The Roles of Access and Health Soon After Birth.” National Bureau of Economic Research Working Paper 15078.

- Chay, Kenneth Y., Jonathan Guryan, and Bhashkar Mazumder.** 2014. “Early Life Environment and Racial Inequality in Education and Earnings in the United States.” National Bureau of Economic Research Working Paper 20539.
- Coile, Courtney C.** 2004. “Health Shocks and Couples’ Labor Supply Decisions.” National Bureau of Economic Research Working Paper 10810.
- Crichton, Sarah, Steven Stillman, and Dean Hyslop.** 2011. “Returning to Work from Injury: Longitudinal Evidence on Employment and Earnings.” *Industrial & Labor Relations Review*, 64(4): 765–785.
- Cuellar, Alison Evans, and Dhaval M. Dave.** 2015. “Causal Effects of Mental Health Treatment on Education Outcomes for Youth in the Justice System.” National Bureau of Economic Research Working Paper 21206.
- Currie, Janet.** 2009. “Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development.” *Journal of Economic Literature*, 47(1): 87–122.
- Currie, Janet, and Brigitte C. Madrian.** 1999. “Health, Health Insurance and the Labor Market.” In *Handbook of Labor Economics*. Vol. 3, , ed. Orley Ashenfelter and David Card, Chapter 50, 3309–3416. Elsevier.
- Currie, Janet, and Enrico Moretti.** 2007. “Biology as Destiny? Short- and Long-Run Determinants of Intergenerational Transmission of Birth Weight.” *Journal of Labor Economics*, 25(2): 231–264.
- Currie, Janet, and Mark Stabile.** 2006. “Child Mental Health and Human Capital Accumulation: The Case of ADHD.” *Journal of Health Economics*, 25(6): 1094–1118.

- Currie, Janet, and Mark Stabile.** 2009. "Mental Health in Childhood and Human Capital." In *The Problems of Disadvantaged Youth: An Economic Perspective.*, ed. Jonathan Gruber, 115–148. University of Chicago Press.
- Currie, Janet, and Rosemary Hyson.** 1999. "Is the Impact of Health Shocks Cushioned by Socioeconomic Status? The Case of Low Birthweight." *American Economic Review Papers & Proceedings*, 89(2): 245–250.
- Currie, Janet, Eric A. Hanushek, E. Megan Kahn, Matthew Neidell, and Steven G. Rivkin.** 2009. "Does Pollution Increase School Absences?" *Review of Economics and Statistics*, 91(4): 682–694.
- Currie, Janet, Joshua Graff Zivin, Jamie Mullins, and Matthew Neidell.** 2014. "What Do We Know About Short- and Long-Term Effects of Early-Life Exposure to Pollution?" *Annual Review of Resource Economics*, 6(1): 217–247.
- Currie, Janet, Mark Stabile, and Lauren Jones.** 2014. "Do Stimulant Medications Improve Educational and Behavioral outcomes for Children with ADHD?" *Journal of Health Economics*, 37(1): 58 – 69.
- Currie, Janet, Mark Stabile, Phongsack Manivong, and Leslie L. Roos.** 2010. "Child Health and Young Adult Outcomes." *Journal of Human Resources*, 45(3): 517–548.
- Cutler, David M., Ellen Meara, and Seth Richards-Shubik.** 2011. "Health Shocks and Disability Transitions Among Near-elderly Workers." National Bureau of Economic Research Retirement Research Center Paper NB11-08.
- Dalsgaard, Søren, Helena Skyt Nielsen, and Marianne Simonsen.** 2014. "Consequences of ADHD Medication Use for Children's Outcomes." *Journal of Health Economics*, 37(1): 137 – 151.

- Dewa, Carolyn S., and Elizabeth Lin.** 2000. “Chronic Physical Illness, Psychiatric Disorder and Disability in the Workplace.” *Social Science & Medicine*, 51(1): 41–50.
- Ding, Weili, Steven F. Lehrer, J.Niels Rosenquist, and Janet Audrain-McGovern.** 2009. “The Impact of Poor Health on Academic Performance: New Evidence Using Genetic Markers.” *Journal of Health Economics*, 28(3): 578–597.
- Disney, Richard, Carl Emmerson, and Matthew Wakefield.** 2006. “Ill Health and Retirement in Britain: A Panel Data-based Analysis.” *Journal of Health Economics*, 25(4): 621 – 649.
- Dobkin, Carlos, Amy Finkelstein, Raymond Kluender, and Matthew J. Notowidigdo.** 2016. “The Economic Consequences of Hospital Admissions.” National Bureau of Economic Research Working Paper 22288.
- Ettner, Susan L., Richard G. Frank, and Ronald C. Kessler.** 1997. “The Impact of Psychiatric Disorders on Labor Market Outcomes.” *Industrial & Labor Relations Review*, 51(1): 64–81.
- Fadlon, Itzik, and Torben Heien Nielsen.** 2015. “Household Responses to Severe Health Shocks and the Design of Social Insurance.” National Bureau of Economic Research Working Paper 21352.
- Feyrer, James, Dimitra Politi, and David N. Weil.** 2016. “The Cognitive Effects of Micronutrient Deficiency: Evidence from Salt Iodization in the United States.” *Journal of the European Economic Association*, Forthcoming.
- Figlio, David, Jonathan Guryan, Krzysztof Karbownik, and Jeffrey Roth.** 2014. “The Effects of Poor Neonatal Health on Children’s Cognitive Development.” *American Economic Review*, 104(12): 3921–3955.

- Fletcher, Jason M.** 2008. “Adolescent Depression: Diagnosis, Treatment, and Educational Attainment.” *Health Economics*, 17(11): 1215–1235.
- Fletcher, Jason M.** 2010. “Adolescent Depression and Educational Attainment: Results Using Sibling Fixed Effects.” *Health Economics*, 19(7): 855–871.
- Fletcher, Jason M.** 2014. “The Effects of Childhood ADHD on Adult Labor Market Outcomes.” *Health Economics*, 23(2): 159–181.
- Fletcher, Jason M., and Barbara Wolfe.** 2008. “Child Mental Health and Human Capital Accumulation: The Case of ADHD Revisited.” *Journal of Health Economics*, 27(3): 794–800.
- Fletcher, Jason M., and Steven F. Lehrer.** 2011. “Genetic Lotteries Within Families.” *Journal of Health Economics*, 30(4): 647–659.
- Frank, Richard G., and Ellen Meara.** 2009. “The Effect of Maternal Depression and Substance Abuse on Child Human Capital Development.” National Bureau of Economic Research Working Paper 15314.
- Frijters, Paul, David W. Johnston, and Michael A. Shields.** 2014. “The Effect of Mental Health on Employment: Evidence from Australian Panel Data.” *Health Economics*, 23(9): 1058–1071.
- García-Gómez, Pilar.** 2011. “Institutions, Health Shocks and Labour Market Outcomes Across Europe.” *Journal of Health Economics*, 30(1): 200–213.
- García-Gómez, Pilar, and Angel López Nicolás.** 2006. “Health Shocks, Employment and Income in the Spanish Labour Market.” *Health Economics*, 15(9): 997–1009.
- García-Gómez, Pilar, Hans van Kippersluis, Owen O’Donnell, and Eddy van Doorslaer.** 2013. “Long Term and Spillover Effects of Health Shocks on Employment and Income.” *Journal of Human Resources*, 48(4): 873–909.

- Graff Zivin, Joshua, and Matthew Neidell.** 2012. “The Impact of Pollution on Worker Productivity.” *American Economic Review*, 102(7): 3652–3673.
- Graff Zivin, Joshua, and Matthew Neidell.** 2013. “Environment, Health, and Human Capital.” *Journal of Economic Literature*, 51(3): 689–730.
- Greve, Jane, Marie Louise Schultz-Nielsen, and Erdal Tekin.** 2015. “Fetal Malnutrition And Academic Success: Evidence From Muslim Immigrants In Denmark.” National Bureau of Economic Research Working Paper 21545.
- Grönqvist, Hans, J Peter Nilsson, and Per-Olof Robling.** 2014. “Early-childhood Lead Exposure and Criminal Behavior: Lessons from the Swedish Phase-out of Leaded Gasoline.” *Mimeo*.
- Gupta, Nabanita Datta, Kristin J. Kleinjans, and Mona Larsen.** 2015. “The Effect of a Severe Health Shock on Work Behavior: Evidence from Different Health Care Regimes.” *Social Science & Medicine*, 136-137(1): 44–51.
- Halla, Martin, and Martina Zweimüller.** 2013. “The Effect of Health on Earnings: Quasi-experimental Evidence from Commuting Accidents.” *Labour Economics*, 24(1): 23–38.
- Hamilton, Vivian H., Philip Merrigan, and Éric Dufrense.** 1997. “Down and Out: Estimating the Relationship Between Mental Health and Unemployment.” *Health Economics*, 6(4): 397–406.
- Hanna, Rema, and Paulina Oliva.** 2015. “The Effect of Pollution on Labor Supply: Evidence from a Natural Experiment in Mexico City.” *Journal of Public Economics*, 122(1): 68–79.

- Heinesen, Eskil, and Christophe Kolodziejczyk.** 2013. “Effects of Breast and Colorectal Cancer on Labour Market Outcomes—Average Effects and Educational Gradients.” *Journal of Health Economics*, 32(6): 1028–1042.
- Heinesen, Eskil, Christophe Kolodziejczyk, Jacob Ladenburg, Ingelise Andersen, and Karsten Thielen.** 2017. “Return to Work After Cancer and Pre-Cancer Job Dissatisfaction.” *Applied Economics*, Forthcoming.
- Isen, Adam, Maya Rossin-Slater, and W. Reed Walker.** 2016. “Every Breath You Take - Every Dollar You’ll Make: The Long-Term Consequences of the Clean Air Act of 1970.” *Journal of Political Economy*, Forthcoming.
- Jeon, Sung-Hee.** 2017. “The Long-Term Effects of Cancer on Employment and Earnings.” *Health Economics*, 26(5): 671–684.
- Karlsson, Martin, Therese Nilsson, and Stefan Pichler.** 2014. “The Impact of the 1918 Spanish Flu Epidemic on Economic Performance in Sweden: An Investigation Into the Consequences of an Extraordinary Mortality Shock.” *Journal of Health Economics*, 36(1): 1–19.
- Kelly, Elaine.** 2011. “The Scourge of Asian Flu: In utero Exposure to Pandemic Influenza and the Development of a Cohort of British Children.” *Journal of Human Resources*, 46(4): 669–694.
- Kessler, Ronald C., Catherine Barber, Howard G. Birnbaum, Richard G. Frank, Paul E. Greenberg, Robert M. Rose, Gregory E. Simon, and Philip Wang.** 1999. “Depression In The Workplace: Effects On Short-Term Disability.” *Health Affairs*, 18(5): 163–171.
- Kolodziejczyk, Christophe, and Eskil Heinesen.** 2016. “Labour Market Participation After Breast Cancer for Employees from the Private and Public Sectors: Educational and Sector Gradients in the Effect of Cancer.” *Economics & Human Biology*, 21(1): 33–55.

- Lavy, Victor, Analia Schlosser, and Adi Shany.** 2016. “Out of Africa: Human Capital Consequences of In Utero Conditions.” National Bureau of Economic Research Working Paper 21894.
- Lavy, Victor, Avraham Ebenstein, and Sefi Roth.** 2014. “The Impact of Short Term Exposure to Ambient Air Pollution on Cognitive Performance and Human Capital Formation.” National Bureau of Economic Research Working Paper 20648.
- Lechner, Michael, and Rosalia Vazquez-Alvarez.** 2011. “The Effect of Disability on Labour Market Outcomes in Germany.” *Applied Economics*, 43(4): 389–412.
- Lee, Chulhee.** 2014. “In Utero Exposure to the Korean War and its Long-Term Effects on Socioeconomic and Health Outcomes.” *Journal of Health Economics*, 33(1): 76–93.
- Lehmann, Hartmut, and Jonathan Wadsworth.** 2011. “The Impact of Chernobyl on Health and Labour Market Performance.” *Journal of Health Economics*, 30(5): 843–857.
- Lindeboom, Maarten, Ana Llana-Nozal, and Bas van der Klaauw.** 2016. “Health Shocks, Disability and Work.” *Labour Economics*, 43(1): 186–200.
- Lin, Ming-Jen, and Elaine M. Liu.** 2014. “Does In Utero Exposure to Illness Matter? The 1918 Influenza Epidemic in Taiwan as a Natural Experiment.” *Journal of Health Economics*, 37(1): 152–163.
- Lin, Ming-Jen, Jin-Tan Liu, and Shin-Yi Chou.** 2007. “As Low Birth Weight Babies Grow, Can Well-Educated Parents Buffer This Adverse Factor? A Research Note.” *Demography*, 44(2): 335–343.
- Lundborg, Petter, Anton Nilsson, and Dan-Olof Rooth.** 2014. “Adolescent Health and Adult Labor Market Outcomes.” *Journal of Health Economics*, 37(1): 25–40.
- Marcotte, Dave E.** 2015. “Allergy Test: Seasonal Allergens and Performance in School.” *Journal of Health Economics*, 40(1): 132–140.

- Marcotte, Dave E., and Virginia Wilcox-Gök.** 2003. "Estimating Earnings Losses due to Mental Illness: A Quantile Regression Approach." *Journal of Mental Health Policy and Economics*, 6(3): 123–134.
- McClellan, Mark B.** 1998. "Health Events, Health Insurance, and Labor Supply: Evidence from the Health and Retirement Survey." In *Frontiers in the Economics of Aging.*, ed. David A. Wise, 301–350. University of Chicago Press.
- Meyer, Bruce D., and Wallace K.C. Mok.** 2013. "Disability, Earnings, Income and Consumption." National Bureau of Economic Research Working Paper 18869.
- Miller, Paul, Charles Mulvey, and Nick Martin.** 2005. "Birth Weight and Schooling and Earnings: Estimates from a Sample of Twins." *Economics Letters*, 86(3): 387–392.
- Mok, Wallace K.C., Bruce D. Meyer, Kerwin Kofi Charles, and Alexandra C. Achen.** 2008. "A Note on "The Longitudinal Structure of Earnings Losses among Work-Limited Disabled Workers"." *Journal of Human Resources*, 43(3): 721–728.
- Møller Danø, Anne.** 2005. "Road Injuries and Long-run Effects on Income and Employment." *Health Economics*, 14(9): 955–970.
- Moran, John R., Pamela Farley Short, and Christopher S. Hollenbeak.** 2011. "Long-term Employment Effects of Surviving Cancer." *Journal of Health Economics*, 30(3): 505–514.
- Nakamuro, Makiko, Yuka Uzuki, and Tomohiko Inui.** 2013. "The Effects of Birth Weight: Does Fetal Origin Really Matter for Long-run Outcomes?" *Economics Letters*, 121(1): 53–58.
- Neelsen, Sven, and Thomas Stratmann.** 2011. "Effects of Prenatal and Early Life Malnutrition: Evidence from the Greek Famine." *Journal of Health Economics*, 30(1): 479–488.

- Neelsen, Sven, and Thomas Stratmann.** 2012. “Long-run Effects of Fetal Influenza Exposure: Evidence from Switzerland.” *Social Science & Medicine*, 74(1): 58–66.
- Niemesh, Gregory T.** 2015. “Ironing Out Deficiencies: Evidence from the United States on the Economic Effects of Iron Deficiency.” *Journal of Human Resources*, 50(4): 910–958.
- Nilsson, J Peter.** 2016. “Alcohol Availability, Prenatal Conditions, and Long-term Economic Outcomes.” *Journal of Political Economy*, Forthcoming.
- Ojeda, Victoria D., Richard G. Frank, Thomas G. McGuire, and Todd P. Gilmer.** 2010. “Mental Illness, Nativity, Gender and Labor Supply.” *Health Economics*, 19(4): 396421.
- Oreopoulos, Philip, Mark Stabile, Randy Walld, and Leslie Roos.** 2008. “Short-, Medium-, and Long-term Consequences of Poor Infant Health: An Analysis Using Siblings and Twins.” *Journal of Human Resources*, 43(1): 88–138.
- Persico, Claudia, David Figlio, and Jeffrey Roth.** 2016. “Inequality Before Birth: The Developmental Consequences of Environmental Toxicants.” National Bureau of Economic Research Working Paper 22263.
- Persson, Petra, and Maya Rossin-Slater.** 2016. “Family Ruptures, Stress, and the Mental Health of the Next Generation.” *American Economic Review*, Forthcoming.
- Rau, Toms, Loreto Reyes, and Sergio S. Urza.** 2013. “The Long-term Effects of Early Lead Exposure: Evidence from a case of Environmental Negligence.” National Bureau of Economic Research Working Paper 18915.
- Reyes, Jessica Wolpaw.** 2015. “Lead Exposure and Behavior: Effects on Antisocial and Risky Behavior among Children and Adolescents.” *Economic Inquiry*, 53(3): 1580–1605.
- Richter, André, and Bhashkar Mazumder.** 2016. “Second Generation Effects of Prenatal Health Shocks: Disentangling Social from Biological Pathways.” *Mimeo*.

- Richter, André, and Per Olof Robling.** 2016. “Multigenerational Effects of the 1918-19 Influenza Pandemic on Educational Attainment: Evidence from Sweden.” *Mimeo*.
- Riphahn, Regina.** 1999. “Income and Employment Effects of Health Shocks - A Test Case for the German Welfare State.” *Journal of Population Economics*, 12(3): 363–389.
- Royer, Heather.** 2009. “Separated at Girth: US Twin Estimates of the Effects of Birth Weight.” *American Economic Journal: Applied Economics*, 1(1): 49–85.
- Salm, Martin, and Daniel Schunk.** 2012. “The Relationship Between Child Health, Developmental Gaps, and Parental Education: Evidence from Administrative Data.” *Journal of the European Economic Association*, 10(6): 1425–1449.
- Sanders, Nicholas J.** 2012. “What Doesn’t Kill You Makes You Weaker: Prenatal Pollution Exposure and Educational Outcomes.” *Journal of Human Resources*, 47(3): 826–850.
- Smith, James.** 2005. “Consequences and Predictors of New Health Events.” In *Analyses in the Economics of Aging*, ed. David A. Wise, 213–240. University of Chicago Press.
- Smith, James P.** 2009. “The Impact of Childhood Health on Adult Labor Market Outcomes.” *Review of Economics and Statistics*, 91(3): 478–489.
- Smith, James P., and Gillian C. Smith.** 2010. “Long-term Economic Costs of Psychological Problems During Childhood.” *Social Science & Medicine*, 71(1): 110–115.
- Trevisan, Elisabetta, and Francesca Zantomio.** 2016. “The Impact of Acute Health Shocks on the Labour Supply of Older Workers: Evidence from Sixteen European Countries.” *Labour Economics*, 43(1): 171–185.
- Van Houtven, Courtney H., and Norma B. Coe.** 2010. “Spousal Health Shocks and the Timing of the Retirement Decision in the Face of Forward- Looking Financial Incentives.” Boston College Center for Retirement Research Working Paper 2010-7.

- von Hinke Kessler Scholder, Stephanie, George L. Wehby, Sarah Lewis, and Luisa Zuccolo.** 2014. "Alcohol Exposure In Utero and Child Academic Achievement." *Economic Journal*, 124(3): 634–667.
- Webbink, Dinand, Sunčica Vujić, Pierre Koning, and Nicholas G. Martin.** 2012. "The Effect of Childhood Conduct Disorder on Human Capital." *Health Economics*, 21(8): 928–945.
- Wu, Stephen.** 2003. "The Effects of Health Events on the Economic Status of Married Couples." *Journal of Human Resources*, 38(1): 219–230.
- Xie, Zong-Xian, Shin-Yi Chou, and Jin-Tan Liu.** 2016. "The Short-Run and Long-Run Effects of Birth Weight: Evidence from Large Samples of Siblings and Twins in Taiwan." *Health Economics*, Forthcoming.