Working Paper No. 316

Investigating survivorship bias: The case of the 1918 flu pandemic

Joël Floris, Laurent Kaiser, Harald Mayr, Kaspar Staub and Ulrich Woitek

Revised version, March 2021

Investigating survivorship bias:
The case of the 1918 flu pandemic*

Joël Floris[a,b], Laurent Kaiser[c,d], Harald Mayr[e], Kaspar Staub[a,f], Ulrich Woitek[g,h]

[a] *Institute of Evolutionary Medicine, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland*
[b] *Department of History, University of Zurich, Karl Schmid-Strasse 4, 8006 Zurich, Switzerland*
[c] *Department of Medicine, Infectious Diseases Service, University Hospital of Geneva and Faculty of Medicine, University of Geneva, Rue Gabrielle-Perret-Gentil 4, 1205 Geneva, Switzerland*
[d] *Department of Genetical and Laboratory Medicine, Virology Laboratory, Laboratory Medicine, University Hospital of Geneva, Rue Gabrielle-Perret-Gentil 4, 1205 Geneva, Switzerland*
[e] *ETH Zurich, Clausiusstrasse 37, 8092 Zurich, Switzerland*
[f] *Institute of History, University of Bern, Länggassstrasse 49, 3012 Bern, Switzerland*
[g] *Department of Economics, University of Zurich, Zürichbergstrasse 14, 8032 Zurich, Switzerland*
[h] *CESifo, Poschingerstr. 5, 81679 Munich, Germany*

March 5, 2021

Conflict of interest statement: The authors have nothing to declare.

---

*Corresponding author: Harald Mayr, ETH Zurich, Clausiusstrasse 37, 8092 Zurich, Switzerland; email: harald.mayr@econ.gess.ethz.ch; phone: +41 44 632 97 15

Investigating survivorship bias:
The case of the 1918 flu pandemic

Abstract

Estimates of the effect of fetal health shocks may suffer from survivorship bias. The fetal origins literature seemingly agrees that survivorship bias is innocuous in the sense that it induces a bias toward zero. Arguably, however, selective mortality can imply a bias away from zero. In the case of the 1918 flu pandemic, a suppressed immune system may have been protective against the most severe consequences of infection. We use historical birth records from the maternity hospital of Bern, Switzerland, to evaluate this possibility. Our results suggest that a careful consideration of survivorship bias is imperative for the evaluation of the 1918 flu pandemic and other fetal health shocks.

1   Introduction

A health shock in utero may kill some fetuses, a situation is often referred to as "culling". The very notion of culling implies a negatively selected "cull" or, conversely, positively selected survivors. It further implies that "[...] *estimates of the effects of fetal health shocks are generally conservative when the shock also increases mortality.*" (Almond and Currie, 2011, p. 165-166) We discuss this assumption in the context of the 1918 flu pandemic.

Almond (2006) uses the 1918 flu pandemic as a natural experiment. His findings—strong negative effects on socioeconomic status, education, and labor market outcomes—proved highly influential. To date, the effects of fetal exposure to the 1918 flu pandemic and other influenza strains were studied comprehensively (recent contributions include e.g. Brown and Thomas, 2018; Schwandt, 2018). A sizable body of research studies various other fetal health shocks (Almond *et al.*, 2017). Survivorship bias concerns this entire literature (Nobles and Hamoudi, 2019).

The 1918 flu pandemic featured unusually high mortality rates among young and healthy individuals. Childhood exposure to related flu strains help explain this pattern (Worobey *et al.*, 2014) and "*vigorous immune responses directed against the virus in healthy young persons could have caused severe disease in 1918*" (Morens *et al.*, 2010, p. e13). The role of the immune system in influenza infection is highly complex and incompletely understood (Iwasaki and Peiris, 2013), but it is plausible that women with a suppressed immune system were partly protected against the most severe consequences of the virus. If this holds true, the estimates of Almond (2006) and similar studies of the 1918 flu pandemic might be inflated.

To shed light on survivorship during the 1918 flu pandemic, we create a new data set from historical birth records in the maternity hospital of Bern, Switzerland. These data provide us with detailed information on all women delivering in the maternity hospital in Bern between 1913 and 1922. We find a 4.8 percentage points higher stillbirth probability at peak exposure in trimester 1 and slightly smaller effects in trimesters 2 and 3. We further find that the 1918 flu pandemic increases stillbirths among married mothers, but not among single mothers. Contemporary reports affirm that marital status predicts mothers'

2

socio-economic status and children's later life outcomes (Kraft, 1908). Thus, our finding provides suggestive evidence for negatively selected survivorship. We conclude that careful consideration of the nature of survivorship is paramount for the evaluation of fetal health shocks.

2 Survivorship during the 1918 flu pandemic

The 1918 flu pandemic was the most devastating of all known pandemics. The influenza A virus infected approximately one-third of the world's population in three waves in 1918 and 1919, with total deaths being estimated at 50 million or more (Taubenberger and Morens, 2010).

The pandemic likely increased maternal mortality, miscarriages, stillbirths, and abortions. Mortality rates were documented at 27% among infected pregnant women and pregnancy termination was reported in more than 50% of infected pregnant patients (Harris, 1919). With pregnancy as an important risk factor for influenza mortality, the scientific community discussed whether abortions should be recommended (Titus and Jamison, 1919). Recent evidence suggests that the 1918 flu pandemic decreased birth rates and increased the risks of miscarriage and stillbirth (Nishiura, 2009; Bloom-Feshbach *et al.*, 2011; Dahal *et al.*, 2018).

In contrast to other flu pandemics, the mortality pattern of the 1918 pandemic was highly unusual. Influenza mortality is usually concentrated among the very young and the very old, but many young and otherwise healthy individuals were affected in 1918 (Taubenberger and Morens, 2010). Worobey *et al.* (2014) argues that certain cohorts were protected by childhood exposure to genetically similar influenza viruses and excessive immune responses may explain the high mortality among young and healthy individuals (Morens *et al.*, 2010). Cytokines play an important role in the human body's innate immune response to influenza A infection. While these proteins are crucial for the human body's protection against influenza A, they can cause serious harm if the immune reaction is unregulated (Fukuyama and Kawaoka, 2011). So-called cytokine storms might have had deleterious effects on the young and healthy who were so strongly affected during the 1918 flu pandemic. Mounting evidence

3

shows that the 1918 virus triggered a vigorous and pathogenic immune response (e. g. Liu *et al.*, 2016).

It is plausible that those with the most vigorous immune system suffered the most. A suppressed immune system during a pregnancy characterized by stress, malnutrition, or other environmental factors might have been lifesaving in 1918. However, It must be noted that pathogenicity of a virus infection is determined in complex ways by both virus and host factors (Fukuyama and Kawaoka, 2011). Moreover, if a suppressed immune system is protective against cytokine storms, it might still be detrimental in other ways, e.g. through increased susceptibility to bacterial infection.

3   Stillbirths at the maternity hospital in Bern

We use transcribed data from the original birth records of the maternity hospital in Bern. These records include the following individual information for all deliveries that took place at the hospital: admission date, birth date, infant sex, singleton/multiple birth, mother's age, parity, marital status, date of last menstruation (used to calculate influenza exposure), and stillbirth/live-birth. We use data from 7,769 deliveries admitted in the years 1913 to 1922.

Exposure to the influenza pandemic is measured by the weekly reported numbers of new influenza cases in the city of Bern. The existing legal obligation for medics to report infectious diseases to the city authorities was extended to influenza no later than July 16, 1918, in the second week of the pandemic (Simonin *et al.*, 1918). Daily values are interpolated from this source. Exposure variables for each trimester are calculated as the sum of all infections during the respective trimester. The resulting measures are normalized, such that the highest possible exposure in each trimester is 1.

We estimate logistic regression models of the following form and present marginal effects.

$$stillbirth = \alpha + \beta e^{trim} + \gamma_1 girl + \gamma_2 multiples + \delta^{age} + \delta^{parity} + \delta^{month} + \delta^{neighborhood} + \epsilon$$

Our outcome of interest is *stillbirth*, a dummy variable equal to 1 for stillbirths and 0 for live births. Each of our models includes one of the three timester exposure variables, e^{trim}.

	Full sample			Married only			Single only		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Weeks 3-13	0.048***			0.054***			0.033		
	(0.017)			(0.020)			(0.034)		
Weeks 14-26		0.041***			0.045**			0.036	
		(0.015)			(0.018)			(0.028)	
Weeks 27-40			0.034**			0.046**			0.001
			(0.015)			(0.018)			(0.037)
Girl	−0.007	−0.007	−0.007	−0.004	−0.004	−0.004	−0.017*	−0.017*	−0.017
	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)	(0.010)	(0.010)	(0.010)
Multiples	0.079***	0.079***	0.080***	0.074***	0.074***	0.074***	0.100***	0.099***	0.099***
	(0.010)	(0.010)	(0.010)	(0.013)	(0.013)	(0.013)	(0.020)	(0.020)	(0.020)
Age 20 to 24	0.056***	0.056***	0.056***	0.059	0.059	0.059	0.053**	0.053**	0.054**
	(0.022)	(0.022)	(0.022)	(0.044)	(0.044)	(0.044)	(0.023)	(0.023)	(0.023)
Age 25 to 29	0.070***	0.069***	0.070***	0.080*	0.079*	0.080*	0.056**	0.055**	0.056**
	(0.022)	(0.022)	(0.022)	(0.044)	(0.044)	(0.044)	(0.024)	(0.024)	(0.024)
Age 30 +	0.087***	0.087***	0.086***	0.097**	0.096**	0.097**	0.075***	0.074***	0.075***
	(0.022)	(0.022)	(0.022)	(0.044)	(0.044)	(0.044)	(0.025)	(0.025)	(0.025)
Parity 2 and 3	−0.015**	−0.015**	−0.015**	−0.026***	−0.025***	−0.026***	0.013	0.014	0.013
	(0.007)	(0.007)	(0.007)	(0.009)	(0.009)	(0.009)	(0.012)	(0.012)	(0.012)
Parity 4 and 5	0.009	0.009	0.009	0.003	0.003	0.003	0.033	0.034	0.034
	(0.009)	(0.009)	(0.009)	(0.011)	(0.011)	(0.011)	(0.024)	(0.024)	(0.024)
Parity 6 +	0.030***	0.030***	0.030***	0.028***	0.028***	0.028***	0.013	0.014	0.012
	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)	(0.031)	(0.030)	(0.031)
Neighborhood	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month of year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7711	7711	7711	5613	5613	5613	1984	1984	1984

Table 1 – Results

The variable *girl* indicates the infant's sex and *multiples* indicates multiple pregnancies. δ^{age}, δ^{parity}, $\delta^{neighborhood}$, and δ^{month} denote fixed effects for age categories, parity categories, neighborhood, and month of the year.

The first three columns in Table 1 present estimates for the full sample, where each model includes one of the three trimester exposure variables. We find substantial effects on the probability of a stillbirth. Mothers at peak exposure in trimester 1 carry a 4.8 percentage points higher risk of stillbirth, as compared with unaffected mothers. The effects are slightly smaller for exposure during the second trimester (4.1 percentage points) and the third trimester (3.4 percentage points).

We aim to assess whether the stillbirths induced by the 1918 flu pandemic are indeed concentrated among those with poor potential outcomes. Of course, potential outcomes are not observable—we will never know what would have happened to those who did not survive. We use marital status as a proxy variable for potential outcomes. Single mothers are more likely to be economically disadvantaged and to have babies with low birth weight, which in turn predicts long-term outcomes like income and education (Aizer and Currie, 2014). Social reform circles in the early 20th century report that single mothers and their children suffered from discrimination, that single motherhood was strongly correlated with poor child health, and that children of single mothers received worse occupational education (Kraft, 1908).

Columns 4 to 6 and 7 to 9 in Table 1 show the estimation results for married and single mothers, respectively. The effects on married mothers are larger than the effects on the overall sample in all three trimesters and statistically significant at the 95% level or higher. Conversely, the effects for single mothers are smaller and statistically insignificant at conventional levels. This finding suggests that fetuses who survived the 1918 flu pandemic are negatively selected, raising concerns that the estimated long-term effects of the 1918 flu pandemic might be inflated.

4   Conclusions

Survivorship bias plays a relatively minor role in the fetal origins literature, as many studies assume positively selected survivors. We argue that this seemingly natural assumption does not necessarily hold for the 1918 flu pandemic. In particular, the medical literature suggests that women may have seen some degree of protection against the virus if their immune system was suppressed. The term "culling" disregards such complexities, as ist implies positively selected survivors. In this sense, the neutral term "survivorship bias" may be preferable.

We collect data from the maternity hospital in Bern, Switzerland, and find that the large effects of the 1918 flu pandemic on stillbirths are driven by married mothers in our sample. We interpret this result as suggestive evidence for negatively selected survivorship. Future work may use richer data to evaluate whether our results hold in full-population samples; whether our results generalize to measures of potential outcomes other than marital status; and whether survivorship bias is a serious concern for the evaluation of other fetal health shocks.

References

Aizer, A. and Currie, J. (2014), "The intergenerational transmission of inequality: Maternal disadvantage and health at birth." *Science* **344**, 856–861.

Almond, D. (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**, 672–712.

Almond, D. and Currie, J. (2011), "Killing Me Softly: The Fetal Origins Hypothesis." *Journal of Economic Perspectives* **25**, 153–172.

Almond, D., Currie, J., and Duque, V. (2017), "Childhood circumstances and adult outcomes: Act II." *NBER Working Paper 23017* .

Bloom-Feshbach, K., Simonsen, L., Viboud, C., Mølbak, K., Miller, M. A., Gottfredsson, M., and Andreasen, V. (2011), "Natality decline and miscarriages associated with the 1918 influenza pandemic: The Scandinavian and United States experience." *The Journal of Infectious Diseases* **204**, 1157–1164.

Brown, R. and Thomas, D. (2018), "On the long term effects of the 1918 influenza pandemic." *Working paper* .

Dahal, S., Mizumoto, K., Bolin, B., Viboud, C., and Chowell, G. (2018), "Stillbirth risk and spatial variation in excess death rates during the 1918-1920 influenza pandemic in Arizona, USA." *American Journal of Epidemiology* .

Fukuyama, S. and Kawaoka, Y. (2011), "The pathogenesis of influenza virus infections: The contributions of virus and host factors." *Current Opinion in Immunology* **23**, 481–486.

Harris, J. W. (1919), "Influenza occuing in pregnant women: A statistical study of thirteen hundred and fifty cases." *JAMA* **72**, 978–980.

Iwasaki, A. and Peiris, M. (2013), "Innate Immunity." In: R. G. Webster, A. S. . Monto, T. J. Braciale, and R. A. Lamb (Eds.), *Textbook of Influenza*, second ed., chap. 17, 269–282, Chichester, Oxford: John Wiley & Sons, Ltd.

Kraft, A. (1908), "Die sozialen Verhältnisse der unehelichen Kinder in ihren Ursachen und Wirkungen." *Jahrbuch der Schweizerischen Gesellschaft für Schulgesundheitspflege* **9**, 291–308.

Liu, Q., Zhou, Y., and Yang, Z. (2016), "The cytokine strom of severe influenza and development of immunomodulatory therapy." *Cellular and Molecular Immunology* **13**, 3–10.

Morens, D. M., Taubenberger, J. K., Harvey, H. A., and Memoli, M. J. (2010), "The 1918 influenza pandemic: Lessons for 2009 and the future." *Critical Care Medicine* **38**, e10–e20.

Nishiura, H. (2009), "Excess risk of stillbirth during the 1918-1920 influenza pandemic in Japan." *Theoretical Epidemiology* **147**, 115.

Nobles, J. and Hamoudi, A. (2019), "Detecting the effects of early-life exposures: why fecundity matters." *Population Research and Policy Review* **38**, 783–809.

Schwandt, H. (2018), "The lasting legacy of seasonal influenza: In-utero exposure and human capital development." *CEPR Discussion Paper 12653* .

Simonin, von Erlach, and Burren (1918), "Verwaltungsbericht der Sanitätsdirektion." *Bericht über die Staatsverwaltung des Kantons Bern* 225–231.

Taubenberger, J. K. and Morens, D. M. (2010), "Influenza: The once and future pandemic." *Public Health Reports* **125**, 16–26.

Titus, P. and Jamison, J. M. (1919), "Pregnancy complicated by epidemic influenza." *JAMA* **72**, 1665–1668.

Worobey, M., Han, G. Z., and Rambaut, A. (2014), "Genesis and pathogenesis of the 1918 pandemic H1N1 influenza A virus." *PNAS* **111**, 8107–8112.