

Pandemic Influenza and the Gender Imbalance: Evidence from Early Twentieth Century Japan*

Kota Ogasawara and Minami Yumitori

May 11, 2020

Abstract

This study uses the 1918–1920 influenza pandemic in Japan with newly digitized and complete census records on births, infant deaths, and sex ratios during childhood to analyze mortality selection *in utero* and its persistency in the gender imbalance. We find that fetal exposure to pandemic influenza during the first trimester of the pregnancy period decreased the proportion of males at birth in this period. We then show that the decline in male births might have been associated with the deterioration of fetal health due to pandemic influenza. Analyses using Population Censuses provide suggestive evidence that fetal influenza exposure had long-term impacts on the sex ratio at 5–12 years old.

Keywords: gender imbalance; fetal death; pandemic influenza; pregnancy outcomes; secondary sex ratio; Trivers–Willard hypothesis

JEL Codes: I19; N35;

*Kota Ogasawara: Department of Industrial Engineering, School of Engineering, Tokyo Institute of Technology, 2-12-1, Ookayama, Meguro-ku, Tokyo 152-8552, Japan (E-mail: ogasawara.k.ab@m.titech.ac.jp). Minami Yumitori: Department of Industrial Engineering, School of Engineering, Tokyo Institute of Technology, 2-12-1, Ookayama, Meguro-ku, Tokyo 152-8552, Japan (E-mail: yumitori.m.aa@m.titech.ac.jp). We wish to thank the participants of the CSG seminar and Tokyo Tech seminars for their helpful comments on this paper. The work was supported by the Tokyo Institute of Technology (grant in 2019). There are no conflicts of interest to declare. All errors are our own.

1 Introduction

A growing body of the literature has found evidence that fetal exposure to adverse health shocks is associated with negative socioeconomic and health outcomes later in their life (see Prinz et al. 2018, for a recent survey).¹ In addition to a sizable long-run literature, the short-run relationship between health shocks, particularly weather shocks, *in utero* and birth outcomes has also been widely studied (Poursafa et al. 2015; Zhang et al. 2017). These studies have found that exposure to weather shocks such as heat and cold waves during pregnancy are associated with lower birthweight (Andalón et al. 2016; Deschênes et al. 2009; Molina and Saldarriaga 2017).² These findings in both the long-run and the short-run strands of the literature are closely related because a lower birthweight can be associated with worse socioeconomic outcomes in later life (Prinz et al. 2018).

In contrast to these studies, however, mortality selection *in utero* has not attracted broad coverage in economics literature. One exception is the study by Valente (2015), who tested a biological proposition named the Trivers–Willard hypothesis, which argues that fetal exposure to adverse health shocks disturbs the gender balance at birth because reproductive success of males is more vulnerable than that of females (Trivers and Willard 1973). Valente found that fetal exposure to civil conflict in Nepal is associated with a higher probability of miscarriage and relatively low male births compared to female births (i.e., lower secondary sex ratio). Although testing the Trivers–Willard hypothesis is not their main object, Sanders and Stoecker (2015) also found that the Clean Air Act Amendments of 1970 in the United States improved fetal health (measured as a higher secondary sex ratio).

Maintaining the natural gender balance is important in an economy because the adult sex ratio imbalance leads to skewed marriage in terms of age and assortative matching as well as other demographic conditions such as out of wedlock fertility (Abramitzky et al. 2011; Andrew 2011; Angrist 2002; Bethmann and Kvasnicka 2012; Brainerd 2017; Chiappori et al. 2002). Considering this scarcity of research, we aim to bridge the gap in the body of knowledge by investigating the associations between fetal exposure to pandemic influenza and mortality selection *in utero*.

¹See also Almond and Currie (2011) and Currie and Vogl (2013) for comprehensive reviews.

²Currie and Rossin-Slater (2013) found that newborns could suffer abnormalities because of stressful events due to weather shocks. According to their analyses, exposure to a hurricane during pregnancy is associated with the use of ventilators and the occurrence of meconium aspiration syndrome. In addition, since climate change is now a global agenda, the consequences of weather shocks on child health have recently begun to attract wider attention (Kousky 2016; Zivin and Shrader 2016)). Carleton and Hsiang (2016) reviewed the social and economic impacts of climate change.

Using a comprehensive dataset of vital statistics in Japan, we find that fetal exposure to pandemic influenza between 1918–1920 decreased the proportion of males at birth. The culling effect was concentrated on exposure during the first trimester of the pregnancy, and the estimated magnitude suggests that such exposure could have decreased the proportion of males at birth by up to 1%, accounting for 60% of one standard deviation. Our results from the analyses using the complete census on annual infant mortality by gender indicate that such a reduction in male births during pandemics might be associated with a “scarring” mechanism under which the distribution of fetal health endowment shifts to the left. We also investigate the persistency of fetal influenza exposure on the sex ratio between 5 and 12 years old using a set of official reports of the Population Censuses conducted in 1925 and 1930. From these exercises, we find evidence that shocks due to pandemic influenza might have persisted into childhood. The estimated magnitude is approximately 0.18% in the maximum case, accounting for 40% of the standard deviation.

This study contributes to the wider literature in the following two ways. First, it is the first to use pandemic influenza as an exogenous shock to test mortality selection *in utero* as well as a set of complete censuses on births in a developing economy, which covers all births in a unit of a prefecture-month cell between 1916 and 1922 in Japan. While previous studies have used survey data to analyze fetal health, Sanders and Stoecker (2015) showed that data from household surveys on fetal losses are more likely to suffer from unobserved selection issues because of the use of small samples.³ Another potential issue with survey samples in developing countries is age heaping (Beckett et al. 2001). In light of these issues, we use the complete prefecture-month-level birth records of prewar Japan to provide new evidence on the association between fetal influenza exposure and the gender imbalance at birth. Using similar comprehensive vital statistics on infant deaths, this study also assesses the mechanism behind male culling before birth.

Second, this study is the first to investigate the persistency of fetal shocks on the sex ratio of children. While previous studies have focused on the associations between fetal shocks and the sex ratio at birth, the later-life gender imbalance due to those fetal shocks has not been studied (Bethmann and Kvasnicka 2014; Sanders and Stoecker 2015; Valente 2015). Investigating the long-term effects of fetal exposure to pandemics on the sex ratio is important given that maintaining the natural gender balance in an economy is preferable, as discussed earlier. Although our data constructed from Population Censuses

³Sanders and Stoecker (2015) used county-year-level birth data from the National Center for Health Statistics’ Vital Statistics Micro-data between 1968 and 1972, which records 50% of all birth certificates in the United States. Valente (2015) used the Demography and Health Surveys conducted in 2001 and 2006, which records nearly 11,000 births.

include children aged up to 12 years old, we find evidence of the persistent effects of fetal influenza exposure on the sex ratio of children.

The structure of the remainder of this paper is as follows. Section 2 introduces the empirical setting. Section 3 provides empirical evidence on the gender imbalance at birth due to pandemic influenza. Section 4 assesses the mechanism behind fetal shocks on the gender imbalance. Section 5 investigates the persistency of fetal influenza exposure on the sex ratio of children. Section 6 concludes the paper.

2 Empirical Setting

2.1 Theoretical Framework

This study investigates the impacts of pandemic influenza on the gender balance. The influential study by Trivers and Willard (1973) in the field of biology proposed a hypothesis about the mechanism behind the determinants of the secondary sex ratio, which has recently attracted attention in the field of health economics (Valente 2015). As Catalano and Bruckner (2006) illustrated, the intuition of the Trivers–Willard hypothesis can be explained using shifts in the distribution of a random variable.

Proportion of Male Births

Let $z_b \sim \mathcal{N}(\theta_b, \sigma^2)$ and $z_g \sim \mathcal{N}(\theta_g, \sigma^2)$ be the initial health endowments of boys and girls *in utero*, respectively. Owing to natural selection, a certain threshold, λ , exists, below which fetuses are culled before birth. Since male fetuses are more vulnerable *in utero* than female fetuses (Kraemer 2000), it is natural to assume that the mean initial health endowment of girls is greater than that of boys: $\theta_g > \theta_b$. When we consider the probability density function for boys ($f_b(\cdot)$) and girls ($f_g(\cdot)$), this initial assumption implies that the number of culled male fetuses is always greater than that of female fetuses because of the following condition:

$$f_b(\lambda) > f_g(\lambda). \quad (1)$$

If fetuses are exposed to health shocks *in utero*, the distribution shifts to the left or the survival threshold moves to the right. The former is called the “scarring” mechanism, whereas the latter is called the “selection” mechanism. In both cases, condition (1) indicates that the male share at birth must decrease, which corresponds to the proposition

implied by the Trivers–Willard hypothesis (Trivers and Willard 1973). This study thus investigates whether this proposition holds for the influenza pandemic in the early 20th century in industrializing Japan.

Mechanism

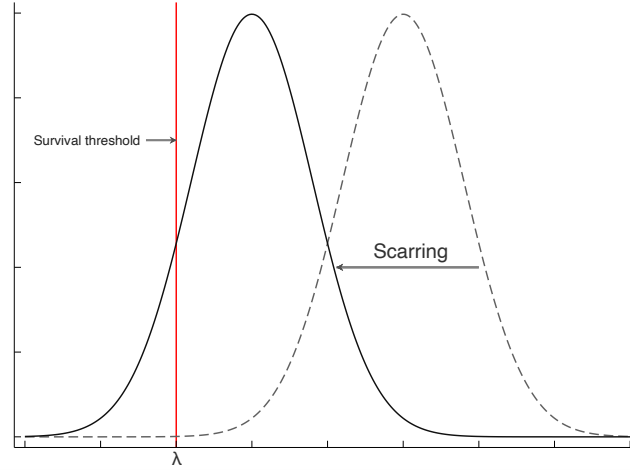
In contrast to culling before birth, the health status of an infant depends on the type of mechanism. If the “scarring” mechanism works, the conditional mean of the truncated normal distribution shifts to the left because the original mean of the distribution moves to the left, as illustrated in Figure 1a. If the “selection” mechanism works instead, the conditional mean of the truncated normal distribution shifts to the right as the survival threshold moves to the right, cutting the lower tail of the original distribution, as illustrated in Figure 1b.⁴ If both mechanisms work at the same time, the health status of an infant should therefore be unchanged.

To test which mechanism is more relevant, we use the infant mortality rate as a proxy for the health status of infants. In the contemporary context, a health measurement at birth such as birthweight can be used (Valente 2015). Although such a measurement at birth was unavailable in early 20th century Japan, the annual vital statistics reports provide complete figures on infant deaths in all prefectures at that time. Since the infant mortality rate can accurately represent health status at birth (Almond 2006), we use data on infant mortality to analyze the mechanism behind the observed secondary sex ratio during the influenza pandemic.

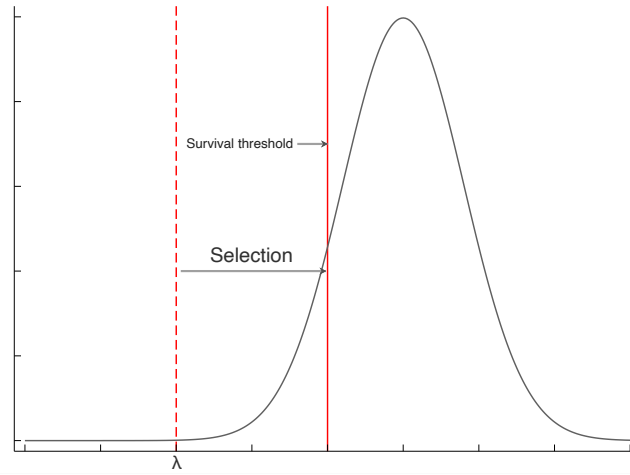
2.2 Maternal Stressor: Pandemic Influenza

The Spanish influenza of 1918 infected 600 million people and killed 20–40 million patients worldwide (Kilbourne 2006; Taubenberger 2006). After only five months of the first reported case of influenza in the United States, Spanish flu hit the Japanese archipelago between August 1918 and July 1920. Figure 2 shows the number of monthly deaths from pandemic influenza between 1918 and 1920 in Japan. Similar to other Asian countries, there were two waves of the pandemic in Japan, with the first and second peaks observed in November 1918 and January 1920, respectively (Hayami 2006). Regarding the intensity of the pandemics, although the influenza mortality rate in Japan (4.5 per 1,000 people)

⁴These illustrations may be understandable intuitively. However, it is easy to show both mechanisms mathematically as the conditional mean of the truncated normal distribution of, for example z_g : $E[z_g | z_g > \lambda] = \theta_g + \sigma f_g(\lambda) / (1 - F_g(\lambda))$, where $F_g(\cdot)$ is the cumulative distribution function.



(a) Scarring mechanism



(b) Selection mechanism

Figure 1: Scarring and selection mechanisms

Notes: In Figure 1a, the original probability density function is the dashed line, whereas the observed probability density function is the solid line. In Figure 1b, the dashed line in red is the original survival threshold, whereas the solid line in red is the observed survival threshold. Source: Created by the authors.

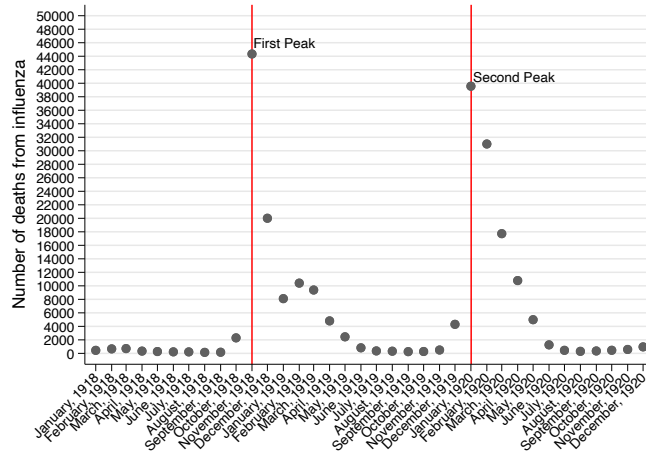


Figure 2: The number of deaths from influenza between January 1918 and December 1920

Notes: The red lines show the first and second peaks of the number of deaths due to influenza.

Source: Created by the authors from Statistics Bureau of the Cabinet (1921a, 1922a, 1923).

between 1918 and 1919 was lower than that of other Asian countries, this rate was in a similar range to that of Western countries (Hayami 2010; Rice and Palmer 1993). Indeed, during the pandemic periods (August 1918–July 1919; September 1919–July 1920), more than one in every five people in Japan became infected with influenza (Central Sanitary Bureau of the Home Ministry 1927).

Moreover, pandemic influenza tended to affect young adult females as well as older adults and children because of its aggressiveness (Almond and Mazumder 2005; Erkoreka 2010; Kawana et al. 2007). Specifically, women aged 20–29 years in Japan were more likely to be affected by the pandemic flu virus than men in the same age range (Hayami 2006; Ogasawara 2017).⁵ Since the average age at first marriage of Japanese women was 23 years old in the 1920s, average age at first birth might have been around 24–25 years old (Statistics Bureau of the Cabinet 1926a, 1928). This means that pandemic influenza affected not only young adult women but also children in utero during the pandemics via maternal infection.

Given these features of pandemics, a growing body of studies has employed pandemic influenza as a natural experiment to identify the long-term effects of fetal shocks on human capital formation (e.g., Almond 2006; Lin and Liu 2014). Indeed, as these previous studies have found, pandemics show a certain random spatiotemporal distribution. Figure 3

⁵Similar gender and age biases in infections were observed in Western countries. Influenza-related mortality rates in the pandemic years were more than five times higher than those in non-pandemic years in the United Kingdom and the United States (Richard et al. 2009). Reid (2005) also reported that such dramatic increases in mortality rates were more obvious in young adult women.

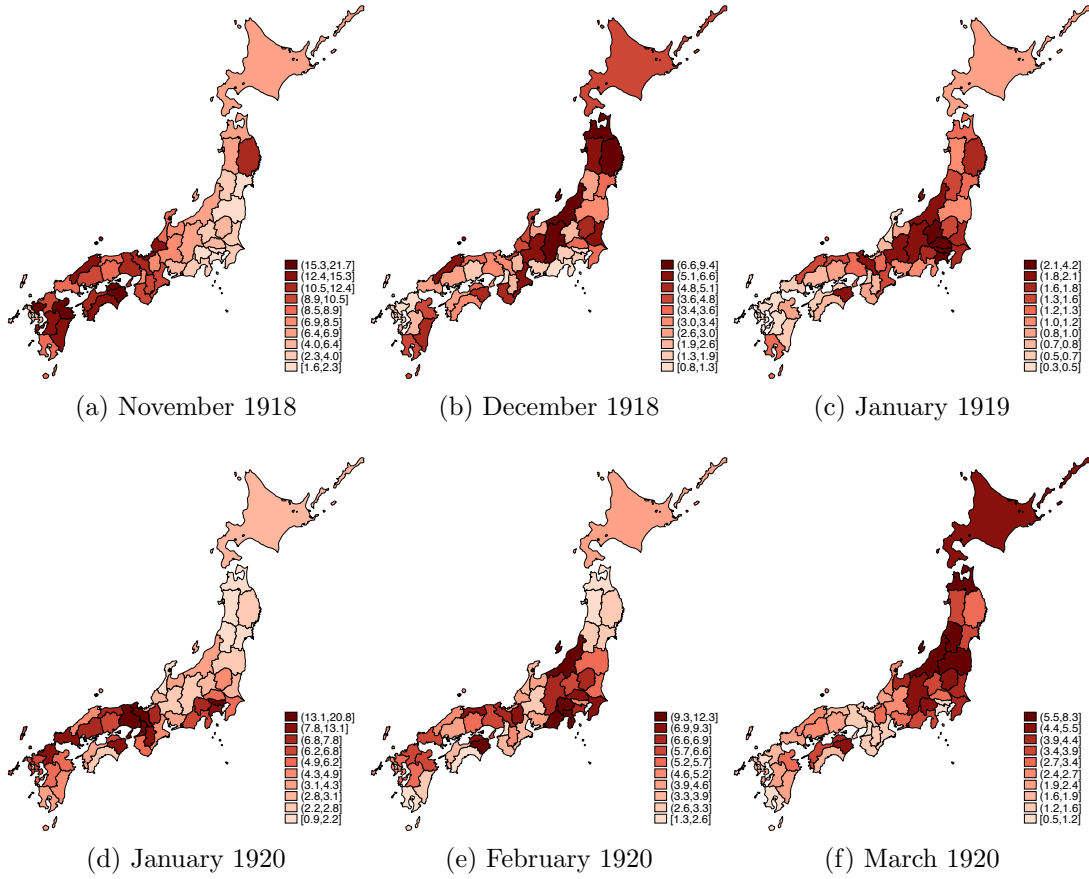


Figure 3: Spatiotemporal distributions of influenza death rates

Notes: Figures 3a–3c illustrate the influenza death rates in the first pandemic period between November 1918 and January 1919, as shown in Figure 2. Figures 3d–3f illustrate the influenza death rates in the second pandemic period between January 1920 and March 1920. Okinawa prefecture is not included in the sample. Sources: Created by the authors from Statistics Bureau of the Cabinet (1921a, 1922a, 1923); Statistics Bureau of the Cabinet (1921b, 1922b, 1924b); Statistical Survey Department, Statistics Bureau, Ministry of Internal Affairs and Communications (database).

illustrates the spatial distribution of influenza death rates in the pandemic months of 1918–1920. In the first wave, the epidemic cluster was generated in the southwestern region (Kyūshū and Shikoku) in November 1918 (Figure 3a); it then jumped to the northern and northeastern regions (Chūbu and Tōhoku) in the next month (Figure 3b) before moving to the central part of the main island (Kantō) (Figure 3c). The second wave exhibits a more straightforward but not persistent transition. The cluster was generated in the western region (Chūgoku) in January 1920 (Figure 3d), and then transited to the northern region (Chūbu) in the next month (Figure 3e). It finally covered a broader region including the northeastern region (Tōhoku) and Hokkaidō, a northeastern island. The foregoing suggests that the patterns of the pandemics were not systematic or concentrated in a specific region.

Potential sorting might be an issue in the identification because this can cause mea-

surement errors in the influenza death rate. However, internal migration as an escaping strategy would not have worked because people could not have predicted the timing and place of the pandemics in early 20th century (Hayami 2006). Moreover, although another potential issue is the immediate response by the Japanese government, the government could not have provided an efficient preventive policy during the pandemics because no vaccination was available at that time (Hayami 2006).

3 Gender Imbalance at Birth

3.1 Data

Proportion of Male Births

Our main analysis uses a unique prefecture-month-level panel dataset on the male share at birth defined as the number of male live births per 100 live births.⁶ We construct the dataset of the proportion of male births using the official vital statistics records published by the Statistics Bureau of the Cabinet. Since these vital statistics have been recorded based on the comprehensive national registration system (*koseki*), the data cover all births during the measured years.⁷ We digitize the 1916–1922 editions of *Nihonteikoku jinkōdōtaitōkei* (Vital Statistics of Empire Japan, hereafter the VSEJ) (Statistics Bureau of the Cabinet 1919b, 1920b, 1921b, 1922b, 1924b,c,d). Online Appendix A shows an example of this vital statistics record.⁸ Panel A of Table 1 lists the summary statistics of the proportion of male births.

Influenza Mortality

We use the influenza death rate, the number of deaths due to influenza per 10,000 people, as the key independent variable that captures the intensity of exposure to pandemic influenza. The data on the monthly death tolls from influenza are obtained from the 1915–1922 editions of *Nihonteikoku shiintōkei* (Statistics of Causes of Death of the Empire of

⁶This is essentially the same as using the secondary sex ratio, namely, the ratio of male live births to female live births (Bethmann and Kvasnicka 2014). Since the Trivers–Willard hypothesis focuses on the vulnerability of male fetuses, the male share at birth may be more useful than the secondary sex ratio for interpreting the results.

⁷Although the quality of Japanese fetal death records that begun in 1900 was not high in the initial stage, the records became reliable around the 1920s (Kawana et al. 2007; Shigeru 1987). See Drixler (2016) for a more in-depth discussion on birth records in prewar Japan.

⁸We confirm the stationarity of our panel dataset. For the proportion of male births used in our analysis, several tests reject the null of unit root non-stationarity. See Online Appendix B.1.

Japan, hereafter the SCDEJ) published by the Statistics Bureau of the Cabinet (Statistics Bureau of the Cabinet 1918b, 1919a, 1920a, 1921a, 1922a, 1923, 1924a, 1925a), whereas the data on the population are taken from the official online database of the Statistical Survey Department, Statistics Bureau, Ministry of Internal Affairs and Communications.⁹ Since fetal influenza exposure matters in our theoretical framework as described, we use the past nine-month average of influenza death rates in the regression analysis. We also investigate the impacts of fetal influenza exposure in each trimester (see the next subsection). Panel B of Table 1 shows the summary statistics of the average influenza death rates.

Additional Control Variables

To control for the observable factors, we include a set of available prefecture-year-level control variables. The first set of controls are the indices of agricultural production. Given the agrarian society at that time, a certain proportion of wealth can be captured by productivity. We herein consider rice yield per hectare, soy yield per hectare, and milk production per capita as measures of potential wealth because these items were the main sources of carbohydrate and protein (Ogasawara et al. 2020). The data on these variables are digitized from *Todōfuken nōgyōkisetōkei* (Basic Statistics of Agriculture in Japanese Prefecture) edited by Nobufumi Kayo (Kayo 1983). Another set of controls is access to medical care. We include the share of medical doctors and midwives to control for access to medical care and related socioeconomic conditions and potential wealth level. To obtain the data on medical access, we digitize the volumes 36–43 of *Nihonteikoku tōkeinenkan* (Statistical Yearbook of the Japanese Empire, hereafter the SYEJ) (Statistics Bureau of the Cabinet 1914–1927).¹⁰

⁹These are publicly available at the official website: <https://www.e-stat.go.jp/stat-search/file-download?statInfId=000000090265&fileKind=0>, accessed on July 31, 2019). In this dataset, the population in month j ($= 1, 2, \dots, 12$) of year t is calculated as $\tilde{P}_t^j = P_t + B_t^j - D_t^j + \{P_{t+1} - (P_t + \sum_{j=1}^{12} B_t^j - \sum_{j=1}^{12} D_t^j)\}/12$, where P_t , B_t^j , and D_t^j are the annual population, number of live births, and number of deaths, respectively. The data on the number of live births and deaths are from the 1915–1922 editions of the VSEJ (Statistics Bureau of the Cabinet 1918d, 1919b, 1920b, 1921b, 1922b, 1924b,c,d).

¹⁰The data on the population used as the denominator are taken from the official database of the Statistical Survey Department, Statistics Bureau, Ministry of Internal Affairs and Communications (<http://www.stat.go.jp/data/chouki/zuhyou/02-05.xls>, accessed on July 13, 2017).

Table 1: Summary statistics

| | Frequency | Mean | Std. Dev. | Min | Max | Observations |
|-------------------------------------------------------------------|-----------|--------|-----------|-------|--------|--------------|
| Panel A: Dependent variables | | | | | | |
| Male births (per 100 births) | Monthly | 51.18 | 1.67 | 41.37 | 61.39 | 3,864 |
| Infant mortality rate (per 1,000 live births) | Annual | 159.18 | 27.28 | 89.90 | 239.17 | 322 |
| Infant mortality rate (boys) | Annual | 166.30 | 27.55 | 96.43 | 243.08 | 322 |
| Infant mortality rate (girls) | Annual | 151.70 | 27.28 | 83.09 | 235.02 | 322 |
| Panel B: Influenza severity | | | | | | |
| Influenza death rate (past nine-month average, per 10,000 people) | Monthly | 0.53 | 0.78 | 0.00 | 3.89 | 3864 |
| Influenza death rate (first trimester average) | Monthly | 0.53 | 1.23 | 0.00 | 9.95 | 3864 |
| Influenza death rate (second trimester average) | Monthly | 0.53 | 1.23 | 0.00 | 9.95 | 3864 |
| Influenza death rate (third trimester average) | Monthly | 0.53 | 1.23 | 0.00 | 9.95 | 3864 |
| Weighted influenza death rate (per 10,000 people) | Annual | 0.04 | 0.05 | 0.00 | 0.22 | 322 |
| Weighted influenza death rate (first trimester average) | Annual | 0.04 | 0.05 | 0.00 | 0.21 | 322 |
| Weighted influenza death rate (second trimester average) | Annual | 0.04 | 0.05 | 0.00 | 0.23 | 322 |
| Weighted influenza death rate (third trimester average) | Annual | 0.05 | 0.06 | 0.00 | 0.27 | 322 |
| Panel C: Control variables | | | | | | |
| Rice yield per hectare (hectoliter) | Annual | 34.26 | 5.93 | 2.39 | 49.47 | 322 |
| Soy yield per hectare (hectoliter) | Annual | 16.34 | 3.14 | 6.82 | 24.95 | 322 |
| Milk production per capita (liter) | Annual | 0.98 | 0.82 | 0.18 | 6.06 | 322 |
| Coverage of doctors (per 100 people) | Annual | 0.07 | 0.02 | 0.04 | 0.19 | 322 |
| Coverage of midwives (per 100 people) | Annual | 0.06 | 0.02 | 0.01 | 0.12 | 322 |

Notes: Panel A reports the summary statistics for the prefecture-month- or prefecture-year-level dependent variables. Panel B reports the summary statistics for the prefecture-month- or prefecture-year-level influenza death rates. Panel C reports the summary statistics for the prefecture-year-level control variables. Sources: Dependent variables are from the VSEJ (1916–1922 editions). Influenza death rates are from the SCDEJ (1915–1922 editions); VSEJ (1915–1922 editions); Statistical Survey Department, Statistics Bureau, Ministry of Internal Affairs and Communications (database). The control variables are from Kayo (1983) and Statistics Bureau of the Cabinet (1914–1927).

3.2 Identification Strategy

We use the difference-in-differences (DID) estimation strategy within the regression framework that compares the proportion of male births among prefectures that experienced different intensities of exposure to the influenza before and after the pandemic. As discussed, there were considerable exogenous variations in the influenza death rates during the pandemic periods. To identify the impacts of fetal exposure to influenza on the sex ratio at birth, we employ a semi-experimental approach using this spatiotemporal variation in influenza death rates. Our baseline specification is given as follows:

$$y_{it} = \alpha + \beta \left(\sum_{j=1}^9 FLUDR_{it-j} \right) / 9 + \mathbf{x}'_{igt} \boldsymbol{\gamma} + \nu_i + \phi_t + e_{it}, \quad (2)$$

where i indexes the prefecture, t indexes the measured year-month, and g_t indicates a group variable for the measured year. The variable y is the proportion of male births, \mathbf{x} is a vector of the prefecture-year-level control variables, ν is the prefecture fixed effect, ϕ is the year-month-specific fixed effect, and e is a random error term. $FLUDR_{it-j}$ is the j -month lagged influenza death rate, and thus our key independent variable is the past nine-month average of influenza death rates. Our parameter of interest is β and its estimate $\hat{\beta}$ captures the marginal effect of the influenza death rate on the proportion of male births. Therefore, we expect $\hat{\beta}$ to be negative and statistically significant.

The first specification in equation 2 assumes that the potential effects of fetal influenza exposure are constant regardless of the timing of exposure. However, medical evidence suggests that fetuses are most susceptible to maternal stress in the first trimester when they experience rapid neuron differentiation and the proliferation of neuronal elements (Moore et al. 2013). This implies that the culling effects on male fetuses are much clearer in the first trimester than in the second and third trimesters. Therefore, our preferred specification is as follows:

$$y_{it} = \pi + \delta_0 \left(\sum_{j=7}^9 FLUDR_{it-j} \right) / 3 + \delta_1 \left(\sum_{j=4}^6 FLUDR_{it-j} \right) / 3 + \delta_2 \left(\sum_{j=1}^3 FLUDR_{it-j} \right) / 3 + \mathbf{x}'_{igt} \boldsymbol{\zeta} + \nu_i + \kappa_t + \epsilon_{it}. \quad (3)$$

The second to fourth terms represented as summations on the right-hand side are the average influenza death rates during the first, second, and third trimesters, respectively. Therefore, we expect the estimates $\hat{\delta}_0$, $\hat{\delta}_1$, and $\hat{\delta}_2$ to be negative; among these, the estimate for the first trimester, $\hat{\delta}_0$, shows the clear adverse effects on the proportion of male births.

Since we use a within estimator for the fixed effect models in equations 2 and 3, the

identification depends on the sharp increases in influenza mortality during the pandemic years (Figure 2). As discussed in Section 2.2, these influenza death rates are plausibly exogenous because no vaccination was available in the pre-war period and internal migration was unrealistic given the rapid spread of the virus. Despite this preferable feature for the identification, we control for a large proportion of the unobservable factors and observable characteristics in the following ways. First, we control for prefecture-specific time-invariant factors such as the baseline wealth level and geographical features using prefecture fixed effects. Second, the macroeconomic shocks and cyclical effects of seasonal epidemics are captured using year-month fixed effects.¹¹ After controlling for these fixed effects, the remaining potential confounding factors included in the error term that might be correlated with the influenza death rate include the time-varying wealth level and access to medical care. To control for these factors, we further include the set of available prefecture-year-level control variables introduced in the previous subsection. Since we employ the regression DID specification, the trends in the proportion of male births are assumed to be similar across prefectures. Since the sex ratio at birth is a biological measure rather than a socioeconomic outcome, this common trend assumption is likely to hold. To relax the common trend assumption, however, we further include the prefecture-specific time trend ($t\mu_i$) in some of the specifications.

To address the potential spatial and prefecture-specific within correlations, we report the cluster-robust variance estimator (CRVE) and cluster standard errors at the 8-area level.¹² To address the small number of clusters in the CRVE, we adopt the wild cluster bootstrap-t method for the statistical inference (Cameron et al. 2008). All the regressions are weighted by the average number of births over the sample period in each prefecture.

3.3 Main Results

Table 2 presents the results. Columns (1)–(4) present the results for the entire period (1916–1922). Column (1) shows the results from the baseline specification in equation 2. The estimate is negative but statistically insignificant. This result is unchanged if we include the prefecture-specific time trend in column (2). Column (3) shows the result from our preferred specification in equation 3. The estimates listed in this column suggest that

¹¹Since we use seven measured years (1916–1922), 83 ($7 \times 12 - 1$) year-month fixed effects are included in the models altogether.

¹²This geographical classification of Japan includes Hokkaidō (northernmost), Tōhoku (eastern), Kantō (east-central), Chūbu (west-central), Kansai (south-central), Chūgoku (westernmost), Shikoku (south-west of the main island), and Kyūshū (southwest island). Our method controls for the correlation and heteroskedasticity within clusters as well as addresses the potential heteroskedasticity across clusters.

fetal influenza exposure in the first trimester has a statistically significantly negative effect on the proportion of male births, whereas that during the second and third trimesters does not have such an effect. As explained, this finding is consistent with the fact that fetuses are more vulnerable in the first trimester than in the other trimesters. In column (4), we find that this result is robust to including the prefecture-specific time trend as expected.

Columns (5) and (6) present the results for non-pandemic years (1916–1917 and 1921–1922), whereas columns (7) and (8) present the results for pandemic years (1918–1920). In columns (5) and (6), we find no statistically significant effects of fetal exposure to influenza on the proportion of male births during non-pandemic years. By contrast, columns (7) and (8) show the clear significant adverse effects of fetal influenza exposure during pandemic years. This implies that while seasonal influenza does not have any significant impacts on the secondary sex ratio, pandemic influenza does have such an effect. The result of this placebo experiment supports the evidence that the identification in our within estimator using the sharp increase in influenza death rates during pandemic years seems to work well and should provide reliable estimates.

If we use the maximum average influenza death rate in the first trimester (Panel B of Table 1) as the reference value to calculate the magnitude, the estimate in column (3) implies that fetal exposure to pandemic influenza decreased the proportion of male births by approximately 1.1% (9.93×0.1092). This magnitude is not very large but still non-negligible given that one standard deviation of the proportion of male births is 1.67 (Panel A of Table 1). Although the estimate in column (7) is slightly larger than that in column (3), the magnitude calculated is in a similar range in both cases (1.1% vs. 1.3%).

Overall, we find that pandemic influenza can disturb the gender balance at birth, consistent with the proposition implied by the Trivers–Willard hypothesis that health shocks *in utero* can decrease the proportion of male births under both the “scarring” and the “selection” mechanisms because male fetuses are more vulnerable than female fetuses.

Table 2: Effects of fetal influenza exposure on the proportion of male births

| | Entire period | | | Non-pandemic years | | | Pandemic years | |
|-----------------------|---------------------------------|---------------------------------|----------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Exposed trimesters | | | | | | | | |
| All trimesters | -0.1584 (0.0774) [0.1490] | -0.1463 (0.0752) [0.1610] | | | | | | |
| First trimester | | | -0.1092*** (0.0331) [0.0030] -0.0442 (0.0323) [0.3010] -0.0036 (0.0325) [0.9350] | -0.1054*** (0.0318) [0.0030] -0.0399 (0.0317) [0.3110] 0.0000 (0.0329) [0.9770] | 0.2069 (0.1242) [0.1870] 0.2808 (0.2032) [0.2110] -0.6460 (0.2994) [0.1910] | 0.2217 (0.1828) [0.3370] 0.3413 (0.2995) [0.2730] -0.5996 (0.3547) [0.2490] | -0.1274*** (0.0364) [0.0030] -0.0542 (0.0425) [0.3330] -0.0029 (0.0316) [0.9470] | -0.1566*** (0.0442) [0.0030] -0.0852 (0.0535) [0.2470] -0.0244 (0.0387) [0.5910] |
| Second trimester | | | | | | | | |
| Third trimester | | | | | | | | |
| Time trend | No | Yes | No | Yes | No | Yes | No | Yes |
| Period | 1916–22 | 1916–22 | 1916–22 | 1916–22 | 1916–17 1921–22 | 1916–17 1921–22 | 1918–1920 | 1918–1920 |
| Observations | 3864 | 3864 | 3864 | 3864 | 2208 | 2208 | 1656 | 1656 |
| Number of prefectures | 46 | 46 | 46 | 46 | 46 | 46 | 46 | 46 |
| Number of clusters | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 |

***, ** and * represent statistical significance at the 1%, 5%, and 10% levels based on the p -values from the wild cluster bootstrap resampling method in brackets, respectively. The number of replications is fixed to 1,000 for all the specifications. Standard errors from the cluster-robust variance estimation reported in parentheses are clustered at the 8-area level.
Notes: All the regressions include controls for rice yield, soy yield, milk production, coverage of doctors, and coverage of midwives. All the regressions are weighted by the average number of births in each prefecture.

4 Mechanism

Next, we investigate the mechanism behind the suggested effects of the pandemics on the proportion of male births by pandemic influenza. To do so, we first digitize the statistics on infant deaths reported in the SCDEJ. We then calculate the weighted annual influenza death rate using the monthly variation in influenza mortality to match the infant death rates observed at the prefecture-year level.

4.1 Data and Specification

As described, we must measure the health of infants to test whether the “scarring” or “selection” mechanism drove the gender imbalance at birth due to pandemic influenza. We digitize the complete censuses of annual infant deaths documented in the SCDEJ to construct the dataset on infant mortality rates between 1916 and 1922 (Statistics Bureau of the Cabinet 1919a, 1920a, 1921a, 1922a, 1923, 1924a, 1925a). The data on the number of annual live births used as the denominator are obtained from the VSEJ (Statistics Bureau of the Cabinet 1919b, 1920b, 1921b, 1922b, 1924b,c,d).¹³

To improve the assignment of the treatments, we calculate a weighted influenza death rate using the monthly variations in the number of influenza deaths and live births. The weighted influenza death rate in prefecture i in year l is defined as follows:

$$Weighted\ FLUDR_{il} = \frac{\sum_{m=Jan}^{Dec} Birth_{ilm} \times \overline{FLUDR}_{ilm}}{12 \sum_{m=Jan}^{Dec} Birth_{ilm}}, \quad (4)$$

where $Birth_{ilm}$ is the number of live births in month m and \overline{FLUDR}_{ilm} is the past nine-month average of influenza mortality in month m .¹⁴ The baseline specification is then given as follows:

$$h_{il} = \psi + \rho Weighted\ FLUDR_{il} + \mathbf{x}'_{il}\boldsymbol{\omega} + \theta_i + \iota_l + u_{il}, \quad (5)$$

¹³Although the VSEJ also documents the number of fetal deaths by month, unfortunately, it does not record any information on the length of gestation period until fetal deaths. This means that it is technically difficult to precisely match the timing of exposure to pandemic influenza with fetal death rates. This sort of crude assignment can attenuate the estimated coefficients on the treatment variables. Despite this difficulty, we also try to run the annual fetal death rate on the weighted influenza mortality rate. As expected, the estimates are statistically insignificant in most cases.

¹⁴This transformation takes both the severity of influenza exposure and the timing of birth into account: \overline{FLUDR}_{ilm} captures the treatment intensity, whereas the weight, $Birth_{ilm}$, coordinates the differences in the timing of birth. Ogasawara (2018) showed evidence that this transformation improves the treatment assignment to a certain extent if we compare it using a simple lagged influenza death rate.

where h is the infant mortality rate, \mathbf{x} is a vector of the same control variables introduced above, θ is the prefecture fixed effect, ι is the year-specific fixed effect, and u is a random error term. Our parameter of interest is ρ and its estimate $\hat{\rho}$ may capture the marginal effect of the influenza death rate on the infant mortality rate. As explained, we expect $\hat{\rho}$ to be negative and statistically significant if the “selection” mechanism works, whereas it should be statistically significantly positive if the “scarring” mechanism is relevant.

In the flexible specification, we consider the weighted influenza death rates for the first, second, and third trimester by replacing \overline{FLUDR}_{ilm} in equation 4 with the average influenza mortality rates for each trimester. The flexible specification is given as follows:

$$h_{il} = \tau + \varrho_0 \text{Weighted } FLUDR_{il}^{First \ Trimester} + \varrho_1 \text{Weighted } FLUDR_{il}^{Second \ Trimester} + \varrho_2 \text{Weighted } FLUDR_{il}^{Third \ Trimester} + \mathbf{x}'_{il}\mathbf{\Gamma} + \xi_i + \varsigma_l + \varepsilon_{il}. \quad (6)$$

This specification allows us to investigate the most sensitive trimester for the impacts of fetal influenza exposure on infants’ health. One must be careful here, as in the regressions using infant mortality as a dependent variable, we do not necessarily expect the first trimester to be the most vulnerable for infants’ health. While fetuses are indeed relatively vulnerable during the first trimester, those affected by any shocks during this trimester are culled before birth. In other words, surviving fetuses are positively selected into birth.¹⁵ Therefore, the observed (i.e., surviving) infants may be more sensitive to shocks during the second and/or third trimesters than those during the first trimester. This natural selection mechanism suggests that the estimates $\hat{\varrho}_1$ and/or $\hat{\varrho}_2$ can be positive (negative) if the “scarring” (“selection”) mechanism works, whereas the estimate $\hat{\varrho}_1$ can be negative or statistically insignificant.

The inferences are conducted in a similar way for the specifications for the proportion of male births. We use the CRVE and cluster the standard errors at the 8-area level to address the potential spatial and prefecture-specific within correlations. The wild cluster bootstrap-t method is employed for the statistical inference. All the regressions are weighted by the average number of live births over the sample period in each prefecture. To relax the common trend assumption, we include the prefecture-specific time ($t\Pi_i$) trend in some of the specifications.

¹⁵We can usually show that the conditional expectation of the truncated normal distribution is always greater than that of the original distribution, as described in Section 2.1.

4.2 Results

Table 3 presents the results. Panels A–C of this table present the results for the infant mortality rates for all infants, boys, and girls, respectively. Columns (1) and (3) show the results from equations 5 and 6, respectively. Columns (2) and (4) add the prefecture-specific time trend for both equations.

Column (1) of Panel A shows that the estimated effect of fetal influenza exposure on the infant mortality rate is positive and statistically significant. This result is unchanged if we consider the prefecture-specific time trend in the infant mortality rate in column (2). This implies that the “scarring” mechanism might have driven the gender imbalance at birth. Column (3) of Panel A indicates that such an effect was concentrated on exposure during the third trimester as expected. This result is still unchanged after controlling for the prefecture-specific time trend in column (4).

Panels B and C of Table 3 show similar results for boys and girls. An interesting gender difference can be highlighted: the estimates for girls are greater in magnitude than those for boys. For example, if we compare column (2) of Panel B with that of Panel C, the estimate for girls is approximately 45 greater than that for boys ($92.6 - 47.8$). We confirm that this difference in magnitude is statistically significant.¹⁶

This gender difference is considered to be consistent with the “scarring” mechanism. As explained, the distribution of the fetal health endowment must shift to the left if the “scarring” mechanism works. Suppose the distributions of both boys and girls shift to the left by same degree and that the survival threshold is fixed at λ . Then, the net shift of the distribution depends only on the degree of the selection effect due to the truncation (at λ). Since the selection effect on the male fetus is always greater than that on the female fetus as condition 1 suggests, the net leftward shift of the distribution of girls can be greater than that of boys.¹⁷ This means that the total shift of the mean caused by the “scarring” mechanism should be greater for girls than for boys, implying that the estimated “scarring” effect of fetal influenza exposure on girls’ infant mortality is greater than that on boys’ infant mortality.

¹⁶To test the gender difference, we pooled the infant mortality rates for boys and girls and interacted all the independent variables including fixed effects with the gender dummy. Online Appendix B.2 summarizes these results.

¹⁷This mechanism can also be explained mathematically. Given that the conditional expectation of the truncated normal distribution of z_g can be written as $E[z_g|z_g > \lambda] = \theta_g + \sigma f_g(\lambda)/(1 - F_g(\lambda))$, the selection effect due to the truncation is expressed as $\sigma f_g(\lambda)/(1 - F_g(\lambda))$ (i.e., the second term). The gender difference (boys minus girls) of this term can then be written as $\sigma(f_b(\lambda) - f_g(\lambda))/((1 - F_b(\lambda))(1 - F_g(\lambda)))$, which is positive because $f_b(\lambda) > f_g(\lambda)$ (condition 1).

Table 3: Effects of fetal influenza exposure on the infant mortality rate

| Exposed trimesters | Dependent variable: Infant mortality rate | | | |
|-----------------------|-------------------------------------------|----------------------------------|----------------------------------|----------------------------------|
| | (1) | (2) | (3) | (4) |
| Panel A: All infants | | | | |
| All trimesters | 64.699*** (25.878) [0.000] | 69.991*** (30.200) [0.000] | | |
| First trimesters | | | −161.513 (81.688) [0.147] | −154.688 (84.095) [0.133] |
| Second trimesters | | | −65.637 (72.453) [0.480] | −72.477 (83.420) [0.467] |
| Third trimesters | | | 229.272** (75.907) [0.013] | 232.289** (81.853) [0.013] |
| Panel B: Boys | | | | |
| All trimesters | 42.043** (22.797) [0.027] | 47.846** (27.644) [0.013] | | |
| First trimesters | | | −131.592 (88.768) [0.187] | −126.951 (90.989) [0.173] |
| Second trimesters | | | −94.240 (76.951) [0.347] | −102.580 (88.546) [0.347] |
| Third trimesters | | | 210.759** (82.214) [0.013] | 217.388** (85.497) [0.013] |
| Panel C: Girls | | | | |
| All trimesters | 87.853*** (31.529) [0.000] | 92.515*** (35.701) [0.000] | | |
| First trimesters | | | −191.828 (77.679) [0.120] | −182.345 (80.924) [0.120] |
| Second trimesters | | | −36.836 (69.936) [0.533] | −42.073 (80.430) [0.533] |
| Third trimesters | | | 247.938** (71.274) [0.013] | 247.215** (80.395) [0.013] |
| Time trend | No | Yes | No | Yes |
| Observations | 322 | 322 | 322 | 322 |
| Number of prefectures | 46 | 46 | 46 | 46 |
| Number of clusters | 8 | 8 | 8 | 8 |

***, **, and * represent statistical significance at the 1%, 5%, and 10% levels based on the p -values from the wild cluster bootstrap resampling method in brackets, respectively. The number of replications is fixed to 150 for all the specifications. Standard errors from the cluster-robust variance estimation reported in parentheses are clustered at the 8-area level.

Notes: All the regressions include controls for rice yield, soy yield, milk production, coverage of doctors, and coverage of midwives. All the regressions are weighted by the average number of live births in each prefecture.

5 Persistency: Evidence from Population Censuses

5.1 Data and Specification

Thus far, we have found that fetal exposure to pandemic influenza decreased the proportion of male births. In this section, we assess whether the gender imbalance at birth persisted into their teens. The Population Censuses conducted in 1925 and 1930 documented the population by age and gender in each prefecture. To investigate the potential lasting effects of fetal influenza exposure on the sex ratio, we digitize the statistics and calculate the proportion of males aged 0–20 years.¹⁸

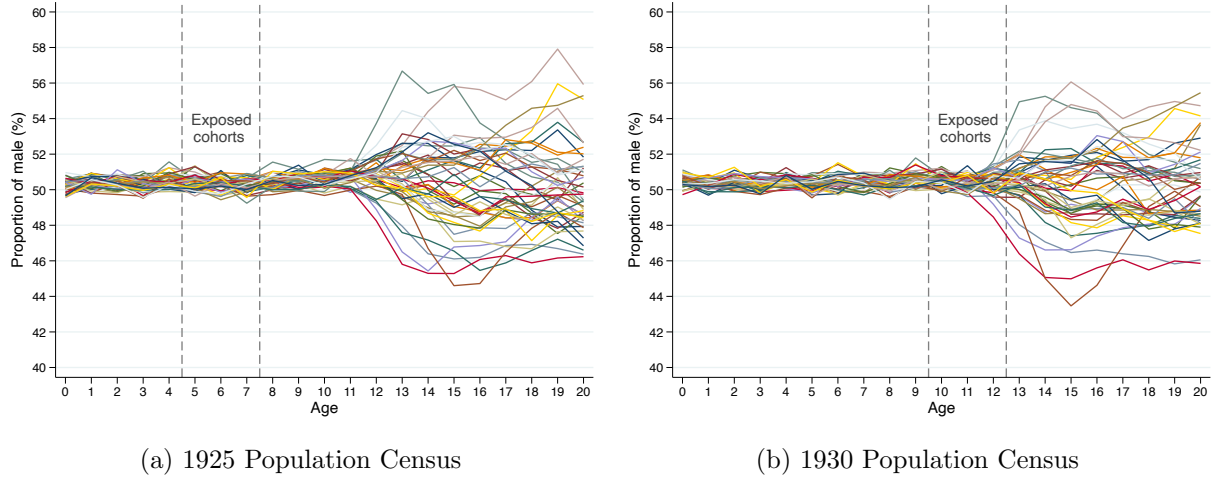


Figure 4: Proportion of males by age (%)

Notes: Figures 4a and 4b illustrate the proportion of males in the 46 prefectures by age in 1925 and 1930, respectively. Okinawa prefecture is not included in the sample. Sources: Created by the authors from Statistics Bureau of the Cabinet (1929) and Statistics Bureau of the Cabinet (1933).

Figure 4 shows the proportion of males in percentage points in each prefecture by age. As shown, the variance in the sex ratio is relatively stable until 12 years old because children graduate from primary school around then (Hijikata 1994). After graduation, while some children go to higher schools, a large part of them begin to work, which creates a gender imbalance due to the flow of migrant workers.¹⁹ This kind of internal migration after primary school age makes it difficult to analyze the potential long-run impacts of fetal influenza exposure on the sex ratio among teen workers because we use

¹⁸Since the prefecture editions of the Population Census were published for each prefecture, we use 92 issues (46 prefectures \times 2 census years) to construct the dataset. For simplicity, we refer to those issues as Statistics Bureau of the Cabinet (1929) and Statistics Bureau of the Cabinet (1933).

¹⁹The school enrollment rate for primary school was near 100% and there were no significant differences in the rates across prefectures at that time in Japan. See (Schneider and Ogasawara 2018) for finer details about primary school students in prewar Japan.

the prefecture-level aggregate dataset that does not have useful information on birthplace. Therefore, we focus on the gender imbalance up to 12 years old. We further trim the ages to improve the DID setting. Since children born between 1918 and 1920 were exposed to pandemic influenza, children aged 5–7 in the 1925 Population Census and 10–12 in the 1930 Population Census are defined as the exposed cohorts, as shown in Figures 4a and 4b. Considering this, we focus on the proportion of boys aged 5–7 and 10–12 years old in 1925 and 1930. This means that our analytical sample includes children aged 5–7 years and 10–12 years born in 1913–1925.²⁰ Panel A of Table 4 lists the summary statistics for the dependent variables used.

²⁰Precisely, those children aged 5–7 years (10–12 years) in 1925 were born in 1918–1920 (1913–1916). Those children aged 5–7 years (10–12 years) in 1930 were born in 1919–1920 (1923–1925). Accordingly, to prepare the weighted influenza death rates in 1913–1925, we additionally digitize the 1912–1914 and 1923–1925 editions of the VSEJ (Statistics Bureau of the Cabinet 1916c, 1917, 1918c, 1924d, 1925d,e, 1926c) and the SCDEJ (Statistics Bureau of the Cabinet 1916a,b, 1918a, 1925b,c, 1926b).

Table 4: Summary statistics: Estimating the persistent effects on the proportion of males aged 5–12

| | Unit | Mean | Std. Dev. | Min | Max | Observations |
|---------------------------------------------------|-----------------------|-------|-----------|-------|-------|--------------|
| Panel A: Dependent variable | | | | | | |
| Proportion of males (per 100 people) | Prefecture-year-age | 50.47 | 0.49 | 48.25 | 53.73 | 552 |
| Panel B: Influenza severity | | | | | | |
| Exposed cohort (dummy variable) | Prefecture-birth year | 0.50 | | | | 552 |
| Weighted influenza death rate (per 10,000 people) | Prefecture-birth year | 0.04 | 0.06 | 0.00 | 0.22 | 552 |
| Panel C: Control variables | | | | | | |
| Rice yield per hectare (hectoliter) | Prefecture-birth year | 33.65 | 6.92 | 1.41 | 77.15 | 414 |
| Soy yield per hectare (hectoliter) | Prefecture-birth year | 16.03 | 3.93 | 4.71 | 54.63 | 414 |
| Milk production per capita (liter) | Prefecture-birth year | 1.07 | 1.10 | 0.15 | 10.87 | 414 |
| Coverage of doctors (per 100 people) | Prefecture-birth year | 3.12 | 1.55 | 0.28 | 12.61 | 414 |
| Coverage of midwives (per 100 people) | Prefecture-birth year | 2.47 | 1.25 | 0.19 | 7.68 | 414 |

Notes: Panel A reports the summary statistics for the proportion of males (%) observed at the prefecture-year-age level. The summary statistics for children aged 5–7 years and 10–12 years in both 1925 and 1930 are listed in this table. Panel B reports the summary statistics for the prefecture-birth year-level indicator variable and weighted influenza death rate. Children aged 5–7 years in 1925 and 10–12 years in 1930 are defined as the exposed cohort. Panel C reports the summary statistics for the prefecture-birth year-level control variables. The number of observations is 414, because there are 46 prefectures \times 9 birth cohorts in our sample. Sources: The dependent variables are from the VSEJ (1912–1925 editions). The influenza death rates are from the SCDEJ (1912–1925 editions); VSEJ (1912–1925 editions); Statistical Survey Department, Statistics Bureau, Ministry of Internal Affairs and Communications (database). The control variables are from Kayo (1983) and Statistics Bureau of the Cabinet (1914–1927).

We begin our analysis by estimating the cohort effects of fetal exposure to pandemic influenza using the following specification:

$$s_{it-a} = \varpi + \chi Exposed_{it-a} + \mathbf{x}'_{it-a} \Delta + \varphi_{it} + \eta_a + o_{it-a}, \quad (7)$$

where i indexes the prefecture, t indexes the measured year, a indexes the age, and thus $t - a$ indexes the cohort (birth year). The variable s is the proportion of male births, *Exposed* is an indicator variable for the exposed cohorts (Figure 4), \mathbf{x} is a vector of the prefecture-birth year-level control variables, φ is the prefecture-year-specific fixed effect, η is the age fixed effect, and o is a random error term. We expect $\hat{\chi}$ to be negative and statistically significant, as it captures the cohort effects of fetal influenza exposure on the proportion of male births.

Our main specification is then designed to estimate the marginal effects of fetal influenza exposure on the proportion of male births:

$$s_{it-a} = \Upsilon + \Xi Weighted\ FLUDR_{it-a} + \mathbf{x}'_{it-a} \Theta + \vartheta_{it} + \Omega_a + \Lambda_{it-a}, \quad (8)$$

where *Weighted FLUDR* is the weighted influenza death rate defined in equation 4. Similarly, we expect $\hat{\Xi}$ to be statistically significantly negative. In this specification, \mathbf{x} includes the same control variables used in equation 2: rice yield, soy yield, milk production, coverage of doctors, and coverage of midwives. An important difference is that we use these variables to control for the variations in the *birth* year (i.e., 1913–1925) rather than the *measured* year (i.e., 1925 and 1930). Therefore, these variables are used to control for the birth year heterogeneities in the potential wealth level and socioeconomic conditions that might be correlated with *Weighted FLUDR*. Panels B and C of Table 4 show the summary statistics for the key and control variables, respectively. On the contrary, the instantaneous effects, namely, any unobserved shocks in the prefecture-*measured* year cells such as local economic shocks, are captured by the prefecture-year-specific fixed effect ϑ . The age fixed effect, Ω , captures the common trend in the proportion of males over time. Thus, the identification assumption is that after controlling for these observed and unobserved factors, *Weighted FLUDR* is uncorrelated with the error term Λ_{it-a} . Together with the randomness of the pandemics, our key variable is thus considered to be plausibly exogenous. However, the specifications of both equations 7 and 8 assume a common trend in the proportion of males across prefectures. To relax this assumption, we therefore allow the trend of the dependent variable to vary across prefectures using the prefecture-specific

trend, say $a\Psi_i$, in some of the specifications.

To address the potential spatial and prefecture-specific within correlations, we use the CRVE and cluster the standard errors at the 8-area level. Since our data are a three-dimensional (i.e., prefecture-measured year-age) panel, this clustering can mitigate the potential correlations across cohorts. To overcome the issue of the small number of clusters, we use the wild cluster bootstrap-t method for the statistical inference. All the regressions are weighted by the average number of children in each prefecture-year cell.

5.2 Results

Table 5: Effects of fetal influenza exposure on the proportion of males (%) between 5 and 12 years old

| | Dependent variable: Proportion of males (%) | | | |
|----------------------------------------|---------------------------------------------|--------------------------------|--------------------------------|--------------------------------|
| | (1) | (2) | (3) | (4) |
| Exposed cohort | −0.100** (0.028) [0.020] | −0.126** (0.036) [0.033] | | |
| Weighted FLUDR | | | −0.655** (0.274) [0.020] | −0.828** (0.279) [0.047] |
| Prefecture-year-specific fixed effect | Yes | Yes | Yes | Yes |
| Age fixed effect | Yes | Yes | Yes | Yes |
| Heterogeneous trend across prefectures | No | Yes | No | Yes |
| Observations | 552 | 552 | 552 | 552 |
| Number of prefectures | 46 | 46 | 46 | 46 |
| Number of clusters | 8 | 8 | 8 | 8 |
| Measured years | 1925 & 1930 | 1925 & 1930 | 1925 & 1930 | 1925 & 1930 |
| Ages | 5–7 & 10–12 | 5–7 & 10–12 | 5–7 & 10–12 | 5–7 & 10–12 |

***, **, and * represent statistical significance at the 1%, 5%, and 10% levels based on the p -values from the wild cluster bootstrap resampling method in brackets, respectively. The number of replications is fixed to 150 for all the specifications. Standard errors from the cluster-robust variance estimation reported in parentheses are clustered at the 8-area level.

Notes: Exposed cohorts are 5–7-year-olds in 1925 and 10–12-year-olds in 1930. All the regressions include the birth year controls for rice yield, soy yield, milk production, coverage of doctors, and coverage of midwives. All the regressions are weighted by the average number of children in each prefecture-year cell.

Table 5 presents the results. Columns (1) and (3) present the estimates from equations 7 and 8, respectively, whereas columns (2) and (4) show the estimates from the specification including the prefecture-specific trend. The estimate in column (1) shows that the exposed cohort, on average, has a 0.1 % lower proportion of males than the surrounding cohorts. This magnitude becomes 0.126% if we relax the common trend assumption across prefectures in column (2), accounting for approximately 26% of the standard deviation of the dependent variable (Panel A of Table 4). Column (3) indicates

that the estimated coefficient on the weighted influenza death rate is negative and statistically significant. This result remains unchanged if we include the prefecture-specific trend in column (4). The estimate in column (4) shows that a one standard deviation (0.06 in Panel B of Table 4) increase in the rate decreases the proportion of males in those children by approximately 0.05% (0.06×0.828). Further, the proportion of males decreases by approximately 0.18% ($0.22 \times 0.828\%$) in the case of exposure to the maximum influenza death rate. This accounts for roughly 40% of the standard deviation of the dependent variable and thus is non-negligible in terms of its magnitude. The foregoing results suggest that fetal exposure to pandemic influenza has lasting effects on the sex ratio after birth, at least for 5–12-year-olds.

6 Conclusion

This study uses the pandemic influenza in pre-war Japan as a natural experiment to investigate mortality selection *in utero*. We find that fetal influenza exposure during the first trimester of the pregnancy period had negative impacts on the proportion of males at birth. Analyses using the infant mortality rate as a proxy of the health status of infants provide evidence that the reduction in male births was associated with the “scarring” mechanism rather than the “selection” mechanism. We also find that the proportion of males in the affected birth cohorts was statistically significantly lower than those of the surrounding cohorts.

As discussed in the Introduction, potential barriers for studies in this strand of the literature include difficulties compiling a set of birth records in developing countries and measurement errors in the observed birth records such as age heaping. Given this issue, industrializing Japan is an ideal study setting because the Registration Act was set in the early stage of its industrialization. Therefore, Japan has comprehensive birth registration records from the beginning of the 20th century. This advantage enables us to investigate in detail the potential impacts of the examined pandemics on the sex ratio. However, using aggregate prefecture-level data makes it difficult to precisely identify the actual assignments of the exposure at the individual level. Although we use an appropriate set of weights for the analyses, the estimated effects found in this study are therefore considered to be the lower bounds rather than the true unobserved treatment effects.

Nevertheless, this study contributes to our understanding of potential mortality selection *in utero* due to the pandemics in early 20th century Japan. It also offers suggestive evidence of the persistency of the pandemics in the gender imbalance into childhood. An-

alyzing the potential long-term effects of fetal exposure to the pandemics on the gender imbalance in adulthood may be a future research avenue.

References

- Ran Abramitzky, Adeline Delavande, and Luis Vasconcelos. Marrying Up: The Role of Sex Ratio in Assortative Matching. *American Economic Journal: Applied Economics*, 3(3):124–157, July 2011.
- Douglas Almond. 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):1–41, 2006.
- Douglas Almond and Janet Currie. Killing Me Softly: The Fetal Origins Hypothesis. *Journal of Economic Perspectives*, 25(3):153–172, August 2011.
- Douglas Almond and Bhashkar Mazumder. Almond and Mazumder 2005 [AERPP]. *American Economic Review*, 95(2):258–262, May 2005.
- Mabel Andalón, João Pedro Azevedo, Carlos Rodríguez-Castelán, Viviane Sanfelice, and Daniel Valderrama-González. Weather Shocks and Health at Birth in Colombia. *World Development*, 82(C):69–82, June 2016.
- Francis M Andrew. Sex Ratios and the Red Dragon: Using the Chinese Communist Revolution to Explore the Effect of the Sex Ratio on Women and Children in Taiwan. *Journal of Population Economics*, 24(3):813–837, 2011.
- Josh Angrist. How do sex ratios affect marriage and labor markets? Evidence from America’s second generation. *Quarterly Journal of Economics*, 117(3):997–1038, August 2002.
- Badi H. Baltagi. *Econometric analysis of panel data, 5th edition*. Wiley, 2013.
- Megan Beckett, Julie Da Vanzo, Narayan Sastry, Constantijn Panis, and Christine Peterson. The quality of retrospective data: an examination of long-term recall in a developing country. *Journal of Human Resources*, 36(3):593–625, 2001.
- Dirk Bethmann and Michael Kvasnicka. World War II, Missing Men and Out of Wedlock Childbearing. *The Economic Journal*, 123(567):162–194, August 2012.

- Dirk Bethmann and Michael Kvasnicka. War, Marriage Markets, and the Sex Ratio at Birth. *The Scandinavian Journal of Economics*, 116(3):859–877, April 2014.
- Elizabeth Brainerd. The Lasting Effect of Sex Ratio Imbalance on Marriage and Family: Evidence from World War II in Russia. *Review of Economics and Statistics*, 99(2): 229–242, May 2017.
- A Colin Cameron, Jonah B Gelbach, and Douglas L Miller. Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, 90(3): 414–427, July 2008.
- T A Carleton and S M Hsiang. Social and economic impacts of climate. *Science*, 353 (6304):aad9837–aad9837, September 2016.
- Ralph Catalano and Tim Bruckner. Secondary sex ratios and male lifespan: Damaged or culled cohorts. *PNAS*, 103(5):1639–1643, 2006.
- Central Sanitary Bureau of the Home Ministry. *Ryūkōsei kanbō (Influenza)*. [in Japanese]. Central Sanitary Bureau of the Home Ministry, Tokyo, 1927.
- Pierre André Chiappori, Bernard Fortin, and Guy Lacroix. Marriage Market, Divorce Legislation, and Household Labor Supply. *Journal of Political Economy*, 110(1):37–72, February 2002.
- In Choi. Unit root tests for panel data. *Journal of International Money and Finance*, 20 (2):249–272, April 2001.
- Janet Currie and Maya Rossin-Slater. Weathering the storm: Hurricanes and birth outcomes. *Journal of Health Economics*, 32(3):487–503, May 2013.
- Janet Currie and Tom Vogl. Early-Life Health and Adult Circumstance in Developing Countries. *Annual Review of Economics*, 5(1):1–36, August 2013.
- Olivier Deschênes, Michael Greenstone, and Jonathan Guryan. Climate Change and Birth Weight. *American Economic Review*, 99(2):211–217, April 2009.
- Fabian F Drixler. Hidden in Plain Sight: Stillbirths and Infanticides in Imperial Japan. *Journal of Economic History*, 76(3):651–696, September 2016.
- Anton Erkoreka. The Spanish influenza pandemic in occidental Europe (1918-1920) and victim age. *Influenza and Other Respiratory Viruses*, 4(2):81–89, March 2010.

- Akira Hayami. *Nihon wo osotta supein infuruenza (Spanish influenza struck down to Japan)*. [in Japanese]. Fujiwara shoten, Tokyo, 2006.
- Akira Hayami. An estimation of Spanish influenza mortality in Imperial Japan: 1918–20. In Satomi Kurosu, Tommy Bengtsson, and Cameron Campbell, editors, *Demographic Responses to Economic and Environmental Crises*, pages 282–300. Reitaku University Press, Kashiwa, 2010.
- Sonoko Hijikata. *Kindainihon no gakkō to chiikishakai (Schools and community in prewar Japan)*. [in Japanese]. University of Tokyo Press, Tokyo, 1994.
- Cheng Hsiao. *Analysis of panel data, 3rd edition*.
- Akihiko Kawana, Go Naka, Yuji Fujikura, Yasuyuki Kato, Yasutaka Mizuno, Tatsuya Kondo, and Koichiro Kudo. Spanish Influenza in Japanese Armed Forces, 1918–1920. *Emerging Infectious Diseases*, 13(4):590–593, April 2007.
- Nobufumi Kayo. *Todōfuken nōgyōkisetōkei (Basic Statistics of Agriculture in Japanese Prefecture)*. [in Japanese]. Nōrintōkeikyōkai, Tokyo, 1983.
- Edwin D Kilbourne. Influenza Pandemics of the 20th Century. *Emerging Infectious Diseases*, 12(1):9–14, January 2006.
- Carolyn Kousky. Impacts of natural disasters on children. *The Future of Children*, 26(1):73–92, 2016.
- Sebastian Kraemer. The fragile male. *British Medical Journal*, 321(7276):1609–1612, December 2000.
- Ming-Jen Lin and Elaine M Liu. Does in utero exposure to Illness matter? The 1918 influenza epidemic in Taiwan as a natural experiment. *Journal of Health Economics*, 37:152–163, September 2014.
- Oswaldo Molina and Victor Saldarriaga. The perils of climate change: In utero exposure to temperature variability and birth outcomes in the Andean region. *Economics and Human Biology*, 24:111–124, February 2017.
- Keith L Moore, Persaud T.V.N. (Vid), and Mark G Torchia. *The Developing Human: Clinically Oriented Embryology, 11th edition*. Elsevier, Saunders, Philadelphia, 2013.

- Kota Ogasawara. Persistence of pandemic influenza on the development of children: Evidence from industrializing Japan. *Social Science & Medicine*, 181:43–53, May 2017.
- Kota Ogasawara. The long-run effects of pandemic influenza on the development of children from elite backgrounds: Evidence from industrializing Japan. *Economics and Human Biology*, 31:125–137, September 2018.
- Kota Ogasawara and Yukitoshi Matsushita. Heterogeneous treatment effects of safe water on infectious disease: Do meteorological factors matter? *Econometrica*, 13(1):55–82, January 2019.
- Kota Ogasawara and Minami Yumitori. Early-life exposure to weather shocks and child height: Evidence from industrializing Japan. *SSM - Population Health*, 7:100317, December 2018.
- Kota Ogasawara, Ian Gazeley, and Eric B Schneider. Nutrition, crowding, and disease among low-income households in Tokyo in 1930. *Australian Economic History Review*, 60(1):73–104, February 2020.
- Parinaz Poursafa, Mojtaba Keikha, and Roya Kelishadi. Systematic review on adverse birth outcomes of climate change. *Journal of Research in Medical Sciences*, 20(4):397–402, April 2015.
- Daniel M Prinz, David D Cutler, and Austin Frakt. Health and economic activity over the lifecycle: Literature review. *NBER Working Paper Series*, 24865:1–128, 2018.
- Alice Reid. The Effects of the 1918–1919 Influenza Pandemic on Infant and Child Health in Derbyshire. *Medical History*, 49:29–54, 2005.
- Geoffrey W Rice and Edwina Palmer. Pandemic influenza in Japan, 1918-19: Mortality patterns and official responses. *Journal of Japanese Studies*, 19(2):389–420, 1993.
- S A Richard, N Sugaya, L Simonsen, M A Miller, and C Viboud. A comparative study of the 1918–1920 influenza pandemic in Japan, USA and UK: mortality impact and implications for pandemic planning. *Epidemiology and Infection*, 137(08):1062–18, February 2009.
- Nicholas J Sanders and Charles Stoecker. Where have all the young men gone? Using sex ratios to measure fetal death rates. *Journal of Health Economics*, 41:30–45, May 2015.

- Eric B Schneider and Kota Ogasawara. Disease and child growth in industrialising Japan: Critical windows and the growth pattern, 1917-39. *Explorations in Economic History*, 69:64–80, July 2018.
- Ito Shigeru. An analysis of prefectural differentials on fertility in early modern Japan [in Japanese]. *Obihirochikusandaigaku gakujyutukenkyūhōkoku I*, 15(2):145–155, 1987.
- Statistics Bureau of the Cabinet. *Nihonteikoku tōkeinenkan (The statistical yearbook of the Japanese Empire, volumes 33–46)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1914–1927.
- Statistics Bureau of the Cabinet. *Nihonteikoku shiintōkei (The statistics of causes of death of the empire Japan, 1912 edition)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1916a.
- Statistics Bureau of the Cabinet. *Nihonteikoku shiintōkei (The statistics of causes of death of the empire Japan, 1913 edition)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1916b.
- Statistics Bureau of the Cabinet. *Nihonteikoku jinkōdōtaitōkei (The vital statistics of empire Japan, 1912 edition)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1916c.
- Statistics Bureau of the Cabinet. *Nihonteikoku jinkōdōtaitōkei (The vital statistics of empire Japan, 1913 edition)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1917.
- Statistics Bureau of the Cabinet. *Nihonteikoku shiintōkei (The statistics of causes of death of the empire Japan, 1914 edition)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1918a.
- Statistics Bureau of the Cabinet. *Nihonteikoku shiintōkei (The statistics of causes of death of the empire Japan, 1915 edition)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1918b.
- Statistics Bureau of the Cabinet. *Nihonteikoku jinkōdōtaitōkei (The vital statistics of empire Japan, 1914 edition)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1918c.

Statistics Bureau of the Cabinet. *Nihonteikoku jinkōdōtaitōkei (The vital statistics of empire Japan, 1915 edition)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1918d.

Statistics Bureau of the Cabinet. *Nihonteikoku shiintōkei (The statistics of causes of death of the empire Japan, 1916 edition)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1919a.

Statistics Bureau of the Cabinet. *Nihonteikoku jinkōdōtaitōkei (The vital statistics of empire Japan, 1916 edition)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1919b.

Statistics Bureau of the Cabinet. *Nihonteikoku shiintōkei (The statistics of causes of death of the empire Japan, 1917 edition)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1920a.

Statistics Bureau of the Cabinet. *Nihonteikoku jinkōdōtaitōkei (The vital statistics of empire Japan, 1917 edition)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1920b.

Statistics Bureau of the Cabinet. *Nihonteikoku shiintōkei (The statistics of causes of death of the empire Japan, 1918 edition)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1921a.

Statistics Bureau of the Cabinet. *Nihonteikoku jinkōdōtaitōkei (The vital statistics of empire Japan, 1918 edition)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1921b.

Statistics Bureau of the Cabinet. *Nihonteikoku shiintōkei (The statistics of causes of death of the empire Japan, 1919 edition)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1922a.

Statistics Bureau of the Cabinet. *Nihonteikoku jinkōdōtaitōkei (The vital statistics of empire Japan, 1919 edition)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1922b.

Statistics Bureau of the Cabinet. *Nihonteikoku shiintōkei (The statistics of causes of death of the empire Japan, 1920 edition)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1923.

- Statistics Bureau of the Cabinet. *Nihonteikoku shiintōkei (The statistics of causes of death of the empire Japan, 1921 edition)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1924a.
- Statistics Bureau of the Cabinet. *Nihonteikoku jinkōdōtaitōkei (The vital statistics of empire Japan, 1920 edition)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1924b.
- Statistics Bureau of the Cabinet. *Nihonteikoku jinkōdōtaitōkei (The vital statistics of empire Japan, 1921 edition)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1924c.
- Statistics Bureau of the Cabinet. *Nihonteikoku jinkōdōtaitōkei (The vital statistics of empire Japan, 1922 edition)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1924d.
- Statistics Bureau of the Cabinet. *Nihonteikoku shiintōkei (The statistics of causes of death of the empire Japan, 1922 edition)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1925a.
- Statistics Bureau of the Cabinet. *Nihonteikoku shiintōkei (The statistics of causes of death of the empire Japan, 1923 edition)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1925b.
- Statistics Bureau of the Cabinet. *Nihonteikoku shiintōkei (The statistics of causes of death of the empire Japan, 1924 edition)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1925c.
- Statistics Bureau of the Cabinet. *Nihonteikoku jinkōdōtaitōkei (The vital statistics of empire Japan, 1923 edition)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1925d.
- Statistics Bureau of the Cabinet. *Nihonteikoku jinkōdōtaitōkei (The vital statistics of empire Japan, 1924 edition)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1925e.
- Statistics Bureau of the Cabinet. *Taishōjyūyōnen Kokuseichōsahōkoku, dainikan (1925 Population Census of Japan, volume 2)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1926a.

- Statistics Bureau of the Cabinet. *Nihonteikoku shiintōkei (The statistics of causes of death of the empire Japan, 1925 edition)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1926b.
- Statistics Bureau of the Cabinet. *Nihonteikoku jinkōdōtaitōkei (The vital statistics of empire Japan, 1925 edition)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1926c.
- Statistics Bureau of the Cabinet. *Taishōkyūnen Kokuseichōsahōkoku, daiikkān (1920 Population Census of Japan, volume 1)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1928.
- Statistics Bureau of the Cabinet. *Taishōjyūyōnen Kokuseichōsahōkoku, fukēhen (1925 Population Census of Japan, prefecture part)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1929.
- Statistics Bureau of the Cabinet. *Shōwagōnen Kokuseichōsahōkoku, fukēhen (1930 Population Census of Japan, prefecture part)*. [in Japanese]. Statistics Bureau of the Cabinet, Tokyo, 1933.
- Jeffery K Taubenberger. 1918 Influenza: the Mother of All Pandemics. *Emerging Infectious Diseases*, 12(1):15–22, January 2006.
- R L Trivers and D E Willard. Natural Selection of Parental Ability to Vary the Sex Ratio of Offspring. *Science*, 179(4068):90–92, January 1973.
- Christine Valente. Civil conflict, gender-specific fetal loss, and selection: A new test of the Trivers–Willard hypothesis. *Journal of Health Economics*, 39:31–50, January 2015.
- Yunquan Zhang, Chuanhua Yu, and Lu Wang. Temperature exposure during pregnancy and birth outcomes: An updated systematic review of epidemiological evidence. *Environmental Pollution*, 225:700–712, June 2017.
- Joshua Graff Zivin and Jeffrey Shrader. Temperature extremes, health, and human capital. *The Future of Children*, 26(1):31–50, May 2016.

Appendices

Appendix A Data Appendix

| 順序 Numéros d'ordre. | 月 Mois. | | 生 産 Naissances vivantes. | | | 死 産 Mort-nés. | | | |
|------------------------|------------|--------------|-----------------------------|--------------|-------------|------------------|--------------|---------------------|-------------|
| | | | 男 Garçons. | 女 Filles. | 計 Total. | 男 Garçons. | 女 Filles. | 不詳 Sexe inconnu. | 計 Total. |
| | | | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| | | | | | | | | | |
| | | | | | | | | | |
| 1. 東京府 Tokio-fu. | | | | | | | | | |
| 1 | 一 月 | Janvier | 6,151 | 6,385 | 12,536 | 349 | 313 | 2 | 664 |
| 2 | 二 月 | Février | 4,348 | 4,400 | 8,748 | 331 | 261 | 4 | 596 |
| 3 | 三 月 | Mars | 5,655 | 5,429 | 11,084 | 329 | 283 | 2 | 614 |
| 4 | 四 月 | Avril | 3,190 | 3,040 | 6,230 | 289 | 245 | 2 | 536 |
| 5 | 五 月 | Mai | 3,107 | 3,013 | 6,120 | 280 | 246 | 1 | 527 |
| 6 | 六 月 | Jun | 2,927 | 2,863 | 5,800 | 284 | 227 | 1 | 512 |
| 7 | 七 月 | Juillet | 3,254 | 3,064 | 6,418 | 338 | 259 | 2 | 599 |
| 8 | 八 月 | Août | 3,559 | 3,399 | 6,958 | 290 | 253 | 5 | 548 |
| 9 | 九 月 | Septembre | 3,782 | 3,646 | 7,428 | 341 | 253 | 1 | 595 |
| 10 | 十 月 | Octobre | 4,054 | 3,754 | 7,808 | 323 | 277 | 1 | 601 |
| 11 | 十一月 | Novembre | 4,281 | 3,944 | 8,225 | 297 | 248 | — | 545 |
| 12 | 十二月 | Décembre | 3,242 | 3,104 | 6,346 | 326 | 292 | 3 | 621 |
| 13 | 不 詳 | Mois inconnu | — | — | — | — | — | — | — |
| | 計 | Total | 47,660 | 46,041 | 93,701 | 3,787 | 3,187 | 24 | 6,998 |

Figure B.1: Example of the vital statistics record from the VSEJ

Notes: This image shows an example of the VSEJ of 1915. Source: Statistics Bureau of the Cabinet (1918d, pp.44).

Figure B.1 shows an example image of the VSEJ that we mainly use to construct the dataset. The SCDEJ takes a similar style to the VSEJ (not reported).

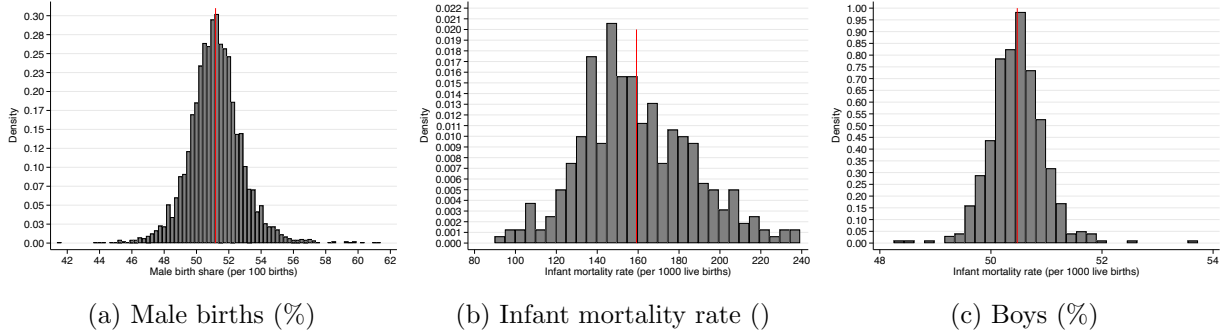


Figure B.2: Distributions of the dependent variables

Notes: Figure B.2a shows the histogram of the proportion of male births (%). Figure B.2b shows the histogram of the infant mortality rates (per 1000 live births). Figure B.2c shows the histogram of the proportion of boys aged 5–7 and 10–12 (%). The red lines indicate the mean values of each variable. Sources: Created by the authors using Statistics Bureau of the Cabinet (1917, 1918c,d, 1919b, 1920b, 1921b, 1922b, 1924b,c,d, 1925d,e, 1926c).

Figure B.2 shows the histograms of the dependent variables used: the proportion of male births (Figure B.2a), infant mortality rate (Figure B.2b, and proportion of boys aged 5–7 and 10–12 (Figure B.2c).

Appendix B Empirical Analysis Appendix

B.1 Testing Stationarity

Table C.1 presents the results of the panel unit root tests for the proportion of male births used in the main empirical analyses, confirming the stationarity of our panel dataset on the secondary sex ratio. We also run regressions for the proportion of male births using the dynamic panel data models in Online Appendix B.3.

Table C.1: Results of the unit root tests for the proportion of male births

| Test statistics | January 1916–December 1922 | | January 1918–December 1920 | |
|------------------------------|----------------------------|--------|----------------------------|--------|
| | (1) | (2) | (3) | (4) |
| P -statistic p -value | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Z -statistic p -value | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| L^* -statistic p -value | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| P_m -statistic p -value | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Number of prefectures | 46 | 46 | 46 | 46 |
| Number of periods | 84 | 84 | 36 | 36 |
| Number of lagged differences | 1 | 3 | 1 | 3 |

Notes: The results of the Fisher-type panel unit root tests based on augmented Dickey–Fuller (ADF) tests are reported in this table. The null hypothesis is that all the panels contain unit roots, whereas the alternative hypothesis is that at least one panel is stationary. In all the specifications, the process under the null hypothesis is assumed to be a random walk with drift. The demeaned data are used to address the effect of cross-sectional dependence. Although the number of lagged differences in the ADF regression equation reported is set as either one or three, the results are not affected by the number of lagged differences. See Choi (2001) for the details of the tests.

B.2 Testing the Gender Difference

In this subsection, we investigate the gender difference in the estimates reported in Table 3. To test the difference, we pool the infant mortality rates for boys and girls and then interact all the independent variables including the fixed effects with the gender dummy. Table C.2 presents the results. As shown, the number of observations is now 644 (322 + 322). Column (1) indicates that the estimated effect of fetal influenza exposure on the infant mortality rate of boys is approximately 45.8 lower than that of girls. This result is largely unchanged if we include the prefecture-specific time trend in column (2). If we disaggregate the exposure variable in columns (3) and (4), however, no such gender difference is observed in a statistical sense.

Table C.2: Effects of fetal influenza exposure on the infant mortality rate:
Testing the gender difference

| Exposed trimesters | Dependent variable: Infant mortality rate | | | |
|---------------------------------|-------------------------------------------|----------------------------------|----------------------------------|----------------------------------|
| | (1) | (2) | (3) | (4) |
| All trimesters | 87.853*** (31.528) [0.007] | 92.515*** (35.695) [0.007] | | |
| All trimesters \times Boys | -45.810** (17.942) [0.033] | -44.669** (19.991) [0.033] | | |
| First trimester | | | -191.282 (77.676) [0.100] | -182.345 (81.909) [0.113] |
| First trimester \times Boys | | | 59.690 (31.092) [0.193] | 55.391 (34.814) [0.180] |
| Second trimester | | | -36.836 (69.934) [0.593] | -42.073 (80.416) [0.633] |
| Second trimester \times Boys | | | -57.404 (24.874) [0.100] | -60.506 (27.292) [0.100] |
| Third trimester | | | 247.938** (71.271) [0.020] | 247.215** (80.381) [0.020] |
| Third trimester \times Boys | | | -37.179 (25.600) [0.247] | -29.827 (27.612) [0.313] |
| Control variables | Yes | Yes | Yes | Yes |
| Control variables \times Boys | Yes | Yes | Yes | Yes |
| Fixed effects | Yes | Yes | Yes | Yes |
| Fixed effects \times Boys | Yes | Yes | Yes | Yes |
| Time trend | No | Yes | No | Yes |
| Time trend \times Boys | No | Yes | No | Yes |
| Observations | 644 | 644 | 644 | 644 |
| Number of prefectures | 46 | 46 | 46 | 46 |
| Number of clusters | 8 | 8 | 8 | 8 |

***, **, and * represent statistical significance at the 1%, 5%, and 10% levels based on the p -values from the wild cluster bootstrap resampling method in brackets, respectively. The number of replications is fixed to 200 for all the specifications. Standard errors from the cluster-robust variance estimation reported in parentheses are clustered at the 8-area level.

Notes: This table shows the results of the specifications using the sample pooling the infant mortality rates for boys and girls. As shown, all the independent variables are interacted with an indicator variable for boys. The control variables include rice yield, soy yield, milk production, coverage of doctors, and coverage of midwives. The fixed effects include both the prefecture and the year fixed effects. The time trend indicates the prefecture-specific linear time trend. All the regressions are weighted by the average number of live births in each prefecture.

Table C.3: Effects of fetal influenza exposure on the proportion of male births:
Dynamic panel data analysis and placebo test

| | Dynamic panel | | Placebo test | |
|----------------------------|---------------------------------|------------------------------------|-----------------------------------|-----------------------------------|
| Exposed trimesters | (1) | (2) | (3) | (4) |
| All trimesters | −0.1633 (0.0740) [0.1490] | | | |
| First trimester | | −0.1108*** (0.0317) [0.0030] | −0.0819** (0.0424) [0.0330] | −0.0754** (0.0273) [0.0390] |
| Second trimester | | −0.0460 (0.0315) [0.2690] | −0.0327 (0.0291) [0.3790] | −0.0266 (0.0290) [0.5390] |
| Third trimester | | −0.0054 (0.0314) [0.8770] | 0.0033 (0.0252) [0.9090] | 0.0081 (0.0262) [0.7410] |
| Lagged dependent variable | 0.0027 (0.0258) [0.9170] | 0.0016 (0.0254) [0.9250] | | |
| Zero trimester (placebo) | | | 0.1397 (0.0601) [0.1310] | 0.1449 (0.0634) [0.1310] |
| Fourth trimester (placebo) | | | 0.0015 (0.0553) [0.9970] | 0.0046 (0.0582) [0.9970] |
| Time trend | No | No | No | Yes |
| Period | 1916–22 | 1916–22 | 1916–22 | 1916–22 |
| Observations | 3818 | 3818 | 3864 | 3864 |
| Number of prefectures | 46 | 46 | 46 | 46 |
| Number of clusters | 8 | 8 | 8 | 8 |

***, **, and * represent statistical significance at the 1%, 5%, and 10% levels based on the p -values from the wild cluster bootstrap resampling method in brackets, respectively. The number of replications is fixed to 1,000 for all the specifications. Standard errors from the cluster-robust variance estimation reported in parentheses are clustered at the 8-area level.

Notes: All the regressions include controls for rice yield, soy yield, milk production, coverage of doctors, and coverage of midwives. All the regressions are weighted by the average number of births in each prefecture.

B.3 Alternative Specifications: Dynamic Panel Data Analysis and Placebo Test

To assess the robustness of our main results, we run regressions using alternative specifications. First, throughout Section 3, we assumed that the regression specifications are static in nature. Since the secondary sex ratio is a highly biological measure, this assumption is likely to hold. However, we herein further consider dynamic panel data models to test the validity of this assumption.²¹ Column (1) of Table C.3 presents the results

²¹The within estimator in a dynamic panel data model becomes consistent as the number of time periods, say T , rises. The number of sampled time periods in columns (1) and (2) is 84 months in total

of a regression including the lagged dependent variable in our baseline specification in equation 2, whereas column (2) shows the result from a regression including the lagged dependent variable in our flexible specification in equation 3. In both columns, the estimated coefficients on the lagged dependent variables are close to zero and statistically insignificant. This result supports the validity of our main specifications using static panel data models.

Second, we include placebo variables to test whether the significant effect of the first trimester is valid in terms of the regression specifications. First, we include a variable named *Zero trimester* representing the average influenza death rates between 10 and 12 months before birth. Since a fetus does not exist *in utero* before conception, exposure to pandemic influenza during this pre-conception period should have no significant impacts on the proportion of males at birth. Second, in a similar way, we include a variable named *Fourth trimester* representing the average influenza death rates between one and three months *after* birth. The estimate of this variable should also be statistically insignificant because fetuses could not be infected during this post-birth period. Column (3) of Table C.4 indicates that the estimated coefficients on these placebo variables are statistically insignificant. This result is unchanged if we include the prefecture-specific time trend in column (4).

B.4 Alternative Specifications: Additional Weather Shock Variables

As discussed in the Introduction, weather shocks can also affect birth outcomes. Moreover, temperature might correlate with the risk of infectious diseases (Ogasawara and Matsushita 2019; Ogasawara and Yumitori 2018). In this subsection, we thus check the robustness of our main results to including measures of heat and cold waves during the sample period.

First, we download temperature data from the official database of the Japanese Meteorological Agency (JMA): <http://www.data.jma.go.jp/gmd/risk/obsdl/index.php> (accessed on March 30, 2018). In this database, the JMA reports information on the number of days with a temperature above or below a certain threshold to record heat and cold waves. We compile monthly meteorological data between 1915 and 1922 for a maximum

(7 years \times 12 months) and thus should be sufficiently large given that the estimated coefficients on the lagged dependent variables are close to zero. See Baltagi (2013, pp. 155–156) and Hsiao, pp. 82–84 for theoretical discussions about dynamic panel data models.

of three weather stations in each prefecture. Following the official definitions provided by the JMA, we then define a heat wave as the annual average number of days on which the maximum temperature exceeds 30°C, whereas a cold wave is the annual average number of days on which the minimum temperature is below 0°C.

Second, we calculate the inverse distance-weighted average of all the valid measurements from these stations in the spirit of Deschênes et al. (2009). Each prefecture’s centroid is set as the city office because a large part of the population lives in the principal city in each prefecture. The weighted average of the weather shock variable for prefecture i in month m is given as follows:

$$\text{IDWS}_{im} = \frac{\sum_{j=1}^3 \frac{w_{imj}}{d_{ij}}}{\sum_{j=1}^3 \frac{1}{d_{ij}}}, \quad (9)$$

where w denotes the weather shock variable and d denotes the geospatial distance from the centroid to station j .²²

Third, we calculate the average number of days of heat and cold waves for all trimesters, as for the average influenza death rates in Section 3. These variables are finally added into our baseline and flexible specifications in equations 2 and 3, respectively.

Table C.4 presents the results. Column (1) indicates that the estimated coefficients on these weather shock variables are statistically insignificant. This result is robust to including the prefecture-specific time trend in column (2). The estimated coefficients on the influenza exposure variables are close to the estimates in columns (3) and (4) of Table 2. This result supports the evidence that our key influenza exposure variables are less likely to be correlated with weather shocks during the pandemics.

²²Data on latitude and longitude are taken from the database of the Geospatial Information Authority of Japan: <http://www.gsi.go.jp/KOKUJYOHO/kenchokan.html>, accessed on August 20, 2017.

Table C.4: Effects of fetal influenza exposure on the proportion of male births:
Additional weather shock variables

| Exposed trimesters | Weather shocks | |
|---------------------------------------|------------------------------------|------------------------------------|
| | (1) | (2) |
| Flu pandemics during first trimester | −0.1236*** (0.0433) [0.0030] | −0.1211*** (0.0415) [0.0030] |
| Flu pandemics during second trimester | −0.0407 (0.0275) [0.2670] | −0.0374 (0.0286) [0.3050] |
| Flu pandemics during third trimester | 0.0094 (0.0225) [0.6930] | 0.0119 (0.0233) [0.6450] |
| Heat waves during first trimester | 0.0058 (0.0059) [0.4370] | 0.0053 (0.0061) [0.4830] |
| Heat waves during second trimester | −0.0048 (0.0059) [0.4930] | −0.0054 (0.0060) [0.4450] |
| Heat waves during third trimester | 0.0004 (0.0047) [0.9090] | −0.0001 (0.0053) [1.0000] |
| Cold waves during first trimester | −0.0006 (0.0051) [0.8590] | −0.0003 (0.0049) [0.9030] |
| Cold waves during second trimester | −0.0040 (0.0045) [0.4390] | −0.0039 (0.0045) [0.4650] |
| Cold waves during third trimester | −0.0052 (0.0087) [0.8110] | −0.0052 (0.0089) [0.8270] |
| Time trend | No | Yes |
| Period | 1916–22 | 1916–22 |
| Observations | 3864 | 3864 |
| Number of prefectures | 46 | 46 |
| Number of clusters | 8 | 8 |

***, **, and * represent statistical significance at the 1%, 5%, and 10% levels based on the p -values from the wild cluster bootstrap resampling method in brackets, respectively. The number of replications is fixed at 1,000 for all the specifications. Standard errors from the cluster-robust variance estimation reported in parentheses are clustered at the 8-area level.
Notes: All the regressions include controls for rice yield, soy yield, milk production, coverage of doctors, and coverage of midwives. All the regressions are weighted by the average number of births in each prefecture.