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

Kota Ogasawara and Minami Yumitori

August 7, 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 evidence suggesting that postnatal influenza exposure had long-term impacts on the sex ratio of children aged 5–12.

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 (2019) and Tokyo Tech seminars (2018–2020) 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] The 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 later in 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).[3]

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

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

[2] Currie 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.

[3] 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).

pandemic influenza and mortality selection *in utero*.

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 1.6%, accounting for roughly 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 analyses, we find evidence that shocks in the postneonatal period due to pandemic influenza might have persisted into childhood. The estimated magnitude is approximately 0.24% in the maximum case, accounting for 50% of the standard deviation.

This study contributes to the wider literature in the following two ways. First, it is the first one to use pandemic influenza as an exogenous shock to test mortality selection *in utero* in the framework of the Trivers–Willard hypothesis. Further, it is the first study to use 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.[4] 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

[4]Sanders 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.

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 include children aged up to 12 years old, we find evidence of persistent effects of postneonatal 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 examines the robustness of the results. Section 7 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:

$$f_b(\lambda) > f_g(\lambda). \tag{1}$$

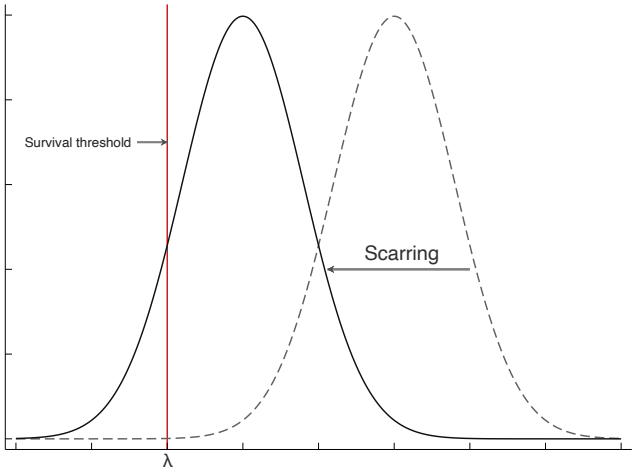
If fetuses are exposed to health shocks *in utero*, the distribution shifts to the left or the survival threshold moves to the right. The former is called the "scarring" mechanism, whereas the latter is called the "selection" mechanism. In both cases, condition (1) indicates that the male share at birth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

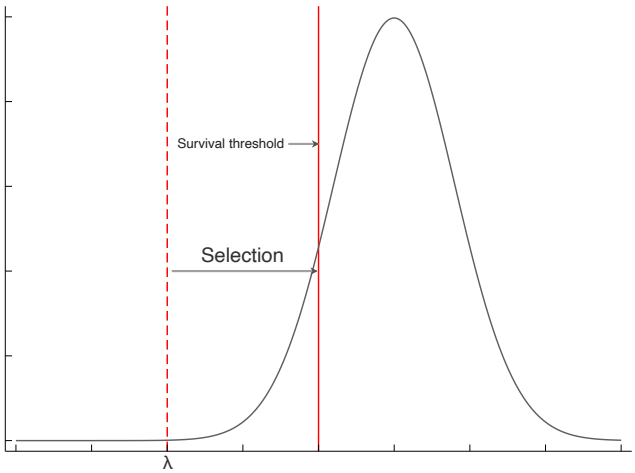
In contrast to culling before birth, the health status of an infant depends on the type of mechanism. If the "scarring" mechanism works, the conditional mean of the truncated normal distribution shifts to the left because the original mean of the distribution moves to the left, as illustrated in Figure 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.[5] 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.

[5]These 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.

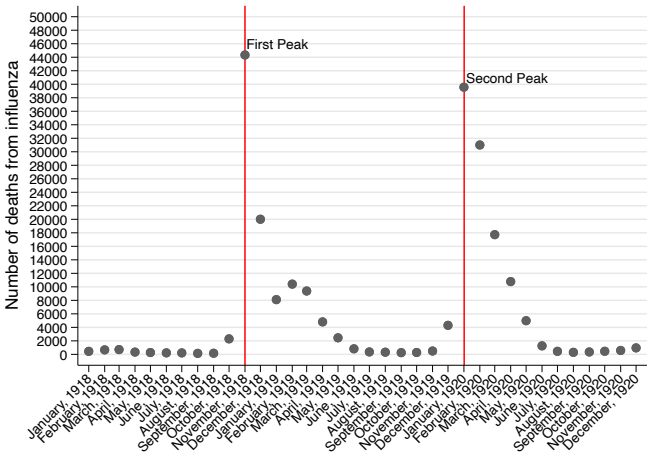


Figure 2: The number of deaths from influenza between January 1918 and December 1920
Notes: The red lines show the first and second peaks of the number of deaths due to influenza.
Source: Created by the authors from Statistics Bureau of the Cabinet (1921a, 1922a, 1923).

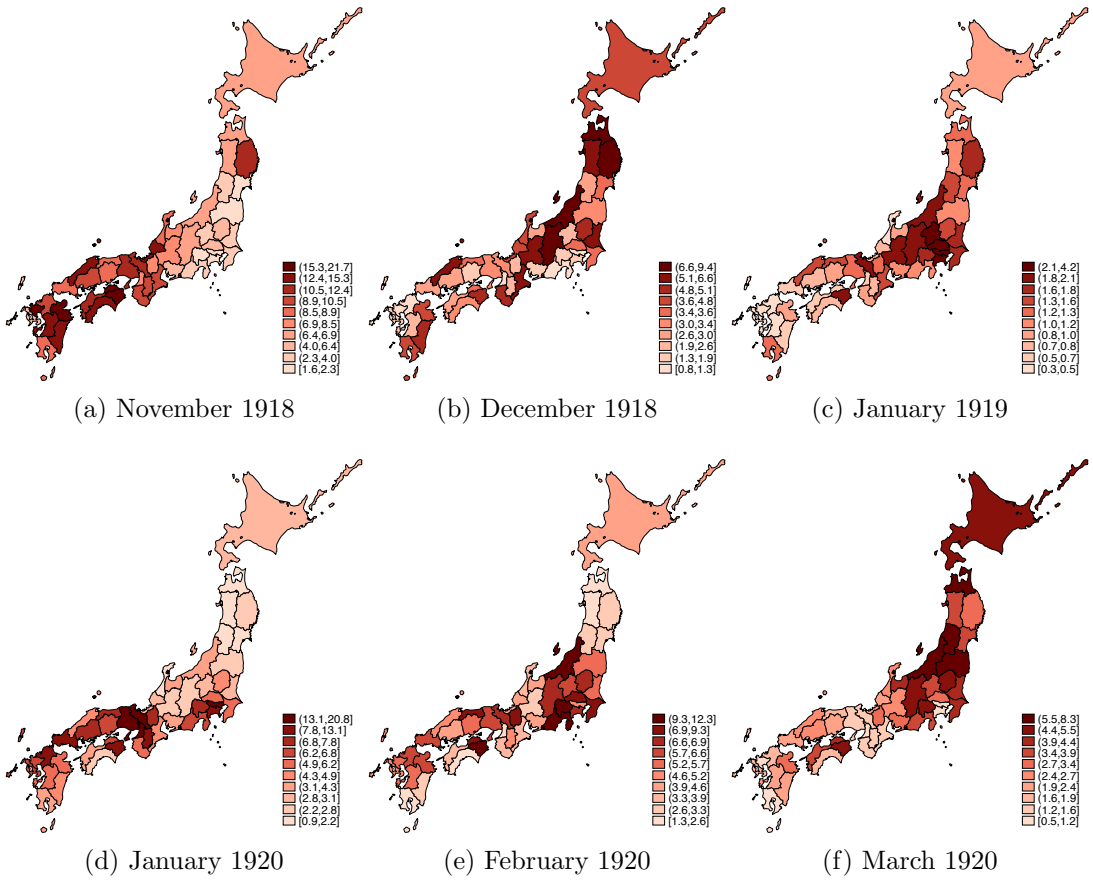
2.2 Maternal Stressor: Pandemic Influenza

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) 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).[6] Since the average age at first marriage of Japanese women was

[6]Similar gender and age biases in infections were observed in Western countries. Influenza-related

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, 1928a). This means that pandemic influenza affected not only young adult women but also children in utero during the pandemics via maternal infection.



(a) November 1918 (b) December 1918 (c) January 1919

(d) January 1920 (e) February 1920 (f) March 1920

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).

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 illustrates the spatial distribution of influenza death rates in the pandemic months of 1918–

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.

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 Tohōku) 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.[7]

Potential sorting might be an issue in the identification because this can cause measurement 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.[8] Using the official vital statistics records published by the Statistics Bureau of the Cabinet, we construct a dataset with the proportions of male births in 46 prefectures between January 1916 and

[7]We have further confirmed that the influenza mortality rates were not positively correlated with the lagged influenza mortality rates during each pandemic wave: November 1918–January 1919 for the first wave and January 1920–March 1920 for the second one. Online Appendix A.3 summarizes these results.

[8]This 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.

December 1922. [9] Since these vital statistics have been recorded based on the comprehensive national registration system (*koseki*), the data cover all births during the measured years.[10] Specifically, 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 A.1 shows an example of this vital statistics record. Panel A of Table 1 lists the summary statistics of the proportion of male births.[11]

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 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.[12] 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.

[9] We excluded Okinawa prefecture from the analyses because its vital statistics sometimes exhibit unnatural values in some cases (Schneider and Ogasawara 2018).

[10] Although 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.

[11] We 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.

[12] These 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 $(= 1, 2, ..., 12)$ 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).

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 nōgyōkisotōkei* (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).[13]

[13]The 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. Prefecture-month-level data include the observations for 46 prefectures between January 1916 and December 1922. Prefecture-year-level data include the observations for 46 prefectures between 1916 and 1922. 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.[14] 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:

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

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

[14] While our monthly panel data on the proportion of male births have a time-series nature, we assumed that the regression specifications are static. This assumption must be made because the secondary sex ratio is a highly biological measure. To test the validity of this assumption, however, we considered a few dynamic panel data models and confirmed that the estimated coefficient of the lagged dependent variable is close to zero and statistically insignificant. This supports the validity of our main specifications using static panel data models. Note that the within estimator in a dynamic panel data model becomes consistent as the number of time periods, T, increases. The number of time periods (months) used is 84 (7 years $\times$ 12 months) and, thus, should be sufficiently large given that the estimated coefficients of 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.

specification is as follows:

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

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.[15] 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. In fact, we have confirmed that the prefectures that experienced different intensities of influenza mortality during the pandemic show very similar trends in the proportion of male births (Online Appendix B.2). In the robustness section, we also confirm the validity of this assumption of a common pretreatment trend (Section 6). To relax the common trend assumption, however, we further include the prefecture-specific time trend ($t\mu_i$) in some of the specifications.

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

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.[16] 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 (January 1916–December 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 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 (January 1916–December 1917 and January 1921–December 1922), whereas columns (7) and (8) present the results for pandemic years (January 1918–December 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.

[16]This geographical classification of Japan includes Hokkaidō (northernmost), Tōhoku (eastern), Kantō (east-central), Chūbu (west-central), Kansai (south-central), Chūgoku (westernmost), Shikoku (southwest 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.

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 (8) implies that fetal exposure to pandemic influenza decreased the proportion of male births by approximately 1.6% (9.93×0.1578). 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).

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 [placebo]		Pandemic years	
Exposed trimesters	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
All trimesters	−0.1609	−0.1528						
	(0.0765)	(0.0740)						
	[0.1490]	[0.1490]						
First trimester			−0.1100***	−0.1072***	0.2099	0.2261	−0.1270***	−0.1578***
			(0.0332)	(0.0323)	(0.1242)	(0.1769)	(0.0355)	(0.0460)
			[0.0030]	[0.0030]	[0.1950]	[0.3310]	[0.0030]	[0.0270]
Second trimester			−0.0452	−0.0423	0.2848	0.3543	−0.0535	−0.0850
			(0.0328)	(0.0324)	(0.2004)	(0.2966)	(0.0412)	(0.0527)
			[0.3010]	[0.3070]	[0.1970]	[0.2610]	[0.3330]	[0.2550]
Third trimester			−0.0045	−0.0024	−0.6410	−0.5861	−0.0022	−0.0230
			(0.0313)	(0.0309)	(0.2974)	(0.3458)	(0.0315)	(0.0390)
			[0.9290]	[0.9670]	[0.1910]	[0.2490]	[0.9470]	[0.6370]
Time trend	No	Yes	No	Yes	No	Yes	No	Yes
Period	Jan. 1916–	Jan. 1916–	Jan. 1916–	Jan. 1916–	Jan. 1916–	Jan. 1916–	Jan. 1918–	Jan. 1918–
	Dec. 1922	Dec. 1922	Dec. 1922	Dec. 1922	Dec. 1917	Dec. 1917	Dec. 1920	Dec. 1920
					Jan. 1921–	Jan. 1921–		
					Dec. 1922	Dec. 1922		
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).[17]

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:

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

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.[18] The baseline specification is then

[17] Although 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.

[18] This 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.

given as follows:

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

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:

$$\begin{aligned} h_{il} = \tau + \varrho_0\, Weighted\ FLUDR_{il}^{First\ \ Trimester} + \varrho_1\, Weighted\ FLUDR_{il}^{Second\ \ Trimester} \\ + \varrho_2\, Weighted\ FLUDR_{il}^{Third\ \ Trimester} + \mathbf{x}'_{il}\boldsymbol{\Gamma} + \xi_i + \varsigma_l + \varepsilon_{il}. \end{aligned} \tag{6}$$

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.[19] 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

[19] We can usually show that the conditional expectation of the truncated normal distribution is always greater than that of the original distribution, as described in Subsection 2.1.

weighted by the average number of live births over the sample period in each prefecture. As regards the common trend assumption, we have confirmed that the prefectures that experienced different intensities of influenza mortality during the pandemic indicate similar trends in infant mortality (Online Appendix B.2).[20] Despite this, to relax the common trend assumption, we include a prefecture-specific time trend ($t\Pi_i$) in some of the specifications.

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). The estimate in column (4) indicates that a one standard deviation increase in the weighted influenza mortality rate increases the infant mortality rate by 11.6 permil (232.289×0.05).

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 permil greater than that for boys ($92.6 - 47.8$). To 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.3 summarizes these results). We then confirm that this difference in magnitude is statistically significant.[21]

[20] We also confirm the validity of this assumption of a common pretreatment trend in the robustness section. See Section 6.

[21] While this gender difference becomes statistically insignificant if we focus on the third-trimester effects reported in columns (3) and (4) of Panel B and C, this must be because the gender differences in

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

	Dependent variable: Infant mortality rate			
Exposed trimesters	(1)	(2)	(3)	(4)
Panel A: All infants				
All trimesters	64.699***	69.991***		
	(25.878)	(30.200)		
	[0.000]	[0.000]		
First trimesters			−161.513	−154.688
			(81.688)	(84.095)
			[0.173]	[0.187]
Second trimesters			−65.637	−72.477
			(72.453)	(83.420)
			[0.480]	[0.480]
Third trimesters			229.272**	232.289**
			(75.907)	(81.853)
			[0.027]	[0.027]
Panel B: Boys				
All trimesters	42.043*	47.846*		
	(22.797)	(27.644)		
	[0.053]	[0.053]		
First trimesters			−131.592	−126.951
			(88.768)	(90.989)
			[0.240]	[0.293]
Second trimesters			−94.240	−102.580
			(76.951)	(88.546)
			[0.333]	[0.320]
Third trimesters			210.759**	217.388**
			(82.214)	(85.497)
			[0.027]	[0.027]
Panel C: Girls				
All trimesters	87.853***	92.515***		
	(31.529)	(35.701)		
	[0.000]	[0.000]		
First trimesters			−191.828	−182.345
			(77.679)	(80.924)
			[0.107]	[0.147]
Second trimesters			−36.836	−42.073
			(69.936)	(80.430)
			[0.573]	[0.587]
Third trimesters			247.938**	247.215*
			(71.274)	(80.395)
			[0.027]	[0.053]
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 (of boys in Panel B; of girls in Panel C) in each prefecture.

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 (in Figure 1). 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.[22] 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.

5 Persistency: Evidence from Population Censuses

5.1 Data and Specification

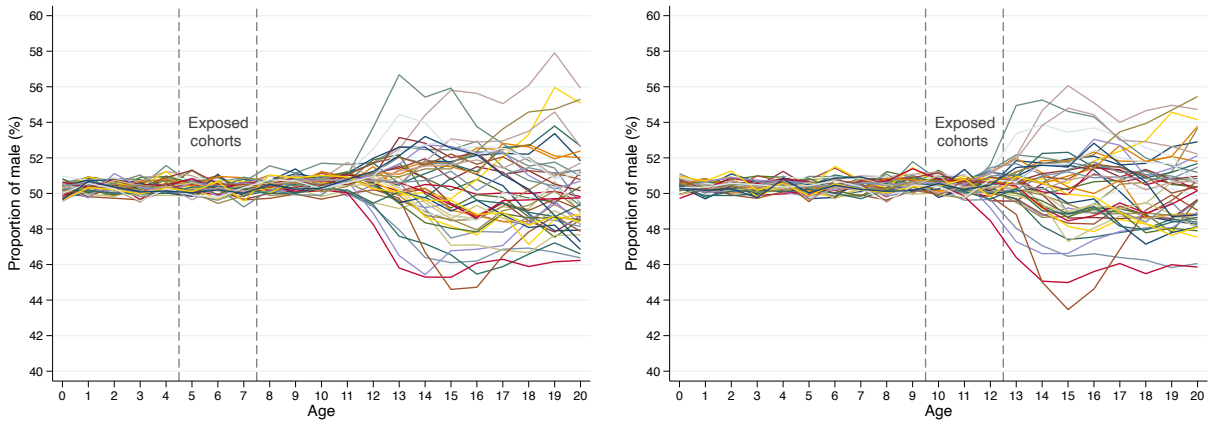
Thus far, we have found that fetal exposure to pandemic influenza decreased the proportion of male births. In this section, we assess whether the gender imbalance at birth persisted into their teens. The Population Censuses conducted in 1925 and 1930 documented the population by age and gender in each prefecture. To investigate the potential lasting effects of fetal influenza exposure on the sex ratio, we digitize the data from some *Kokuseichōsahōkoku* (Reports of the Population Census) and calculate the proportion of males aged 0–20 for each prefecture.[23]

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

the effects are generated by the cumulative effects of all trimesters.

[22]This 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).

[23]Since 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).



(a) 1925 Population Census

(b) 1930 Population Census

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).

creates a gender imbalance due to the flow of migrant workers.[24] 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 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. Specifically, 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.[25] Panel A of Table 4 lists the summary statistics for the dependent variables used.

[24]The 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.

[25]Precisely, 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 1918–1920 (1923–1925). Accordingly, to prepare the weighted influenza death rates in 1913–1925, we additionally digitize the 1912–1914 and 1923–1927 editions of the VSEJ (Statistics Bureau of the Cabinet 1916c, 1917, 1918c, 1924d, 1925d,e, 1926c, 1927b, 1928c) and the SCDEJ (Statistics Bureau of the Cabinet 1916a,b, 1918a, 1925b,c, 1926b, 1927a, 1928b).

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 FLUDR (exposed *in utero*)	Prefecture-birth year	0.05	0.06	0.00	0.22	552
Weighted FLUDR (exposed at age 0)	Prefecture-birth year	0.05	0.05	0.00	0.23	552
Weighted FLUDR (exposed at age 1)	Prefecture-birth year	0.03	0.04	0.00	0.23	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 (per 10,000 people). Regarding the indicator variable, children aged 5–7 in 1925 and 10–12 in 1930 are defined as the exposed cohort. "Weighted FLUDR (exposed *in utero*)" refers to the 9-month weighted average of the influenza mortality rates before a birth. "Weighted FLUDR (exposed at age 0)" refers to the 12-month weighted average of the influenza mortality rates after a birth. "Weighted FLUDR (exposed at age 1)" refers to the 13- to 24-month weighted average of the influenza mortality rates after a birth. 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 (1912–1927 editions); VSEJ (1912–1927 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:

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

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 males.

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

$$\begin{aligned} s_{it-a} = \Upsilon + \Xi_0\, Weighted\ FLUDR^{in\ utero}_{it-a} + \Xi_1\, Weighted\ FLUDR^{Age\ 0}_{it-a} + \Xi_2\, Weighted\ FLUDR^{Age\ 1}_{it-a} \\ + \mathbf{x}'_{it-a}\mathbf{\Theta} + \vartheta_{it} + \Omega_a + \Lambda_{it-a}, \end{aligned} \tag{8}$$

where $Weighted\ FLUDR^{in\ utero}$ is the weighted influenza death rate defined in equation 4. $Weighted\ FLUDR^{Age\ 0}$ and $Weighted\ FLUDR^{Age\ 1}$ are the 12-month and 13- to 24-month weighted average of the rates after a birth, respectively. We consider these rates because the sex ratio at primary school ages might have been affected by postnatal exposures to the pandemic influenza rather than by prenatal exposure, unlike the impacts on the secondary sex ratio and on the infant mortality rates. 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)	(5)	(6)	(7)	(8)
Exposed cohort	−0.100**				−0.126**			
	(0.028)				(0.036)			
	[0.020]				[0.033]			
Weighted FLUDR (exposed *in utero*)		−0.647**	−0.170	−0.134		−0.821**	−0.277	−0.258
		(0.273)	(0.322)	(0.423)		(0.276)	(0.297)	(0.417)
		[0.020]	[0.527]	[0.567]		[0.047]	[0.273]	[0.353]
Weighted FLUDR (exposed at age 0)			−1.077**	−1.179**			−1.275**	−1.326**
			(0.243)	(0.465)			(0.398)	(0.483)
			[0.020]	[0.020]			[0.033]	[0.060]
Weighted FLUDR (exposed at age 1)				0.166				0.085
				(0.716)				(0.799)
				[0.993]				[1.000]
Prefecture-year-specific fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Heterogeneous trend across prefectures	No	No	No	No	Yes	Yes	Yes	Yes
Observations	552	552	552	552	552	552	552	552
Number of prefectures	46	46	46	46	46	46	46	46
Number of clusters	8	8	8	8	8	8	8	8
Measured years	1925 & 1930	1925 & 1930	1925 & 1930	1925 & 1930	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	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. Column (1) presents the estimates from equation 7, whereas column (5) shows the estimates from the specification including the prefecture-specific trend. The estimate in column (1) shows that the exposed cohort, on average, exhibits a 0.1% lower proportion of males than the surrounding cohorts. This becomes 0.126% if we relax the common trend assumption across prefectures in column (5), accounting for approximately 26% of the standard deviation of the dependent variable (Panel A in Table 4).

The estimated coefficient of $Weighted\ FLUDR^{in\ utero}$ in column (2) is negative and statistically significant. This result remains unchanged if we include the prefecture-specific trend in column (6), suggesting that *in utero* exposure to the pandemic influenza can decrease the sex ratio of children at primary school ages. However, the estimated coefficient of $Weighted\ FLUDR^{in\ utero}$ in column (3) is negative but statistically insignificant whereas the estimated coefficient of $Weighted\ FLUDR^{Age\ 0}$ is negative and statistically significant. This result remains unchanged if we include the prefecture-specific trend in column (7). Column (4) indicates that exposure to the pandemic influenza at age 1 did not have such a negative impact on the proportion of males in children of primary school ages. Finally, column (8) confirms that this result is largely unchanged if we partially relax the common trend assumption.

The estimate in column (8) shows that a one standard deviation (0.05 in Panel B in Table 4) increase in the weighted influenza death rate (at age 0) decreases the proportion of males in these children by approximately 0.07% (0.05×1.326). Further, the proportion of males decreases by approximately 0.24% (0.23×1.326) in the case of exposure to the maximum influenza death rate. This accounts for approximately 50% of the standard deviation of the dependent variable and thus, is considered to be non-negligible in terms of its magnitude.

The foregoing results suggest that exposure to pandemic influenza during the first 12 months after birth has lasting effects on the sex ratio after birth, at least for children aged 5–12. This implies that postnatal exposure to pandemic influenza might have increased the infant mortality rate, especially that of the exposed boys. This is considered to be consistent with the results in the previous section. In the last section, we found that *in utero* exposure (note: not postnatal exposure) to the pandemic influenza had increased the infant mortality rate, especially that of the girls, under the "scarring mechanism."

This means that the surviving female infants might have been healthier than their male counterparts. In other words, male infants could be more vulnerable to postnatal exposures. This type of interpretation seems reasonable because there were two waves of epidemics in the case of pre-war Japan (Figure 2): the infants exposed to the first wave *in utero* were at age 0 when the second wave arrived. To test this potential mechanism, we regress infant mortality on the measured influenza mortality using the 1920–1921 subsample that included the infants who were exposed to the first wave and who might have also been impacted by the second wave. Table 6 shows the results. The estimated coefficients are positive in both columns (1) and (2) but the estimates in column (2) are statistically insignificant. This implies that the second wave might have increased the infant mortality rates of boys.

Table 6: Testing mechanism behind the postnatal exposure effects

	Infant mortality rate (‰) in the 2nd wave	
Exposed trimesters	(1) Boys	(2) Girls
Influenza mortality rate (in measured years)	0.398*	0.468
	(0.152)	(0.246)
	[0.093]	[0.200]
Measured years	1920–21	1920–21
Observations	96	96
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 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. The regressions in columns (1)–(2) are weighted by the average number of live births (of boys in column (1); of girls in column (2)) in each prefecture.

6 Robustness Checks

We provide two types of evidence: (1) evidence to examine the sensitivity of the results to potential omitted variables and (2) evidence in support of the common trend assumption. First, we examine the robustness of our main results to the inclusion of measures of heat and cold waves during the sample period. As discussed in the Introduction, weather shocks can affect birth outcomes, and temperature might correlate with the risk of infectious

diseases during epidemics.[26] To control for heat and cold waves, we draw temperature data from the official database of the Japanese Meteorological Agency (JMA).[27] In this database, the JMA reports 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 of three weather stations in each prefecture. Using the official definitions provided by the JMA, we define a heat wave as the annual average number of days on which the maximum temperature exceeded 30°C and a cold wave as the annual average number of days on which the minimum temperature was below 0°C. Then, following Deschênes et al. (2009), we calculate the inverse distance-weighted average of all the valid measurements from these.[28] As regards the average influenza death rates in Section 3, we calculate the average number of days of heat and cold waves for all trimesters by using the weighted average of the weather shock variable. These weather shock variables are finally added to equations 3, 6, 7, and 8.

Columns (1), (3),(4), (7), and (8) in Table 7 show the results. All specifications include the baseline control variables and the heterogeneous trend terms. Clearly, the estimated coefficients of the influenza exposure variables are close to the corresponding estimates in Tables 2, 3, and 5. This result provides evidence that our key influenza exposure variables are less likely to be correlated with weather shocks during the pandemics.

[26] A few recent studies investigate the relationship between temperature and risk of infectious diseases in pre-war Japan (Ogasawara and Matsushita 2019; Ogasawara and Yumitori 2018)

[27] The data are publicly available from the JMA database at http://www.data.jma.go.jp/gmd/risk/obsdl/index.php (accessed on March 30, 2018).

[28] 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:

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

where w denotes the weather shock variable and d denotes the geospatial distance from the centroid to station j. Data 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 7: Robustness: Effects of exposure to the influenza on gender imbalance

	Proportion of male births (%)		Infant mortality rate (‰)				Proportion of males (%)		
Exposed trimesters	(1)	(2)	(3) Boys	(4) Girls	(5) Boys	(6) Girls	(7)	(8)	(9)
First trimester	−0.1759**	−0.1182*	−120.655	−184.901*	−108.397	−160.947			
	(0.0552)	(0.0569)	(67.905)	(72.896)	(55.049)	(59.868)			
	[0.0370]	[0.0610]	[0.213]	[0.067]	[0.187]	[0.107]			
Second trimester	−0.0965	−0.0535	−69.637	−15.574	−19.840	29.125			
	(0.0547)	(0.0550)	(88.793)	(85.161)	(69.968)	(79.083)			
	[0.1950]	[0.3570]	[0.520]	[0.867]	[0.840]	[0.720]			
Third trimester	−0.0170	0.0158	188.255	229.790*	195.504*	218.856*			
	(0.0372)	(0.0474)	(94.771)	(89.830)	(84.101)	(84.036)			
	[0.6930]	[0.7250]	[0.120]	[0.080]	[0.080]	[0.080]			
"Zero" trimester (placebo)		0.1311			−96.441	−72.898			
		(0.0715)			(33.358)	(29.424)			
		[0.2130]			[0.107]	[0.120]			
"Fourth" trimester (placebo)		0.0171			−18.904	−39.510			
		(0.0676)			(27.235)	(22.717)			
		[0.8290]			[0.573]	[0.120]			
Exposed cohort							−0.129*		
							(0.044)		
							[0.073]		
Weighted FLUDR (exposed *in utero*)								−0.309	
								(0.475)	
								[0.393]	
Weighted FLUDR (exposed at age 0)								−1.353*	−1.518**
								(0.494)	(0.373)
								[0.060]	[0.020]
Weighted FLUDR (exposed at age 1)								0.175	0.254
								(0.802)	(0.677)
								[0.967]	[0.833]
Weighted FLUDR (placebo: exposed before conception)									−0.180
									(0.593)
									[0.713]
Heat and cold waves	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period	Jan. 1918–Dec. 1920	Jan. 1918–Dec. 1920	1916–22	1916–22	1916–22	1916–22	1925 & 1930	1925 & 1930	1925 & 1930
Observations	1656	1656	322	322	322	322	552	552	552
Number of prefectures	46	46	46	46	46	46	46	46	46
Number of clusters	8	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. "'Zero' trimester" refers to the average influenza death rates between 10 and 12 months before a birth. "'Fourth' trimester" refers to the average influenza death rates between one and three months after a birth. "Weighted FLUDR (exposed *in utero*)" refers to the 9-month weighted average of the influenza mortality rates before a birth. "Weighted FLUDR (exposed at age 0)" refers to the 12-month weighted average of the influenza mortality rates after a birth. "Weighted FLUDR (exposed at age 1)" refers to the 13- to 24-month weighted average of the influenza mortality rates after a birth. "Weighted FLUDR (exposed before conception)" refers to the 10- to21-month weighted average of the influenza mortality rates before a birth. The regressions in columns (1) and (2) are weighted by the average number of births in each prefecture. The regressions in columns (3)–(6) are weighted by the average number of live births (of boys for columns (3) and (5); of girls for columns (4) and (6)) in each prefecture. The regressions in columns (7)–(9) are weighted by the average number of children in each prefecture-year cell.

Second, we further include a pretreatment variable to test whether the identifying assumption of common pretreatment trends is valid. In column (2), we include a variable named "'Zero' trimester," representing the average influenza death rates between 10 and 12 months *before* a birth. Since a fetus does not exist *in utero* before conception, a future mother's exposure to pandemic influenza during this pre-conception period should have no significant impacts on the proportion of males at birth. Similarly, we include a variable named "'Fourth' trimester," representing the average influenza death rates between one and three months *after* a birth. The estimated coefficient of this variable should also be statistically insignificant because *fetuses* could not be infected during this post-birth period. As shown, the estimated coefficients on these placebo variables are statistically insignificant. We also find a similar result for infant mortality rate in columns (5) and (6). In column (9), we conduct a similar analysis for the proportion of males using the 10- to 21-month weighted average of the influenza mortality rates before a birth as a pretreatment variable.[29] The estimated coefficient of the pretreatment variable is close to zero and statistically insignificant. The results of these analyses provide evidence that our exposure variables are not capturing any secular pretrends in the outcome variables.

7 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 birth cohorts who were exposed to the pandemic influenza in their postneonatal period 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,

[29]Note that, in this specification, we exclude the *in utero* exposure variable ($Weighted\ FLUDR^{in\ utero}$) due to the strong collinearity with the pretreatment variable.

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. Analyzing 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 nōgyōkisotōkei (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? *Cliometrica*, 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. *Taishōjyūyonen 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. *Nihonteikoku shiintōkei (The statistics of causes of death of the empire Japan, 1926 edition). [in Japanese].* Statistics Bureau of the Cabinet, Tokyo, 1927a.

Statistics Bureau of the Cabinet. *Nihonteikoku jinkōdōtaitōkei (The vital statistics of empire Japan, 1926 edition). [in Japanese].* Statistics Bureau of the Cabinet, Tokyo, 1927b.

Statistics Bureau of the Cabinet. *Taishōkyūnen Kokuseichōsahōkoku, daiikkan (1920 Population Census of Japan, volume 1). [in Japanese].* Statistics Bureau of the Cabinet, Tokyo, 1928a.

Statistics Bureau of the Cabinet. *Nihonteikoku shiintōkei (The statistics of causes of death of the empire Japan, 1927 edition). [in Japanese].* Statistics Bureau of the Cabinet, Tokyo, 1928b.

Statistics Bureau of the Cabinet. *Nihonteikoku jinkōdōtaitōkei (The vital statistics of empire Japan, 1927 edition). [in Japanese].* Statistics Bureau of the Cabinet, Tokyo, 1928c.

Statistics Bureau of the Cabinet. *Taishōjyūyonen Kokuseichōsahōkoku, fukenhen (1925 Population Census of Japan, prefecture part). [in Japanese].* Statistics Bureau of the Cabinet, Tokyo, 1929.

Statistics Bureau of the Cabinet. *Shōwagonen Kokuseichōsahōkoku, fukenhen (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.

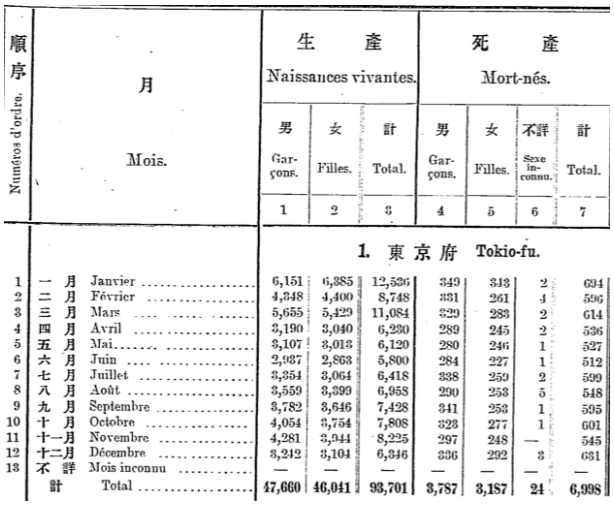
Yunquan Zhang, Chuanhua Yu, and Lu Wang. Temperature exposure during pregnancy and birth outcomes: An updated systematic review of epidemiological evidence. *Environmental Pollution*, 225:700–712, June 2017.

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

A.1 Image of documents



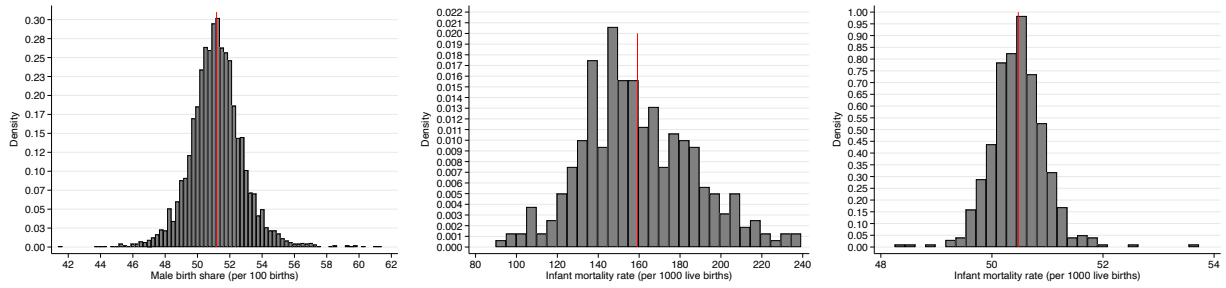
順序 Numéros d'ordre.	月 Mois.	生産 Naissances vivantes. 男 Garçons.	女 Filles.	計 Total.	死産 Mort-nés. 男 Garçons.	女 Filles.	不詳 Sexe inconnu.	計 Total.
		1	2	3	4	5	6	7
		1. 東京府 Tokio-fu.						
1	一月 Janvier	6,151	6,385	12,536	349	343	2	694
2	二月 Février	4,348	4,400	8,748	331	261	4	596
3	三月 Mars	5,655	5,429	11,084	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 Juin	2,937	2,863	5,800	284	227	1	512
7	七月 Juillet	3,354	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	336	292	3	631
13	不詳 Mois inconnu	—	—	—	—	—	—	—
	計 Total	47,660	46,041	93,701	3,787	3,187	24	6,998

Figure A.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 A.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).

A.2 Histograms of the dependent variables



(a) Male births (%) (b) Infant mortality rate (‰) (c) Boys (%)

Figure A.2: Distributions of the dependent variables

Notes: Figure A.2a shows the histogram of the proportion of male births (%). Figure A.2b shows the histogram of the infant mortality rates (‰). Figure A.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 A.2 shows the histograms of the dependent variables used: the proportion of

male births (Figure A.2a), infant mortality rate (Figure A.2b, and proportion of boys aged 5–7 and 10–12 (Figure A.2c).

A.3 Dynamics of the influenza mortality during the pandemic

Table A.1: Dynamic relationships in the influenza mortality rates during the pandemics

Dependent variable: Influenza mortality rate	(1) November 1918–January 1919	(2) January 1920–March 1920
Lagged influenza mortality rate	-0.0430	0.1287
	(0.0986)	(0.0742)
	[0.6800]	[0.4000]
Number of prefectures	46	46
Number of months	3	3
Number of observations	92	92
R-squared	0.6614	0.7450

The p-values from the wild cluster bootstrap resampling method are in brackets. The standard errors from the cluster-robust variance estimation reported in parentheses are clustered at the 8-area level. The number of replications is fixed at 150 for all the specifications.
Notes: There are no control variables but both prefecture and year-month fixed effects are included. R-squared is obtained from the least-squares dummy variable regression. All the regressions are weighted by the average number of people in each prefecture.

In Figure 3 in Subsection 2.2, we have confirmed that there were no systematic spatiotemporal correlations among the influenza mortality rates during the first and second pandemic waves. In this subsection, we further test whether the influenza mortality rates were correlated with the lagged rates in each wave. If there were statistically significantly positive correlations between the influenza mortality rates and their lagged values, the randomness in the influenza mortality during the pandemics is considered weak, and thus there might have been systematic endogeneity.

Table A.1 presents the estimation results from the simple dynamic panel data analysis. For each pandemic wave, we regress the influenza mortality rate on the lagged influenza mortality rate. Both prefecture and year-month fixed effects are included in both specifications because we also use those fixed effects in our models above (see equations 2 and 5). Columns (1) and (2) in Table A.1 show the results for the first and second waves illustrated in Figure 3, respectively. The estimated coefficients of the lagged influenza mortality rates are statistically insignificant in both cases. This provides evidence that our key exposure variable, the influenza mortality rate, has a certain random nature.

Appendix B Empirical Analysis Appendix

B.1 Testing Stationarity

Table B.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.

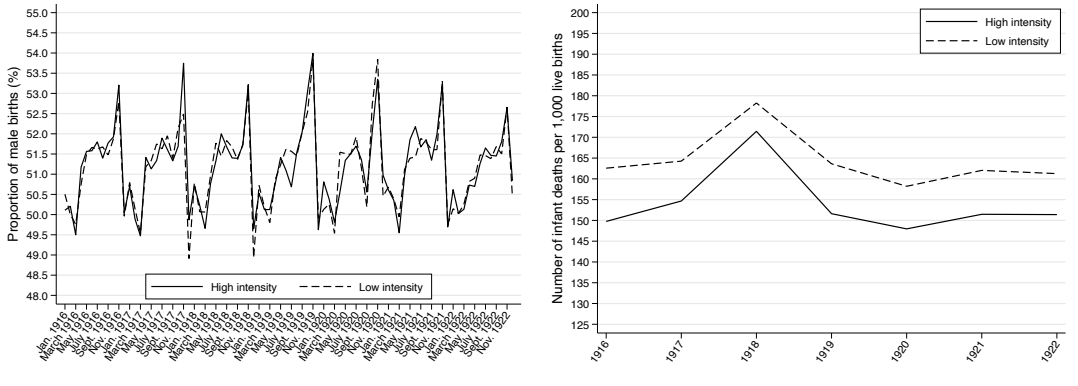
Table B.1: Results of the unit root tests for the proportion of male births

	January 1916–December 1922		January 1918–December 1920	
Test statistics	(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 Time-series plots of the proportion of male births and infant mortality

Figures B.1a and B.1a present the time-series plots of the proportion of male births and infant mortality rates between 1916 and 1922. In Figure B.1a, "High intensity" ("Low intensity") refers to the prefectures that experienced a monthly average influenza mortality above (below) the median, that is, 4.72 permyriad during the first and second waves (November 1918–January 1919 and January 1920–March 1920, respectively). As shown, the two series overlap in most year-month cells and the trends in both groups can be considered to be very similar. The same is true for the infant mortality shown in Figure B.1a. In Figure B.1a, "High intensity" ("Low intensity") refers to the prefectures that experienced a monthly average influenza mortality above (below) the median, that is, 1.27 permyriad in 1918 and 1920. Clearly, both series exhibit very similar trends. Note that the seasonality in the proportion of male births (shown in Figure B.1a) is effectively



(a) Male births (%) (b) Infant mortality rate (‰)

Figure B.1: Time-series plots of the male births (%) and infant mortality (‰)

Notes: Figure B.1a shows the monthly average time-series plots for the proportion of male births (%). In Figure B.1a, "High intensity" ("Low intensity") refers to the prefectures that experienced the monthly average influenza mortality above (less than) median: 4.72 permyriad during the first and second waves. The first and second waves are the periods between November 1918 and January 1919 and between January 1920 and March 1920, respectively (see Figure 3). Figure B.1b shows the annual average time-series plots of the infant mortality rates (‰). In Figure B.1b, "High intensity" ("Low intensity") refers to the prefectures that experienced an annual average influenza mortality above (below) the median, that is, 1.27 permyriad in 1918 and 1920. Sources: Created by the authors using data from Statistics Bureau of the Cabinet (1917, 1918c,d, 1919b, 1920b, 1921b, 1922b, 1924b,c,d, 1925d,e, 1926c).

captured by using the year-month fixed effects in our specifications (Subsection 3.2).

B.3 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 B.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. This may be because the gender differences in the effects are generated by the cumulative effects of all trimesters.

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

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