HIAS-E-102

DEMOGRAPHIC SHOCKS AND WOMEN'S LABOR MARKET PARTICIPATION: EVIDENCE FROM THE 1918 INFLUENZA PANDEMIC IN INDIA

JAMES FENSKE, BISHNUPRIYA GUPTA, AND SONG YUAN

University of Warwick

December, 2020

Hitotsubashi Institute for Advanced Study, Hitotsubashi University
2-1, Naka, Kunitachi, Tokyo 186-8601, Japan
tel:+81 42 580 8668 http://hias.hit-u.ac.jp/

DEMOGRAPHIC SHOCKS AND WOMEN'S LABOR MARKET PARTICIPATION: EVIDENCE FROM THE 1918 INFLUENZA PANDEMIC IN INDIA

JAMES FENSKE†, BISHNUPRIYA GUPTA‡, AND SONG YUAN⋆

Abstract. How did the 1918 influenza pandemic affect female labor force participation in India over the short run and the medium run? We use an event-study approach at the district level and four waves of decadal census data in order to answer this question. We find that districts most adversely affected by influenza mortality saw a temporary increase in female labor force participation in 1921, an increase that was concentrated in the service sector. By 1931, this increase had been reversed. We find suggestive evidence that distress labor supply by widows and rising wages help account for these results.

1. Introduction

Women's labor force participation is an important driver of economic development and gender equality (Duflo, 2012; Jayachandran, 2015). In this paper, we study the short and medium run effects of a major demographic shock on women's employment: the 1918 influenza pandemic in India. Demographic shocks, even when mortality is roughly equal by gender, may affect female labor supply. The Black Death, for example, is seen to have led to the growth of a labor market for women in North-Western Europe (Broadberry, 2013; De Moor and Van Zanden, 2010). Empirical evidence for this pattern is scarce, however, particularly for regions outside the Western world. Where mortality is more unbalanced by gender, effects may be more pronounced; the World Wars were another shock to the labor market that had consequences for women's labor market participation (Boehnke and Gay, 2019; Fernández et al., 2004). We focus on a different economic, social, and cultural context – that of India – in order to gain a different perspective on the drivers of women's economic participation. Both historically and in the present, India has had low levels of participation of women in economic activity outside the home. A recent literature (e.g. Fletcher et al.

†University of Warwick
‡University of Warwick
†University of Warwick
E-mail addresses: J.Fenske@warwick.ac.uk, B.Gupta@warwick.ac.uk, S.Yuan.2@warwick.ac.uk.
Date: July 30, 2020.
We are grateful to the Economic History Society Carnevali Small Research Grants Scheme and the Centre for Competitive Advantage in the Global Economy for their funding. We also thank Debraj Ray and audiences at the Delhi School of Economics, the Economic History Society, the European Historical Economics Society, NYU Abu-Dhabi, and the University of Warwick for their comments.

(2017)) on female labor force participation (FLFP) in India has highlighted a declining trend, despite rising incomes and greater education of women. Early marriage, social conservatism, and limited participation of women in certain types of economic activities are among the explanations that have been cited. In recent times, FLFP has declined as economic growth has increased.

From 1918 to 1919, a deadly influenza pandemic hit India. It first appeared in Bombay and then spread to the North and the West. Regions that received less than normal rainfall were more affected (Chandra and Kassens-Noor, 2014; Hill, 2011). Mortality estimates for India are wide-ranging, from 12 to 20 million (Chandra et al., 2012; Davis, 1951; Patterson and Pyle, 1991). Mortality was as high as 15% in some districts. This exogenous demographic shock might have affected female labor force participation in both the short run and the long run. The pandemic led to a shortage of labor and created a situation where women could potentially participate in the labor force and substitute for men in activities previously performed by men. This would generate a short-run increase in female labor force participation during the 1920s and 1930s, similar to what Goldin (1991) identifies after the Second World War. Although the population eventually reverted back to previous levels, the impact of this demographic shock on female labor supply could have persisted if women's working behaviour was transmitted to subsequent generations, for example by updating their beliefs toward the role of women (Fernández, 2013; Fernández et al., 2004). By contrast, if norms against women's work are strong and difficult to change, low levels of FLFP could be durable despite changes in the economic environment (Fernandez, 2007). The influenza epidemic provides a natural experiment with which to look at the effect of a demographic shock on FLFP in India. Was there a response, and if so, was it transitory or permanent?

We use district-level influenza mortality data for India from the Sanitary Reports of 1918 and 1919 and historical data on FLFP from the decadal censuses of 1901 through 1931 to create district-level panel data on women's labor market participation. We consider both aggregate participation and participation in the specific sectors of agriculture, industry, and services. We estimate the short and medium-run effects by adopting the event study method, which compares the change in women's employment in districts that were exposed to different levels of influenza mortality, before and after the influenza pandemic. We consider the 1931 census data to see if the changes in women's labor market participation lasted beyond the immediate short run. We find that FLFP increased only in services and only in the 1921 census; there is no evidence of a sustained response. Our results suggest that the mortality shock experienced by a typical district led FLFP to be 2.3 percentage points higher in 1921 than it would have been otherwise.

We test two potential mechanisms that could potentially drive the short run results. On the one hand, women might have needed to enter the labor market in order to mitigate the economic shock due to the death of their husbands or other male relatives during the pandemic. On the other hand, the general shortage of labor could have driven up wages and induced more women to join the labor force. Reduced population may, however, have also reduced demand, lessening the availability of jobs. Over the medium term, the same cultural norms that explain sectoral and regional variations in FLFP prior to the epidemic could help explain why the response to the epidemic was transitory. By contrast, if FLFP were initially low due to frictions in the labor market, the increase could have been permanent. We show that the share of women who were widowed rose in districts more greatly affected by the pandemic, though there is no similar increase in the share of never-married women. Population density fell in more affected districts and wages rose, including in services. Districts with greater shares of widows saw greater FLFP, both generally and in services. Similarly, more densely populated districts saw lower FLFP in services. However, while both these channels can help explain our results, they do not fully mediate them. We find little evidence of a response of male labor force participation to the influenza pandemic, suggesting that women's reactions were not part of a more generalized increase in labor force participation in response to higher wages.

We subject our results to several robustness exercises. The availability of labor force participation data in 1901, two census waves prior to the epidemic, allows us to show positive evidence that the parallel trends assumption of the difference-in-differences estimator holds conditional on controls. We show additional robustness to controlling for alternative sources of mortality, differential time trends due to agricultural potential and other geographic characteristics, and time-varying urbanization and land use. We show that our results hold using the male influenza death rate as an alternative measure of treatment.

1.1. **Contribution.** This research project is related to several strands of literature. The first considers the role of women in South Asian economies. Recent work has emphasized the response of Indian FLFP to economic incentives, including agro-ecological zones (Chen, 1989), cropping patterns (Gulati, 1975; Reddy, 1975), deep tillage (Carranza, 2014), and caste (Eswaran et al., 2013; Luke and Munshi, 2011). This literature has emphasized the role of culture and the high relative returns to home production in explaining why female labor force participation is both low and declining. In addition to evaluating the response of Indian female labor force participation to the influenza pandemic, our paper makes a first-order contribution in terms of data digitization and description. Ours is the first paper of which we are aware to document district-level variations in FLFP over the first three decades of the 20th century for the Indian subcontinent.

The second strand of literature to which we contribute studies the impact of demographic shocks on women's labor supply. Existing research focuses mainly on the Second World War, with a particular emphasis on the United States. Acemoglu et al. (2004) and Goldin and

Olivetti (2013) use exogenous variation in mobilization rates across states and find that the impact of the Second World War on female labor force participation was still present in the 1960s. Fernández et al. (2004) find an effect on female labor supply that persists through the 1980s. To our knowledge, ours is the first paper to put together a data set on women's labor force participation at the district level in colonial India and study the role of a historical epidemic. The historical literature has discussed the consequences of the Black Death on women's work in Europe, but little is known about the consequences of a major epidemic in economies outside Europe, especially in the context of different family structures and social norms. Both gender-biased and gender-neutral demographic shocks might increase FLFP, though these responses may be constrained by social norms and other factors.

The third strand of literature to which we contribute examines the relationship between FLFP and economic development. One branch of recent work focuses on how FLFP evolves with economic development (Boserup, 1970; Goldin, 1995; Mammen and Paxson, 2000) and under the constraints of culture (Alesina et al., 2013; Giuliano, 2014; Grosjean and Khattar, 2019). Another has shown that FLFP increases in response to a shortage of male labor in the context of other shocks such as wars, immigration, and the slave trade (Boehnke and Gay, 2019, 2020; Fogli and Veldkamp, 2011; Luo, 2017; Teso, 2019). We connect these two strands in the literature, providing novel evidence from a developing country and considering a more gender-balanced demographic shock.

The paper that is most closely related to ours is Donaldson and Keniston (2016). They show that the influenza pandemic in India had effects consistent with a Malthusian model: the greater land-labor ratios induced by mortality led to increased agricultural output per worker. This was then used to support both greater child quantity and greater child quality. Fertility rose in affected districts, and children born in these districts achieved both greater adult heights and literacy. We differ in focus, as our paper is concerned more narrowly with the labor force outcomes of women. However, their results complement ours.

The rest of this paper proceeds as follows. Section 2 outlines the historical context, identification strategy, and data sources. Section 3 presents our main results. Section 4 evaluates the empirical evidence for the mechanisms explaining our results. Section 5 examines the robustness of our results. In section 6, we conclude.

2. Context, Identification, and Data

2.1. **Context.**

2.1.1. *Influenza in India.* The Influenza pandemic that began in 1918 killed perhaps 50 million people globally. Over 13 million of those who died were in India – close to 5% of the total population (Arnold, 2019; Chandra et al., 2012; Hill, 2011). The pandemic began in the Bombay region and spread to the north and east. Regions that experienced less humidity compared to the seasonal average were more affected (Chandra and Kassens-Noor, 2014), and mortality was greater where population density and long-distance railroad travel were greater and rainfall was lower (Reyes et al., 2018). The pandemic disproportionately killed individuals of working age (20-45) and was concentrated in districts that experienced lower-than-normal levels of humidity. The most-affected districts of India experienced mortality rates of 15% during the pandemic (Donaldson and Keniston, 2016; Schultz, 1967).

The economic impacts of the pandemic varied across countries. In the United States, areas that were hardest-hit experienced rapid catch-up growth (Brainerd and Siegler, 2003). The pandemic reduced manufacturing employment in the US through both demand and supply side channels. (Barro et al., 2020; Correia et al., 2020). In Sweden, by contrast, the areas most affected saw increased poverty and reduced returns to capital (Karlsson et al., 2014). In addition, there has been considerable work in both economics and economic history on the long-term effects of exposure to the pandemic; cohorts exposed in early life achieved worse adult outcomes in terms of health and mortality (Almond, 2006; Lin and Liu, 2014).

2.1.2. *Women and work in India.* The literature on FLFP finds an inverted-U relationship between women's labor market participation and development (Boserup, 1970; Goldin, 1995). The conventional explanation is that this is driven by structural change. In early stages of economic development, agriculture is the dominant sector, income is low, and women work in agriculture within a household-based production process. With industrialization, rising income, and marketization of work, fewer women work in factories outside the home as social norms discourage participation in the labor market. Women's labor force participation falls with industrialization. Fertility and social conventions regarding women's work reduce married women's participation in particular. The rise of the male breadwinner family during the industrial revolution in England is a commonly-cited example of falling FLFP with rising incomes (Horrell and Humphries, 1995). A similar argument in Goldin (1995) suggests that social stigma against men whose wives work in blue collar jobs reduces women's labor market participation. On the demand side, blue collar jobs are mainly in male-dominated activities. In this explanation, women's labor force participation increases with further increases in income, expansion of education, and a structural shift towards service sector white collar jobs (Mammen and Paxson, 2000).

In present-day India, the trend has been quite different. For its level of economic development, India is an outlier in terms of FLFP. In 1970, Roy (2005, p. 135-6) claims that the participation rate was 12%, while peer countries in East Asia and Latin America had a FLFP rate of 30% . Despite economic growth, rising literacy, and a structural shift to services, FLFP has declined in recent decades. Unpacking this trend shows that the decline has mainly been in rural India among women with secondary education (Fletcher et al., 2017). In the present, cross-sectional comparisons suggest higher levels of FLFP for Indian women aged 15-24, but that they withdraw from the labor force after marriage, only to return later in life (Klasen and Pieters, 2015). Women without education are typically in families with low income and work to supplement family income. Women with tertiary education increase their labor market participation. It is the women in the middle who drop out of the labor force.

Changes in the sectoral structure of employment have contributed to the recent reduction in FLFP rates. Most employment growth has occurred in construction and low-skilled services, while expansion of employment in white-collar services has not been sufficient to absorb the growing female population of working age (Klasen and Pieters, 2015). Preference for men in clerical jobs lowers the participation of women with secondary education in the labor market (Chatterjee et al., 2018).

Various socio-cultural factors also reduce women's participation. Staying out of the labor market signals social status (Chen and Drèze, 1992). Education of high-caste women increases their marriage market returns, but reduces participation in the labor market due to social barriers (Eswaran et al., 2013). Low-caste women have higher rates of labor market participation (Klasen and Pieters, 2015). Women with more education marry more educated men and have higher family income, which reduces the incentives for female labor market participation; Afridi et al. (2016) find that, from 1987 to 2009, rising education among rural men and women increased the returns to women's work in home production relative to work in the labor market. Women with secondary education increase investment in the education of their children and withdraw from the labor market (Behrman et al., 1999). While participation of married women has declined, participation of widows in the labor market has increased and participation of never married women has stayed the same (Afridi et al., 2016).

Women's labor market participation in colonial India was close to 30% (Thorner and Thorner, 1962). This fits in with the general pattern of higher participation rates in agricultural societies and within the household economy. In 1911 just under one-third of the workers were women in agriculture and industry. Within the service sector, the share of women was as high as 38-40% in a sector referred to as "insufficiently described occupations" in official sources. Agriculture was the main activity for most women. In 1911, 73% of women workers were in agriculture and related activities, 11% in industry, mining and construction, 12% in trade, transport and other services, and 4% in insufficiently defined activities. (Thorner and Thorner, 1962).

In colonial India, Roy (2005) identifies gender segmentation in the labor market. The work done by women was concentrated in specific industries and services, including bidi rolling, hand spinning, basket weaving, grain processing by hand, and stone cutting. Some industries in colonial India rarely hired women. These included metal working, chemicals, and printing. Some industries employed both men and women, though specific tasks were largely segregated by gender, as in mining, cotton textiles, and quarries. In table 1, we show the distribution of workers in selected sectors reported by Thorner and Thorner (1962). There were strong regional differences in FLFP as well. Our data show regional variation in 1911 – FLFP was relatively high in the Central Provinces, Madras, Bihar and Orissa, and in the United Provinces. It is relatively low in Punjab and Bengal. This pattern fits with the regional variation found after independence. In 1961, for example, women's labor force participation was 18% in Uttar Pradesh, 6%, in Punjab, and 9% in West Bengal, while it was closer to 30% in Assam, 38% in Maharashtra and 43% in Himachal Pradesh (Gulati, 1975). From 1961, Roy's estimates show a sharp decline in FLFP.

2.2. **Identification.** In order to test for short-run and medium-run effects of the 1918 Influenza epidemic on female labor supply in India, we compare the change in women's employment in districts that were exposed to different levels of influenza mortality, before and after the influenza epidemic. This allows us to control for unobservable characteristics of districts that do not change over time, and for unobserved variables in specific time periods that affect all of India equally. In particular, we use an event-study approach. We estimate:

$$FLFP_{dt} = \beta_t Influenza_d + (x'_d \times \eta_t)\gamma + \delta_d + \eta_t + \varphi_p \times t + \epsilon_{dt}. \quad (1)$$

Here, $FLFP_{dt}$ is female labor force participation in district d in year t. We compute this as the ratio of women working to the population of women in the district.[1] In various specifications we either use total FLFP or FLFP for the separate sectors of agriculture, industry, or services. $Influenza_d$ is the influenza mortality rate in district d in 1918 and 1919. β_t is a separate coefficient for each year, equivalent to interacting $Influenza_d$ with year fixed effects. The omitted category is 1911, the last pre-treatment year. An insignificant coefficient for 1901 validates the parallel trends assumption. The coefficients for 1921 and 1931 are our coefficients of interest.

$x'_d \times \eta_t$ includes the interactions of three key district-specific variables with our year fixed effects. These are humidity, latitude and longitude. These are selected because they are predictive of influenza mortality and ensure that there are parallel trends in FLFP between high-exposure and low-exposure districts prior to the 1918 pandemic. These controls also help avoid the possibility that our results are driven by correlates of influenza mortality, rather than influenza mortality itself. In particular, previous writers have stressed that humidity helps explain the variation in the severity of the epidemic across space in India (Chandra and Kassens-Noor, 2014; Hill, 2011). In addition, in some specifications we control

[1]Because of the existence of labor by both children and the elderly, as well as the sometimes incomplete enumeration of the population by age, we use the total female population as the denominator.

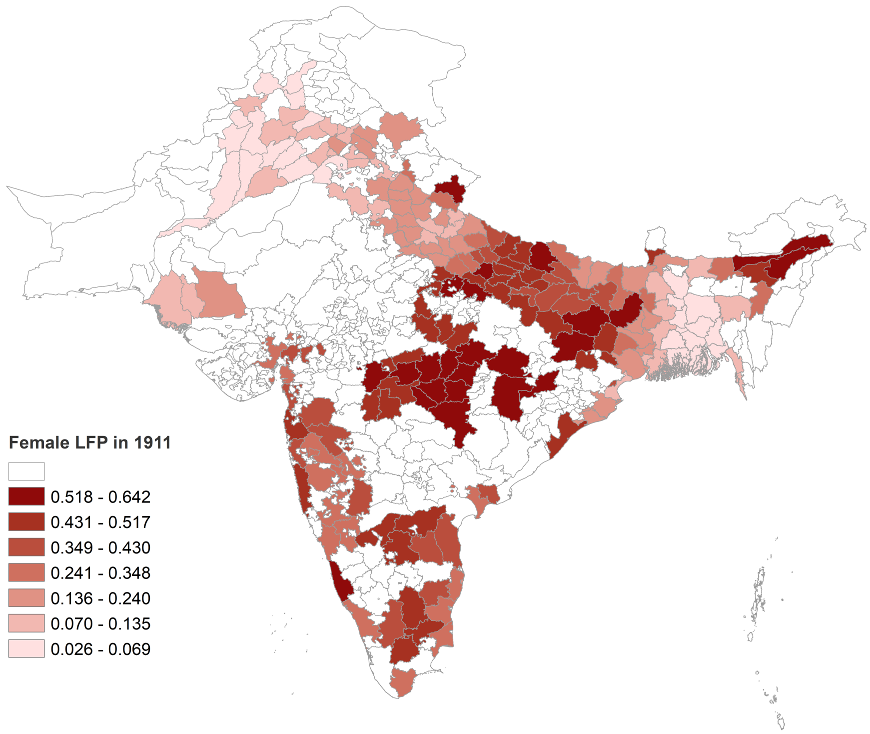
for the time-varying urbanization rate. We do not include it in our baseline set of controls, as it might both respond to influenza mortality and proxy for a critical channel through which the pandemic may have influenced women's work outcomes, e.g. in capturing the thickness of the potential labor market or the abundance of labor relative to other factors of production, most notably land. In robustness specifications, we will include additional time-invariant controls interacted with year fixed effects and additional time-varying controls. δ_d and η_t are district and year fixed effects, respectively. $\varphi_p \times t$ is a province-specific time trend, i.e. an interaction of province fixed effects with linear year variables. Standard errors are clustered by district.

2.3. **Data.** Historical data on our dependent variable, female labor force participation, exists in the colonial censuses collected every ten years between 1901 and 1931. To our knowledge, we are the first to digitize these data at the district level and disaggregated into the major categories of industry, agriculture, and services. For each district in each census year, the census reports counts of men and women working in several occupations. We have categorized these ourselves as agricultural, industrial, or service-based. We begin with the broad classifications already employed by the colonial census. These are:

- Sub-Class I. Exploitation of Animals and Vegetation
- Sub-Class II. Exploitation of Minerals
- Sub-Class III. Industry
- Sub-Class IV. Transport
- Sub-Class V. Trade
- Sub-Class VI. Public Force
- Sub-Class VII. Public Administration
- Sub-Class VIII. Professions and Liberal Arts
- Sub-Class IX. Persons Living Principally on Their Income
- Sub-Class X. Domestic Service
- Sub-Class XI. Insufficiently Described Occupations
- Sub-Class XII. Unproductive Labour

We classify as "agriculture" all occupations in Sub-Class I. We classify as "industry" all occupations in Sub-class III. We classify as "services" all others apart from Sub-Class II, which we have not included in our analysis. Mining was only more than 3% of employment in four districts. In the census, Sub-Class XI is sometimes referred to as "General Labour," while Sub-Class XII includes the category of beggars, vagrants, and prostitutes. It also includes inmates of jails, asylums and almshouses. In our primary results, we also include these in "Services", we remove them in a robustness exercise. To convert these to labor-force participation rates, we divide these counts by the total population of the district of a given gender. We exclude two districts from the analysis whose data on FLFP contain obvious errors. For Garhwal, recorded FLFP is between 61% and 69% in 1901, 1921, and 1931, but is only 8% in 1911. For Narsinghpur, the number of female workers recorded in 1931 is more than three times the female population. We show maps of FLFP in 1911 and its change from 1911 to 1921 at the district level in Figures 1 and 2, respectively.

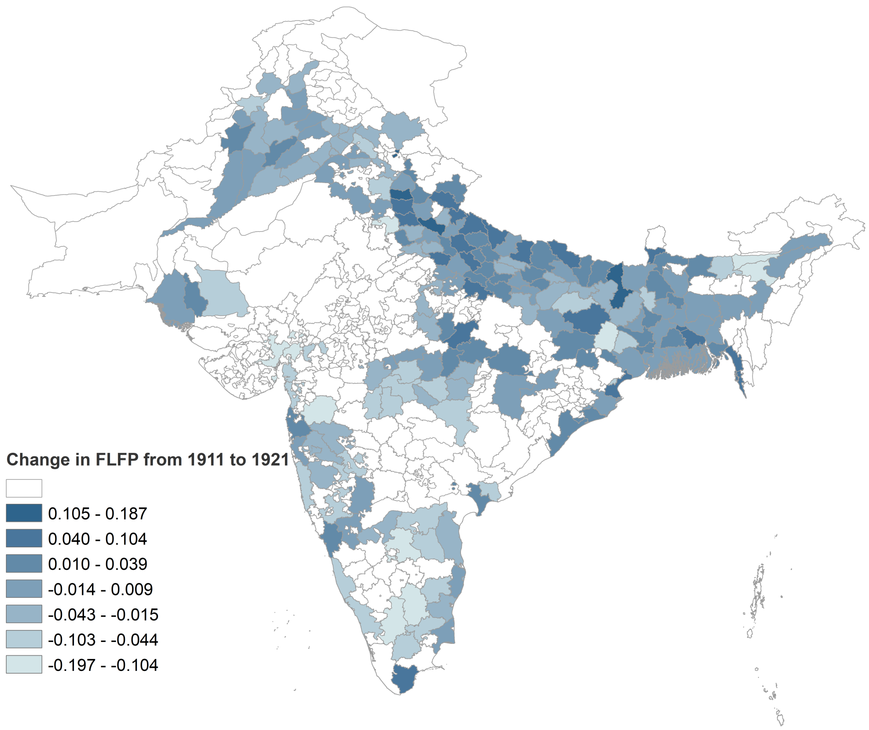
Figure 1. FLFP in 1911



We have collected and entered the 1918 influenza mortality data for India from the Sanitary Reports for the years 1918 and 1919. These provide deaths by cause and gender for each district in the sample. Because these Sanitary Reports are available only for the British-ruled districts of India and exclude the princely states, and because not every district reports occupational data in every year, our baseline sample includes 196 districts over four periods in time, giving us a maximum of 784 observations in each regression.

While the inclusion of district and year fixed effects in our regressions implies that we need not, and cannot, control for time-invariant controls, we do interact three variables with year fixed effects in order to discipline the pre-pandemic district trends in our data. In particular, these are latitude, longitude, and humidity. To compute each of these, we begin by creating a shapefile map of these districts based on the paper map in the 1931 census. Following the procedure in Fenske and Kala (2017), we correspond each colonial district mapped in the 1931 census with all current sub-district units (e.g. tehsils) that intersect the erstwhile historical districts, and treat the union of these polygons as the polygon corresponding to

Figure 2. Change in FLFP from 1911 to 1921



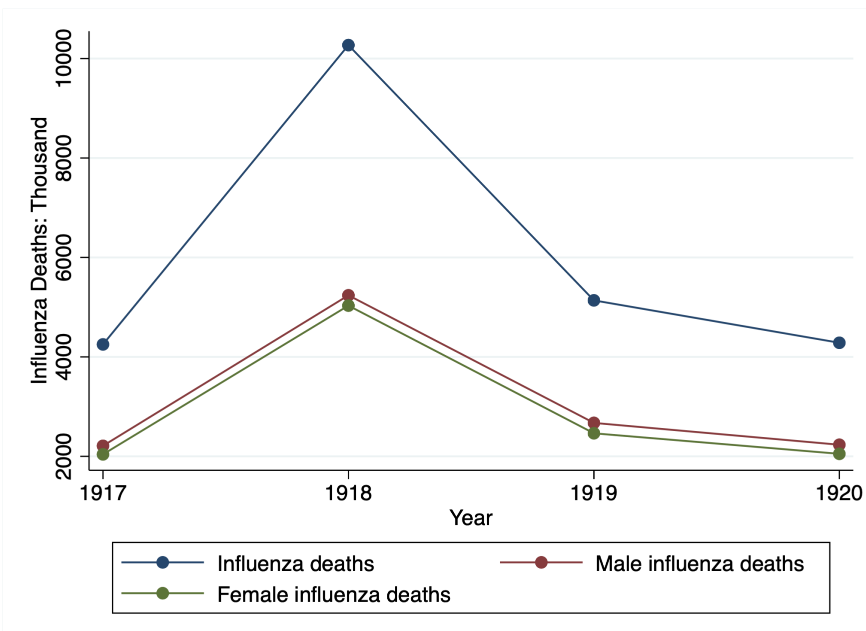
that colonial district. We use the centroid of this polygon to compute the latitude and longitude of the district. For variables originally available in raster format, such as humidity or crop suitabilities, we average over raster points in a district. Our humidity data are taken from the Climatic Research Unit at the University of East Anglia.

In some specifications, we control for population density and urbanization rates. These are computed using the areas, populations, and populations of cities from each census wave. In robustness checks, we employ data on crop suitabilities. These are taken from the Food and Agriculture Organization of the United Nations Global Agro-Ecological Zones's project (FAO-GAEZ), and are reported as expected yields in kilograms per hectare under low levels of inputs. We focus on crops that are important in Indian agriculture, in particular banana, chickpea, cocoa, coffee, cotton, groundnut, wet and dry rice, oil palm, onion, soybean, sugar, tea, potato and wheat. We will also control, for robustness, for the areas planted to major crops during each census year. We have digitized these from the Agricultural Statistics of British India, and focus on rice, wheat, sugarcane, jute, opium, tea, tobacco, and cotton.

In our analysis of possible mechanisms, we use two additional variables that we compute using the population counts by district, gender and marital status in the colonial census. These are the fraction of the female adult population that is unmarried and the fraction that is widowed. We also test whether greater FLFP is explained by sector-specific increases in wages due to labor scarcity arising from the epidemic. To do this, we assemble an unbalanced panel of more than 6,500 wage observations from a number of sources, the most important of which are *Wages and Prices in India* and the wage censuses of each province. We classify these by sector (agriculture, industry, and services) and by skill intensity (skilled or unskilled).[2]

In Figures 3, 4 and 5, we show the timing and geographical distribution of the influenza pandemic. The first two figures depict the number of deaths and mortality rates (deaths as a share of the total population), respectively. The final figure is a map of the death rate by district. Two patterns are clear immediately. The first is the sharp concentration of mortality in 1918; influenza deaths spiked in that year but had largely returned to their pre-pandemic levels by 1919 and were back at their 1917 levels by 1920. Second, while mortality rates were lower in southern India and in Assam than elsewhere, there was geographic dispersion in mortality rates both within and across states.

Figure 3. Influenza Mortality: Deaths

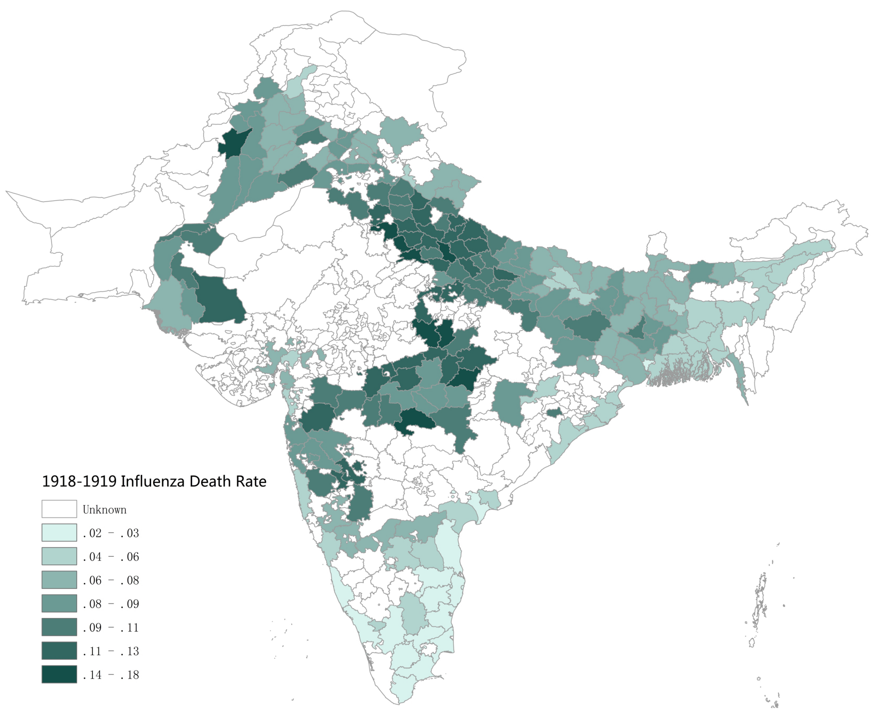


We present summary statistics for our results in Table 2. Looking at mean FLFP by year, the rate of women working fell from roughly 31% in 1901 to 29% in 1911, where it remained

[2]Our classifications by sector are as follows. In agriculture: agricultural, harvester, sower, weeder, reaper, ploughman, farm servant. In industry: blacksmith, carpenter, earth worker, fitter, mason, cotton weaver, hardware, metal worker, mill hand. In services: syce, thatcher, boatmen, domestic servant, coolie, carter, fireman, general labor. Our classifications by skill are as follows. Skilled: blacksmith, carpenter, mason, cotton weaver, hardware, metal worker, fitter, fireman. Unskilled non-agricultural: syce, earth worker, thatcher, boatmen, domestic servant, coolie, carter, general labor, mill hand. Our sources of wage data are listed in the appendix.

FIGURE 4. Influenza Mortality: Rates

FIGURE 5. Influenza Mortality: Distribution



in 1921. By 1931, it had fallen below its 1901 value. The decline to 1931 has many causes. Partly, this is about how the census counted secondary occupations, which we expect to have a uniform level effect across the districts in our data and that we expect to be uncorrelated

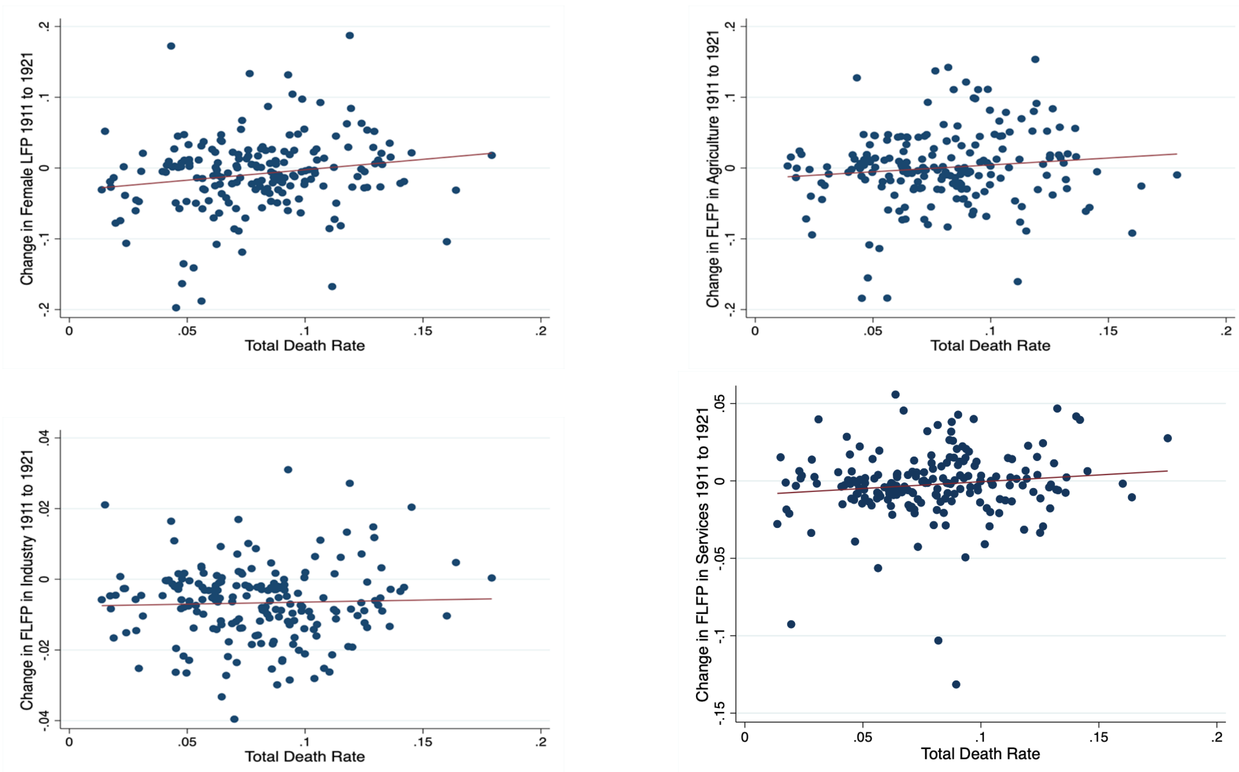


Figure 6. Scatterplots: Influenza mortality and changes in FLFP across districts, by sector

with the extent of influenza mortality. That is, we expect it would be accounted for by a year fixed effect. To the extent that is was due to the Great Depression, we account for this in our robustness exercises by allowing for flexible time trends by crop suitabilities (accounting flexibly for possible shocks to global commodity markets).

3. Results

3.1. **Graphical analysis.** Before presenting our main event study results, we document that the same trends we will uncover in these specifications are also apparent in the raw data. In Figure 6, we show four scatterplots. In each scatterplot, each dot represents a district in our data. The x axis of each plot is the same: the death rate due to influenza during the pandemic. The y axis in each figure is a different measure of the change in FLFP between 1911 and 1921, i.e. through the pandemic. It is clear from the figure that there is a positive correlation between influenza mortality and the change in overall FLFP from 1911 to 1921. Modest upward-sloping relationships are also apparent for the agricultural and services sectors, and there appears to be no correlation for industrial FLFP. Our event study results will verify the relationships for overall FLFP and services below.

3.2. **Main results.** We present our main results in Table 3. In column (1), we report a sparser specification with only district fixed effects, year fixed effects, and influenza mortality interacted with the year fixed effects. We find a statistically significant increase in female labor force participation in more influenza-affected districts in 1921, followed by a decline in 1931. In column (2), we add province-specific linear time trends. In column (3), we control as well for variables that help discipline the differences in pre-trends across districts: latitude, longitude, and humidity, each interacted with year fixed effects. Finally, in column (4), we control for time-varying urbanization. Across columns, we continue to find a statistically significant increase in female labor force participation in 1921 in districts that experienced greater mortality during the pandemic, and that this is reversed in 1931.

To put the magnitude of these results in context, the influenza death rate has a mean of 0.0796 across districts and a standard deviation across districts of 0.032.[3] The coefficient of approximately 0.294 implies that a typical mortality shock increased FLFP by 2.3 percentage points in 1921, with a comparable reverse effect almost three times as large in 1931. Alternatively, a one standard deviation increase in mortality led to a 0.94 percentage point increase in FLFP in 1921, and a reduction almost three times as large in 1931.

There are several ways to put the magnitude of these results in perspective. One is to compare them with other effects of influenza in India. In their study of the pandemic, Donaldson and Keniston (2016) find that a 10% death rate raises birth rates by 18% in the decade afterwards while reducing marriage rates by 10%. Another approach is to contrast them with other studies of demographic shocks and female labor force participation. For France, moving from the 25th to the 75th percentile of the male death rate during the First World War increased FLFP by 3.5 percentage points during the interwar period (Boehnke and Gay, 2019), an effect that has risen to 5 percentage points since the 1960s (Boehnke and Gay, 2020). Finally, we can compare them to modern correlates of FLFP in India. For example, the rural-urban gap in labor force participation among women with low levels of education is 8 percentage points in the present. Today, 22% of women with a secondary education are in the labor force in India. This is lower than for less educated women (Fletcher et al., 2017). Our results are modest in comparison, though of a similar order of magnitude.

3.3. **Results by Sector.** We show results by sector in Table 4. That is, we now compute FLFP for agriculture, industry, and services. In 1921, FLFP in services increases in the most adversely affected districts. In 1931, it falls relative both to its baseline mean and relative to other districts where the mortality rate was lower. We find no significant patterns of response in agriculture or industry.

[3]The smaller standard deviation in Table 2 reflects the variation across both districts and census years, which is less because the variable is not time-varying.

What are the most quantitatively important sectors in which women worked, and that have the potential to drive our results? In Table 5, we consider the five sub-sectors that each accounted for at least 1% of the female population in 1911: cultivators, farm servants, textiles, trade, and domestic service. Here, our results can only be taken as suggestive. There is a significant, though small, differential increase in FLFP in textiles in more pandemic-affected districts in 1921. There is a large differential increase in female participation in domestic service as well, but here the significant coefficient on the interaction of influenza mortality and the 1901 dummy implies that we cannot rule out violations of the parallel trends assumption.

In Table 5, we also consider a category of FLFP that can be easily interpreted as "distress' participation: that of beggars and prostitutes. We find no evidence that FLFP in this category rose differentially in 1921 in districts that experienced greater influenza mortality.

4. Mechanisms

To evaluate the mechanisms that could potentially explain our results, we begin by considering three possible mediating variables: the share of widows in the female population, the share of never-married women in the female population, and population density. The share of widows is likely to capture distress labor by women who have lost husbands and other relatives during the pandemic. The share of what the census refers to as "single" women may capture a similar effect, working through the unavailability of potential husbands. Population density is interpreted here as a measure of the overall availability of labor. In a more competitive labour market, employers may give priority to male workers, leading women to withdraw from the labor force. The increase in FLFP may have been driven by an aggregate shortage of labor in the economy and a corresponding rise in wages.

We begin by testing whether these variables respond to the influenza pandemic in Table 6. In columns (1) and (2), we show there is indeed a significant increase in the prevalence of widows in the female population in 1921 in districts that experienced higher mortality rates during the pandemic. Though the coefficient for 1931 suggests that there were still more widows in these districts more than a decade later, this is not statistically significant. In columns (3) and (4), by contrast, we see no change in the share of never-married women. In columns (5) and (6), we do show that there is a statistically significant decline in 1921 in population density in districts with more mortality. In addition to suggesting that labor became more scarce in these districts, it also validates the mortality data, which are taken from the Sanitary Reports, and not from the census.

In Table 7, we then ask the degree to which these variables mediate our main results, including them as additional controls. Columns (1) and (2) are analogous to columns (3) and (4) of Table 3. Districts with a higher share of widows in the population have greater female labor force participation, though the magnitudes of the coefficients on the interactions of the influenza death rate and the year fixed effects have only fallen slightly. The same is true in columns (7) and (8), which are analogous to columns (5) and (6) in Table 4. So, while distress labor by widows may help explain the temporary increase in FLFP in more influenza-affected districts, it is not a complete explanation.

In Table 8, we perform a similar exercise, but for population density. In columns (1) and (2), we find no evidence that the overall supply of potential labor drove aggregate FLFP. In columns (7) and (8), we do find a negative coefficient on population density, suggesting that female participation in services grew when labor was more scarce generally. Again, however, this does little to diminish the magnitudes of the coefficients on the interactions of the influenza death rate and the year fixed effects. The labor scarcity induced by the pandemic is, then, only a partial explanation, and a sector-specific one.

In a related exercise, we use Table 9 to demonstrate that greater mortality did increase wages. Following a procedure similar to that in Donaldson and Keniston (2016), we estimate the following event study specification using our unbalanced panel of wage data:

$$\begin{aligned}\ln(w_{odpt}) =& \beta_1 Influenza_d \times Year_{1901} + \beta_2 Influenza_d \times Year_{1902-11} \\ &+ \beta_3 Influenza_d \times Year_{1919-21} + \beta_4 Influenza_d \times Year_{1922-28} \\ &+ z'_{ot}\rho + (x'_d \times \eta_t)\gamma + \delta_d + \eta_t + \varphi_p \times t + \epsilon_{odpt}. \end{aligned} \tag{2}$$

Here, w_{odpt} is the log wage for occupation o in district d in province p in year t. As in (1), $Influenza_d$ is the mortality rate from influenza in district d during the pandemic. Because the wage data come as an unbalanced panel, we collapse our time periods together into broader groups: 1901 and earlier, 1902 to 1911, 1919 to 1921, and 1922 to 1928. Insignificant estimates of the coefficients β_1 and β_2 validate the parallel trends assumption. β_3 is an estimate of the immediate impact of influenza mortality on wages, while β_4 estimates whether this impact persisted throughout the 1920s.

z_{ot} includes occupational characteristics – whether the wage is urban or rural, whether it is paid in cash or kind, and fixed effects for occupation. Other terms in (2) are defined as in (1); $x'_d \times \eta_t$ is the interaction of district-specific controls with year fixed effects, δ_d is district fixed effects, η_t is year fixed effects, and $\varphi_p \times t$ is province-specific time trends. As before, we cluster standard errors by district.

Pooling all wage observations together in Column (1) of Table 9, we confirm the results from Donaldson and Keniston (2016) – greater influenza mortality created a scarcity of labor, increasing wages throughout the 1920s. Our results allow us to show results by sector that Donaldson and Keniston (2016) do not report. We find no wage effects in agriculture. Rather, it is in industry, services, general labor, and the labor of skilled workers that wage increases are apparent. For industry and skilled work, these increases persist throughout the 1920s. Because our wage data form an unbalanced panel, we cannot control for wages in a mediation analysis similar to Table 8.

If aggregate labor scarcity were to explain women's responses, male labor force participation (MLFP) would increase in 1921 in a manner similar to FLFP. In Table A1, in the appendix, however, we show that we find no evidence of such an effect. One possible explanation of this result is that MLFP was largely universal and remained so after the pandemic.

Finally, to evaluate whether the impact of the influenza pandemic was greatest in districts where there was a greater pent up supply of underutilized female labor, Table 10 separates the sample by initial FLFP. That is: we divide the sample into districts with above and below median levels of FLFP in 1901, the first year in our data. We find that the FLFP response was, rather, driven by districts with initially greater levels of FLFP. Rather than the influenza pandemic working to break the constraints imposed by social norms on FLFP, these results suggest that the response itself was constrained by the presence of these norms or by the prevalence of activities in which women did not work.

5. Robustness

In this section, we outline robustness exercises that we report in the appendix. We begin by showing that our results survive controlling for cholera mortality in 1918, interacted with year fixed effects, in Table A2. We do this to demonstrate that our results are not driven by differential trends in districts that have high levels of mortality for reasons unrelated to the pandemic.

Because of the possibility that our results for 1931 are confounded with India's exposure to the Great Depression, we add as controls the interactions of sutiabilities for major crops in India with year fixed effects. These are computed using data from the Food and Agriculture Organization of the United Nations' Global Agro-Ecological Zones (FAO-GAEZ) project, and represent expected yields under rainfed agriculture with low input use based on climate and soil productivity. These are widely used in the literature as exogenous predictors of crop yields (Donaldson and Storeygard, 2016). Note that this is equivalent to controlling flexibly for how the world prices of these crops affect these districts based on an exogenous determinant of their propensity to produce them. These results are reported in Table A3. The general result remains: districts that were most affected by the influenza pandemic saw an increase in female labor force participation in services that was reversed by 1931.

For similar reasons, we control in Table A4 for the time-varying areas planted to major crops recorded in the Agricultural Statistics of British India in our census years. In particular, we consider rice, wheat, sugarcane, jute, opium, tea, tobacco, and cotton. Results remain similar to our baseline analysis.

Because it is possible that influenza mortality during the pandemic was correlated with geographical characteristics that might predict differential trends in FLFP even in the absence of the pandemic, we use Table A5 to show that our results survive controlling for a number of additional geographic controls interacted with year fixed effects: temperature, precipitation, and altitude, all computed using data from the FAO-GAEZ data, the Nunn and Puga (2012) ruggedness index, and the Kiszewski et al. (2004) index of the stability of malaria transmission.

So far, we have considered the overall rate of influenza mortality, combining genders. Gender-specific mortality rates were strongly correlated across districts during the pandemic: the male and female death rates have a correlation coefficient of 0.97, and each has a correlation coefficient of at least 0.99 with the overall death rate. Although this gives us limited scope to evaluate whether FLFP responded differently to male mortality, we replace the overall death rate with male mortality in Tables A6, A7, and A8. The table analogous to our main results, Table A6, shows patterns similar to our baseline. In Table A7, which is analogous to Table 6, we show that the share of widows in the female population rises in response to male influenza deaths, as it had to overall mortality. In Table A8, which is analogous to Table 7, we again show evidence that the increase in widowhood due to the pandemic helps explain the FLFP response, but does not completely mediate it. In Table A9, we show that using the female death rate gives results that are similar in size to our baseline, but only significant at conventional levels for FLFP in services.

Finally, we demonstrate that the smaller categories of beggars and prostitutes and inmates of jails, asylums and almshouses do not drive our results. In Table A10, we show that excluding these categories from the numerators of both total FLFP and FLFP in services does not change our main results. In Table A11, we do the same for our potential mediating variables.

6. Conclusion

In this paper, we have asked whether the influenza pandemic in India led to an increase in FLFP. Our answer is yes, but that the increase is temporary. We observe a short-lived increase, concentrated in services, which is reversed within a decade. Population density (the aggregate land-labor ratio) is not a sufficient statistic for these results. Influenza increased the share of widows in the population. This explains part of the short-run effect, but not all of it. Mortality also raised wages, including in services. The FLFP response to a major demographic shock in India was confined to services and was not durable. This contrasts with other historical cases. One possible explanation of this contrast is cultural norms. It may be that the ability of women to increase their participation in the labor market is limited in this context to periods of economic distress.

References

Acemoglu, D., Autor, D. H., and Lyle, D. (2004). Women, war, and wages: The effect of female labor supply on the wage structure at midcentury. *Journal of Political Economy*, 112(3):497–551.

Afridi, F., Dinkelman, T., and Mahajan, K. (2016). Why are fewer married women joining the work force in India? A decomposition analysis over two decades. *IZA Discussion Paper No. 9722*.

Alesina, A., Giuliano, P., and Nunn, N. (2013). On the Origins of Gender Roles: Women and the Plough. *The Quarterly Journal of Economics*, 128(2):469–530.

Almond, D. (2006). Is the 1918 influenza pandemic over? Long-term effects of in utero influenza exposure in the post-1940 US population. *Journal of Political Economy*, 114(4):672–712.

Arnold, D. (2019). Death and the Modern Empire: The 1918-19 Influenza Epidemic in India. *Transactions of the Royal Historical Society*, 29:181–200.

Barro, R. J., Ursúa, J. F., and Weng, J. (2020). The coronavirus and the great influenza pandemic: Lessons from the "Spanish flu" for the coronavirus's potential effects on mortality and economic activity.

Behrman, J. R., Foster, A. D., Rosenweig, M. R., and Vashishtha, P. (1999). Women's schooling, home teaching, and economic growth. *Journal of Political Economy*, 107(4):682–714.

Boehnke, J. and Gay, V. (2019). The Missing Men: World War I and Female Labor Force Participation. *Toulouse School of Economics Working Paper*.

Boehnke, J. and Gay, V. (2020). The missing men world war i and female labor force participation. *Forthcoming, Journal of Human Resources*.

Boserup, E. (1970). *Woman's role in economic development*. Routledge.

Brainerd, E. and Siegler, M. V. (2003). The economic effects of the 1918 influenza epidemic. *CEPR Discussion Paper No. 3791*.

Broadberry, S. (2013). Accounting for the great divergence. *Working Paper: London School of Economics and Political Science*.

Carranza, E. (2014). Soil endowments, female labor force participation, and the demographic deficit of women in India. *American Economic Journal: Applied Economics*, 6(4):197–225.

Chandra, S. and Kassens-Noor, E. (2014). The evolution of pandemic influenza: evidence from India, 1918–19. *BMC infectious diseases*, 14(1):1–10.

Chandra, S., Kuljanin, G., and Wray, J. (2012). Mortality from the influenza pandemic of 1918–1919: the case of India. *Demography*, 49(3):857–865.

Chatterjee, E., Desai, S., and Vanneman, R. (2018). Indian paradox: Rising education, declining women's employment. *Demographic Research*, 38:855–881.

Chen, M. and Drèze, J. (1992). Widows and health in rural north India. *Economic and Political Weekly*, 27(43/44):WS81–WS92.

Chen, M. A. (1989). Women's work in Indian agriculture by agro-ecologic zones: Meeting needs of landless and land-poor women. *Economic and Political Weekly*, 24(43):WS79–WS89.

Correia, S., Luck, S., and Verner, E. (2020). Pandemics depress the economy, public health interventions do not: Evidence from the 1918 flu. *Available at SSRN: https://ssrn.com/abstract=3561560 or http://dx.doi.org/10.2139/ssrn.3561560*.

Davis, K. (1951). *The Population of India and Pakistan*. Princeton University Press.

De Moor, T. and Van Zanden, J. L. (2010). Girl power: The European marriage pattern and labour markets in the North Sea region in the late medieval and early modern period. *The Economic History Review*, 63(1):1–33.

Donaldson, D. and Keniston, D. (2016). Dynamics of a Malthusian Economy: India in the Aftermath of the 1918 Influenza. *Working Paper*.

Donaldson, D. and Storeygard, A. (2016). The view from above: Applications of satellite data in economics. *Journal of Economic Perspectives*, 30(4):171–98.

Duflo, E. (2012). Women empowerment and economic development. *Journal of Economic Literature*, 50(4):1051–79.

Eswaran, M., Ramaswami, B., and Wadhwa, W. (2013). Status, caste, and the time allocation of women in rural India. *Economic Development and Cultural Change*, 61(2):311–333.

Fenske, J. and Kala, N. (2017). Linguistic distance and market integration in India. *Centre for Competitive Advantage in the Global Economy Working Paper 331*.

Fernandez, R. (2007). Women, work, and culture. *Journal of the European Economic Association*, 5(2-3):305–332.

Fernández, R. (2013). Cultural change as learning: The evolution of female labor force participation over a century. *American Economic Review*, 103(1):472–500.

Fernández, R., Fogli, A., and Olivetti, C. (2004). Mothers and sons: Preference formation and female labor force dynamics. *The Quarterly Journal of Economics*, 119(4):1249–1299.

Fletcher, E., Pande, R., and Moore, C. M. T. (2017). Women and work in India: Descriptive evidence and a review of potential policies. *HKS Working Paper No. RWP18-004*.

Fogli, A. and Veldkamp, L. (2011). Nature or nurture? Learning and the geography of female labor force participation. *Econometrica*, 79(4):1103–1138.

Giuliano, P. (2014). Female labor force participation: Persistence and evolution. *The New Palgrave Dictionary of Economics, Edited by Steven N. Durlauf and Lawrence E. Blume.*

Goldin, C. (1995). The U-shaped female labor force function. *"Economic Development and Economic History" in Investment in Women's Human Capital, TP Schultz (ed.) Chicago: The University of Chicago Press.*

Goldin, C. and Olivetti, C. (2013). Shocking labor supply: A reassessment of the role of World War II on women's labor supply. *American Economic Review*, 103(3):257–62.

Goldin, C. D. (1991). The role of World War II in the rise of women's employment. *The American Economic Review*, 81(4):741–756.

Grosjean, P. and Khattar, R. (2019). It's raining men! Hallelujah? The long-run consequences of male-biased sex ratios. *The Review of Economic Studies*, 86(2):723–754.

Gulati, L. (1975). Female work participation: a study of inter-state differences. *Economic and Political Weekly*, 10(1/2):35–42.

Hill, K. (2011). Influenza in India 1918: excess mortality reassessed. *Genus*, 67(2):9–29.

Horrell, S. and Humphries, J. (1995). Women's labour force participation and the transition to the male-breadwinner family, 1790-1865. *Economic History Review*, 48(1):89–117.

Jayachandran, S. (2015). The roots of gender inequality in developing countries. *Annual Review of Economics*, 7(1):63–88.

Karlsson, M., Nilsson, T., and Pichler, S. (2014). The impact of the 1918 Spanish flu epidemic on economic performance in Sweden: An investigation into the consequences of an extraordinary mortality shock. *Journal of Health Economics*, 36:1–19.

Kiszewski, A., Mellinger, A., Spielman, A., Malaney, P., Sachs, S. E., and Sachs, J. (2004). A global index representing the stability of malaria transmission. *The American journal of tropical medicine and hygiene*, 70(5):486–498.

Klasen, S. and Pieters, J. (2015). *What explains the stagnation of female labor force participation in urban India?* The World Bank.

Lin, M.-J. and Liu, E. M. (2014). Does in utero exposure to illness matter? The 1918 influenza epidemic in Taiwan as a natural experiment. *Journal of Health Economics*, 37:152–163.

Luke, N. and Munshi, K. (2011). Women as agents of change: Female income and mobility in India. *Journal of Development Economics*, 94(1):1–17.

Luo, P. (2017). The Other Gender Gap: Female Entrepreneurship after World War II. *Working Paper, Harvard Business School.*

Mammen, K. and Paxson, C. (2000). Women's work and economic development. *Journal of Economic Perspectives*, 14(4):141–164.

Nunn, N. and Puga, D. (2012). Ruggedness: The blessing of bad geography in Africa. *Review of Economics and Statistics*, 94(1):20–36.

Patterson, K. D. and Pyle, G. F. (1991). The geography and mortality of the 1918 influenza pandemic. *Bulletin of the History of Medicine*, 65(1):4–21.

Reddy, D. N. (1975). Female Work Participation: A Study of Inter-State Differences: A Comment. *Economic and Political Weekly*, 10(23):902–905.

Reyes, O., Lee, E. C., Sah, P., Viboud, C., Chandra, S., and Bansal, S. (2018). Spatiotemporal patterns and diffusion of the 1918 influenza pandemic in British India. *American journal of epidemiology*, 187(12):2550–2560.

Roy, T. (2005). *Rethinking economic change in India: labour and livelihood*. Routledge.

Schultz, T. W. (1967). Significance of India's 1918-19 losses of agricultural labour-A reply. *The Economic Journal*, 77(305):161–163.

Teso, E. (2019). The long-term effect of demographic shocks on the evolution of gender roles: Evidence from the transatlantic slave trade. *Journal of the European Economic Association*, 17(2):497–534.

Thorner, D. and Thorner, A. (1962). *Land and labour in India*. Bombay: Asia Publishing House.

Table 1. Changes in the share of female workers in the total workforce

	1881	1901	1911	1921	1931
Total	30	31	32	31	29
Agriculture, fishing, forestry	27	31	32	32	28
Manufacturing, Mining, Construction	39	34	35	32	29
Trade	18	30	33	30	27
Transport and Other Services	13	20	22	21	24
General labour	42	42	43	44	45

Source: Calculated from Thorner and Thorner (1962), Chapter VI Table 1, pp.78-81

Table 2. Summary Statistics

	(1) mean	(2) min	(3) max	(4) sd	(5) N
1901					
Female labor force participation rate	0.3079	0.0286	0.6754	0.1932	196
Female LFP in agriculture	0.2114	0.0046	0.6073	0.1774	196
Female LFP in industry	0.0368	0.0012	0.0998	0.0207	196
Female LFP in services	0.0594	0.0087	0.1937	0.0316	196
1911					
Female labor force participation rate	0.2946	0.0258	0.6416	0.1805	196
Female LFP in agriculture	0.2139	0.0055	0.6153	0.1730	196
Female LFP in industry	0.0401	0.0028	0.1116	0.0218	196
Female LFP in services	0.0400	0.0084	0.1856	0.0266	196
1921					
Female labor force participation rate	0.2861	0.0101	0.6793	0.1764	196
Female LFP in agriculture	0.2143	0.0020	0.6582	0.1667	196
Female LFP in industry	0.0334	0.0016	0.1208	0.0204	196
Female LFP in services	0.0377	0.0026	0.1517	0.0236	196
1931					
Female labor force participation rate	0.2616	0.0196	0.6677	0.1725	196
Female LFP in agriculture	0.1674	0.0027	0.6538	0.1363	196
Female LFP in industry	0.0268	0.0023	0.2198	0.0216	196
Female LFP in services	0.0669	0.0060	0.4305	0.1004	196
1918 death rate	0.0796	0.0138	0.1792	0.0314	784
Latitude	24.3433	8.8242	34.4261	5.4145	784
Longitude	80.4551	68.0142	94.9060	6.0040	784
Humidity	62.6667	45.1817	79.5181	8.4238	784
Population density	0.6328	0.0267	36.2647	2.4714	784
Urbanization	0.1093	0.0052	1.0000	0.1239	784

Table 3. Main Results

	(1)	(2)	(3)	(4)
	Female labor force participation rate			
Influenza X 1901	-0.098	-0.215	-0.027	-0.031
	(0.181)	(0.182)	(0.261)	(0.262)
Influenza X 1921	0.294**	0.411***	0.386**	0.388**
	(0.124)	(0.134)	(0.189)	(0.188)
Influenza X 1931	-1.080***	-0.846***	-0.763**	-0.764**
	(0.264)	(0.262)	(0.341)	(0.342)
Urbanization				0.047
				(0.170)
District and year fixed effects	Yes	Yes	Yes	Yes
Province-specific time trends		Yes	Yes	Yes
Latitude x year fixed effects			Yes	Yes
Longitude x year fixed effects			Yes	Yes
Humidity x year fixed effects			Yes	Yes
Observations	784	784	784	784
R-squared	0.901	0.926	0.933	0.933

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by district in parentheses.

Table 4. Results by Sector

	(1)	(2)	(3)	(4)	(5)	(6)
	FLFP in agriculture		FLFP in industry		FLFP in services	
Influenza X 1901	0.153	0.157	-0.060	-0.062	-0.120	-0.126
	(0.217)	(0.217)	(0.052)	(0.053)	(0.074)	(0.077)
Influenza X 1921	0.093	0.091	0.028	0.029	0.266***	0.269***
	(0.146)	(0.148)	(0.036)	(0.035)	(0.094)	(0.093)
Influenza X 1931	-0.411	-0.409	-0.062	-0.063	-0.292**	-0.294**
	(0.290)	(0.289)	(0.054)	(0.054)	(0.125)	(0.125)
Urbanization		-0.050		0.028		0.062
		(0.146)		(0.023)		(0.064)
District and year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Province-specific time trends	Yes	Yes	Yes	Yes	Yes	Yes
Latitude x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Longitude x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Humidity x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	784	784	784	784	784	784
R-squared	0.939	0.939	0.788	0.788	0.846	0.847

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by district in parentheses.

Table 5. Results by Sub-Sector

	(1)	(2)	(3)	(4)	(5)	(6)
	FLFP for cultivators		FLFP for farm servants		FLFP in textiles	
Influenza X 1901	0.059	0.080	-0.170	-0.170	0.022	0.021
	(0.249)	(0.242)	(0.178)	(0.183)	(0.022)	(0.022)
Influenza X 1921	0.087	0.075	-0.059	-0.059	0.034*	0.035*
	(0.125)	(0.130)	(0.112)	(0.113)	(0.018)	(0.018)
Influenza X 1931	-0.685*	-0.679*	0.155	0.155	0.040	0.040
	(0.353)	(0.348)	(0.190)	(0.190)	(0.028)	(0.028)
Urbanization		-0.267**		0.007		0.017
		(0.111)		(0.119)		(0.012)
	FLFP in trade		FLFP in domestic service		FLFP in beggars and prostitutes	
Influenza X 1901	0.018	0.017	-0.158**	-0.163**	-0.009	-0.010
	(0.029)	(0.029)	(0.061)	(0.063)	(0.010)	(0.010)
Influenza X 1921	0.046	0.046	0.181***	0.184***	0.010	0.011
	(0.028)	(0.028)	(0.056)	(0.055)	(0.007)	(0.007)
Influenza X 1931	-0.021	-0.021	-0.267**	-0.269**	0.021	0.021
	(0.023)	(0.023)	(0.115)	(0.116)	(0.019)	(0.019)
Urbanization		0.009		0.062		0.021*
		(0.008)		(0.060)		(0.011)
District and year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Latitude x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Longitude x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Humidity x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Province-specific time trends	Yes	Yes	Yes	Yes	Yes	Yes
Observations	784	784	784	784	784	784

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by district in parentheses.

Table 6. Influenza and Possible Mediators

	(1)	(2)	(3)	(4)	(5)	(6)
	Female widowed share		Female single share		Population Density	
Influenza X 1901	-0.089	-0.112	0.031	0.026	1.218	1.305
	(0.135)	(0.136)	(0.045)	(0.045)	(1.073)	(1.107)
Influenza X 1921	0.211*	0.224*	-0.005	-0.002	-0.953**	-0.999**
	(0.123)	(0.116)	(0.052)	(0.052)	(0.465)	(0.476)
Influenza X 1931	0.137	0.130	0.010	0.009	-7.512	-7.487
	(0.116)	(0.116)	(0.044)	(0.042)	(5.761)	(5.748)
Urbanization		0.288*		0.062**		-1.074*
		(0.164)		(0.025)		(0.586)
District and year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Latitude x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Longitude x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Humidity x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Province-specific time trends	Yes	Yes	Yes	Yes	Yes	Yes
Observations	784	784	784	784	784	784
R-squared	0.821	0.826	0.925	0.926	0.984	0.984

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by district in parentheses.

Table 7. Widow Share as Mediator

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Female LFP		FLFP in agriculture		FLFP in industry		FLFP in services	
Influenza X 1901	-0.002	0.001	0.166	0.175	-0.055	-0.057	-0.112	-0.117
	(0.256)	(0.256)	(0.216)	(0.216)	(0.051)	(0.051)	(0.073)	(0.075)
Influenza X 1921	0.326*	0.324*	0.063	0.057	0.017	0.018	0.248***	0.251***
	(0.187)	(0.188)	(0.144)	(0.147)	(0.038)	(0.038)	(0.093)	(0.092)
Influenza X 1931	-0.802**	-0.801**	-0.430	-0.429	-0.069	-0.069	-0.304**	-0.304**
	(0.331)	(0.330)	(0.287)	(0.285)	(0.052)	(0.052)	(0.122)	(0.122)
Female widowed share	0.284***	0.287***	0.144***	0.153***	0.051**	0.050**	0.088***	0.084***
	(0.065)	(0.069)	(0.045)	(0.047)	(0.021)	(0.022)	(0.024)	(0.026)
Urbanization		-0.036		-0.099		0.014		0.047
		(0.158)		(0.141)		(0.025)		(0.063)
District and year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Latitude x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Longitude x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Humidity x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-specific time trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	784	784	784	784	784	784	784	784
R-squared	0.935	0.935	0.939	0.939	0.791	0.791	0.848	0.848

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by district in parentheses.

Table 8. Population Density as Mediator

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Female LFP		FLFP in agriculture		FLFP in industry		FLFP in services	
Influenza X 1901	-0.021	-0.025	0.147	0.151	-0.060	-0.062	-0.109	-0.114
	(0.260)	(0.261)	(0.217)	(0.217)	(0.052)	(0.053)	(0.076)	(0.078)
Influenza X 1921	0.382**	0.384**	0.098	0.096	0.027	0.029	0.258***	0.260***
	(0.189)	(0.189)	(0.146)	(0.148)	(0.036)	(0.036)	(0.094)	(0.093)
Influenza X 1931	-0.798**	-0.799**	-0.375	-0.375	-0.065	-0.065	-0.360***	-0.360***
	(0.344)	(0.345)	(0.298)	(0.298)	(0.055)	(0.055)	(0.104)	(0.104)
Population density	-0.005	-0.005	0.005	0.005	-0.000	-0.000	-0.009**	-0.009**
	(0.009)	(0.009)	(0.006)	(0.006)	(0.000)	(0.000)	(0.004)	(0.004)
Urbanization		0.044		-0.048		0.027		0.063
		(0.173)		(0.150)		(0.023)		(0.063)
District and year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Latitude x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Longitude x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Humidity x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-specific time trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	784	784	784	784	784	784	784	784
R-squared	0.933	0.933	0.939	0.939	0.788	0.788	0.849	0.849

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by district in parentheses.

Table 9. Influenza and Wages

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	ln Wage						
	All occupations	Agriculture	Industry	Services	General Labour in Services	Skilled	Unskilled non-agricultural
Influenza X Year before 1901	4.041	4.157					
	(3.778)	(3.572)					
Influenza X Year 1902 - 1911	-0.088	0.578	-0.164	-0.385	1.591	-0.379	0.329
	(0.523)	(0.654)	(0.738)	(0.944)	(1.036)	(0.709)	(0.895)
Influenza X Year 1919 - 1921	1.095**	-0.117	1.368*	2.609***	3.048***	1.398*	1.487
	(0.498)	(0.911)	(0.710)	(0.996)	(1.060)	(0.764)	(1.041)
Influenza X Year 1922 - 1928	1.288**	0.507	1.972***	-0.599	-0.160	1.997***	0.374
	(0.514)	(0.667)	(0.581)	(1.491)	(1.627)	(0.608)	(1.116)
Urbanization	0.203	-1.371*	1.251	2.557***	2.557***	1.285**	1.421
	(0.544)	(0.711)	(0.769)	(0.855)	(0.903)	(0.618)	(1.667)
Occupation characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District and year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Latitude x Year dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Longitude x Year dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Humidity x Year dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-specific time trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,806	2,557	3,498	751	343	3,622	627
R-squared	0.833	0.771	0.787	0.785	0.824	0.788	0.851

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by district in parentheses.

Table 10. Heterogeneity by Initial FLFP

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Female LFP		FLFP in agriculture		FLFP in industry		FLFP in services	
Panel A. Above-median FLFP in 1901								
Influenza X 1901	0.451	0.430	0.386	0.362	0.086	0.092	-0.019	-0.022
	(0.376)	(0.377)	(0.332)	(0.335)	(0.070)	(0.071)	(0.134)	(0.135)
Influenza X 1921	0.736***	0.741***	0.015	0.021	0.094**	0.093**	0.631***	0.632***
	(0.242)	(0.242)	(0.226)	(0.223)	(0.044)	(0.044)	(0.154)	(0.155)
Influenza X 1931	-1.397***	-1.390***	-0.853**	-0.845*	0.147	0.145	-0.689***	-0.688***
	(0.524)	(0.528)	(0.423)	(0.428)	(0.090)	(0.089)	(0.199)	(0.200)
Observations	392	392	392	392	392	392	392	392
R-squared	0.850	0.850	0.881	0.882	0.769	0.770	0.865	0.865
Panel B. Below-median FLFP in 1901								
Influenza X 1901	-0.034	-0.087	0.118	0.097	-0.050	-0.055	-0.102	-0.128
	(0.303)	(0.300)	(0.250)	(0.244)	(0.061)	(0.062)	(0.085)	(0.086)
Influenza X 1921	-0.046	-0.028	-0.037	-0.030	-0.050	-0.049	0.039	0.048
	(0.272)	(0.267)	(0.217)	(0.217)	(0.055)	(0.053)	(0.077)	(0.077)
Influenza X 1931	-0.607	-0.604	-0.437	-0.436	-0.207***	-0.207***	0.034	0.035
	(0.502)	(0.507)	(0.401)	(0.404)	(0.078)	(0.078)	(0.161)	(0.161)
Observations	392	392	392	392	392	392	392	392
R-squared	0.854	0.854	0.857	0.857	0.861	0.861	0.837	0.839
Urbanization	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District and year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Latitude x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Longitude x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Humidity x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-specific time trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by district in parentheses.

Appendix: Not for publication

Appendix A. Sources of Wage Data

Our sources of wage data are as follows:

- Wage Census of Bengal taken in 1911, 1916, 1925
- Statement of Rural and Urban wages prevailing in the Central Provinces and Berar for 1910-1923
- Wage Census of Bombay Presidency taken in 1911, 1916
- Report on an enquiry into agricultural wages in the Bombay presidency
- Report on the Wage Census of Eastern Bengal and Assam 1911
- Madras Wage Census report 1908, 1911
- Wage Census of the North-west Frontier Province taken in 1912, 1917, 1923, 1928
- Report on the Regular Wage Survey of the Punjab 1912, 1917, 1922, 1927
- The wage census of the United Provinces taken in 1906, 1911, 1916, 1928
- Prices and Wages in India. 1919 (1911, 1916)

These provide wage observations for the following provinces and years:

- Assam: 1911, 1916
- Bengal: 1911, 1916, 1925
- Bihar and Orissa: 1911, 1916
- Bombay: 1900-1922
- Central Provinces and Berar: 1910-1923
- Madras: 1908, 1911, 1916
- Punjab: 1909, 1911, 1912, 1916, 1917, 1922, 1927
- United Provinces: 1906, 1911, 1916, 1928
- North-west Frontier Province: 1911, 1912, 1916, 1917, 1923, 1928

Appendix B. Additional Tables

Table A1. Male Labor Force Participation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Male LFP		MLFP in agriculture		MLFP in industry		MLFP in services	
Influenza X 1901	0.164	0.158	0.106	0.123	0.083	0.078	-0.027	-0.044
	(0.156)	(0.158)	(0.136)	(0.137)	(0.052)	(0.051)	(0.125)	(0.133)
Influenza X 1921	0.084	0.087	0.186	0.177	0.010	0.013	-0.114	-0.105
	(0.125)	(0.125)	(0.130)	(0.127)	(0.043)	(0.044)	(0.123)	(0.116)
Influenza X 1931	-0.017	-0.019	0.052	0.057	0.020	0.018	-0.088	-0.093
	(0.159)	(0.160)	(0.165)	(0.161)	(0.053)	(0.051)	(0.123)	(0.124)
Urbanization		0.073		-0.209**		0.070*		0.209***
		(0.073)		(0.089)		(0.036)		(0.053)
District and year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Latitude x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Longitude x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Humidity x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-specific time trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	784	784	784	784	784	784	784	784
R-squared	0.802	0.803	0.911	0.912	0.940	0.941	0.925	0.927

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by district in parentheses.

Table A2. Control for Cholera X Year Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Female LFP		FLFP in agriculture		FLFP in industry		FLFP in services	
Influenza X 1901	-0.047	-0.051	0.120	0.125	-0.059	-0.062	-0.108	-0.115
	(0.259)	(0.261)	(0.215)	(0.215)	(0.052)	(0.053)	(0.076)	(0.079)
Influenza X 1921	0.396**	0.398**	0.091	0.088	0.024	0.026	0.282***	0.286***
	(0.189)	(0.188)	(0.146)	(0.148)	(0.036)	(0.035)	(0.091)	(0.089)
Influenza X 1931	-0.785**	-0.785**	-0.447	-0.446	-0.060	-0.060	-0.279**	-0.280**
	(0.337)	(0.339)	(0.282)	(0.281)	(0.054)	(0.055)	(0.128)	(0.128)
Urbanization		0.048		-0.062		0.029		0.079
		(0.163)		(0.139)		(0.023)		(0.066)
Cholera X year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District and year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Latitude x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Longitude x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Humidity x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-specific time trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	784	784	784	784	784	784	784	784
R-squared	0.934	0.934	0.940	0.940	0.789	0.789	0.848	0.848

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by district in parentheses.

Table A3. Control for Crop Suitabilities X Year Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Female LFP		FLFP in agriculture		FLFP in industry		FLFP in services	
Influenza X 1901	0.284	0.290	0.387	0.385	-0.044	-0.041	-0.054	-0.049
	(0.286)	(0.287)	(0.254)	(0.255)	(0.057)	(0.057)	(0.083)	(0.084)
Influenza X 1921	0.368**	0.369**	0.147	0.146	0.066*	0.067*	0.153*	0.154*
	(0.162)	(0.163)	(0.161)	(0.161)	(0.038)	(0.038)	(0.082)	(0.082)
Influenza X 1931	-0.281	-0.278	0.011	0.010	-0.099*	-0.097*	-0.190	-0.188
	(0.307)	(0.308)	(0.254)	(0.254)	(0.057)	(0.058)	(0.138)	(0.138)
Urbanization		0.067		-0.019		0.037		0.048
		(0.185)		(0.153)		(0.025)		(0.069)
Crop suitabilities X year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District and year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Latitude x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Longitude x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Humidity x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-specific time trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	784	784	784	784	784	784	784	784
R-squared	0.945	0.945	0.950	0.950	0.827	0.828	0.890	0.890

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by district in parentheses.

Table A4. Control for Agricultural Commodity Areas

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Female LFP		FLFP in agriculture		FLFP in industry		FLFP in services	
Influenza X 1901	-0.033	-0.035	0.173	0.180	-0.077	-0.079	-0.130*	-0.135*
	(0.264)	(0.264)	(0.222)	(0.222)	(0.054)	(0.055)	(0.074)	(0.077)
Influenza X 1921	0.330*	0.332*	0.059	0.053	0.001	0.003	0.272***	0.277***
	(0.195)	(0.195)	(0.150)	(0.153)	(0.039)	(0.039)	(0.097)	(0.096)
Influenza X 1931	-0.808**	-0.807**	-0.437	-0.439	-0.070	-0.070	-0.301**	-0.300**
	(0.343)	(0.344)	(0.294)	(0.292)	(0.054)	(0.055)	(0.128)	(0.128)
Urbanization		0.021		-0.078		0.026		0.071
		(0.172)		(0.146)		(0.024)		(0.066)
Agricultural commodity areas	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District and year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Latitude x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Longitude x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Humidity x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-specific time trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	784	784	784	784	784	784	784	784
R-squared	0.934	0.934	0.940	0.940	0.796	0.796	0.848	0.848

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by district in parentheses. Agricultural commodities include rice, wheat, sugarcane, jute, opium, tea, tobacco, and cotton.

Table A5. Control for Geographic Controls X Year Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Female LFP		FLFP in agriculture		FLFP in industry		FLFP in services	
Influenza X 1901	0.039	0.040	0.130	0.130	-0.069	-0.068	-0.024	-0.023
	(0.257)	(0.258)	(0.224)	(0.224)	(0.052)	(0.053)	(0.070)	(0.071)
Influenza X 1921	0.319*	0.324*	0.111	0.109	0.050	0.052	0.158**	0.162**
	(0.169)	(0.169)	(0.146)	(0.147)	(0.036)	(0.036)	(0.078)	(0.077)
Influenza X 1931	-0.673**	-0.671**	-0.455*	-0.456*	-0.055	-0.054	-0.165	-0.163
	(0.323)	(0.324)	(0.275)	(0.274)	(0.050)	(0.050)	(0.135)	(0.135)
Urbanization		0.075		-0.036		0.034		0.075
		(0.176)		(0.153)		(0.024)		(0.059)
Geographic controls X year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District and year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Latitude x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Longitude x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Humidity x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-specific time trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	784	784	784	784	784	784	784	784
R-squared	0.935	0.935	0.942	0.942	0.796	0.796	0.876	0.876

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by district in parentheses. Geographic controls include temperature, precipitation, altitude, ruggedness, and malaria.

Table A6. Robustness to Using the Male Influenza Death Rate

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Female LFP		FLFP in agriculture		FLFP in industry		FLFP in services	
Male Influenza X 1901	-0.078	-0.082	0.063	0.068	-0.051	-0.053	-0.088	-0.094
	(0.260)	(0.262)	(0.215)	(0.215)	(0.053)	(0.054)	(0.081)	(0.084)
Male Influenza X 1921	0.483**	0.486**	0.146	0.143	0.028	0.029	0.309***	0.313***
	(0.189)	(0.188)	(0.147)	(0.149)	(0.038)	(0.037)	(0.100)	(0.098)
Male Influenza X 1931	-0.873**	-0.873**	-0.533*	-0.532*	-0.060	-0.060	-0.280**	-0.281**
	(0.356)	(0.357)	(0.301)	(0.300)	(0.056)	(0.056)	(0.137)	(0.136)
Urbanization		0.051		-0.051		0.028		0.072
		(0.168)		(0.145)		(0.023)		(0.064)
District and year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Latitude x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Longitude x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Humidity x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-specific time trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	784	784	784	784	784	784	784	784
R-squared	0.934	0.934	0.939	0.939	0.787	0.788	0.847	0.847

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by district in parentheses.

Table A7. Male Death and Marital Outcomes

	(1)	(2)	(3)	(4)
	Female widowed share		Female single share	
Male Influenza X 1901	-0.087	-0.113	0.049	0.044
	(0.145)	(0.146)	(0.047)	(0.047)
Male Influenza X 1921	0.232*	0.248**	0.014	0.017
	(0.124)	(0.117)	(0.053)	(0.053)
Male Influenza X 1931	0.168	0.166	-0.003	-0.004
	(0.124)	(0.123)	(0.046)	(0.044)
Urbanization		0.290*		0.062**
		(0.165)		(0.025)
District and year fixed effects	Yes	Yes	Yes	Yes
Province-specific time trends	Yes	Yes	Yes	Yes
Latitude x Year dummy	Yes	Yes	Yes	Yes
Longitude x Year dummy	Yes	Yes	Yes	Yes
Humidity x Year dummy	Yes	Yes	Yes	Yes
Observations	784	784	784	784
R-squared	0.821	0.826	0.925	0.926

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by district in parentheses.

Table A8. Male Death and Widow Share as Mediator

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Female LFP		FLFP in agriculture		FLFP in industry		FLFP in services	
Male Influenza X 1901	-0.053	-0.050	0.076	0.085	-0.046	-0.047	-0.080	-0.085
	(0.255)	(0.254)	(0.214)	(0.214)	(0.051)	(0.052)	(0.079)	(0.081)
Male Influenza X 1921	0.417**	0.415**	0.113	0.105	0.016	0.017	0.289***	0.293***
	(0.188)	(0.190)	(0.145)	(0.148)	(0.040)	(0.039)	(0.099)	(0.098)
Male Influenza X 1931	-0.920***	-0.920***	-0.557*	-0.557*	-0.068	-0.068	-0.295**	-0.295**
	(0.343)	(0.343)	(0.297)	(0.295)	(0.054)	(0.054)	(0.133)	(0.133)
Female widowed share	0.283***	0.286***	0.142***	0.151***	0.051**	0.050**	0.088***	0.084***
	(0.065)	(0.069)	(0.045)	(0.047)	(0.022)	(0.022)	(0.024)	(0.026)
Urbanization		-0.032		-0.095		0.014		0.047
		(0.157)		(0.141)		(0.025)		(0.063)
District and year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-specific time trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Latitude x Year dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Longitude x Year dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Humidity x Year dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	784	784	784	784	784	784	784	784
R-squared	0.935	0.935	0.940	0.940	0.791	0.791	0.848	0.848

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by district in parentheses.

Table A9. Robustness to Using the Female Influenza Death Rate

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Female LFP		FLFP in agriculture		FLFP in industry		FLFP in services	
Female Influenza X 1901	0.027	0.025	0.232	0.236	-0.066	-0.068	-0.141**	-0.145**
	(0.258)	(0.258)	(0.219)	(0.219)	(0.050)	(0.050)	(0.066)	(0.068)
Female Influenza X 1921	0.275	0.277	0.041	0.039	0.026	0.027	0.209**	0.211**
	(0.185)	(0.185)	(0.145)	(0.146)	(0.033)	(0.033)	(0.086)	(0.085)
Female Influenza X 1931	-0.617*	-0.618*	-0.269	-0.267	-0.062	-0.063	-0.288***	-0.290***
	(0.318)	(0.319)	(0.272)	(0.271)	(0.050)	(0.051)	(0.109)	(0.109)
Urbanization		0.043		-0.058		0.029		0.070
		(0.172)		(0.146)		(0.023)		(0.067)
District and year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Latitude x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Longitude x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Humidity x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-specific time trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	784	784	784	784	784	784	784	784
R-squared	0.932	0.932	0.939	0.939	0.788	0.788	0.846	0.846

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by district in parentheses.

Table A10. Results excluding beggars and prostitutes

	(1)	(2)	(3)	(4)
	Female LFP		FLFP in service	
Influenza X 1901	-0.019	-0.021	-0.112	-0.116
	(0.258)	(0.259)	(0.074)	(0.076)
Influenza X 1921	0.376**	0.377**	0.256***	0.258***
	(0.187)	(0.187)	(0.092)	(0.091)
Influenza X 1931	-0.785**	-0.785**	-0.314***	-0.315***
	(0.337)	(0.338)	(0.114)	(0.114)
Urbanization		0.027		0.051
		(0.165)		(0.061)
District and year fixed effects	Yes	Yes	Yes	Yes
Province-specific time trends	Yes	Yes	Yes	Yes
Latitude x year fixed effects	Yes	Yes	Yes	Yes
Longitude x year fixed effects	Yes	Yes	Yes	Yes
Humidity x year fixed effects	Yes	Yes	Yes	Yes
Observations	784	784	784	784
R-squared	0.934	0.934	0.851	0.851

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by district in

Table A11. Mediation analysis excluding beggars and prostitutes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Female LFP				FLFP in service			
Influenza X 1901	0.007	0.012	-0.016	-0.018	-0.104	-0.106	-0.104	-0.107
	(0.254)	(0.253)	(0.258)	(0.259)	(0.072)	(0.074)	(0.075)	(0.077)
Influenza X 1921	0.315*	0.312*	0.374**	0.375**	0.237***	0.238***	0.249***	0.251***
	(0.185)	(0.187)	(0.187)	(0.187)	(0.091)	(0.090)	(0.092)	(0.091)
Influenza X 1931	-0.824**	-0.823**	-0.801**	-0.801**	-0.326***	-0.326***	-0.363***	-0.363***
	(0.327)	(0.326)	(0.343)	(0.344)	(0.112)	(0.112)	(0.102)	(0.103)
Female widowed share	0.287***	0.292***			0.091***	0.088***		
	(0.063)	(0.066)			(0.022)	(0.024)		
Population density			-0.002	-0.002			-0.007*	-0.006*
			(0.009)	(0.009)			(0.003)	(0.003)
Urbanization		-0.057		0.024		0.025		0.044
		(0.153)		(0.165)		(0.059)		(0.060)
District and year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Latitude x Year dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Longitude x Year dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Humidity x Year dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-specific time trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	784	784	784	784	784	784	784	784
R-squared	0.936	0.936	0.934	0.934	0.852	0.852	0.852	0.852

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by district in parentheses.