LONG-TERM EFFECTS OF EARLY-LIFE DEVELOPMENT: EVIDENCE FROM THE 1959-1961 CHINA FAMINE

Douglas Almond†
Lena Edlund†
Hongbin Li§
Junsen Zhang¶

June 16, 2008

*We would like to thank Janet Currie, Andrew Gelman, Hilary Hoynes, Robert Kaestner, Mark Rosenzweig, David St. Clair, Jane Waldfogel, and David Wise for helpful comments. Holly Ho Ming and Hongyan Zhao provided outstanding research assistance. Almond and Edlund thank Russell Sage Foundation and Almond the Fulbright Program for financial support.

†Department of Economics, Columbia University and NBER.
‡Department of Economics, Columbia University.
§Department of Economics, Tsinghua University.
¶Department of Economics, Chinese University of Hong Kong.
Abstract

This paper estimates the effects of maternal stress and malnutrition using the 1959-1961 Chinese famine as a natural experiment. Observed forty years later in the 2000 China Census (1% sample), Famine survivors showed impaired literacy, labor market, wealth, and marriage market outcomes. In addition, maternal malnutrition reduced the sex ratio (males to females) in two generations – those prenatally exposed and their children – presumably through heightened male mortality. This tendency toward female cohorts is interpretable in light of the Trivers-Willard (1973) hypothesis, according to which parents in poor condition should skew the offspring sex ratio toward daughters. Hong Kong Natality micro data from 1984-2004 further confirm this pattern. The persistence of poor nutrition in China – particularly in rural areas and among girls – suggests that health and economic outcomes will be compromised well into the 21st century.
1 Introduction

The dramatic success of China’s One Child Policy in reducing fertility cata-
pults the question of population aging to center stage. As China’s dependency ratio increases, the health and productivity of those of working age will play key roles. So far, attention has generally focussed on investments in these “working age” cohorts that occur after birth (e.g. educational investments). This paper focuses instead on the prenatal environment and its impact on health and economic outcomes in adulthood, exploiting the 1959-1961 Chinese famine (henceforth “the Famine”) as a natural experiment in maternal stress and nutrition.

While starvation on the scale of the Famine may seem remote, maternal malnutrition is not. In the twenty years following the Famine, average nutrition was little improved from the 1930s (White, 1991). Smil (1981) noted that “Chinese food availability has remained virtually static for at least half a century.” Meat remained scarce and diets were heavily reliant on grains, which accounted for 90% of energy and 80% of protein (Smil, 1981). Disruptions in grain production brought “permanent malnutrition to at least 200 million peasants” since the mid-1960s (Smil, 1981). Food rationing, first introduced in 1953, was used as a tool to encourage compliance with the One Child Policy (Li & Cooney, 1993).

Even after the precipitous decline in fertility during the late 1970s and 1980s, poor nutrition persisted, especially in rural areas and among girls. Between 1987 and 1992, the height of children in urban areas increased five times as fast as rural areas, attributable in part to “more inequitable distribution of the economic resources for nutrition” (Shen & Chang, 1996). Similarly, Hesketh et al. (2002) found that diets were less varied and nutritional depri-
vation more common in rural areas of eastern China — anaemia (Hb $\leq 110$ gl$^{-1}$) was 50% more common than in the rapidly-developing cities. Moreover, more than three quarters of those with anaemia in rural areas were girls: 19% were anaemic versus 4.8% for rural boys. Fully 55% of rural girls were moderately anaemic (haemoglobin concentrations below 120 gl$^{-1}$), versus 21% of rural boys.

Our inquiry is motivated by a growing literature finding the pre- and perinatal periods critical to morbidity and lifespan. Pioneered by Barker (1992), the “fetal origins hypothesis” linked cardiovascular mortality to maternal nutritional status. Later research has honed in on maternal stress as triggering biological responses in the growing fetus that programs for a life in a resource poor environment (Gluckman & Hanson, 2004). However, the bulk of empirical evidence derives from animal experiments; evidence for humans is surprisingly scarce (see e.g., Rasmussen (2001); Walker et al. (2007)). Omitted factors (e.g., parental abilities and attitudes) can generate positive associations between measures of fetal health and adult socioeconomic outcomes in the absence of a controlled experiment. Therefore, the “most compelling examinations of the fetal origins hypothesis look for sharp exogenous shocks in fetal health that are caused by conditions outside the control of the mother” Currie (2007, page 27).

Observing cohorts born 1956-64 in the 2000 Chinese Population Census (1% sample), we find that men were 9% more likely to be illiterate, 6% less likely to work, and 6.5% less likely to be married if exposed to the Famine in utero. Women were 7.5% more likely to be illiterate and 3% less likely to work, and tended to marry men with less education, if exposed in utero. We also find fetal exposure to the Famine substantially reduced the the cohort’s
sex ratio (fewer males), suggesting greater male vulnerability to maternal mal-nutrition. Perhaps most intriguingly, we find an “echo effect” of the Famine on the next generation: children whose mothers were exposed prenatally also register Famine impacts. In particular, Famine-exposed mothers were more likely to give birth to daughters. To our knowledge, ours is the first study to trace the offspring sex ratio to the \textit{in utero} environment of the parent.

To test the robustness of our findings, we pursue two additional approaches. First, we utilize geographic variation in Famine severity to generate comparisons \textit{within} birth cohorts. Here, estimates of Famine damage will be confounded by events experienced later in life (e.g. the Cultural Revolution 1966-76) insofar as these events replicated the geographic variation in Famine intensity \textit{and} differentially impacted those cohorts \textit{in utero} during the Famine. Second, while the Famine was endemic in mainland China (affecting both urban and rural areas), Hong Kong, then a British colony, was spared. The Famine resulted in a large inflow of mainland Chinese into Hong Kong. We can therefore observe whether children of mainland emigrants exposed to the Famine register intergenerational damage using Hong Kong’s natality data, derived from the universe of Hong Kong birth certificates.\footnote{These certificates record country of birth of the mother. Among Hong Kong mothers who emigrated from the mainland, those exposed to the Famine \textit{in utero} had worse birth outcomes than other mainland emigrants.} Results from these two additional approaches corroborate the findings from across-cohort comparisons in the Census data: damage to a broad spectrum of outcomes persists 40 years after the Famine.

In addition to the potential for remedial investments (Heckman, 2007), two factors lead us to believe our estimates of long-term damage are conservative (i.e., biased toward zero). First, the selective effects of the Famine are likely
to cull the relatively weak. Second, the comparison group was also affected by the Famine: older cohorts experienced it directly and younger cohorts were the children of Famine survivors. Assuming that these adjacent cohorts were also negatively affected by the Famine, our estimated effects are of the incremental effect of acute maternal malnutrition, as opposed to, e.g., starvation while an infant or toddler or from being born to a mother who starved prior to her pregnancy.

China is experiencing rapid economic growth and, perhaps ironically, this rapid transition may exacerbate the health consequences of maternal (or grand maternal, see below) malnutrition as the “thrifty phenotype” finds itself in a resource rich environment. One reason for long lasting, even inter-generational effects, is that a girl is born with all her eggs, which means that daughters and the eggs for their children, future grand-children, share in utero environment. Another reason is that gene expression is affected by the early life environment, and therefore, the mother’s status (health and otherwise) has epigenetic effects. Therefore, while rapid economic growth holds the promise of greater access to education and health care, this new found affluence also pose health challenges akin to those faced by (especially) minority populations in the U.S.: obesity, type II diabetes and hypertension.

The remainder of the paper is organized as follows. Sections 1.1-1.3 describe the background of the Famine and reviews the related literature. Section 2 describes the 2000 Chinese Population Census and the 1984-2004 Hong Kong Natality files. Section 3 reports descriptive and regression results, along with a discussion of potential biases. Section 4 concludes.
1.1 Famine background

The Famine ranks as the worst in recorded history. Between 18 and 30 million died due to the “systemic failure” of Mao’s Great Leap Forward (Li & Yang, 2005). The Famine began in the fall of 1959 and impacted all regions of China. Grain output dropped 15% in 1959 and another 15% in 1960 (Li & Yang, 2005, page 846). By 1962, birth and death rates had returned to normal levels.

While weather conditions contributed to the Famine, the radical economic policies of the Great Leap Forward were chiefly to blame (Lin, 1990; Li & Yang, 2005). In a breakneck attempt to overtake Britain and eventually the U.S., labor was diverted from agriculture to industry while grain procurement from rural areas was increased. At the same time, collectivization of agricultural production resulted in shirking and falling productivity (Lin, 1990). The political climate encouraged provincial leaders to overstate grain production and despite widespread starvation, China was a net grain exporter throughout 1960 (Yao, 1999; Lin & Yang, 2000).

Famine intensity varied by region (Peng, 1987). Rural death rates rose to 2.5 times pre-Famine levels. Urban residents fared better but were not spared, death rates in the peak year 1960 were 80% above pre-Famine levels (China Statistical Press, 2000). Central provinces such as Anhui, Henan and Sichuan were the worst hit, while northeastern provinces such as Heilongjiang and Jilin were relatively spared. By 1961, death rates had returned to normal in more than half of the provinces, but remained high in, for instance, the southern provinces Guangxi and Guizhou (close to Hong Kong).
1.2 Famine studies: Epidemiology

The best epidemiological evidence to date linking maternal nutritional deprivation to subsequent adult outcomes derives from the cohort in utero during the 1944-45 Dutch famine. While the seminal study found limited effects at age 18 (Stein et al., 1975), at middle age, this cohort exhibited a broad spectrum of damage including: self-reported health (Roseboom et al., 2001b), coronary heart disease morbidity (Roseboom et al., 2001b; Bleker et al., 2005), and adult antisocial personality disorders (Neugebauer et al., 1999). These, and studies of the 1866-1868 Finnish Famine and the Nazi Seige of Leningrad, have focused exclusively on health outcomes. Epidemiological findings from the Chinese Famine include heightened risk of schizophrenia (Clair et al., 2005) and obesity among women (Luo et al., 2006).

1.3 Famine studies: Economics

A number of recent studies evaluating the Famine’s impact on the socioeconomic outcomes of survivors have used the China Health and Nutrition Surveys (CHNS) (Chen & Zhou, 2007; Meng & Qian, 2006; Gorgens et al., 2005). The CHNS is a panel dataset that began in 1989 of health and economic outcomes of approximately 4,000 Chinese households from nine provinces (out of 31 provinces or province level administrative regions). The small sample size combined with the collapse of fertility during the Famine necessitates the inclusion of ages well after birth as “treated.” However, broad “early childhood” hypotheses make it difficult to reject alternative explanations. The possibility that events at other ages – for instance the subsequent Cultural Revolution

---

2105 rural CHNS respondents and 62 urban CHNS respondents were born in 1960, with 66 and 45 respectively in 1961 (Chen & Zhou, 2007, table 2).
and the forced “rustification” of students in outlying areas – confounds results is a concern.

Chen & Zhou (2007) considered those up to age 6 as treated. They proxied Famine intensity by the province level death rate in 1960 and found the Famine thus measured to have resulted in stunting of those born in 1955, 1957, 1959, 1960 and 1962, with the largest height reductions for the 1959, 1960, and 1962 birth cohorts. Moreover, they found reduced labor supply of those born in 1959 and 1960, and lower wealth as measured by the size of residence for birth cohorts 1958 and 1959.

Meng & Qian (2006) considered the following birth cohorts as potentially affected: 1952-54, 1955-58, 1959-60, with cohorts born 1961-64 as the reference group. Using reductions in cohort size as a proxy for Famine severity (assumed to occur through Famine mortality), their OLS estimation returned mixed results, and little evidence for a particularly strong effect for the 1959-60 cohort. Instrumenting for cohort size, using per capita grain production in 1997, they found a small negative effects on education, but a substantial (25%) reduction in hours worked for the 1959-60 cohort.

Gorgens et al. (2005) studied adult heights of cohorts exposed to the Famine in childhood. They argued both that children who survived the Famine did not show any stunting and that stunting did occur. They reconcile these two arguments by a third: Famine mortality was concentrated among shorter people. The net effect of stunting and selection, the authors argued, made the height of survivors appear unchanged. However, the claim that no stunting is observed among survivors is controversial (Chen & Zhou, 2007; Yan, 1999; Morgan, 2006).
2 Data

Our primary data set is 2000 Population Census of China.\(^3\) The 1\% sample includes more than 11 million records and has not (to our knowledge) been used to evaluate long-term effects of the Famine.\(^4\) Outcomes include educational attainment, labor market status, and residence information of respondents. Demographic information includes sex, birth year and month, marriage and fertility information (see the Appendix).

Unlike preceding Census surveys and the CHNS data, the 2000 Census records the province of birth, eliminating the potential for confounding due to internal migration.\(^5\) The 2000 Census captures Famine cohorts near age 40, and therefore near the flat portion of their occupation and earnings profile. Moreover, it is the first Census to capture near-complete fertility histories of women born during the Famine.\(^6\) We restrict the analysis to those born 1956-1964, a sub-sample which includes three pre-Famine years and three post-Famine years (death rates peaked in 1960 but were elevated in 1959 and 1961 as well, see Figure 1). Our relatively narrow birth interval is intended to increase the similarity of the unobserved later-life factors and their effects on Census outcomes.

Our second data source is the natality microdata for Hong Kong (1984-2004), derived from the universe of birth certificates. These data include information on maternal country of birth. Restricting the sample to mothers

---

\(^3\)Conducted by the Chinese National Bureau of Statistics for mainland China.  
\(^4\)Shi (2006) used a 0.1\% sub-sample of the 2000 Census.  
\(^5\)6\% of those born 1956-64 reported moving from another province since birth, with another 10\% relocating towns within the province of birth.  
\(^6\)A mere 0.3\% of women born in 1960 reported having a child between November 1999 and October 2000. For comparison, 14.8\% for women born in 1976 had born a child in the same period.
of singletons either born in mainland China or Hong Kong in the years 1957-1965 yields some 600,000 records, approximately one-third of whom emigrated from the mainland. The Hong Kong data provide an important control group since all of mainland China was afflicted by the Famine (Cai & Feng, 2005).

2.1 Measuring the Famine

We use two measures of famine intensity: death rates and average month of birth.

**Death rates** We use the all-age death rate (China Statistical Press, 2000) by year and province to calculate two (mortality-based) proxies of Famine intensity. We have data for 29 out of the 31 provinces (or province level divisions).

- First, for every person, we calculate the weighted average of the death rate in the province of birth for the duration of the fetal period, henceforth “weighted death rate” or \( wdr_{jt} \). For example, a person born in January 1960 in Beijing is assigned 1/9th of Beijing’s 1960’s mortality rate and 8/9th of Beijing’s 1959’s mortality rate. This weighted death rate ranged from 0.005 to 0.069 (per person).

- Second, we collapse this weighted death rate by month of birth, thus calculating a population weighted national average for each month and year, henceforth “aggregate weighted death rate” or \( awdr_t \). During the study period, this measure ranged from 0.010 (in 1963) to 0.022 (at the end of 1960), a difference of 0.012. Thus measured, those born towards the end of 1960, and early 1961, were exposed to the greatest Famine intensity *in utero* (Figure 1).
Average month of birth In the northern hemisphere, famines tend to be most severe during the winter months. This reduces fertility disproportionately in the later half of the calendar year, thereby lowering the average month of birth (Stein et al., 1975). This proxy applied to the 2000 Census indicates 1960 as the worst year for mainland China (Figure 2), i.e. consistent with the mortality data. Because emigrants to Hong Kong were a highly selected group, both geographically (the Famine hit bordering provinces later) and due to the particular migration policies in place (further described in Section 3.5), we cannot rely on mainland mortality data in the Hong Kong analysis. To obtain a proxy for when the Famine peaked for this group of immigrants, we use average month of birth. This proxy indicates 1961 as the worst Famine year for Hong Kong mothers born on the mainland (Figure 3). As expected of the “control group,” there was no corresponding change for Hong Kong natives (Figure 4).

3 Results

3.1 Descriptive Results

We begin by presenting unadjusted outcomes by quarter of birth (for all Chinese) in the four panels of Figure 5. These figures indicate that those born around 1960 had worse socio-economic outcomes than the cohort trend would predict. Recall that this cohort was in utero during the period with the highest death rate, as measured by the weighted death rate (Figure 1). In 2000,

7Authors' tabulation of appendix table 4 data in Stein et al. (1975).
8Natality data for Hong Kong identify Mainland immigrants, but not their province of birth (nor province of last residence).
the 1960 birth cohort was more likely to be: (1) not working at the time of
the Census; (2) supported by other household members; (3) living in a smaller
home and; (4) parents of female children. For some of these outcomes, de-
partures from the cohort trend appear in the adjacent cohorts as well. This
pattern mirrors the 1959-1961 duration of the Famine, with a peak in 1960.

### 3.2 Regression Results

To investigate systematically how adult outcomes vary with prenatal Famine
exposure, we focus on the cohorts born 1956-1964 and estimate by OLS:

$$y_{it} = \beta_0 + \theta \cdot \text{awdr}_t + \beta_1 \cdot \text{YOB} + \beta_2 \cdot \text{YOB}^2 + \beta_3 \cdot \text{YOB}^3 + \lambda_{\text{province}} + \epsilon_{it}, \quad (1)$$

where $y_{it}$ denotes the outcome for individual $i$ born in period $t$, $\text{awdr}_t$ denotes
the aggregate weighted death rate by birth year and month of birth $t$, and
$\text{YOB}$ denotes birth year. We enter $\text{YOB}$ as a cubic to control for the non-
linear cohort/age effects apparent in the four panels of Figure 5. Finally, we
include a vector of province dummies, $\lambda_{\text{province}}$. Thus (1) allows for a flexible
cohort profile within a narrowly-defined birth interval, and assesses whether
the prenatal death rate contributes additional explanatory power, as reflected
by $\hat{\theta}$. We estimate equation (1) separately for men and women. We do not
include dummies for the month of birth, given its apparent endogeneity in
Figure 2. (However, inclusion of month of birth dummies does not alter the
basic results from estimating (1) and (2); results are available on request.)

Results from estimating (1) for 2000 Census outcomes are reported in Ta-
bles 1-3. Table 1 shows a consistent deleterious effect of prenatal Famine
exposure on labor market outcomes. Greater famine intensity is associated
with a higher likelihood of being illiterate and not working. During the

---

*aSee Section 2.1.*
Famine, awdr increased by 1.2 percentage points, implying, e.g., that the most Famine exposed cohorts were 7.5% (0.5052×0.012/0.081) [women] and 9% (0.1585×0.012/0.021) [men] more likely to be illiterate; 3% (0.4714×0.012/0.189) [women] and 5.9% (0.4017×0.012/0.082) [men] more likely to not work; and 13% (0.0448×0.012/0.004) [women] more likely to be disabled. Men in utero during Famine were 9% more likely to be supported financially by other household members (“Dependent”), and the figure for women was 4%.

The census does not have any direct measure of earnings, but there is information on housing, which may serve as a wealth proxy. Thus measured, greater fetal Famine exposure reduced adult economic status (Table 1, last column).

We also estimate equation (1) for marriage market outcomes (Table 2). While marriage was nearly universal for women, inspection of who they married reveals that Famine exposed women married men with less education. For men, both the extensive and intensive margins were affected. Men were 6.5% (0.4902×0.012/0.09) more likely to be unmarried and 8.2% (0.2676×0.012/0.039) more likely to never have married. Moreover, they married at older ages (1.5 months) and were 0.7 % (0.5145×0.012/0.87) less likely to head their households.

The poor marriage market outcomes are unlikely to be driven by conventional supply and demand factors. As cohorts born during the Famine were substantially smaller than adjacent cohorts, the “marriage squeeze” would work in their favor.10

Prenatal famine exposure also raised male (relative to female) mortality as

10For both men and women, the three smallest cohorts 1950-1970 were those born 1959-1961.
evidenced by survival around age 40. The most exposed cohort was 1.5%-age points (1.3147×0.012) more female (Table 3, column 1). The most striking finding, however, is that prenatally exposed women bore more girls, the offspring of the most Famine exposed were 0.4 percentage points (0.3194×0.012) less male (column 2).\textsuperscript{11} To anticipate results, the Hong Kong data (derived from birth certificates) corroborate this pattern.

### 3.3 Geographic variation in Famine intensity

The second test of our hypothesis isolates the geographic variation in the Famine and makes comparisons exclusively within (annual) birth cohorts. This approach reduces the potential for confounding from later-life events with age-specific effects (e.g. if the Cultural Revolution, launched in 1966, delayed school entry among six-year olds). Here, confounding by such later-life events would require their geographic variation to mirror the Famine (while also replicating the Famine’s cohort effects).\textsuperscript{12} We estimate by OLS:

\[
y_{itj} = \beta_0 + \theta \cdot w_{dr_{it}} + \gamma_{yob} + \lambda_{province} + \varepsilon_{itj}, \tag{2}
\]

where \(\theta\) is the parameter of interest, \(t\) denotes year and month of birth and \(j\) the province of birth. The mortality rate is the weighted death rate \((w_{dr_{it}})\) previously described for the individual’s birth date (year and month) and province of birth. As in equation (1), we include vectors of province of birth dummies \((\lambda_{province})\), and, as the goal is to isolate the geographic variation in

\textsuperscript{11}Similar results are obtained when the logit transform of the proportion of male children is the dependent variable.

\textsuperscript{12}In contrast to the Famine, urban residents were more affected by the Cultural Revolution than rural residents. In addition, the Cultural revolution lasted ten years and therefore impacted a broader span of birth cohorts.
health induced by the Famine, we absorb the average differences for each birth year by including a vector of year of birth dummies ($\gamma_{yob}$).

Results from estimating (2) provide qualitatively distinct evidence of Famine damage: regional differences in outcomes for the Famine cohort line up with regional differences in malnutrition (Tables 4-6). Table 4 shows that local famine severity indeed corresponds to the magnitude of damage in Census outcomes. Women born in high-Famine areas had larger increases in disability rates and larger reductions in house sizes. For men, differences in literacy, work status, disability, and house size correspond to Famine severity in the expected direction.

The magnitude of damage obtained from estimating (2) is generally either similar to that found with (1), or somewhat smaller. Famine-exposed women were again about 13% ($0.0418 \times 0.012/0.004$) more likely to be disabled, and the corresponding figure for men was 12% ($0.0582 \times 0.012/0.006$). As for housing, the Famine is estimated to reduce the residence size by slightly under 1 square meter ($58.95 \times .012$), with a similar effect for men. For men, illiteracy increased 7% and the likelihood of not working increased 2.4%.

Again, men from high-Famine areas were less likely to be married (3.5%), more likely to never have married (5%), married older (.8 months), and were less likely to head their households (.7%) (Table 5). For women, the point estimates have the expected signs, but are not statistically significant. Finally, Table 6 shows that coefficients for the sex ratio are significant in the expected direction, but roughly one-third the size of the corresponding estimates in Table 3.
Rural versus urban We also estimate the above models separately for those born in rural versus urban regions. We find a Famine effect on the labor and marriage market outcomes for both areas, although the effects for the rural sample were larger (presumably reflecting the greater severity of the Famine in rural areas). For both rural and urban areas, we find that the Famine reduced the sex ratio of the *in utero* cohort and again in the next generation (results available from authors).

Province of residence Finally, we note that estimates reported in Tables 1-6 are essentially unchanged when fixed effects for the 2000 province of residence are included along with the province of birth dummies.

3.4 Potential Biases

As the Famine both raised mortality and reduced fertility, Famine cohorts were approximately 25-50% smaller than neighboring cohorts in the 2000 Census. To the extent that Famine-induced mortality was negatively selective, as would seem most plausible (especially insofar as health is concerned), estimates of damage to survivors are downward biased.

Negative selection into fertility is a greater potential concern, since this could generate the appearance of effects absent any true damage (i.e., upward bias). However, historical evidence suggests that the Famine, unlike the subsequent Cultural Revolution, hit poorer individuals the hardest (see, e.g., Cai & Feng (2005)). The Dutch Famine provides further evidence: fathers of children conceived in the winter of 1944-45 were more likely to have non-manual occupations (Stein *et al.*, 1975).

Direct evidence on selection into fertility is available from the China Fer-
tility surveys (conducted in 1985 and 1988), which include information on the respondent’s mother’s educational attainment (further information in the Appendix). Plotting the share of women whose mothers had no education, primary or less, secondary or more, or who did not know their mother’s education, the 1959-61 birth cohorts do not appear any worse than adjacent cohorts (Figure 6). If anything, maternal education for the 1959-61 birth cohorts was better than for adjacent cohorts.

Cohorts born after the Famine may constitute a better control group than those born in the 1950s (who were exposed to higher mortality rates and malnutrition in childhood). Re-estimating equations (1) and (2) on the sample restricted to birth cohorts 1959-1964, we obtain similar, if not slightly stronger, results (available on request).

Another possible source of bias is that those born during famines may be born to more fecund women or parents who favor offspring quantity over quality. Whereas we cannot control for parental preferences (other than note, as above, that the maternal education of the Famine cohorts was if anything better than that of adjacent cohorts), we can investigate sibship size using a recent survey: The 2005 Urban Chinese Education and Labor Survey conducted by the Ministry of Education in 12 cities in China, covering some 10,000 households.\(^\text{13}\) The 1959-61 cohorts do not appear to have more siblings (Figure 7). Rather, these birth cohorts are on a negative trend (linear and decreasing in year of birth).\(^\text{14}\)

\(^\text{13}\)The 2000 census does not have information on sibship size. Neither can it be inferred from the relationship variable for a household, since most adult siblings live in different households. Finally, the earliest publicly available Chinese census was conducted in 1982, when the 1959-1961 cohorts were in their early 20s.

\(^\text{14}\)This is confirmed by a regression of sibsize on a dummy for birth cohorts 1959-61, controlling for a linear trend in birth year. The coefficient on this dummy is about zero,
3.5 Birth Outcomes in Hong Kong

A shortcoming of the analysis using the 2000 (mainland) Census is the want of a truly unexposed control group. Hong Kong Natality data offer a potential solution to this problem. Communist China severely restricted out-migration, a policy that was temporarily and dramatically suspended during a six-week period in the spring of 1962 when a large number of mainlanders entered Hong Kong (Burns, 1987). Among the refugees were mainland born children, who themselves show up as parents in the 1984-2004 Hong Kong Natality files. The migration of mainland residents to Hong Kong, during and in the years after the Famine, provides a common environment for those affected by the Famine (mainland immigrants) and those who were not (Hong Kong born).

The Hong Kong Natality microdata allow us to focus on second generation birth outcomes, specifically low birth weight and sex. Low birth weight may be a negative outcome because it is a correlate of poor adult health and economic performance. As for sex of offspring, a daughter may not be a poor outcome. Still, it may signal poor parental condition; see Section 4.

We estimate a modified version of equation (1) separately on the sub samples of mainland born and Hong Kong born mothers giving birth in Hong Kong 1984-2004. That is, among Hong Kong mothers who emigrated from the mainland, we compare the birth outcomes of mothers exposed to the Famine in utero to other mainland emigrants born before or after the Famine. While migrants are clearly a select group, our identifying assumption here is not that migrants are a random sample, but instead that this selection in to migration did not change discontinuously for the cohort of migrants in utero during the Famine.

with a very large standard error (not reported).
Dating famine exposure for migrants require some care. The Hong Kong natality files do not record province of birth for mainland born mothers, rendering the application of year and province level mortality rates impossible. Therefore, we date Famine exposure by the average month of the immigrant cohorts. Month of birth drops dramatically for mainland born mothers born in 1961 (Figures 3 and 4). Consequently, we substitute the dummy variable $I(1961)$, which takes on the value $1/100$ for those born in 1961, for the death rate $(awdr_t)$. A later year for the immigrants to Hong Kong is consistent with the likely geographic selection (more migrants likely from the south, an area that was hit later) and the timing of the migration policy. Again, we do not include month of birth given its apparent endogeneity.

A dummy for the sex of the child is also included when the dependent variable is birth weight since males are on average heavier than females. The birth interval is shifted forward one year from the mainland Census regressions, that is, we focus on births to parents themselves born 1957-1965.\(^{15}\) Furthermore, we restrict the sample to singleton births. We find that mothers born in 1961 were 8% (0.247/0.030) more likely to give birth to a child of low birth weight (less than 2,500 grams) and 1.2% (0.00629/0.52) less likely to give birth to a son than mothers born in adjacent years (Table 7). No significant effects were detected for the Hong Kong born mothers, despite their greater numbers.

\(^{15}\)Clearly, mainland-born mothers born after 1962 could not have been part of the Famine induced wave of immigration in the spring of 1962. It is reassuring that restricting the sample to 1957-1961 strengthens our results (available from authors on request).
4 Summary and Discussion

We use the Chinese Famine 1959-61 as a natural experiment in maternal stress and malnutrition. Despite some 40 years of potential catch up, cohorts exposed in utero registered substantial damage in the 2000 Census. Higher Famine intensity – by virtue of either time or place of birth – was associated with greater risk of being illiterate, out of the labor force, marrying later (men), and marrying spouses with less education (women).

Osmani & Sen (2003) argued that maternal malnutrition “rebounds on the society as a whole in the form of ill-health of their offspring – male and females alike – both as children and as adults.” Despite its importance, the nutritional status of girls continues to lag that of boys (Hesketh et al., 2002). Our results suggest that male-biased nutritional allocations handicap not only future health outcomes, but also future economic outcomes.

Similarly, our findings offer fresh perspective on current health and socioeconomic outcomes among adults, positively correlated at both the individual and national levels (see, e.g., Case et al. (2002); Cutler et al. (2006)). The mechanism behind this “dual relationship” (Smith, 1999) has proved difficult to unravel empirically. Our findings suggest that poor fetal health conditions of the past may be at the nexus of the relationship. Indeed, historical nutritional deprivation in developed countries may also undermine outcomes in cohorts born prior to major nutrition-assistance programs for the poor.

Perhaps the most intriguing finding is that Famine exposure lowered the

---

16 Almond et al. (2007) found improvements in birth outcomes (including birth weight) with the introduction of the Food Stamps Program during the 1960s in the U.S., particularly among Black infants. These cohorts also manifest improved health and educational outcomes in adulthood (Almond & Chay, 2006).
sex ratio of not only the first but also the second generation—prenatally exposed women were themselves more likely to bear daughters. This pro-female effect is all the more noteworthy given the well-documented prevalence of son preference in mainland China. Famine-induced reductions in the sex ratio are consistent with empirical work finding lower sex ratios for unmarried or poorly educated mothers (Almond & Edlund, 2007). While the magnitude of the Famine’s effect on the sex ratio may appear small, it is several times larger than that associated with marital status in U.S. natality data (Almond & Edlund, 2007) and is similar to differences found in survey data between mothers living with a partner around the time of conception and those who were not (Norberg, 2004). Thus, small changes in the sex ratio can reflect large differences in maternal circumstance.

Trivers & Willard (1973) proposed that evolution would favor parental ability to vary the sex ratio of offspring according to condition: parents in poor condition would favor daughters and parents in good condition would favor sons. Their argument was based on the observation that while the average number of offspring to males and females equalizes, the reproductive success of a male offspring tends to be more resource-sensitive. Maternal malnutrition has been observed to correlate with more female births (see, e.g., Andersson & Bergström (1998)). Pathways include heightened rates of male fetal deaths, as was found to be the case during the Dutch famine (Roseboom et al., 2001a). Another possibility is that starvation affects early cell division of male and female embryos differentially (Cameron, 2004). Fetal “predictive adaptive responses” (to use the terminology of Gluckman & Hanson (2004)) set parameters for the adult individual, for instance her height, which means that maternal constraints affect not only her children, but also her daughters’
children.

To our knowledge, ours is the first large scale quasi-experimental evidence of a Trivers-Willard effect in human populations. It is also the first evidence (quasi-experimental or otherwise) of an intergenerational “echo-effect” of maternal status on the sex ratio (to our knowledge). Low offspring sex ratios in two generations underscore the long term impact of maternal health.
References


Appendix

Variable Definitions:

Census 2000

\textbf{wdr} Weighted death rate for the gestation period, assuming 9 month gestation, and province of birth. For example, a person born in January 1960 in Beijing is assigned $1/9$th of Beijing’s 1960’s mortality rate and $8/9$th of Beijing’s 1959’s mortality rate.

\textbf{awdr} Aggregate weighted average death rate, the \textbf{wdr} collapsed by month and birth year. Thus, it is the population weighted mean of \textbf{wdr} by month and year of birth.

\textbf{mean} Mean of dependent variable.

\textbf{Province} The province of birth. Our results are robust to inclusion of dummies for province of residence.

\textbf{Illiterate} Dummy indicating that the respondent was either illiterate or semi-literate.

\textbf{Don’t work} Dummy indicating that the person did not work for more than 1 hour between October 25 and October 31 (in 2000). This includes those who are on leave from a job, as well as non-workers.

\textbf{On leave from job} Not working because on leave, training, or seasonal lay-off.

\textbf{Supported by other HH members/Dependent} Main income source was support by other household members.

\textbf{Disabled} Dummy indicating that the person does not work because he/she has “lost ability to work.”

\textbf{House area} Area of home, in square meters.

\textbf{Unmarried} Dummy indicating that the respondent was unmarried at the time of the census.

\textbf{Never married} Dummy indicating that the respondent had never married.

\textbf{Spousal education} Includes head-spouse couples only. Education is in years.

\textbf{Marriage age} Age in months at time of first marriage.

\textbf{Household head} Dummy indicating that the respondent was household head. Includes only respondents living in “family type” households (as opposed to “collectives”).
Male  Dummy indicating that the respondent is male.

Sons/Kids  Fraction sons among ever borne children. Excludes women who had not borne any children.

No kid  Dummy indicating that the woman had borne no children.

Child mortality  Number of children ever borne minus number of surviving children (at the time of the census) divided by the number of children ever borne, by year and quarter of birth of mother.

Hong Kong Natality data

I(1961)  Dummy indicating that the mother was born in 1961, scaled by 1/100.

Low BWT  Low birth weight. Dummy indicating that child weighed less than 2,500 grams at birth.

China  Dummy for whether born in mainland China.

China Fertility Surveys

The China Fertility surveys were carried out in 1985 and 1987 in the following provinces: Hebei, Shaanxi, Liaoning, Guangdong, Guizhou, Gansu; and the municipalities of Beijing and Shanghai. (We have not been able to access data for Shandong.) In total, some 46,000 ever-married women between 15-49 years of age were interviewed, providing detailed information on pregnancy history. These data are available from the Office of Population Research, Princeton University, http://opr.princeton.edu/Archive/cdfs/.
Table 1: **2000 Census: Labor and Housing Outcomes for 1956-1964**

<table>
<thead>
<tr>
<th>Birth Cohorts</th>
<th>Illiterate</th>
<th>Don’t work</th>
<th>Disabled</th>
<th>Dependent</th>
<th>House area</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Women</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>0.081</td>
<td>0.189</td>
<td>0.004</td>
<td>0.119</td>
<td>87.162</td>
</tr>
<tr>
<td><strong>awdr</strong></td>
<td>0.5052**</td>
<td>0.4714***</td>
<td>0.0448*</td>
<td>0.3972***</td>
<td>-220.1528***</td>
</tr>
<tr>
<td></td>
<td>[0.2169]</td>
<td>[0.1530]</td>
<td>[0.0250]</td>
<td>[0.1354]</td>
<td>[48.4753]</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>786156</td>
<td>786156</td>
<td>786156</td>
<td>786156</td>
<td>772260</td>
</tr>
<tr>
<td><strong>Men</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>0.021</td>
<td>0.082</td>
<td>0.006</td>
<td>0.019</td>
<td>83.933</td>
</tr>
<tr>
<td><strong>awdr</strong></td>
<td>0.1585*</td>
<td>0.4017***</td>
<td>0.0657</td>
<td>0.1399**</td>
<td>-104.7566**</td>
</tr>
<tr>
<td></td>
<td>[0.0784]</td>
<td>[0.1131]</td>
<td>[0.0426]</td>
<td>[0.0674]</td>
<td>[38.3963]</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>818103</td>
<td>818103</td>
<td>818103</td>
<td>818103</td>
<td>790342</td>
</tr>
</tbody>
</table>

**awdr** – aggregate weighted death rate by birth year and month.

**mean** – mean of dependent variable.

Standard errors clustered at province of birth in square brackets. * significant at 10%; ** significant at 5%; *** significant at 1%.
Table 2: **2000 Census, Marriage Market Outcomes, 1956-1964 Birth Cohorts**

<table>
<thead>
<tr>
<th></th>
<th>Unmarried</th>
<th>Never married</th>
<th>Spousal ed.(^a)</th>
<th>Marriage age (^b)</th>
<th>Household head (^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Women</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>0.061</td>
<td>0.004</td>
<td>9.057</td>
<td>269.237</td>
<td>0.118</td>
</tr>
<tr>
<td>\texttt{awdr}</td>
<td>0.2608</td>
<td>-0.0013</td>
<td>-6.3342**</td>
<td>67.4994**</td>
<td>-0.0998</td>
</tr>
<tr>
<td></td>
<td>[0.1632]</td>
<td>[0.0249]</td>
<td>[2.4652]</td>
<td>[28.5417]</td>
<td>[0.1633]</td>
</tr>
<tr>
<td>\texttt{N}</td>
<td>786156</td>
<td>786156</td>
<td>685989</td>
<td>783015</td>
<td>786156</td>
</tr>
<tr>
<td><strong>Men</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>0.090</td>
<td>0.039</td>
<td>8.060</td>
<td>290.898</td>
<td>0.870</td>
</tr>
<tr>
<td>\texttt{awdr}</td>
<td>0.4902***</td>
<td>0.2676**</td>
<td>-0.1692</td>
<td>125.1309***</td>
<td>-0.5145**</td>
</tr>
<tr>
<td></td>
<td>[0.1285]</td>
<td>[0.1035]</td>
<td>[2.5349]</td>
<td>[28.5395]</td>
<td>[0.2302]</td>
</tr>
<tr>
<td>\texttt{N}</td>
<td>818103</td>
<td>818103</td>
<td>683041</td>
<td>785927</td>
<td>818103</td>
</tr>
</tbody>
</table>

\texttt{awdr} – aggregate weighted death rate by birth year and month.

mean – mean of dependent variable.

\(^a\) Includes head-spouse couples only. Education is in years.

\(^b\) Marriage age is in months.

\(^c\) Includes those residing in family units (i.e., excludes those residing in collectives).

Standard errors clustered at province of birth in square brackets. * significant at 10%; ** significant at 5%; *** significant at 1%.
Table 3: **2000 Census: Sex Ratio outcomes, 1956-1964 Birth Cohorts**

<table>
<thead>
<tr>
<th></th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male(^a)</td>
</tr>
<tr>
<td>mean</td>
<td>0.51</td>
</tr>
<tr>
<td>awdr</td>
<td>-1.3147***</td>
</tr>
<tr>
<td></td>
<td>[0.2651]</td>
</tr>
<tr>
<td>N</td>
<td>1604259</td>
</tr>
</tbody>
</table>

awdr – aggregate weighted death rate by birth year and month.

mean – mean of dependent variable.

\(^a\) Dummy – equals 1 if respondent was male.

\(^b\) Pertains to children borne.

Standard errors clustered at province of birth in square brackets. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 4: **2000 Census: Cross-sectional Variation in Famine Severity, Labor Market and Housing Outcomes, 1956-1964 Birth Cohorts**

<table>
<thead>
<tr>
<th></th>
<th>Illiterate</th>
<th>Don’t work</th>
<th>Disabled</th>
<th>Dependent</th>
<th>House area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>wdr</td>
<td>0.1659</td>
<td>0.0953</td>
<td>0.0418***</td>
<td>0.0755</td>
<td>-58.9501**</td>
</tr>
<tr>
<td></td>
<td>[0.1269]</td>
<td>[0.1657]</td>
<td>[0.0116]</td>
<td>[0.0917]</td>
<td>[22.0095]</td>
</tr>
<tr>
<td>N</td>
<td>764786</td>
<td>764786</td>
<td>764786</td>
<td>764786</td>
<td>751352</td>
</tr>
</tbody>
</table>

|       |            |            |          |           |            |
| Men   |            |            |          |           |            |
| wdr   | 0.1231*    | 0.1628**   | 0.0585***| 0.0321    | -52.1040*  |
|       | [0.0688]   | [0.0666]   | [0.0170] | [0.0376]  | [28.5949]  |
| N     | 795408     | 795408     | 795408   | 795408    | 768522     |

Standard errors clustered at province of birth in square brackets. * significant at 10%; ** significant at 5%; *** significant at 1%.
Table 5: **2000 Census: Cross-sectional Variation in Famine Severity, Marriage Market Outcomes, 1956-1964 Birth Cohorts**

<table>
<thead>
<tr>
<th>Unmarried</th>
<th>Never married</th>
<th>Spousal ed. $^a$</th>
<th>Marriage age $^b$</th>
<th>Household head $^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Women</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>wdr</td>
<td>0.0505</td>
<td>0.0217</td>
<td>0.0794</td>
<td>14.7224</td>
</tr>
<tr>
<td></td>
<td>[0.0623]</td>
<td>[0.0130]</td>
<td>[1.5906]</td>
<td>[19.9378]</td>
</tr>
<tr>
<td>N</td>
<td>764786</td>
<td>764786</td>
<td>668672</td>
<td>761879</td>
</tr>
<tr>
<td><strong>Men</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>wdr</td>
<td>0.2666***</td>
<td>0.1555**</td>
<td>1.5938</td>
<td>67.6296***</td>
</tr>
<tr>
<td></td>
<td>[0.0696]</td>
<td>[0.0634]</td>
<td>[1.3770]</td>
<td>[22.8696]</td>
</tr>
<tr>
<td>N</td>
<td>795408</td>
<td>795408</td>
<td>665857</td>
<td>764670</td>
</tr>
</tbody>
</table>

**wdr** Weighted average death rate for the gestation period, assuming 9 month gestation. Varies by province and month and year of birth.

$^a$ Includes head-spouse couples only. Education is in years.

$^b$ Marriage age is in months. $^c$ Includes those residing in family units (i.e., excludes those residing in collectives).

Standard errors clustered at province of birth in square brackets. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 6: **2000 Census: Cross-sectional Variation in Famine Severity, Sex ratio outcomes, 1956-1964 Birth Cohorts**

<table>
<thead>
<tr>
<th></th>
<th>Male $^a$</th>
<th>Sons/Kids</th>
<th>No child</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Women $^b$</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>wdr</td>
<td>-0.3264**</td>
<td>-0.1693**</td>
<td>0.0325</td>
</tr>
<tr>
<td></td>
<td>[0.1390]</td>
<td>[0.0797]</td>
<td>[0.0251]</td>
</tr>
<tr>
<td>N</td>
<td>1560194</td>
<td>752418</td>
<td>764786</td>
</tr>
</tbody>
</table>

$^a$ Dummy – equals 1 if respondent was male.

$^b$ Pertains to children borne.

Standard errors clustered at province of birth in square brackets. * significant at 10%; ** significant at 5%; *** significant at 1%.
Table 7: 1984-2004 Natality outcomes in Hong Kong: mainland vs Hong Kong born mothers

<table>
<thead>
<tr>
<th>Mother born:</th>
<th>Mainland</th>
<th></th>
<th></th>
<th>Hong Kong</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low BWT&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Son&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td>Low BWT&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Son&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>0.031</td>
<td>0.52</td>
<td></td>
<td>0.039</td>
<td>0.517</td>
<td></td>
</tr>
<tr>
<td>I(1961)</td>
<td>0.247**</td>
<td>-0.629***</td>
<td></td>
<td>0.014</td>
<td>-0.009</td>
<td></td>
</tr>
<tr>
<td>[0.099]</td>
<td>[0.121]</td>
<td></td>
<td>[0.037]</td>
<td>[0.074]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>198452</td>
<td>198452</td>
<td></td>
<td>393419</td>
<td>393419</td>
<td></td>
</tr>
</tbody>
</table>

mean – mean of dependent variable.
a – dummy, equals 1 if birth weight was less than 2,500 grams.
b – dummy, equals 1 if child male.
Regression results from estimating equation 1 where I(1961) substitutes for awdr and without the province dummies. The birth weight regressions also include a dummy for the sex of the child.
Standard errors clustered by year of birth in square brackets. * significant at 10%; ** significant at 5%; *** significant at 1%.
Figure 1: Aggregate Weighted Death Rate by Year and Month of Birth, Mainland China

Note: Authors’ calculations based on all age death rates by year and province as reported by China Statistical Press (2000).

Figure 2: Average Month of Birth, Mainland China

Source: 2000 census.
Figure 3: Average month of birth, Hong Kong Mothers born in Mainland

Figure 4: Average month of birth, Hong Kong Mothers born in Hong Kong

Source: Hong Kong Natality microdata.
Figure 5: Census outcomes by year and quarter of birth

Source: 2000 census.
Figure 6: Mother’s education by child’s (respondent) year of birth

Source: China Fertility surveys 1985/87.

Note: For mother’s education unknown, the universe is all respondents. For the remainder, the universe is those who knew their mother’s education.
Figure 7: Number of siblings by respondent’s year of birth