# Web Appendix to Missing Women: Age and Disease, by Siwan Anderson and Debraj Ray

An objective of our paper is to assess missing women by disease, as well provide an estimate of the number of missing women attributable to the changing composition of disease. The importance of this question cannot be overstated. Indeed, if disease composition does play a significant role, it becomes difficult to describe such women as missing: the "similar care" counterfactual does not apply, and explanations based on "discrimination" cannot be easily invoked in those circumstances. Conversely, if most excess female deaths are explicable on a disease-by-disease basis (and not a change in the disease mix), then it becomes that much harder — though not impossible, as discussed in the main paper — to reject the discrimination viewpoint.

This appendix provides some background on how the disease composition typically changes with development and the relative disease patterns in our countries of interest. It also provides additional evidence to support our claim that the epidemiological transition plays a minor role in accounting for the missing women in developing countries.

## 1. The Epidemiological Transition

The *epidemiological transition* (Omran (1971)) refers to a transformation in the cause composition of mortality by which acute infectious and deficiency diseases are gradually replaced by chronic and degenerative diseases as the leading causes of death. Briefly, infectious disease and nutritional or reproductive health problems predominate in poorer, high-mortality populations, while chronic and degenerative ailments predominate in low-mortality populations. The implicit linearity in this definition is likely too simplistic: several infectious diseases, such as tuberculosis, can certainly make a reappearance at "advanced" stages of economic development. Moreover, entirely new infectious diseases, such as AIDS, can also appear.

Historically, improvements in public health (from the mid-nineteenth century onwards) began in Western Europe and the United States with the major breakthroughs of the germ theory of disease, and with the discovery of penicillin and other vaccines. This led to a substantial decline in mortality. Up to around 1940 most of this progress was confined to rich countries, but the benefits began to percolate thereafter to all countries. The resulting demographic transition is, of course, well-known. What we want to focus on here is the associated *compositional* change in the nature of disease.

## 2. India, China and sub-Saharan Africa

Table 1 describes the distribution of death by disease (for the year 2000) in the three developing regions of interest. There is clear evidence for the various phases that describe the epidemiological transition.

Sub-Saharan Africa exhibits all the pre-transition characteristics, with Group 1 diseases heading the list (bolstered by HIV/AIDS). India, China and the group of developing countries we consider furnish more snapshots of the transition on the cross-section, with China significantly closer to the developed regions than it is to either India or sub Saharan Africa.

To be sure, the process is far from entirely linear. For instance, it is true that in many ways,

| Disease | sSAfrica | India | China | Dev |
|---|---|---|---|---|
| **1. Communicable, Maternal, Perinatal, Nutritional** | **72** | **40** | **12** | **7** |
| A. Infectious and parasitic | 53 | 20 | 5 | 2 |
| *1. Tuberculosis* | *3* | *4* | *3* | *0* |
| *2. HIV/AIDS* | *20* | *3* | *0* | *0* |
| *3. Diarrhoeal* | *7* | *4* | *1* | *0* |
| *4. Childhood Clusters* | *7* | *2* | *0* | *0* |
| *5. Malaria* | *10* | *0* | *0* | *0* |
| B. Respiratory Infections | 10 | 10 | 3 | 5 |
| C. Maternal | 2 | 1 | 0 | 0 |
| D. Perinatal | 5 | 7 | 3 | 0 |
| **2. Noncommunicable** | **21** | **50** | **77** | **87** |
| A. Malignant neoplasms | 4 | 7 | 19 | 26 |
| B. Cardiovascular | 10 | 27 | 33 | 38 |
| C. Respiratory | 2 | 6 | 16 | 6 |
| D. Digestive | 2 | 3 | 4 | 4 |
| **3. Injuries** | **7** | **10** | **11** | **6** |

TABLE 1. DEATHS BY DISEASE, 2000 (%). Source: *Global Burden of Disease (2002).*

China has almost completed its epidemiological transition. The pre-revolutionary health situation was very poor, and there have been remarkable improvements since then. By the 1990s, life expectancy had doubled. Indeed, during the Deng era 1960–1980, China moved through the late stages of the transition utilising advances in public health care and medicine (Cook and Dummer (2004)). Yet it is possible to argue that China's rapid economic growth in the last two decades has come at a heavy environmental price, with serious implications for the incidence of respiratory disease. This is apparent in the table above where close to 20% of deaths in China may be attributed to respiratory problems. It is a safe prediction that India will exhibit no different a pattern as it continues to urbanize.[1]

Nevertheless, the main text argues that the epidemiological transition plays a minor role in accounting for missing females in developing countries. As we shall now see, information on gender-based death rates by disease is consistent with this argument.

[1]In a similar vein, cardiovascular disease is quickly becoming a burden in *developing* countries even before they have rid themselves of infectious diseases (Reddy and Yusuf (1998)). This is evident from Table 1 in the case of India, where heart disease far outstrips cancer as a cause of death compared to the developed world. Improvements in nutrition and health status as well as the successful eradication of major killer diseases have contributed to the ongoing epidemiological transition in India (Gupte et. al. (2001)). Nevertheless, as observed by the World Health Organization, the burden of disease as measured by "premature death" there is second only to sub-Saharan Africa.

| Region | Early Childhood, 0–4 | | | All Ages | | |
|---|---|---|---|---|---|---|
| | Male | Female | M/F | Male | Female | M/F |
| Developed, Gr. 1 | 0.69 | 0.55 | **1.25** | 0.60 | 0.62 | **0.97** |
| China, Gr. 1 | 4.59 | 6.56 | **0.70** | 0.85 | 0.79 | **1.07** |
| India, Gr. 1 | 16.77 | 17.99 | **0.93** | 4.01 | 3.68 | **1.09** |
| sub-S Africa, Gr. 1 | 39.44 | 36.47 | **1.08** | 11.88 | 10.93 | **1.09** |
| Developed, Gr. 2 | 0.50 | 0.44 | **1.30** | 7.61 | 7.46 | **1.02** |
| China, Gr. 2 | 1.44 | 1.65 | **0.87** | 5.41 | 5.42 | **1.00** |
| India, Gr. 2 | 1.36 | 1.85 | **0.73** | 4.93 | 4.80 | **1.03** |
| sub-S Africa, Gr. 2 | 1.12 | 0.95 | **1.19** | 3.39 | 3.46 | **0.98** |

TABLE 2. DEATH RATES PER THOUSAND BY DISEASE GROUP, 2000. Source: *Global Burden of Disease* (2002). Note: Groups 1 and 2 as described in Table 1, except that maternal mortality is excluded.

Consider a simple breakdown into Group 1 and Group 2 diseases. Table 2 lists death rates (and relative death rates) by gender under these two headings for the regions of interest: the developed countries, China, India and sub-Saharan Africa. The definitions follow exactly the description in Table 1, except that maternal mortality is excluded from Group 1.

The figures in boldface are death rates for males relative to females; all other figures in the table are death rates per thousand. Of course, Table 2 agrees with Table 1 in that the death rates for Group 1 diseases increase dramatically as we move from the developed regions, via China and India, to sub-Saharan Africa. (There isn't comparable variation for Group 2 diseases, which is also to be expected.) What is striking, however, are the relative death rates by gender. The last column reports overall mortality rates. The developed countries do, indeed, conform to the view that male-female relative death rates are higher for Group 2 diseases. But *none* of the developing regions follow suit. In fairly sharp contrast, males appear to die relatively *more*, rather than less, within the group of communicable diseases.

A comparison of early childhood mortality (the first of the two boldface colums) is somewhat more in line with the compositional hypothesis. Except for India where the expected rankings continue to be reversed, female death rates are relatively higher for Group 1 diseases.

In any case, this compositional effect across disease groups is comprehensively swamped by the fact that male death rates are so much higher (relative to that for females) in richer countries. Compare, for instance, the early-childhood relative mortality rates (for either group of diseases) in India and China with the same rates for developed regions. Or simply look at the overall figures in the last column for developed regions, and compare these to any of the other figures in that column. *This*, and not a comparison of Group 1 and Group 2 diseases, is the dominant feature of Table 2.

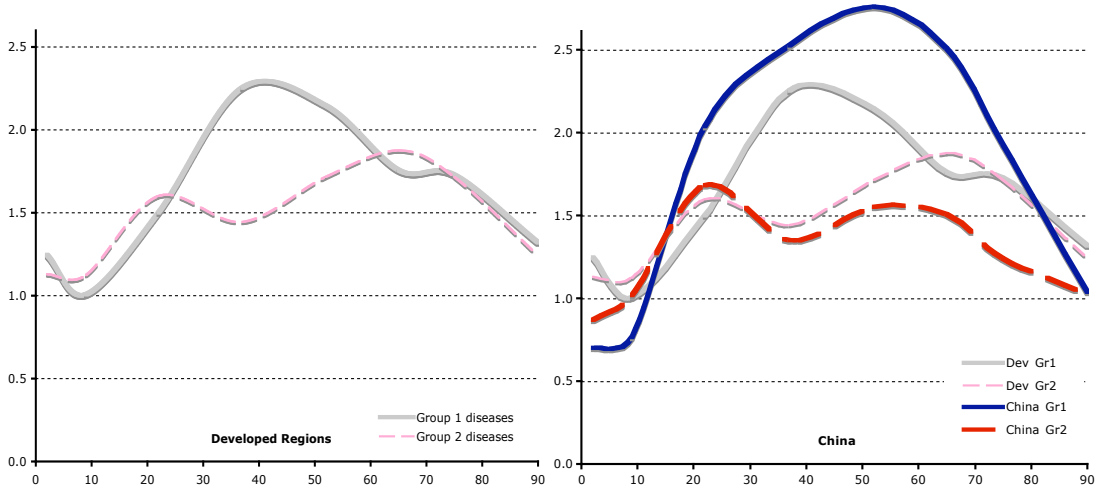


FIGURE 1. MALE-FEMALE DEATH RATIOS BY AGE FOR GROUP 1 AND GROUP 2 DISEASES; DEVELOPED REGIONS AND CHINA. Source: *Global Burden of Disease*. Note: Maternal mortality is excluded from Group 1.

Figures 1 and 2 reinforce this point in the starkest possible way. The figures compare relative death rates for diseases in Groups 1 and 2 across the various regions of interest, for a variety of ages. The first panel of Figure 1 shows relative death rates at different ages for developed regions; in this (as in all panels) the solid line refers to Group 1 diseases, excluding maternal mortality, while the dotted line refers to Group 2 diseases. As already discussed, there is little to choose between the two groups as far as relative mortality is concerned. Moreover, both the lines lie above the 50-50 mark (a ratio of 1.0 in the diagram), displaying the familiar bias in favor of male mortality at all ages.

The second panel, as well as the two panels in Figure 2, conduct exactly the same exercise for China, India, and sub-Saharan Africa. In each of these panels, we've reproduced the graphs for developing regions as pale lines for easy comparison.

At least for India and sub-Saharan Africa, the results speak for themselves. For each disease group, the graph representing relative male mortality rates by age lies almost uniformly below its counterpart for the developed regions. At the same time, in line with Table 2, there is little to choose *across* the graphs for disease groups in any one region. Put another way, these diagrams reinforce our suspicion that the bulk of missing women are due to relative mortality differences *disease by disease*, and not to a change in the composition of disease.

The China panel in Figure 1 is somewhat misleading and requires interpretation. In line with the findings for India and sub-Saharan Africa, the graph for Group 2 diseases lies uniformly below its counterpart for developed regions. But the corresponding graph for Group 1 does not. It appears that in China, and roughly from adolescence onwards, males die more than females do from Group 1 diseases, relative to developed countries. But this

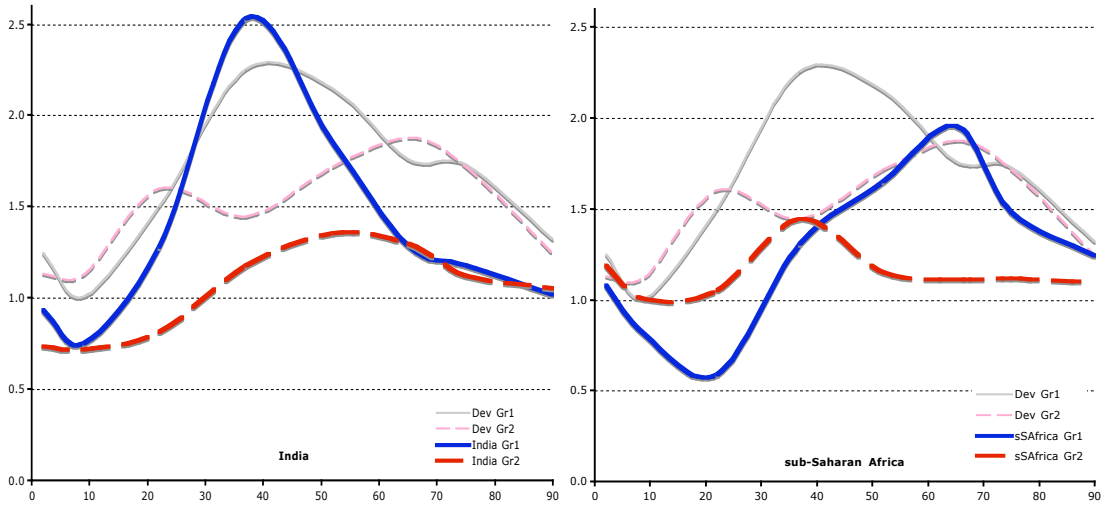


FIGURE 2. MALE-FEMALE DEATH RATIOS BY AGE FOR GROUP 1 AND GROUP 2 DISEASES; INDIA AND SUB-SAHARAN AFRICA. Source: *Global Burden of Disease*. Note: Maternal mortality is excluded from Group 1.

reversal is relatively unimportant: in early childhood and in old age, where the bulk of Group 1 deaths lie, the relationship is entirely the same as in India and sub-Saharan Africa, and indeed, Table 2 confirms that the *overall* average of relative death rates from Group 1 for China is almost the same as for India and sub-Saharan Africa.

## 3. THE UNITED STATES IN 1900

Table 3 augments Table 1 by including the United States in 1900. The table also adds other indicators such as life expectancy. As is evident from the table, the United States in 1900 fits in rather snugly somewhere between sub-Saharan Africa and India on a number of indicators, and in particular is fully in line with the epidemiological transition.

The epidemiological transition — in the cross section — leaves a clear footprint through the table. We see that the two leading causes of death in the United States in 1900 were infectious disease (primarily tuberculosis) and respiratory infections. Indeed, we know that death rates from heart disease, tuberculosis, and pneumonia were highly comparable between 1910 and 1917. Then for a short period, 1918 to 1920, pneumonia returned as the main cause of death, after which heart disease became the leading killer, a trend that persists to this day. Between 1900 and 1940 in the United States, the rate of infectious disease and respiratory infection both decreased fourfold, maternal deaths and infancy-related disease both decreased twofold, whereas the death rates from cancer and heart disease both more than doubled.

As we did in the case of the developing regions, we must now ask whether this transition might account for a sizable proportion of missing women in 1900 United States. The main text tentatively answers in the affirmative. The following analysis backs up this assertion.

| Disease Deaths (%) / Other Indicators | sSA | US 1900 | India | China | Dev |
|---|---|---|---|---|---|
| **Group 1** | **72** | **53** | **40** | **12** | **7** |
| A. Infectious and parasitic | 53 | 33 | 20 | 5 | 2 |
| *1. Tuberculosis* | *3* | *11* | *4* | *3* | *0* |
| *2. HIV/AIDS* | *20* | *0* | *3* | *0* | *0* |
| *3. Diarrhoeal* | *7* | *11* | *4* | *1* | *0* |
| *4. Childhood Clusters* | *7* | *5* | *2* | *0* | *0* |
| *5. Malaria* | *10* | *1* | *0* | *0* | *0* |
| B. Respiratory Infections | 10 | 15 | 10 | 3 | 5 |
| C. Maternal | 2 | 1 | 1 | 0 | 0 |
| D. Perinatal | 5 | 2 | 7 | 3 | 0 |
| **Group 2** | **21** | **35** | **50** | **77** | **87** |
| A. Malignant neoplasms | 4 | 3 | 7 | 19 | 26 |
| B. Cardiovascular | 10 | 11 | 27 | 33 | 38 |
| C. Respiratory | 2 | 1 | 6 | 16 | 6 |
| D. Digestive | 2 | 4 | 3 | 4 | 4 |
| E. Neuropsychiatric | 1 | 7 | 2 | 1 | 5 |
| **Injuries** | **7** | **6** | **10** | **11** | **6** |
| Male Life Expectancy (yrs) | 48 | 48 | 62 | 71 | 72 |
| Female Life Expectancy (yrs) | 50 | 51 | 64 | 74 | 79 |
| Overall Death Rate (per 1000) | 16 | 17 | 7 | 9 | 11 |

TABLE 3. DEATHS BY DISEASE (%) AND MISCELLANEOUS INDICATORS, THREE DEVELOPING REGIONS (2000) AND THE UNITED STATES (1900). Source: Global Burden of Disease (2002) and the Historical Vital Statistics of the United States.

Figure 3 is the analogue of Figures 1 and 2. For easy reference, the figure recalls relative male-female death rates in developed countries for both Groups 1 and 2. The two corresponding plots for the United States in 1900 lie significantly below their developed-country counterparts, just as in the case of modern India and sub-Saharan Africa.[2] That suggests a large disease-by-disease effect, with females dying relatively quickly at all age groups and for both groups of disease, and this effect is line with what we've seen earlier. What *is* different, however, is that there also appears to be a compositional effect at both low age groups (under 20) and high age groups (above 60) that is eminently in line with the hypothesis suggested by the epidemiological transition. In particular, in *both* these age ranges (which are the age ranges that really matter for death anyway), the male-female relative death rates are lower for Group 1 diseases, compared to Group 2 diseases. A perusal of Figures 1 and 2 reveals that a similar compositional effect is to be found at the very youngest age ranges for China, and at young ages for sub-Saharan Africa (though not

[2]The same is also true of China, except for the middle age segment of Group 1 in which overall death rates are very low anyway.

anywhere for India), but the historical United States displays this effect at both low and high ages.[3]

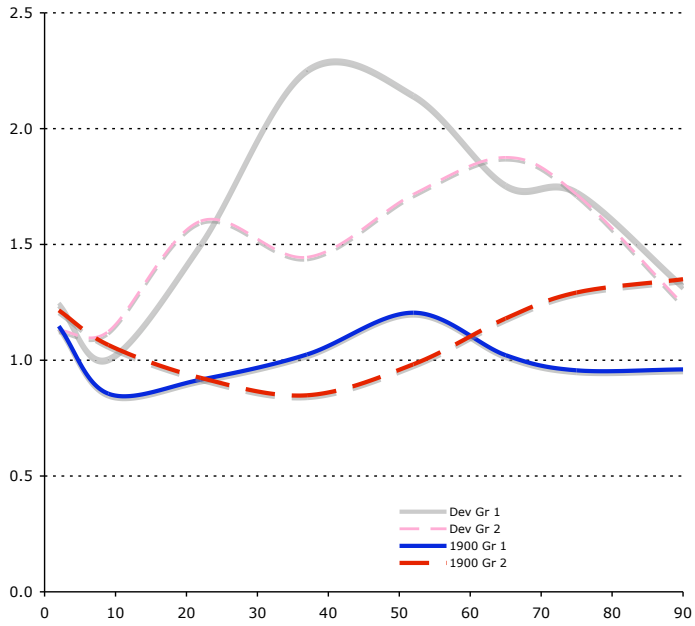


FIGURE 3. MALE-FEMALE DEATH RATIOS BY AGE FOR GROUP 1 AND GROUP 2 DISEASES; THE UNITED STATES, 1900, AND DEVELOPED REGIONS, 2000. Source: *Global Burden of Disease* and Historical Vital Statistics of the United States. Note: Maternal mortality is excluded from Group 1.

To reinforce this point, we also present an extended form of Table 2, now expanded to include the United States in 1900. This is Table 4. Just as we've seen before, the historical United States settles in quite comfortably between India and sub-Saharan Africa, if we use the overall death rate, whether it is for Group 1 or Group 2.[4]

Focus on the last column, which lists relative male-female mortality rates (over all ages) for Group 1 and Group 2 diseases. As discussed in the previous section, each of our developing regions contradicts the idea that males and females are more equally afflicted by Group 1 diseases. But the information we have for the United States in 1900 proclaims it to be quite different: we do indeed see the hypothesized pattern, with death ratios equal to 1.05 under Group 1 and significantly higher — 1.15 — under Group 2. That pattern is also robustly

[3]Moreover, in the case of the sub-Saharan Africa the bulk of this effect is due to HIV/AIDs. This is *not*, however, a disease that has "disappeared" over the epidemiological transition, so strictly speaking it cannot be attributed to a compositional effect.

[4]An outlier in the overall pattern, however, is the 1900-US rate of child mortality from Group 2 disease, which is way too high relative to what we see today in developing countries. This is almost surely a matter of classification and/or inadequate diagnosis. It turns out that childhood deaths under the "neuropsychiatric" heading account for close to 50% of all deaths in the 0–4 category. This is enormous. The corresponding number for, say, modern India is under 5%. It is clear that several perinatal or respiratory conditions leading to sudden death could easily have been clumped under the neuropsychiatric category.

| Region | Early Childhood, 0–4 | | | All Ages | | |
|---|---|---|---|---|---|---|
| | Male | Female | M/F | Male | Female | M/F |
| Developed, Gr. 1 | 0.69 | 0.55 | **1.25** | 0.60 | 0.62 | **0.97** |
| China, Gr. 1 | 4.59 | 6.56 | **0.70** | 0.85 | 0.79 | **1.07** |
| India, Gr. 1 | 16.77 | 17.99 | **0.93** | 4.01 | 3.68 | **1.09** |
| US 1900, Gr. 1 | 25.39 | 22.25 | **1.14** | 7.17 | 6.84 | **1.05** |
| sub-S Africa, Gr. 1 | 39.44 | 36.47 | **1.08** | 11.88 | 10.93 | **1.09** |
| | | | | | | |
| Developed, Gr. 2 | 0.50 | 0.44 | **1.30** | 7.61 | 7.46 | **1.02** |
| China, Gr. 2 | 1.44 | 1.65 | **0.87** | 5.41 | 5.42 | **1.00** |
| India, Gr. 2 | 1.36 | 1.85 | **0.73** | 4.93 | 4.80 | **1.03** |
| US 1900, Gr. 2 | 7.69 | 6.33 | **1.21** | 5.13 | 4.45 | **1.15** |
| sub-S Africa, Gr. 2 | 1.12 | 0.95 | **1.19** | 3.39 | 3.46 | **0.98** |

TABLE 4. DEATH RATES PER THOUSAND BY DISEASE GROUP, 2000. Source: *Global Burden of Disease* (2002) and Historical Vital Statistics of the United States. Note: Groups 1 and 2 as described in Table 1, except that maternal mortality is excluded.

maintained in the early childhood category (the ratios are 1.14 and 1.21 for Groups 1 and 2 respectively). As we have noted earlier, China and sub-Saharan Africa also display this pattern in the 0–4 age category. However, the historical United States is the only one of our four regions of interest that unambiguously exhibits the common wisdom regarding the epidemiological transition, both in childhood and across all age categories. In this single respect, the United States is unlike our three developing regions and is more like developed regions today.

In particular, this last finding is supportive of the suggestion in the main text that there is a significant effect of disease composition in accounting for missing women in the historical United States.

## REFERENCES


[1] Cook, I. and T. Dummer (2004) "Changing Health in China: Re-Evaluating the Epidemiological Transition Model" *Health Policy* **67**, 329–343.

[2] Gupte, M D, V. Ramachandran, and R K Mutatkar (2001) "Epidemiological profile of India: Historical and contemporary perspectives" *Journal of Bioscience*, 26(4), 437-464.

[3] Omran Abdel (1971), "The Epidemiological Transition: A Theory of the Epidemiology of Population Change," *Milbank Quarterly* **49**, 509–537.

[4] Reddy, K S and S. Yusuf (1998) "Emerging epidemic of cardiovascular disease in developing countries" *Circulation*, 97, 596-601.