CrossMark

Morbidity in India since 1944

**T. N. Srinivasan[1,4] · V. R. Muraleedharan[2] ·
Bhanu Pratap[3]**

Abstract Surveys in countries at all stages of development have founded their work on health-status and morbidity, on self-reported health status by individual members of households who feel sick. Doubts have been raised related to cross-population comparisons on the objectivity of a person's judgement of his/her health. Amartya Sen (Objectivity and position, University of Kansas, Department of Philosophy, Kansas, 1992, Philos Public Affair 126–145, 1993) has written on the philosophy of objectivity and, in Sen (Br Med J 324:860, 2002), compared morbidity data across Indian States, and countries like the United States. His discussion helps formulating and testing a null hypothesis that an Individual's self-reported health-status (SRH) and morbidity (SRM) do not depend on his/her socio-economic status (SES) as well the socio-economic environment in which he/she lives. The test rejects the null hypothesis in favour of an alternative that there is a positive association between the two using data from the 71st Round (January–June 2014) survey of the National Sample Survey Office (NSSO). This means that lower the SES, the lower will be the health-status (reported as having higher morbidity); the higher the SES, higher will be the health-status (reported as having low morbidity). We also explore a linear probability model with constraints on the error term for ensuring that the estimated probabilities lie within the closed unit interval [0, 1].

Keywords Development · Health · Self-reported health and morbidity · Logistic regression · Linear probability model

✉ T. N. Srinivasan
t.srinivasan@yale.edu

[1] Yale University, New Haven, CT, USA

[2] Indian Institute of Technology, Madras, Chennai, India

[3] Madras School of Economics, Chennai, India

[4] IFMR, 24 - Kothari Road, Nungambakkam, Chennai 600034, India

1 Introduction

One of the important indicators of development of a nation is progressive improvement in the health status of its population. Nobel laureate Angus Deaton uses the term 'wellbeing' to refer to all the things that are good for a person that makes for a good life. Wellbeing includes material wellbeing, such as income and wealth; physical and psychological wellbeing, represented by health and happiness; education; and the ability to participate in civil society through democracy and the rule of law. Deaton (2013) further argues, "Not the least of the health problems faced by the poor countries of the world today is the lack of good information…[there are] invented and interpolated numbers from international agencies…these are not an adequate basis for policy or for thinking about or assessing external aid. The need to do something tends to trump…and without data, anyone who does anything is free to claim success".

Undoubtedly, it is good to improve health services, and to make sure that those who are in medical need are looked after. A life saved today is a future asset for the economy. The concerns about public health as well as about lack of reliable data in India date to the pre-independence era. Broadly, this paper has two parts. After a brief review of literature, the first part covered in Sects. 3 and 4 provides a historical perspective of the evolution of National-level surveys carried out by NSSO on individuals' health status across states of India, from 7th Round (1953–1954) to 71st Round (2014).[1] We provide an account of the relative positions of major states of India on their (ill)-health status with respect to rural/urban, male/female categories. The measure of health status is the self-report by the ailing person of his/her health status. This is the common practice across countries.

We also provide a brief critique of the reliability and inherent biases in the "self-reported health" (SRH) status as a method for assessing morbidity conditions of a given population. Despite the weaknesses of the SRH approach, and despite the variations in the definitions of measures of ill-health over the years, we find that Kerala, Tamil Nadu, West Bengal, Punjab, Andhra, maintain their relative rankings (with respect to self-reported illness status) much higher than national average. While Assam, Bihar, Madhya Pradesh, Rajasthan, Uttar Pradesh (also Haryana) have lowest rankings (i.e., their self-reported ailments are far below the national average) across the years.

---

[1] The very first survey of public health was by All-India Institute of Hygiene and Public Health at Kolkata in 1944 by Lal and Seal (1949) in the village of Singur, West Bengal. The next was a pilot methodological survey by Poti et al. (1955) of Indian Statistical Institute, Kolkata in 1955. The National Sample Survey Organization (NSSO) made its first attempt to collect morbidity data in its 7th Round (1953–1954) and subsequently, in the 11th (1956–1957), 12th (1957), 13th (1957–1958) and 17th (1961–1962) Rounds. Most of these were exploratory pilots to supply the methodology for future studies on a large scale. The first full scale survey was in 28th Round (1973–1974). After the 28th Round, there was a pause of more than a decade. Then, NSSO resumed the collection of morbidity data under the rubric of social consumption in 42nd (1986–1987), 52nd (1995–1996), 60th (2004) and 71st (2014) Round surveys.

The above observations led to a view, partly "influenced" by Amartya Sen, that higher reported morbidity in states like Kerala and countries like United States is due to individuals' higher social and economic status (SES), and that socio-economically disadvantaged individuals (such as those in Bihar) will tend to fail to observe presence of illnesses.

This takes us to Part II of our paper, covered in Sect. 5, where we ask the question, whether the above observations (views) can be empirically examined using the unit-level household data. A review of the literature with specific reference to India, showed one published article by Subramanian et al. (2009) providing an empirical examination of Sen's view. We summarize Subramanian et al's findings in contrast to Sen's null hypothesis. Our paper contributes to this literature by empirically examining the unit-level data from the 71st Round NSSO survey (NSSO 2015) by estimating logistic regression model as well as a linear probability model with Heckman-corrections to ensure that the estimated probabilities lie within the unit interval, thus extending Subramanian et al's paper both in scope and empirical rigour for examination of relationship between SES and self-reported morbidity (SRM). In Sect. 6, we conclude by stating some of the limitations of our study and the need for further research on this complex subject.

2 Brief Review of Literature on Self-Reported Health Assessment (SHA)

There are two alternative assessment procedures of a person's health status: (a) by himself or herself (SRH), as a member of a sample survey; and in contrast, (b) diagnostic assessment by a team of clinicians (DAC). It is evident that in general questioning or getting a person to fill a questionnaire relating to health is likely to be considerably cheaper than examining that person by a team of clinicians. However, the information elicited by the questionnaire has to be analytically linked to that elicited by the clinicians, for example through an appropriately designed experiment of doing both to a set of individuals, establishing an effective link, for the survey method to be cheaper overall. Such a link is yet to be established.

Prinja et al. (2012) point out, "Self-reports have been used extensively in both developed and developing countries. Large scale demographic health surveys (DHS) have used self-reported morbidity (SRM) for estimating prevalence of maternal and childhood diseases in India… Community based surveys have also used self-reports for assessment of risk factors leading to ill-health. Self-reports have also been used in evaluating interventions in clinical and community settings, using a pre- and post-intervention design. In spite of the large-scale use of self-reports…validity of SRH and SRM have been continuously put to question". Prinja et al. (2012) suggest two approaches for improving the interpretation of self-reports: use of case studies and vignettes and the use of econometric techniques such as 'decomposition' analysis. Our paper is one such attempt.

Amartya Sen (2002) highlighted the limitations of self-reports in Indian States. We shall comeback to Sen (2002) below on self-reports which in principle are

'subjective', and to his two philosophical papers (Sen 1992, Sen 1993) on objectivity.

Subramanian et al. (2009) tested whether the association between self-reported poor health/morbidities and socioeconomic status (SES) in India, followed the expected direction—that poor socio-economic status is associated with a lower perception of illness and low health status, or not. If a positive (or a null) association between SES and self-reports of poor health/morbidities is observed, such that high SES individuals report higher (or the same) prevalence of ill-health compared to low SES individuals, then this casts a doubt on the use of self-reported measures of health or disease status in population-based surveys. The authors carried out cross-sectional logistic regression analyses on a nationally representative population-based sample from the 1998 to 1999 Indian National Family Health Survey (INFHS); and 1995–1996 and 2004 Indian National Sample Survey (INSS).

The key individual predictor variable of interest for Subramanian et al. (2009) was educational attainment. It was measured in terms of years of education for every individual, and was grouped using the following conventional benchmarks: illiterate (no formal education), primary (less than 5 years), secondary (6–12 years), and post-secondary (more than 13 years). Using attained education level as a proxy for education, they consider it to be a reasonable indicator of an individuals' level of awareness and health expectation, besides being a chronic marker of social disadvantage. As per their results, individuals who had no formal education reported significantly higher levels of any self-reported morbidity (OR 1.49, 95% CI 1.42–1.56) compared to those with more years of educational attainment.

The association between SES and self-reporting of morbidity thus followed an inverse gradient; as educational attainment decreases the average odds of reporting morbidities increase and confidence intervals widen. Individuals with no education were found to be 2.5 times (95% CI 2.34–2.63) and 50% (95% CI 1.36–1.54) more likely to report sick in the last 15 or 365 days, respectively, compared to those with post-secondary education. Contrary to the hypothesis—that there is a positive (or null) association between measures of SES and self-reported poor health/morbidities in less-developed countries, it was found that those with less education are more likely to report specific morbidities, sicknesses and overall poor health status in India.

Dilip (2002) examined the prevalence of ailments and hospitalization in Kerala using data from the 52nd Round NSSO survey on healthcare (NSSO 1998). Analyzing data in a logistic regression setup, he found that age and seasonality had considerable effects on the morbidity of individuals. The odds ratios of 2.04, 2.03 and 4.27 observed for age groups 0–14, 40–59 and those above 60 years, respectively, were statistically significant, often at 1% level, and confirmed a 'J-shaped' relation between age and morbidity. The burden of ill-health was higher in rural areas than in urban areas, with people living in rural areas 31% more likely to report an illness than those living in urban areas. People who were more likely to have a better 'lifestyle' had a higher level of morbidity and hospitalization. His analysis showed an inverse relation between monthly per capita expenditure (MPCE) and a person's health status. People from the highest MPCE category were 41% more likely to report illness than those in the lowest MPCE category. Regional differences were seen, with levels of morbidity and hospitalization higher in the

comparatively developed regions of Southern Kerala than in Northern Kerala. He finally concludes that factors like physical accessibility of health care services and capacity to seek health care services could create artificial differences in morbidity and hospitalization among different subgroups of the population in Kerala.

3 Subjectivity of Self-Reporting

3.1 Amartya Sen on Objectivity and Position

In his Lindley Lecture at the University of Kansas (Sen 1992) and a follow-up (Sen 1993), Amartya Sen delves into the philosophical foundations of "objectivity and position". It is worth describing briefly his arguments. The opening paragraph succinctly lays out the basic question, "The subject of this paper is the relationship between the inescapable positionality of observations and the demands of objectivity in science and practical reason. What we observe depends on our position vis-à-vis and the object of observation, and that positionality relates to a number of parameters—locational and others—that influence acts of observation. But even though observations are parametrically variable with positions, they are central to our understanding of the world, and thus to science, decisions and ethics. Objectivity would seem to demand some kind of *invariance* with respect to particular characteristics of the observer and her circumstances. But the question is: which characteristics should figure in the invariance conditions—and no less importantly, which must not so figure?" (Sen 1992).

Sen views Thomas Nagel's 'A View from Nowhere' as an "excellent example of the fruitfulness of this approach in seeing objectivity". Sen points out that Nagel's approach "is nevertheless misleading in some crucial aspects", which concerns him, and leads himself to distinguish between two concepts of objectivity—positional and trans-positional objectivity. He finds positional objectivity to be of interest in itself, and as the crucial building block of trans-positional objectivity and discusses, "the relevance of positional perspectives on objectivity in, respectively, science decision theory, ethics, and public affairs. Given the topic of this paper, we will discuss Sect. 8 from Sen (1992) on perceptions, health and well-being.

Sen (1992) points to the problem of ill health, and particularly the contrast between (1) self-perception of health and (2) examination by doctors. "In some contexts, self-perception is part of the ailment. Having a head-ache, or experiencing nausea or dizziness, is part of the ill-health itself and not just a symptom of it. In these cases, priority of self-perception would seem to be hard to escape in arriving at a position-independent assessment…Methodical use of medical services both (1) reduces one's morbidity, and (2) increases the self-perception of morbidity." (ibid., p. 12).

3.2 Perceptions, Health and Well-Being

In another very short editorial piece of British Medical Journal, Sen (2002) points out that the critical scrutiny of public health and medical strategy, inter alia depends on how individual state of health are assessed. Sen argues that "one of the complications arises

from the fact that a person's own understanding of his or her own health may not accord with the appraisal of medical experts". Sen continues that "more generally there is a conceptual contrast between 'internal' view of health (based on a person's own perceptions) and 'external' view (based on observation of doctors or pathologists)". According to Sen, this 'external' view has come under considerable criticism recently. This is no surprise, since the debate about the role of participant—observer has been under discussion in anthropological circles, for a long time. Quoting him in full, "consider the different states of India, which have very diverse medical conditions, mortality rates, educational attainments and so on. The state of Kerala has the highest levels of literacy… and longevity. But it also has by a very wide margin, the highest rate of reported morbidity among all Indian states… At the other extreme, states with low longevity, with woeful medical and educational facilities such as Bihar, have the lowest rates of reported morbidity in India." Sen asks, "why such dissonance arises?" and argues that, "there is much evidence that people in states that provide more education and better medical and health facilities are in a better position to diagnose and perceive their own particular illnesses than are the people in less advantaged states, where there is less awareness of treatable conditions (to be distinguished from "natural" states of being). The medically ill-served and substantially illiterate population of Bihar may have a very low perception of illness, but that is no indication that there is little illness to perceive. This interpretation is supported also by comparing the reported morbidity rates in the Indian states and in the United States. In disease by disease comparison, while Kerala has much higher reported morbidity rates than the rest of India, the United States has even higher rates for the same illnesses".

This argument suggests that in testing the null hypothesis of no association, the appropriate alternative is that an individual's health status depends positively on his/her socio-economic status and the socio-economic environment in which he/she lives. We test Sen's null hypothesis of no association against the alternative of positive association, using data from the 71st Round survey of the NSSO (NSSO 2015). We reject it, against the alternative that an individual's SRH depends positively, on his/her own socio-economic status as well as the society in which she lives. This means that lower the SES, the lower will be the health-status (reported as having higher morbidity); the higher the SES, higher will be the health-status (reported as having low morbidity). Subramanian et al. (2009), cited earlier, also arrive at the same conclusion using different data sets including from NSSO.

Almost all surveys in India and elsewhere capture only self-reported morbidity which is by definition 'subjective'—being dependent on the responses by a member of the household who need not necessarily be an ill person, particularly in the case of children whose parents respond for them. In the Indian case, particularly, the surveys seem to involve subjectivity at two levels: first, at the respondent level and the second in the definition of illness. Reporting for an ill person by another who is not the person possibly creates biases. Bias may also arise from unobservable and implicit standards being used in the definition of illness. There is the issue of a person's socio-economic status as perceived by him/her that may influence a person's response to questions about his/her health state. Not only one's own perceived socio-economic status, but also his/her external socio-economic, physical environment could influence a person's responses.

For the second case of bias due to unobserved and implicit standards for defining illness; we note that until the 13th Round (NSSO 1961), observable standards such as being unable to engage in regular or normal activities or having restricted diet etc., were used as indicators of illness. From the 13th Round onwards, a rather subjective concept—a person is deemed ill, if he or she is in a mental state that deviates from being normal, is being used. The reason is that in early surveys after the Special Study on Morbidity (NSSO 1969), i.e., the 17th Round (November 1960–October 1961, NSSO 1968) until 28th Round (October 1973–June 1974 NSSO 1980), ailment or illness during a particular period [i.e. the reference period] was defined as any deviation of the state of physical and mental well-being with a specific cause (i.e. a person was sick if he felt sick— (NSSO 1968). In the 71st Round (NSSO 2015) survey, "ailment, i.e. Illness or injury, meant any deviation from the state of physical or mental well-being" (NSSO 2015). The report emphasizes, "Note that the identification of ailment is necessarily subjective as it depends on the feeling or perception of the person concerned. This is a problem inherent in all surveys of general illness or morbidity" (ibid, footnote 2). But this report as well as reports of earlier surveys that adopt the definition of illness as being a deviation in the mental state of the person from that of his/her well-being does not note that the mental state of well-being is not observable. On the other hand, in the early surveys prior to the Special Study on Morbidity) (NSSO 1969), morbidity data were collected from persons who also deviated from their normal behaviour, "by being confined to bed for at least 24 h or were unable to attend to normal activities" (NSSO 1961), in principle—normal diet and normal activities, as well as the state of being confined to bed are observable. As such there is no layer of subjectivity in addition to that of self-reporting. Unfortunately, Report 119 (NSSO 1969) offers no reason for change in the definition of illness after the special study. Perhaps ensuring conformity with definitions of WHO might have been one of the reasons.

The reference period in the earlier surveys was 15 days prior to the date of the visit of NSS investigator to the household except for hospitalization for which the period was 365 days prior to the visit. In the later surveys also, the reference periods for ailments and hospitalization were essentially the same as in earlier surveys. The differences between the two sets of surveys with respect to the estimated average current population (the denominator for prevalence and incidence rates) seem to be relatively unimportant.

4 Descriptive Analysis of NSS Data from 71st and Earlier Rounds

Sen (2002) does not analyse the data in detail. In what follows we do, using NSS morbidity data from the earliest in 7th Round (1953–1954) to 71st Round (January–June 2014). The data are in three sets. The first set covers Rounds 7th (1953–1954), 11th, 12th and 13th (September 1957–May 1958) Rounds (NSSO 1961); the second from 17th (September 1961–June 1962, NSSO 1969) till 28th (October 1973–June 1974, NSSO 1980) Rounds and the third consists of quinquennial rounds 42nd (1986–1987, NSSO 1989) to 71st (January–June 2014, NSSO 2015) in which morbidity is covered under social consumption. Concepts of illness and indicators

such as Incidence Rates (IR), prevalence rates (PR) and Proportion of Ailing Persons (PAP) at the State and National levels differ across rounds. We will ignore these differences under the strong assumption that a pure numerical indicator, the percentage deviation of units of any indicator for a state from its corresponding national average is comparable. However, the variations in concepts and the lengths and seasons of rounds, calls for caution in our interpretation.

NSS reports, after noting "that the patterns of Rural–Urban, Male–Female and Age-group differences are similar across rounds" but cautions that "keeping in mind the differences in the definition of sickness and recall periods, that the incidence rates in the developed countries were much higher than those reported for India." It turns out, that the caution is true for prevalence rates as well (Report 49, NSSO (1961))".

4.1 Incidence Rate, Prevalence Rate and PAP: Inter-state and Inter-temporal Variations in Morbidity

4.1.1 The First Set

In the first set (see Appendix D, Table 4), the incidence (prevalence) rate for Rural India as a whole, fell from 49.14 (64.77) per 1000 persons in the 7th Round (October 1953–March 1954) to 26.39 (31.66) per 1000 persons in the 11th Round (August 1956–Jan 1957) and rose to 37.38 (56.73) per 1000 persons in the 12th Round (February 1957–July 1957). On the other hand, the incidence (prevalence) rate for Urban India as a whole, fell from 45.58 (61.75) per 1000 persons in the 7th Round to 32.25 (35.55) in the 11th Round, then rose to a peak of 53.19 (81.43) per 1000 persons in the 12th Round and then fell to 34.81 (42.01) in the 13th Round.[2]

Report 49, NSSO (1961) on Morbidity covered 7th, 11th, 12th and the 13th Round. It notes that the patterns of rural–urban, male–female and age group differences are similar across rounds. However, the report compares prevalence and incidence rates by the NSS definition of the first set with those in surveys of developed countries of Canada, England and Wales and Denmark and notes that, "keeping in mind the differences in the definition of sickness and recall periods, that the incidence rates in the developed countries were much higher than those reported for India." It turns out that the same is true for prevalence rates as well.

Report 49, Chapter 6 is devoted to average duration of sickness calculated by excluding meaningfully the likely long duration categories of those "beginning before the reference period but ending within it" and then those beginning before the reference period and continuing on the date of the survey" (Report 49, Section 6.1). The chapter also calculates "days of incapacity per person" defined as the product of the prevalence rate per person and average duration of sickness per spell (Report 49, Section 6.6). The chapter compares average duration and days of incapacity per person in India with developed countries of Canada, England and Wales and Denmark.

Section 2.9 of Report 49 (NSSO 1961) and entire chapter 7 contain a wealth of information. Apparently, those who designed and executed the later surveys not

---

[2] Only Urban areas were covered in the 13th Round (Sept 1957–March 1958) NSSO Survey.

only seemed to be unaware of this fact and are also under the mistaken beliefs that even for All-India estimates only the quinquennial round's large sample sizes are adequate. They did not recognize that the reliability of survey estimates depends on the absolute size of the sample and for a long time the absolute sizes of All-India Samples have been large enough to yield adequately reliable estimates at the All-India level, perhaps also at the level of large states. They reason as to why Indian concepts and definitions might lead to lower incidence rates but higher average durations of illness as compared to developed countries. These remarks are suggestive of studies that need to be undertaken.

4.1.2 The Second Set

From the second set, Report 119 (NSSO 1969) on the Special Study on Morbidity and Report 129 (17th Round, NSSO (1968)) on Pilot Enquiry on Morbidity provide comparisons on incidence and prevalence rates etc., responses of self vs proxy respondents, rural and urban areas, different recall periods, categories of sickness, prevalence rate by type of disability, response to probing questions and their implications. In particular, investigators were instructed not to attempt correcting what may appear to be naturally inconsistent responses and many others. Unfortunately, it is not possible to calculate overall All-India prevalence and incidence rates during the 17th round (NSSO 1968) from the reports. All one can say from Tables 6.1 and 6.2 of Report 129 is that incidence rates in rural and urban areas were between 18.37 and 26.62 and between 18.39 and 29.09, respectively. The prevalence rates in rural (and urban) areas varied between 54.07 and 81.94 (45.90 and 79.15); both rates being normalized to per 1000 of estimated population exposed to risk (see Appendix D, Table 5).

Report 292 (NSSO 1980) Report 292 on the 28th Round (October 1973–June 1974) is the last of the second set of surveys. Its Table 1 presents data on Incidence rates and prevalence rates by states and India as a whole for rural and urban areas. Interestingly, footnotes to Table 1 note that "the incidence rate and prevalence rates of morbidity of NSS 28th Round as estimated are somewhat lower than the rates observed in the 17th Round of the NSS. A Seminar meeting…held under the auspices of NSSO to examine…morbidity rates of the 28th Round. The consensus was that the morbidity rates could be released even though the estimates of NSS 28th Round appear somewhat lower than those of the Pilot Enquiry on Morbidity [in the 17th Round]" (Report 292, Chapter 3, NSSO (1980)).

In fact, the All-India Rural (Urban) incidence rates in the 28th Round of 12.57 (13.53) were substantially lower than the lower limits of 18.37 (18.39) in the 17th Round. Similarly, the All India Rural (Urban) prevalence rate 22.46 (32.18) were substantially lower than the lower limits 54.07 (45.90) in the 17th Round (see Appendix D, Table 5).

It should be noted that none of the time spans of the rounds in the second set covered the whole year but covered different months of the year. It is possible that the differences in incidence and prevalence rates across were confounded by seasonal factors, (for example, the seasons of the year covered by the round) and inter-temporal factors, since the 28th Round came more than a decade after the 17th

Round (1960–1961). Report 364 (NSSO 1989) includes data on morbidity and utilisation of medical services in the 42nd Round which covered a full year (July 1986–June 1987) but does not provide data on any morbidity-related rates.

4.1.3 The Third Set

The third set consisting of the quinquennial surveys of the 42nd (1986–1987, NSSO 1989), 52nd (1995–1996, NSSO 1998), 60th (January–June 2004, NSSO 2006) and 71st (January–June 2014, NSSO 2015) Rounds. This set as noted earlier uses the same definition of illness, namely, a deviation from the mental state of well-being, in itself subjective, depending on the ill-person's implicit judgement of his/her own state of well-being in addition to the usual subjecting of self-reported or proxy respondent's report of illness.

The first and second sets presented data on prevalence and incidence rates. In the third set there are no analogues of either. The report of the 71st Round states that "the morbidity rate presented in the document gives the estimated proportion of persons reporting ailment at any time during the 15-day reference period and are not strictly the prevalence rates as recommended by the Expert Committee on Health Statistics of the WHO" (NSSO 2015).

The Report lists possible inherent limitations of the subjectivity of the identification of ailments [Section 4.1 of (NSSO 2015) and its footnotes]. The concepts and definitions in its Appendix B refer mostly to other data collected on the surveys such as 'nature of treatment', 'level of care in institutions with provision for admission of sick persons as in in-patients for treatment', 'ailment and other terms', 'medical expenditure for treatment', 'non-medical expenditure', 'total expenditure' and finally of 'final expenses'.

It is clear, that the third set contains relevant information about morbidity, not all of which were included in the first two sets. It is surprising that NSS does not seem to have an institutional memory—the third set does not even mention any of the surveys in the first two sets let alone comment on them. Even more surprisingly while the Report on the 42nd Round provides detailed data on utilisation of hospitals and wards as well as sources of financing of medical expenditure, Report on the 71st Round which also covered details of hospitalization, treatments and their costs does not refer to Report 364 (NSSO 1989) of the 42nd Round at all.

4.1.4 Some Comparisons of the Three Sets

The incidence and prevalence rates of the first two sets do not have exact comparison categories in the third set. Still there is some overlap and also some exclusion in comparing either the prevalence or incidence rates of the second set with the morbidity rates. The last row of Table 1 of Report 292 (NSSO 1980) shows that in Rural India, the prevalence rates in Rural (Urban) India were 23.46 and 22.77 while the incidence rate in Rural (Urban) India were 12.57 and 13.53.

For the later surveys, Table 5 (Appendix D) gives the estimates of morbidity rates. It shows that in the 52nd Round, there was no significant Gender or Rural–Urban differences in the Percentage of Ailing Persons (PAP) per 1000 of living

persons. However, for the 60th and 71st Rounds, PAP exhibits substantial increases across all categories. Now, Gender and Rural/Urban differences emerge. Thus, PAP for Rural India as a whole goes up from 55 in the 52nd to 88 and 89 respectively in the 60th and 71st Rounds. In Urban India, the PAP almost doubles and goes up from 54 in 52nd to 99 and 118 respectively in the 60th and 71st Rounds with females exhibiting larger rises than males.

The report speculates that "the increase in PAP over time is probably due to increasing health consciousness over time and consistent improvement in the reporting of ailments by the informants especially for urban section" NSSO (2015). In this speculation, the fact that while the 52nd (and 42nd) Round covered the entire year 1995–1996 (1986–1987), the 60th and 71st Rounds covered only the January–June period, is not taken into account. Hypothetically, if in the unobserved second halves of the 60th and 71st Rounds, the PAP were to be lower than in the first half, PAP for the entire year of 2004 and 2014 would have been lower too so that instead of an increase there would have been a decrease over time in PAP. Thus, other than being a speculation without supporting evidence it confounds intra-year and inter-year shifts in PAP.

Going back to the second set, a similar confounding of possible intra-year and inter-year shifts was noted. Unravelling the confounding requires formal empirical modelling and statistical testing of the self-reported responses and the factors influencing them. We do not attempt any theorizing or modelling.

Although incidence rate and PAP of any state in any given time are not comparable arguably by taking percentage deviation for either of each state from the corresponding National Average, one gets a unit-free, pure number that can be deemed comparable across states.[3]

Figure 1 shows for Rural (and Urban areas), the percentage deviations (from the respective National Average) of incidence rates of States in the 28th Round; and Fig. 5 shows inter-state deviations in PAP in the 71st Round, in increasing order. It is seen that in the 28th Round, for Rural (Urban) areas of 16 selected states, 6 (7) out of 16 had incidence rates above the National Average as shown in Fig. 1. Ranking them by the percentage deviations from the National Average, the top-five with positive deviations in Rural areas were Kerala as the first with almost two times the national average, followed by Andhra Pradesh as a distant second, with Tamil Nadu, Punjab and Maharashtra being the next three in decreasing order of deviations in rural incidence rates. The bottom-five among those, with negative deviations in rural areas, were Gujarat, Bihar, Uttar Pradesh, Karnataka and Assam in increasing order of incidence rates.

In the 71st Round, in Rural (Urban) areas, 8 (5) states out of 16 had incidence rates above the National Average as shown in Fig. 5. Ranking them by the percentage deviations from the National Average, the top-five with positive deviations in Rural areas were, Kerala as the first, followed by West Bengal, Punjab,

---

[3] Since Bihar, Madhya Pradesh and Uttar Pradesh had been divided after the 28th Round into the states of Bihar and Jharkhand; Madhya Pradesh and Chhattisgarh; Uttar Pradesh and Uttarakhand, their Rural (and Urban) PAP values for the undivided states were computed as the Rural (and Urban Population) weighted averages of PAP values of their component states after division.

Tamil Nadu and Andhra Pradesh. The bottom-five among those with negative deviations were Assam, Madhya Pradesh, Rajasthan, Bihar and Haryana.

In the 28th Round in Urban areas, also shown in Fig. 1, Kerala was again at the top among all the states with positive deviations in incidence rates with almost one and a half times the National Average followed by a distant Tamil Nadu with Maharashtra, Andhra Pradesh and West Bengal in decreasing order among top-five. Four out of top five are the same in Rural and Urban areas. The bottom-five in urban areas were Gujarat, Assam, Bihar, Uttar Pradesh and Karnataka in increasing order. Remarkably the top five and the bottom-five states are the same, though not in the same order, in Rural and Urban areas.

In the 71st Round, in Urban areas (Fig. 5), the top among states with positive deviations in PAP was Kerala with Tamil Nadu, West Bengal, Punjab and Andhra Pradesh following it among the top. The bottom-five were, Jammu and Kashmir, Assam, Madhya Pradesh, Maharashtra and Haryana. Once again, for PAP too, the top five states are the same and include the Southern states of Kerala, Andhra Pradesh and Tamil Nadu in Rural and Urban areas. However, among the bottom five Assam, Madhya Pradesh and Haryana are present in both Rural and Urban areas. If we consider states with negative deviations which include the bottom five and more, they include the so called BIMARU[4] [Bihar, Madhya Pradesh, Rajasthan and Uttar Pradesh] states!

5 Econometric Analysis of NSS Data from 71st Round

Before turning to the econometric analysis of data, a few words of its motivation are in order. Recall that Sen's hypothesis was entirely motivated by a comparison of Bihar with Kerala and Rest of India to argue that the differences in socio-economic environment of Bihar and Kerala, not the absence of diseases in Bihar and morbidity in the others led to Kerala being the most morbid state in India.

More generally, the Sen's alternative hypothesis is in conformity with the finding that personal socio-economic status of an individual resident of a state and the general socio-economic environment of the state, increasingly influence an individual's response to questions about his health. A more rigorous socio-economic and medical analysis of the issue than what we have been able to do is necessary, but it was not possible to do so for a variety of methodological and measurement issues involved and given the extremely poor information base we have on various dimensions, besides our capability and resources at disposal.

It is clear from the several figures that we present, on percentage deviations of various morbidity measure(s) from their corresponding national averages, from 28th Round (1973–1974) and the 71st Round (January–June 2014),[5] whichever way one slices the data, by rural/urban, male/female and so on, almost in all figures, Kerala

---

[4] The Economic Demographer Ashis Bose, imaginatively named Bi(har), Ma(dhya Pradesh), R(ajasthan) and U(ttar Pradesh) as BIMARU states evoking the Hindi word 'BIMAR' for illness.

[5] Figures 2, 3, 4, 6 and 7 show inter-state deviations of incidence rate/PAP by Gender and Age-groups for 28th and 71st Round survey.

has remained the most morbid state in India over decades,[6] not just when Sen happened to hypothesize it.

Moreover, there seems to be a pattern in the ranking of states by morbidity. Among top-five morbid states, besides Kerala at the very top, more often than not, one or more of Andhra Pradesh, Tamil Nadu, West Bengal and Punjab appear. Interestingly, the old 'BIMARU' states, so named by the economic demographer Ashis Bose, often make their appearances among the bottom-5 or least morbid states. The inter-state stability in morbidity pattern over decades requires a far deeper and causal analysis than we have done.

Moving to other plausible and related indicators of morbidity, that are available for analysis, the average duration of sickness per person (only for individuals reporting an ailment during the reference period) defined in number of days, the picture is strikingly in contrast to what we observe for morbidity. Any person reporting an ailment(s) during the survey was further asked to report the total duration of ailment(s) in days for each case of ailment separately. We were able to put together information on the average duration of sickness from the 28th Round and the 71st Round only, and we present this in Figs. 8, 9, 10 and 11. Ranking each state by its percentage deviation of average duration of sickness from the national average, the 'BIMARU' states, -now occupy the top positions while states like Kerala, Tamil Nadu, Punjab etc. take the lowest positions.

5.1 Data and Methodology

The dataset used for our analysis comes from the 71st Round (January–June 2014) NSSO survey (NSSO 2015). In this survey or round; 65,932 households and 333,104 individuals were surveyed. Out of the entire sample of HHs, about 36,480 were rural HHs whereas 29,452 were urban HHs. The NSSO obtained information on morbidity based on the survey respondents' answer (yes or no) to the following questions separately

1. Have you been suffering from any chronic ailment?
2. Have you been suffering from any other ailment anytime during the last 15 days?
3. Have you been suffering from any other ailment on the day before the date of the survey?

The respondent, usually the head of the household, answered the presence or absence of morbidity for themselves (self-reporting) as well as for other household members (proxy-reporting). We focus on the second question above and model morbidity based on the binary response to the same question. Further, each person reporting an ailment, whether chronic or acute, was asked to report the total duration of each case of ailment separately. The status of ailment (A, B, C or D; see Appendix A) was also noted. Utilizing this information, we calculate the average duration of sickness per person. Additionally, in a sub-sample of elderly population

---

[6] As revealed by ranking of states by deviations of PAP (per 1000 persons) from National Average for 52nd Round (1993–1994) and the 60th Round of NSSO survey. Available with the authors upon request.

aged 60 and above, the NSSO also recorded overall health perception of the individual by explicitly asking them their own perception about the current state of their health—excellent/very good; good/fair; poor. We created a binary self-reported poor health variable for every individual that was equal to 1 if the individual reported poor, 0 otherwise. Information on physical mobility (considered immobile, if confined to bed, confined to home, able to move outside but only in a wheel chair; otherwise physically mobile) was also recorded.

The NSSO also obtained a plethora of socio-economic information about the surveyed individuals, which was of primary interest to us. Along with an individual's age and gender, we also utilize information on her social background i.e. religion—Hindu, Islam, Christian and Others (Jain, Sikhism etc.) and social group—SC, ST, OBC and Others; her educational status—Illiterate (no formal education), Primary (less than or equal to 5 years), Secondary (6–12 years) and Post-secondary (more than 12 years); her area—Rural, Urban and State of residence and her economic status measured through the household's 'Usual Monthly Consumption Expenditure (UMCE). Unless stated otherwise, each of these variable is treated as a categorical variable with each category represented by a binary variable (1 or 0), except the household's monthly consumption expenditure (UMCE). It is converted into a per-capita measure 'Usual Monthly Per-Capita Consumption Expenditure (UMPCE) using the household's size and is treated as a continuous variable.

Thus, we employ three different regression techniques on the full- and sub-sample data to study both morbidity and average duration of sickness (see Appendix-B for details on economic model(s)). First, logistic regression approach is used to analyse morbidity in the full-sample of NSSO data and the sub-sample of elderly population aged 60 and above. The response to the question (2) above is used as a dependent variable for the full-sample, while the response to one's own perception of current health is used similarly for the sub-sample. To facilitate interpretation, we report the results in terms of odds-ratio (OR) and the estimated standard errors (see Table 1, Appendix C).

Second, in a similar sample and dependent variable setting, we estimate a simple linear probability model (LPM). The case of employing a linear probability model (LPM) to analyse the data appealed to us because of its simplicity and transparency. Each variable including UMPCE (divided into quintiles), is treated as a categorical variable this time. Every estimated coefficient for a categorical variable in an LPM, is interpreted as the probability of occurrence of the event (in our case, person 'i' reporting any other ailment), keeping all other things constant (see Table 2, Appendix C). The estimated coefficients in an LPM often lie outside the closed unit interval [0, 1] violating their range from a probabilistic standpoint. Here, we constrain the estimated coefficients so that they fall in the [0, 1] and name the constraints 'Heckman-type' analogous to Heckman (1976) bounds. As in the Heckman (1976) case, the analogue of a Probit or Logit model is the linear probability model. The analogue of wage equation, that the wage is observed only for labour force participants, are the constraints that the probability is the estimated value only if the [0, 1] constraints are met, that is the estimate falls in the closed interval [0, 1]. Otherwise, the estimate is replaced by the probability that the term inclusive of the coefficient and the error term meets the constraints. This probability

Table 1 Logistic regression models: odds-ratio

Independent variables	Any-ailment[a]	Health-status[b]
Gender		
Male	0.926[c]	0.912[c]
Female	1	1
Age		
0–6 years	1.910[c]	1.047
6–61 years	0.722[c]	
61 years and above	1	
Physical mobility[b]		
Physically immobile	–	11.06[c]
Mobile	–	1
Religion		
Hindu	1	1
Muslim	1.043	1.384[c]
Christian	1.114[d]	0.787[d]
Others	0.957	0.896
Social groups		
SC	1	1
ST	0.863[c]	0.795[d]
OBC	0.963	0.943
Other social groups	0.920[d]	0.857[d]
Area of residence		
Rural	1	1
Urban	1.021	1.056
Education		
Illiterate	1	1
Primary	0.905[c]	0.812[c]
Secondary	0.719[c]	0.610[c]
Post-secondary	0.589[c]	0.428[c]
Reporting		
Self	1.676[c]	1.234[c]
Proxy	1	1

Table 1 continued

Independent variables	Any-ailment[a]	Health-status[b]
UMPCE (in 100's INR)	1.006[c]	1.002[c]
State of residence		
Jammu & Kashmir	0.301[c]	0.911
Himachal Pradesh	0.369[c]	0.290[c]
Punjab	0.968	0.399[c]
Uttarakhand	0.570[c]	0.452[c]
Haryana	0.410[c]	0.390[c]
Rajasthan	0.313[c]	0.522[c]
Uttar Pradesh	0.488[c]	0.691[c]
Bihar	0.315[c]	0.719[c]
Assam	0.186[c]	0.648[c]
West Bengal	0.736[c]	0.763[c]
Jharkhand	0.399[c]	0.828
Odisha	0.726[c]	0.971
Chhattisgarh	0.456[c]	0.438[c]
Madhya Pradesh	0.425[c]	0.455[c]
Gujarat	0.496[c]	0.195[c]
Maharashtra	0.433[c]	0.378[c]
Andhra Pradesh	0.533[c]	0.713[c]
Karnataka	0.559[c]	0.527[c]
Goa	0.852	0.137[c]
Tamil Nadu	0.730[c]	0.288[c]
Telangana	0.459[c]	0.355[c]
Kerala	1	1
Union Territories	0.475[c]	0.348[c]
North-East states	0.235[c]	0.556[c]
Constant	−2.229[c]	−3.768[c]

[a] Person 'i' reporting any ailment during last 15-days (yes = 1, no = 0)

[b] Person 'i' reporting own perception about current health (poor = 1, good/excellent = 0)

[c] Significant at <1% level

[d] Significant at <5% level

[e] Significant at <10% level

is estimated by assuming that the error term is normally distributed with mean zero and a standard deviation equalling the estimated standard error for large samples, which is true in our setting. In our case, when a coefficient failed to meet the constraints, we add (subtract) the lowest multiple '×' of the standard error which equates the coefficient to the lower (upper) bound of [0, 1]. Instead of the estimated coefficient, we then report this minimum multiple '×' for the particular variable (see Appendix B). It should be noted that for each categorical variable where this term is reported, it has to be interpreted as follows—the higher the multiple '×' for the variable, lower the predicted probability associated with that variable.

Table 2 Heckman-type linear probability models

Independent variables	Any-ailment[a]	Health-status[b]
Sex		
Male	0.02*[c]	0.03*[d]
Female	1	1
Age		
0–6 years[a]/60–75 years[b]	0.04*[c]	1
6–61 years[a]/75–90 years[b]	0.08*[c]	0.11[c]
Above 61 years[a]/above 90 years[b]	1	0.15[c]
Physical mobility[b]		
Physically immobile	–	0.52[c]
Mobile	–	1
Religion		
Hindu	1	1
Muslim	0.003*[d]	0.05[c]
Christian	0.005*[d]	0.09*[d]
Others	0.012	0.03*
Social groups		
SC	1	1
ST	0.03*[c]	0.03*[d]
OBC	0.01*[d]	0.02*
Other social groups	0.02*[c]	0.05*[d]
Area of residence		
Rural	1	1
Urban	0.003*	0.01[d]
Education		
Illiterate	1	1
Primary	0.02*[c]	0.11[c]
Secondary	0.07*[c]	0.07[c]
Post-secondary	0.10*[c]	0.03[c]
Reporting		
Self	0.02*[c]	0.03*[c]
Proxy	1	1
UMPCE quintiles		
0–20	1	1
20–40	0.008[c]	0.035[c]
40–60	0.009[c]	0.020[c]
60–80	0.014[c]	0.023[c]
80–100	0.020[c]	0.020[c]

Third, to study the contrasting pattern observed in the inter-state distribution of average duration of sickness, we use the ordinary least-squares (OLS) estimation approach. The sub-sample of all individuals reporting sicknesses is taken and

Table 2 continued

Independent variables	Any-ailment[a]	Health-status[b]
State of residence		
Jammu & Kashmir	0.30[c]	0.06
Himachal Pradesh	0.27[c]	0.50[c]
Punjab	0.03[e]	0.40[c]
Uttarakhand	0.18[c]	0.36[c]
Haryana	0.25[c]	0.42[c]
Rajasthan	0.30[c]	0.30[c]
Uttar Pradesh	0.21[c]	0.20[c]
Bihar	0.29[c]	0.18[c]
Assam	0.35[c]	0.23[c]
West Bengal	0.11[c]	0.15[c]
Jharkhand	0.25[c]	0.11[d]
Odisha	0.11[c]	0.02
Chhattisgarh	0.22[c]	0.37[c]
Madhya Pradesh	0.24[c]	0.36[c]
Gujarat	0.22[c]	0.61[c]
Maharashtra	0.24[c]	0.42[c]
Andhra Pradesh	0.20[c]	0.18[c]
Karnataka	0.19[c]	0.31[c]
Goa	0.07[d]	0.67[c]
Tamil Nadu	0.12[c]	0.51[c]
Telangana	0.24[c]	0.46[c]
Kerala	1	1
Union Territories	0.23[c]	0.44[c]
North-East states	0.34[c]	0.28[c]
Constant	0.10[c]	0.20[c]

* Coef. is adjusted for zero lower bound

[a] Person 'i' reporting any ailment during last 15-days (yes = 1, no = 0)

[b] Person 'i' reporting own perception about current health (poor = 1, good/excellent = 0)

[c] Significant at <1% level

[d] Significant at <5% level

[e] Significant at <10% level

divided into three sub-samples—chronic ailments only, acute ailments only and all ailments. Age and UMPCE are taken as continuous variables. While estimating the OLS model for 'all ailments' sample, the status of ailment (type A, B, C or D; see Appendix A) is used as a control variable in addition to all the other socio-economic variables (see Table 3, Appendix-C). We briefly discuss the results in the next section and conclude with appropriate remarks.

Table 3 Linear regression model: avg. duration of sickness per person

Independent variables	Chronic ailments only[a]	Acute ailments only[b]	All ailments[c]
Sex			
Male	−30.1	−2.4[d]	−1.0
Female	1	1	1
Age (in years)	16.8[d]	1.9[d]	10.2[d]
Religion			
Hindu	1	1	1
Muslim	−50.3	5.6	−42.2
Christian	19.4	5.8	14.5
Others	−130.6	9.8	−105.7[e]
Social groups			
SC	1	1	1
ST	10.5	4.2	23.7
OBC	94.7	6.9	45.9[d]
Other social groups	98.1	15.4[e]	75.5[e]
Area of residence			
Rural	1	1	1
Urban	147.6[d]	18.0[d]	92.5[d]
Education			
Illiterate	1	1	1
Primary	63.4[f]	−6.6	−1.0
Secondary	84.9[f]	−21.3[e]	−24.8
Post-secondary	39.7[f]	−15.5	−50.3
Reporting			
Self	−70.6[e]	−4.4	−103.0[d]
Proxy	1	1	1
UMPCE (In 100 s INR)	5.6	−0.1	4.8[d]
Status of ailment			
Type-A	–	–	983.8[d]
Type-B	–	–	32.5
Type-C	–	–	−37.4
Type-D	–	–	1

Table 3 continued

Independent variables	Chronic ailments only[a]	Acute ailments only[b]	All ailments[c]
State of residence			
Jammu & Kashmir	−1146.8[d]	−55.9[d]	−734.5[d]
Himachal Pradesh	−197.7	−64.2[d]	−251.3[d]
Punjab	−669.7[d]	−45.8[d]	−516.0[d]
Uttarakhand	−612.3[d]	−84.9[d]	−527.5[d]
Haryana	−363.1[d]	−36.0[e]	−336.9[d]
Rajasthan	−408.8[d]	−73.7[d]	−364.3[d]
Uttar Pradesh	−704.6[d]	12.0	−564.8[d]
Bihar	−866.3[d]	−28.3[e]	−527.1[d]
Assam	−1665.8[d]	−74.1[d]	−812.8[d]
West Bengal	−387.4[d]	−56.0[d]	−285.6[d]
Jharkhand	−671.6[d]	−37.9[e]	−436.3[d]
Odisha	−1150.0[d]	−71.0[d]	−607.3[d]
Chhattisgarh	−1049.9[d]	−61.3[d]	−538.4[d]
Madhya Pradesh	−901.1[d]	−67.4[d]	−633.0[d]
Gujarat	−641.2[d]	−70.8[d]	−418.1[d]
Maharashtra	−746.5[d]	−77.2[d]	−478.5[d]
Andhra Pradesh	−381.0[d]	−36.5[d]	−231.4[d]
Karnataka	−541.4[d]	−79.4[d]	−358.0[d]
Goa	−1478.3[d]	−85.7[d]	−810.2[d]
Tamil Nadu	−349.7[d]	−81.0[d]	−228.1[d]
Telangana	−854.4[d]	−75.5[d]	−544.5[d]
Kerala	1	1	1
Union Territories	−170.0[e]	−58.7[d]	−187.7[d]
North-East states	−1571.3[d]	−82.5[d]	−663.4[d]
Constant	842[d]	36.1[e]	55.7

[a]For Person 'i' reporting chronic ailment(s)

[b]For Person 'i' reporting acute ailment(s)

[c]For Person 'i' reporting any-chronic or acute ailment(s)

[d]Significant at <1% level

[e]Significant at <5% level

[f]Significant at <10% level

5.2 Results

Taking our results from the logistic regressions first, it was found that males were slightly less likely to report sickness than their female counterparts in both the full- and sub-sample. Age of an individual followed an interesting pattern–children (0–6 years), more likely their proxies than themselves, were highly likely to report sickness while people between the age of 6–61 years, had lower odds of reporting morbidity. Keeping 'illiterate' or people with no formal education, as the reference category, we find that individuals who had higher educational attainment levels

reported lower levels of any self-reported morbidity. Individuals with the highest level (post-secondary level) of education were about 41% (57% in the case of elderly population) less likely to report sickness than a person with no formal education. If one assumes educational attainment to be a proxy for socio-economic status (Subramanian ([2009])), then there is an evidence of an inverse relationship between education attainment and the odds of reporting any morbidity. We analyse the issue of biases arising through 'proxy reporting' by controlling for it in our model, finding that individuals had significantly higher odds (OR 1.67) of reporting morbidity when reporting for one-self compared to when reporting for others. The effect was lower (OR 1.23), yet significant, in the elderly sample. Area of residence shows no significance when we control for State of residence as well. Religion and social group of an individual had no significant bearing on her odds of reporting an ailment, except for in the case ST's and Muslims in the full and elderly sample respectively.

Employing the linear probability model with Heckman-corrections also yields us similar results. Educational attainment is again found to be inversely related to the probability of reporting an ailment in both the samples. In contrast to the findings from the logistic regression, social group and religion significantly impact the prevalence of an ailment, especially in the case of ST's, OBC's and Muslims. There is also evidence of UMPCE positively and significantly impacting the probability of reporting an ailment; however, the top quintile only had a 2% higher chance of reporting ailment compared to the lowest quintile. For the elderly sample, an indicator of physical immobility is also used. A physically immobile individual had a 52% chance of reporting being sick stressing the importance of recording observable indicators of health.

Lastly, the results from the OLS-based models on the sub-samples of individuals reporting any sickness reveal important patterns about the prevalence of chronic and acute ailments. It is no surprise that, different socio-economic variables have different impact depending on the type of ailment. We note that, duration of sickness for both type of ailments, is not impacted significantly by either religion or the social group to which the person belongs. Age positively impacted the duration of sickness, significantly for both type of ailments while the income of an individual does not seem to be affecting the duration of sickness. On average, urban residents report being sick for a significantly higher number of days compared to their rural counterparts. For the sample of 'all ailments' controlled by status of ailment, a key result is the one for proxy reporting. There is evidence of substantial under-reporting of duration of sickness, when a 'proxy' respondent answers the survey questions for an individual. Ailments that started more than 15-days ago and still continue on the day of the survey (see appendix A; can also be interpreted as chronic ailments), do not seem to go 'unnoticed' given its highly significant coefficient. Duration of sickness was at best unrelated to education attainment of an individual, being significant only at 10% level of significance, that too only for chronic-type ailments.

6 Conclusions

In several ways, our interest on morbidity originated from Nobel Laureate Angus Deaton's emphasis on health achievements and their unequal distribution in a population. The seminal article by Sen (2002), his discussion and the hypotheses emanating from his discussion, drove us to delve deeper into the subjectivities and biases associated with the self-reporting of health status (SRH) and morbidity (SRM) by ill-persons, plus the controversies it generated.

At this stage, we can only say that our results along with those from Subramanian et al. (2009), (a) do reject the null hypothesis that SRH and SRM are independent of the individual's own health status and also of the socio-economic framework of the community in which he or she lives, and (b) support the alternative that self-reported health-status (SRH) and individual socio-economic status (SES) are positively correlated. This also indicates that, the lower the SES, the higher the self-reported morbidity and vice-versa. However, given the problems with the data and analysis, our results and the alternative that we therefore propose, should be taken as tentative. It would be fair to conclude that a full-fledged model of infections, illnesses, understanding of diseases by individuals; and the associative personal and medical diagnostic reactions and treatment options, is yet to be done.

What we have done is essentially scratching the surface. It is also clear that, the socio-economic variables at both personal and societal levels impacts responses of individuals, be they ill-persons or their proxies. Although, NSS over its several rounds has collected information relevant for a more thoroughgoing analysis, there are problems in analysing it. First, for the rounds prior to 42nd, unit-level data are not available. Second, the concepts and definitions change over rounds significantly. Yet, with unit-level data there should be enough observations to conceptualize a more satisfactory model. Dealing with lack of observations and/or methodological issues in combining aggregate data from the earlier rounds with unit-level data from the later NSSO rounds poses a significant challenge. We hope it would encourage scholars to research on this vital issue of health for our population.

Having sounded the cautionary notes, we take the opportunity to emphasize some of our findings. It is worth noting that, even though Kerala in India (always), and other southern states as well as West-Bengal and Punjab, most often show high morbidity rates in terms of inter-state percentage deviations from their corresponding All-India average; United Sates as well as other developed countries show even higher morbidity rates.

In order to supplement our analysis of morbidity, we seem to be the first to examine the inter-state variations in the average duration of sickness (in days). Very interestingly, in contrast to the incidence rates (prevalence rates as well as PAP), Kerala is no longer at the top but among the bottom with respect to average duration. Indeed, with a bit of hyperbole, one can say that the states included in the BIMARU category by the economic demographer Ashis Bose, often have higher duration of sickness than Kerala or Tamil Nadu or Andhra Pradesh. Unfortunately, we do not yet have a tried and tested clinical theory of ailments, their incidence, prevalence and duration for explaining their variations based on relevant exogenous variations.

In economic development, 'demographic transition' refers to a developing country transiting from a regime of high fertility, high mortality and low population growth to a state of low fertility, low mortality and low population growth, and indeed ideally to a regime of zero steady-state growth. Some countries like Japan, and those in Europe have unexpectedly moved to a regime of negative population growth with fertility falling to levels below the replacement levels. Interestingly many of the states in India, particularly in the south and the west, that have low morbidity rates and duration of sickness, have already reached or are close to their replacement rates of complete fertility. In our view, this calls for formulating a theory of joint health and demographic transition, and it is no coincidence that the two transitions seem contemporaneous.

Similarly, 'demographic dividend' refers to the gain in total factor productivity in countries yet to transit to a low fertility regime; and hence with positive rates of labour force growth, such countries can gain faster economic development through investment in health, education and skills of their labour force, leading to lower morbidity overall. The dynamic processes involved could be and in many countries, are contemporaneous and lead to total factor productivity growth. In India, the share of persons educated up to secondary-level seems to be lower while the rate of growth of labour force higher in the northern, rather than southern states, thus lowering the potential demographic dividend of the country.

Finally, we would like to stress again that the issues of morbidity and health we discuss involve complicated and interrelated dynamic processes that call for future research. We conclude by reiterating the increasing opportunity for theoretical and empirical work in related areas of international trade and migration, demography, health, and development on a national, regional and multilateral and global basis.

Appendix A: Concepts & Definitions Relating to NSSO Survey(s)

The primary purpose of this note is to summarise and review comparatively the findings on Indian morbidity from a pre-independence survey by All India Institute of Hygiene and Public Health in Singur, West Bengal in 1944 (Lal and Seal (1949)) and another by the Indian Statistical Institute, Kolkata, West Bengal in 1955 (Poti et al. 1955). In the National Sample Survey (NSS) "information on morbidity began to be collected from the 7th Round (October 1953–March 1954) onwards as an

exploratory measure to supply the methodology for future studies on a large scale. Information on morbidity collected in the NSS from 7th to the 13th Round (September 1957–May 1958),[7,8] but due to the emphasis on other aspects of information, the sampling intensity for morbidity had to be necessarily small and also the analysis of the data could not be taken up due to lack of resources" (NSS Report 49, 1961, Section 1.3, emphasis added). In fact, NSS collected morbidity data after the 13th Round until the 28th Round (October 1973–June 1974). In the 42nd Round (1986–1987), data on morbidity and expenses on medical services seem to have been collected. After a long hiatus, NSS again began collecting morbidity data as part of its surveys from 1995 to 96 on Social Consumption, 52nd Round. Since then, two more rounds 60th (January–June 2004) and 71st Round (January–June 2014) has been completed.

NSS Surveys from 7th round (October 1953–March 1954) till 28th Round (October 1973–June 1974)

- The following four (A, B, C and D) categories of sickness according to time of commencement and of termination were adopted in the NSSO. These are in common with the international practices. Thus, the category 'A' relates to sicknesses beginning and ending within the reference period; category 'B' to sicknesses beginning within the reference period and continuing on the date of survey; category 'C' to sicknesses beginning before the reference period but ending within it, and, category 'D' to sicknesses beginning before the reference period and continuing on the date of survey.
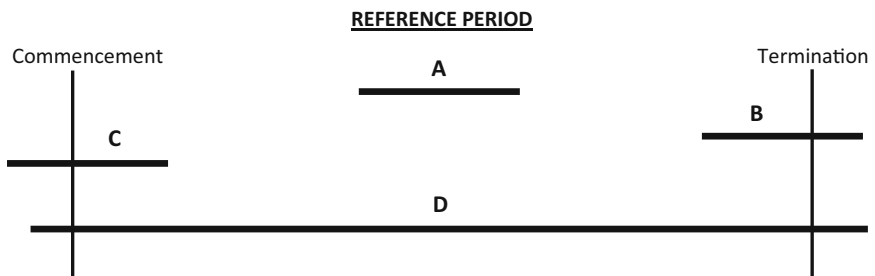
- Spell of sickness: A person was considered to be under one single spell of sickness, if the interval between the successive periods of sickness was less than three days with the same causes. [t]he word 'spell' and 'sicknesses are taken as equivalent.
- Incidence rate: the incidence rate recommended by the expert committee on health statistics of the world health organization (who) to be defined as, "the

---

[7] The results of the 7th Round morbidity data are in NSS Draft Report 47. They are quoted for comparison and new results in NSS Report 49 which gives the results of the examination of the morbidity data from 7th to 13th Rounds.

[8] Only Urban sector was surveyed in the 13th Round for morbidity data.

measurement of frequency of illness commencing during a defined period", was computed from:

$$= \frac{Number\ of\ sicknesses\ beginning\ during\ the\ reference\ period}{Average\ Population\ during\ the\ survey} \times 1000$$

$$= \frac{A + B}{Average\ Population\ during\ the\ survey} \times 1000$$

It is usual to calculate the rate per average population at risk, that is, the average of the population between the two limits of the reference period. In the NSS, however, the population as obtained on the date of survey was taken as the base for simplicity of calculation.

- Prevalence rate: the rate was recommended by the world health organization (who) expert committee on health statistics "to be used to describe the measurement of frequency of illness in existence at any time during a defined period (that is, a year, a month, a week)". In the NSS, the prevalence rate were calculated as number of cases of sickness experienced during the reference period per 1000 population:

$$= \frac{A + B}{Average\ Population\ during\ the\ survey} \times 1000$$

- Average duration of sickness: the average duration of sickness calculated in this report was defined as the total weeks of sickness for a certain category divided by the number of spells in the category.

Information was collected on the sex, age, marital status, and industry (or activity) status, cause of sickness and duration of sickness (in weeks). The reference period for entering information on morbidity was the last month, i.e. 30 days preceding the date of the survey. Information on morbidity was also collected for persons who died during the reference period of a month, i.e., for other persons also who, if alive, would have been treated as members of the household. Information on morbidity was collected for persons who:

- Were confined to the bed for at least 24 h; or
- Abstained from taking the normal diet, i.e. had to live on sick diet appropriate to the nature of sickness, for at least 24 h; or
- Were unable to attend the normal duties and activities for at least 24 h; due to illness or injury.

The following cases were excluded, namely, pregnancy, delivery, puerperium and menstruation, not receiving any medical attention; "handicapped" conditions with fixed symptoms, and myopia, hyper-metropia and astigmatism; but injuries and accidents were included.

Quinquennial *'Social-Consumption'* Surveys of NSS Round(s) 52nd (1995–1996), 60th (January–June, 2004) and 71st (January–June 2014)

- Ailment: Ailment, i.e. illness or injury, meant any deviation from the state of physical and mental well-being. In these rounds whether a person suffered an ailment during a particular period, it was judged by some deviation from physical or mental well-being was felt by the person during the period subject to the following inherent limitations:

  - An ailment may not cause any necessity of hospitalisation, confinement to bed or restricted activity.
  - An ailment may be untreated or treated.

- For the purpose of this survey, ailments are inclusive of:

  - All types of injuries, such as cuts, wounds, haemorrhage, fractures and burns caused by an accident, including bites to any part of the body
  - Cases of abortion—natural or accidental

- However, the following:

  - Cases of sterilisation, insertion of IUD, getting MTP etc.
  - A state of normal pregnancy without complications
  - Cases of pre-existing visual, hearing, speech, locomotor and mental disabilities were not included in ailment

- Spell of ailment: a spell is a continuous period of sickness due to a specific ailment.

Appendix B: Econometric Modelling

To analyse the interaction between the socio-economic status of an individual and his/her surrounding environment with that of self-reported morbidity, the following model(s) were estimated.

Logistic Regression Model

Given y_i^* is a latent variable, defined as the propensity of a person 'i' (member or head of household) to correctly report an illness for herself (or for other household

members), it is assumed to be linked to the observed socio-economic variables through the following structural model:

$$y_i^* \{P(A) = 1\} = \beta_0 + \beta_i \cdot X_i + \varepsilon_i.$$

The latent variable is linked to the observed binary y_i, which is equal to 1, if an illness is reported, or 0 otherwise, through the following measurement equation:

$$y_i = \left\{ \begin{array}{ll} 1 & if\ y_1^* > \tau \\ 0 & otherwise \end{array} \right\}$$

where τ is the threshold point for event 'A' to occur. Event 'A' is defined as the occurrence of a person reporting an illness to the surveying researcher. X_i is vector of socio-economic variables of interest—gender, age, social group, religion, educational attainment and so on. β_i is a vector of parameters. The error term ε_i is assumed to follow a standard logistic distribution with mean 0 and variance $\pi^2/3$. Maximum likelihood estimation procedure is employed to estimate the above model resulting into consistent, asymptotically normal and asymptotically efficient estimates. In order to facilitate interpretation, coefficient estimates are converted into odds-ratio (OR) as follows:

$$P(y = 1|X_i) = \frac{\exp(\beta_i \cdot X_i)}{1 + \exp(\beta_i \cdot X_i)}.$$

Linear Probability Model (LPM) and OLS-based Model for Average Duration of Sickness

The linear probability model is a regression model applied to a binary variable y_i, which is equal to 1, if an illness is reported, or 0 otherwise:

$$y_i = \beta_0 + \beta_i \cdot X_i + \varepsilon_i$$

where X_i is vector of socio-economic variables (all categorized) and β_i is a vector of parameters. ε_i, the error term is assumed to follow a normal distribution with mean 0 and variance σ^2. The β vector is interpreted as the predicted probability of person 'I' reporting an ailment if he/she belongs to the said category, keeping all other things constant. The model is estimated through a simple OLS-based procedure.

To study average duration of sickness (in days), we use a similar structural model as defined above with y_i now representing avg. duration of sickness (in days) for individual 'i'. The parameter vector is now interpreted as in the case of a simple linear regression model.

Appendix C: Econometric Results

Appendix D: Tables & Figures

See Tables 4, 5, 6 and Figs. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11.

Table 4 Morbidity rates from first set of surveys

Sector	Round	Survey period	Incidence rate			Prevalence rate		
			Male	Female	Total	Male	Female	Total
Rural	7th	Oct 1953–Mar 1954	50.18	48.08	49.14	68.17	61.29	64.77
	11th	Aug 1956–Jan 1957	30.20	22.21	26.39	35.23	27.93	31.66
	12th	Feb 1957–Jul 1957	39.50	35.23	37.38	59.92	53.47	56.73
Urban	7th	Oct 1953–Mar 1954	46.51	44.57	45.58	65.01	58.16	61.75
	11th	Aug 1956–Jan 1957	35.19	28.99	32.25	38.60	32.17	35.55
	12th	Feb 1957–Jul 1957	56.53	49.62	53.19	83.91	78.78	81.43
	13th	Sept 1957–May 1958	34.98	34.62	34.81	41.59	42.49	42.01

Table 5 Morbidity rates from second set of surveys

Sector	Round	Survey period	Incidence rate				Prevalence rate			
			By weeks prior to survey				By weeks prior to survey			
			1	2	3	4	1	2	3	4
Rural	17th	Sept 1961–Jul 1962	18.37	26.62	22.13	18.49	81.92	83.94	69.96	54.07
Urban	17th	Sept 1961–Jul 1962	23.95	29.09	23.67	18.39	79.19	77.34	62.77	45.9
			By gender				By gender			
			Male	Female	Total		Male	Female	Total	
Rural	28th	Oct 1973–Jun 1974	13.53	11.55	12.57		24.11	20.70	22.46	
Urban	28th	Oct 1973–Jun 1974	13.87	13.14	13.53		35.49	28.3	32.18	

Table 6 Morbidity rates from third set of surveys

| Sector	Round	Survey period	PAP per 1000 persons		
			Male	Female	Total
Rural	42nd	1986–1987	64	63	64
	52nd	1995–1996	54	57	55
	60th	Jan–Jun 2004	83	93	88
	71st	Jan–Jun 2014	80	99	89
Urban	42nd	1986–1987	30	33	31
	52nd	1995–1996	51	58	54
	60th	Jan–Jun 2004	91	108	98
	71st	Jan–Jun 2014	101	135	118
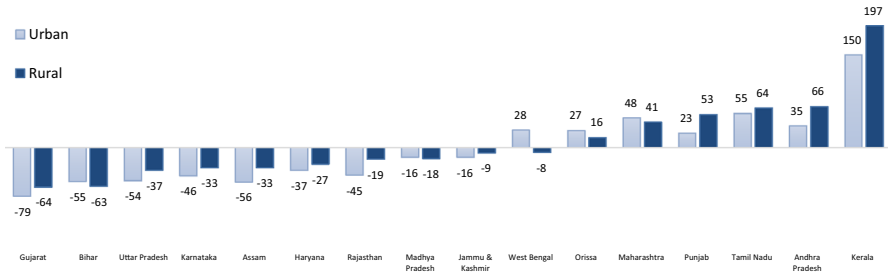
Fig. 1 Inter-state percentage deviation of incidence rate from all-India level: NSSO 28th round (1973–1974)
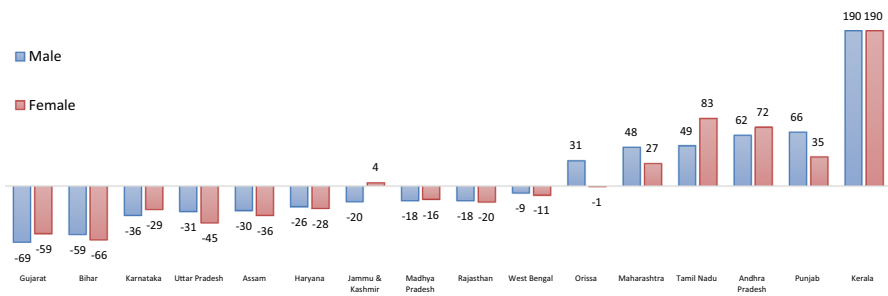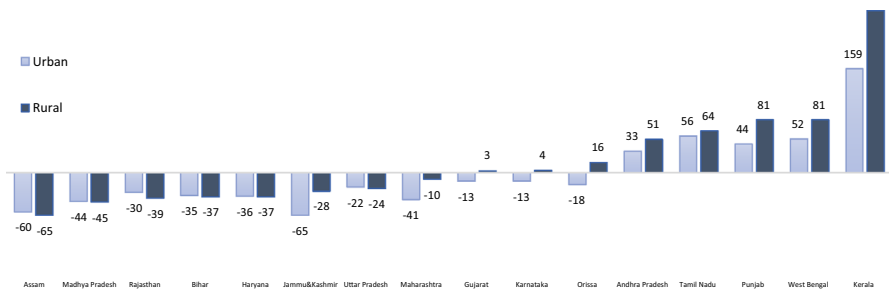

Fig. 2 Inter-state percentage deviations of rural incidence rate from all-India level: by gender; NSSO 28th round (1973–1974)
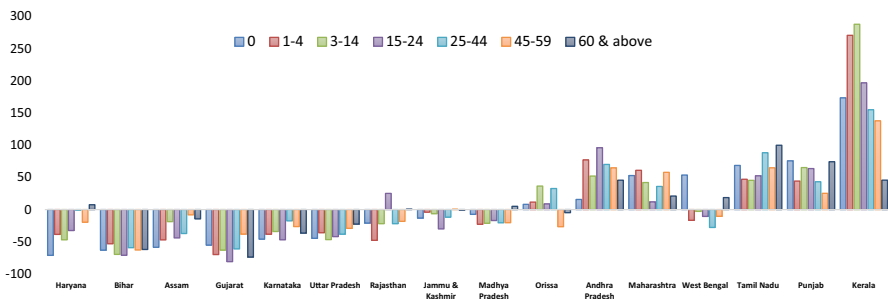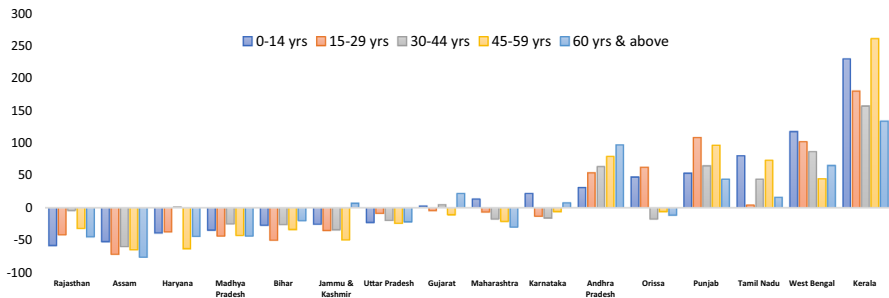

Fig. 3 Inter-state percentage deviations of rural incidence rate from all-India level: by age; NSSO 28th round (1973–1974)
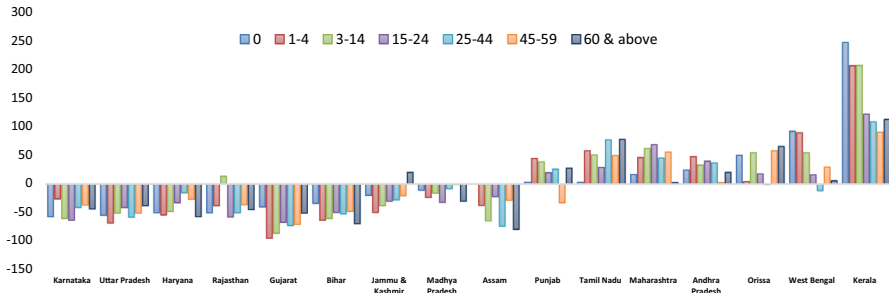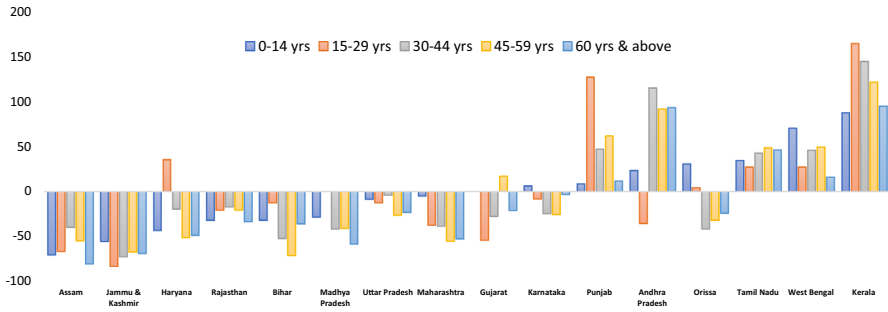
Fig. 4 Inter-state percentage deviations of urban incidence rate from all-India level: by age; NSSO 28th round (1973–1974)
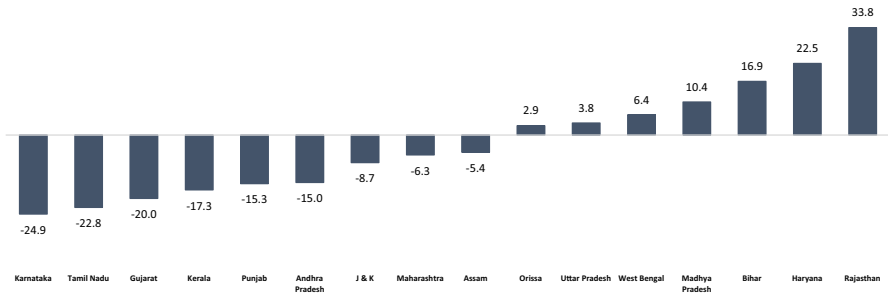

Fig. 5 Inter-state percentage deviation of PAP from all-India Level: NSSO 71st round (2014)
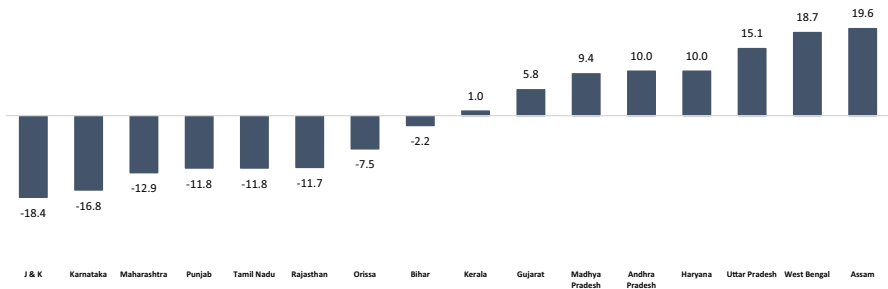

Fig. 6 Inter-state percentage deviations of rural PAP from all-India level: by age; NSSO 71st round (2014)
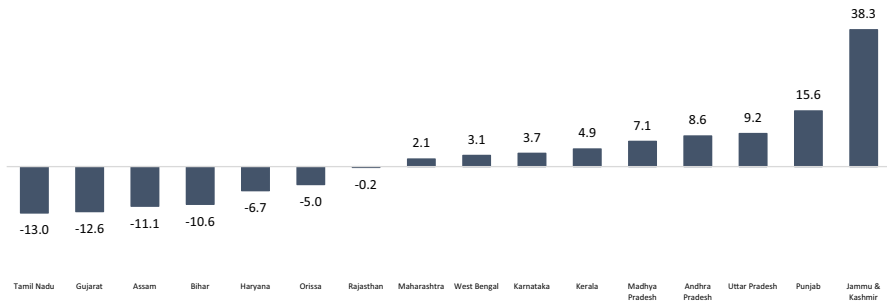
Fig. 7 Inter-state percentage deviations of urban PAP from all-India level: by age; NSSO 71st round (2014)
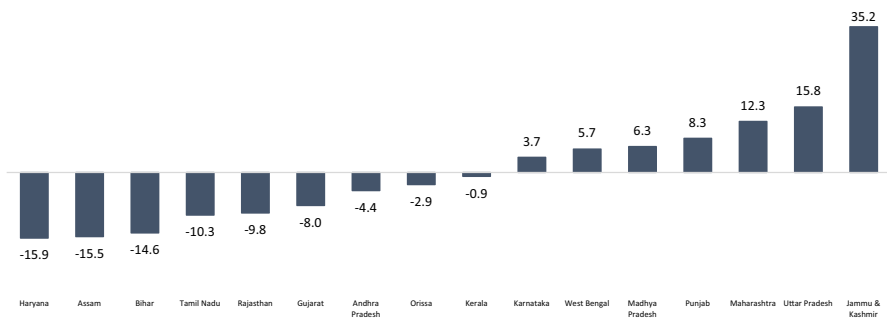

Fig. 8 Percentage deviations from national average: avg. duration (days) of temporary ailments terminating into recovery for rural HHs; NSSO 28th round (1973–74)

Fig. 9 Percentage deviations from national average: avg. duration (days) of temporary ailments terminating into recovery for urban HHs; NSSO 28th round (1973–1974)

Fig. 10 Percentage deviations from national average: avg. duration (days) of temporary ailments terminating into recovery for rural HHs; NSSO 71st round (2014)

Fig. 11 Percentage deviations from national average: avg. duration (days) of temporary ailments terminating into recovery for urban HHs; NSSO 71st round (2014)

References

Deaton, A. (2013). *The great escape: Health, wealth, and the origins of inequality*. Princeton: Princeton University Press.

Dilip, T. R. (2002). Understanding levels of morbidity and hospitalization in Kerala, India. *Bulletin of the World Health Organization, 80*(9), 746–751.

Heckman, J. J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. In Annals of Economic and Social Measurement, 5(4), 475–492. NBER.

Lal, R. B., Seal, S. C., & All India Institute of Hygiene and Public Health. (1949). *General rural health survey, Singur Health Centre,* 1944. Calcutta: Printed by Govt. of India Press.

National Sample Survey Office (1961). Report on morbidity (Report No. 49), NSSO, New Delhi.

National Sample Survey Office (1968). Report on pilot enquiry on morbidity (NSSO 17th Round, September 1961–July 1962, Report No. 129), NSSO, New Delhi.

National Sample Survey Office (1969). Special study on morbidity (November 1960–October 1961, Report No. 119), NSSO, New Delhi.

National Sample Survey Office (1980). Tables with notes on morbidity (NSSO 28th Round, October 1973–June 1974, Report No. 292), NSSO, New Delhi.

National Sample Survey Office (1989). Survey on utilisation of medical services [NSSO 42nd Round, July 1986–June 1987, Report No. 364 (42/25.7/1)], NSSO, New Delhi.

National Sample Survey Office (1998). Morbidity and treatment of ailments (NSSO 52nd Round, July 1995–June 1996, Report No. 441 (52/25.0/1)], NSSO, New Delhi.

National Sample Survey Office (2006). Morbidity, health care and the condition of the aged [NSSO 60th Round, January–June 2004, Report No 507(60/25.0/1)], NSSO, New Delhi.

National Sample Survey Office (2015). Key Indicators of Social Consumption in India: Health (NSSO 71st Round, January–June 2014), NSS KI (71/25.0), NSSO, New Delhi.

Poti, S., Raman, M., Biswas, S., & Chakraborty, B. (1959). A Pilot Health Survey in West Bengal: 1955. Sankhyā. *Indian Journal of Statistics (1933-1960), 21*(1–2), 141–204.

Prinja, S., Jeet, G., & Kumar, R. (2012). Validity of self-reported morbidity. *Indian Journal of Medical Research, 136*(5), 722.

Sen, A. (1992). *Objectivity and position*. Kansas: University of Kansas, Department of Philosophy.

Sen, A. (1993). Positional objectivity. *Philosophy & Public Affairs, 22*(2), 126–145.

Sen, A. (2002). Health: perception versus observation: Self-reported morbidity has severe limitations and can be extremely misleading. *BMJ. British Medical Journal, 324*(7342), 860.

Subramanian, S. V., Nandy, S., Irving, M., Gordon, D., Lambert, H., & Davey Smith, G. (2006). The mortality divide in India: The differential contributions of gender, caste, and standard of living across the life course. *American Journal of Public Health, 96*(5), 818–825.

Subramanian, S. V., Subramanyam, M. A., Selvaraj, S., & Kawachi, I. (2009). Are self-reports of health and morbidities in developing countries misleading? Evidence from India. *Social Science and Medicine, 68*(2), 260–265.