USC Dornsife Center for Economic and Social Research

USC Schaeffer Leonard D. Schaeffer Center for Health Policy & Economics

Education, Gender, and State-Level Gradients in the Health of Older Indians: Evidence from Biomarker Data

Jinkook Lee, Mark E. McGovern, David E. Bloom, P. Arokiasamy, Arun Risbud, Varsha Kale, Jennifer O'Brien, Peifeng Hu

Paper No: 2014-007

CESR-SCHAEFFER WORKING PAPER SERIES

cesr.usc.edu **healthpolicy.usc.edu**

Education, Gender, and State-Level Gradients in the Health of Older Indians:

Evidence from Biomarker Data

Jinkook Lee[1], Mark E. McGovern[2], David E. Bloom[2], P. Arokiasamy[3],

Arun Risbud[4], Varsha Kale[4], Jennifer O'Brien[2] & Peifeng Hu[5]

This project is funded by the National Institute on Aging/National Institutes of Health (R21AG034443, R01 AG030153)

[1] Contact author, Professor, Davis School of Gerontology and Dornsife Center for Economic and Social Research, University of Southern California, and Senior Economist, RAND Corporation, jinkook.lee@usc.edu, Tel. 213-740-1557, 3715 McClintock Ave., Los Angeles, CA 90089

[2] Harvard School of Public Health

[3] International Institute for Population Sciences, India

[4] National AIDS Research Institute, India

[5] Division of Geriatric Medicine, University of California, Los Angeles

Abstract

This paper examines health disparities in biomarkers among a representative sample of Indians aged 45 and older, using data from the pilot round of the Longitudinal Aging Study in India (LASI). We document an educational gradient in hemoglobin (Hb) level, a marker used for diagnosing anemia. Survey respondents with no schooling have substantially lower Hb levels (0.6 g/dL less in the adjusted model) than those with some formal education. Individuals among the oldest old and those with greater body-mass index (BMI) had higher levels of C-reactive protein (CRP), an indicator of inflammation and a risk factor for cardiovascular disease. We find no evidence of educational or gender differences in CRP but do find respondents living in rural areas have CRP levels that are 0.7 to 0.8 mg/L lower than urban ones. We also find state-level disparities, with Kerala residents exhibiting the lowest CRP levels (1.96 mg/L compared to the overall mean of 3.28 mg/L in Rajasthan). This finding is consistent with data on access to health-care services across states. We use the Oaxaca-Blinder decomposition approach to explain group-level differences, and find that state-level gradients in CRP are mainly due to heterogeneity in the effect of the observed characteristics of respondents, as opposed to differences in the distribution of endowments across the sampled state populations.

Research Highlights

- There have been few efforts to analyze objective population health data on non-communicable diseases in India. Using such data for India, we explore levels of hemoglobin and C-reactive protein.
- We find a strong education gradient in hemoglobin, a marker for nutrition-based anemia.
- C-reactive protein, a marker of cardiovascular risk, is highest among the oldest and in rural areas.
- We do not find an education gradient in C-reactive protein, suggesting that cardiovascular risk may be a health concern for older Indians across all socioeconomic statuses.
- We find state-level differences with respect to cardiovascular risk.

Keywords: Biomarkers, Health Disparities, Gender Differences, Oaxaca-Blinder Decomposition, India, Aging, Cardiovascular disease, Anemia

1. Introduction

Many developing countries are undergoing rapid demographic and economic transitions. In particular, the proportion of individuals over the age of 50 in low and middle income nations is expected to rise rapidly in coming decades (Shetty, 2012). This is the result of an ongoing epidemiological transition in which life expectancy is rising and mortality is shifting more towards later life (Prentice, 2006). The main risks for premature death are no longer solely the well-studied problems associated with poverty, such as malnutrition and poor sanitation. In India and China non-communicable diseases (NCDs), such as cardiovascular disease (CVD), which had been largely limited to higher-income countries, are becoming the main causes of premature mortality (Kearney et al., 2005). For example, NCDs now account for 60% of all deaths in India (WHO, 2014). The contribution to mortality of non-communicable relative to communicable diseases is expected to rise substantially in coming years. Given that India and China comprise one third of the world's population, how these changes affect well-being in these two countries will greatly affect global welfare. As well as being inherently important health outcomes, these diseases have important economic consequences. Non-communicable diseases exert large monetary and non-monetary impacts on society (Bloom et al., 2013). NCDs also make a substantial contribution to morbidity, accounting for 54% of healthy life years (DALYs) lost in India (Murray et al., 2012). At the same time, malnutrition and communicable diseases remain substantial health threats in low and middle income countries (Narayan et al., 2010). For example, 70% of women and children in India currently suffer from anemia (Balarajan et al., 2011).

Rapid economic growth has contributed to better living standards, but has also led to changing patterns of urbanization, diet, and other modifiable risk behaviors (Prentice, 2006). Coupled with existing excess mortality due to communicable disease, these differences in demographic, behavioral, and economic circumstances create a challenging environment for public health officials in developing countries. This challenge is exacerbated by difficulties in measuring the extent of health inequality. There are major differences in access to health care by gender, region, and level of education within

countries. Understanding group-level differences in health outcomes is therefore important for establishing policy priorities.

Given that NCDs are increasingly affecting a growing portion of the population (Lopez et al., 2006), their impact warrants greater attention. Cardiovascular disease is now the second-most important contributor to mortality in India, accounting for 28% of deaths (WHO, 2005). Yet CVD has been largely ignored in public discourse due to the perception that heart disease is mainly a problem of the urban rich. While CVD increasingly affects all socioeconomic groups, the problem of inadequate nutrition remains a problem for a substantial proportion of the less well-off in India, especially for women (Bentley and Griffiths, 2003). India has the highest incidence of anemia in the world, with levels that have remained static for the past decade, despite economic growth (Balarajan et al., 2011).

Both cardiovascular disease and malnutrition thus have a substantial impact on public health in India, and quite likely in other low and middle-income countries too. However, good evidence on the prevalence of malnutrition and CVD in India can be difficult to obtain, particularly at the regional level, where there is varying access to medical services and diagnosis. There are two reasons for this.

First, self-reporting on health in India is problematic due to differential state-level access to care, diagnosis, and treatment. This access affects the extent to which individuals are aware of their health status, their recall bias, and differences in how they perceive their health (Johnston et al., 2009). This can result in heterogeneity in the thresholds used by respondents for indicating that they suffer from a medical condition (Sen, 2002). For example, there are large differences in hypertension diagnosis evident in self-reported and measured hypertension (Lee et al. 2012). Figure 1 illustrates these by educational level. These discordances differ significantly by state as well.

[FIGURE 1 HERE]

Second, among lower- and middle-income countries, there is a dearth of nationally representative data on the objective biological markers of malnutrition and CVD. In particular, there is little existing

evidence from these countries about the risks for NCDs among persons aged 45 years and older, the population most likely to be affected by these conditions (Chaves et al., 2005). If health service records are available, they do not provide a complete picture of population health because they only provide information on people seeking diagnosis or treatment. Such individuals may not be a representative sample of the population, especially if access to health care is low.

Although much analysis focuses on country as a whole, regional differences within them are important as well. For example, India is a union of 29 states and 6 territories, which vary greatly in their economic development, cultures, education levels, and policies (Lee and Smith, 2014). These differences may lead to cross-state variation in risks for these health outcomes.

Assessing health outcomes accurately is important for understanding the successes and failures of alternative policies and environments, as well as for understanding where to target interventions. Differential health outcomes are evident when comparing urban and rural groups, states, and genders. Table 1 illustrates some of these disparities in four Indian states. While all four of these states saw urban consumption increase substantially between the 1960s and the 1990s, there was little improvement in rural consumption between 1960-1961 and 1993-1994 with the exception of Kerala. Kerala, also had more favorable statistics for females, including a higher ratio of girls to boys and higher school attendance rates.

[TABLE 1 HERE]

Using biomarkers to directly assess risks for particular outcomes provides a potential solution to the lack of good health information, while also providing an immediate assessment of objective health disparities for both individuals and group. CVD and anemia appear to be especially important public health issues in India, given the increasing importance of heart disease and the well-documented lack of adequate nutrition (Bentley and Griffiths, 2003). This is particularly true for older individuals (Carmel, 2001, Chaves et al., 2005).

Two attractive candidates for targeted biomarkers data collection are C-reactive protein (CRP) and hemoglobin (Hb). CRP is a biomarker for inflammation, which is associated with increased risk of cardiovascular disease (Vikram et al., 2003). Elevated levels of CRP are also associated with hypertension and diabetes, with thresholds defined by the Center for Disease Prevention and Control and the American Heart Association (Myers et al., 2004). The Hb biomarker can be used to evaluate prevalence of anemia (Balarajan et al., 2011). Anemia results from a lack of either red blood cells or hemoglobin, and leads to weakness or fatigue (Aguayo et al., 2003, Beghé et al., 2004, Denny et al., 2006). Anemia is often associated with iron or vitamin deficiencies due to poor nutrition.

Data collection on the prevalence of anemia has mainly focused on preschool-age children, pregnant women, and non-pregnant women of reproductive age, particularly in India (Balarajan et al., 2013, Bentley and Griffiths, 2003, Ghosh, 2009). The WHO does not report country-level estimates for school-age children, men, and the elderly for this reason (WHO, 2008). Efforts such as the Longitudinal Aging Study in India (LASI), which recently collected data on a variety of biomarkers from the elderly, are helping shift the focus of data collection. LASI is designed to be representative of both India as a whole and of its constituent states, and will ultimately follow more than 50,000 respondents longitudinally (Arokiasamy et al., 2012). The survey includes respondents aged 45 and older, as well as their spouses (regardless of age).

In 2010, LASI collected pilot data in four states: Punjab, Rajasthan, Kerala, and Karnataka, interviewing 1,683 eligible individuals. Of these, 78% (1,305 respondents) provided a dried blood spot (DBS) sample. Individual and household micro data are publically available, and the biomarker data are also accessible through an application for a restricted-data file (at the USC Gateway to Global Aging Data: www.g2aging.org). One of LASI's main contributions is to assess risks for CRP and anemia using objective biomarker data from nationally and in-state representative samples of older Indians. We use these data to address a number of important research questions.

First, given the public discourse on socioeconomic status (SES) and cardiovascular disease, we assess whether there is an education gradient in CRP. We also examine differences between urban and rural areas. Previous research found no evidence of an SES gradient in CRP in Costa Rica (Rosero-Bixby and Dow, 2009), another country experiencing demographic, economic, and epidemiologic transitions. We hypothesize that we will find no SES gradient in CRP in India.

Second, given the varying growth and social policies of Indian states in recent decades, we examine state-level variation in CRP. We aim to determine whether differences in economic growth and social policies contribute to state-level variation in CRP.

Third, we establish whether there is an education gradient for hemoglobin. Given the existing literature, we hypothesize the existence of socioeconomic disparities for this outcome (Bentley and Griffiths, 2003).

Fourth, we examine the risk factors associated with anemia separately for men and women, given the much higher incidence of anemia among women (Balarajan et al., 2011).

In Section 2, we present our data and analytic approach. In Section 3, we present our results. We present our conclusions in Section 4.

2. Data and Analytic Approach

The 2010 LASI pilot sample was drawn using a stratified, multistage, area probability sampling design based on the 2001 Indian Census. From each state, we randomly chose two Census 2001 districts. We then randomly selected eight primary sampling units (PSUs) from each district to match the urban/rural share of the state population. Finally, we selected 25 community-residing households through a random sampling from each PSU.

LASI has two main modules: the household and individual interview, and the biomarker collection. The household interview asks about physical environment and household finances. The individual

interview asks about demographics, family, social activities, health and health behaviors, and work and pension. For the collection of dried blood spots (DBS), respondents provided separate consent permitting interviewers to prick their finger and place five drops of blood on a Whatman 903 Protein Saver card. The collected DBS cards were left to air dry for at least 4 hours, then sent to the National AIDS Research Institute (NARI) in Pune, India, where they were stored under -20^{o}C and later assayed.

Both CRP concentrations and hemoglobin levels in the DBS specimens were measured using a validated enzyme-linked immunosorbent assay (ELISA) method. CRP concentrations were measured using the method developed by McDade et al. (2004), and hemoglobin levels were measured using the method developed by O'Broin and Gunter (1999). To ensure quality, all samples, standards, and controls were measured in duplicate. Internal quality controls were run on every plate, and plates with out-of-range quality-control values were re-run. We further ensured the quality of laboratory assay results through periodic use of external quality control samples prepared by the USC/UCLA Center on Biodemography and Population Health.

We also externally validated work at the NARI laboratory (for further details of the LASI biomarker data collection and the external validation work done at the NARI laboratory, please see: http://www.rand.org/pubs/working_papers/WR1043.html). For CRP assays, we compared NARI's results on 32 validation samples with DBS-based values from the reference laboratory in the United States (at the University of Washington). The correlation coefficient was 0.95. In addition to DBS-based hemoglobin levels from NARI, 33 validation samples had venous-based results from the UCLA Clinical Laboratory. The Pearson's correlation coefficient was 0.78. In general, NARI had higher Hb values than the corresponding venous-based results. The average difference was 0.55 gram/dL (standard deviation: 0.86 gram/dL).

For measures of both Hb and CRP, DBS specimens were run in duplicate. We combined the duplicate measures into an average. Both were very highly correlated, and using either the first or second

measure on its own had little effect on the results. Table 2 presents descriptive statistics for the analysis of the two dependent variables (hemoglobin, measured in grams per deciliter, and CRP, measured in milligrams per liter) and independent variables (age, gender, state, caste, urban/rural residency, education, smoking, body mass index, and self-reported diagnosis of diabetes, stroke, hypertension, or a heart condition). We focus on respondents over the age of 45, excluding a small number of spouses under this age. This left 1,150 observations in total. There were a small number of missing values for some covariates and outcomes but these did not exceed 6% of observations. For example of the 1,150 total respondents, 1,077 had information on Hb. The mean Hb in the sample was 14.3 g/dL, and mean CRP was 2.7 mg/L. Table 2 also shows the proportion of respondents with anemia (20%) as determined by Hb levels (below 12 g/dL for women and below 13 g/dL for men), as well as those at high risk for cardiovascular disease (30%), as determined by CRP levels of more than 3 mg/L.

Nearly half the sample (46%) was between 45-54 years of age, nearly half (48%) were male, and more than half (55%) had received some formal schooling. There were roughly equal numbers of respondents in each state (ranging from 253 in Karnataka to 329 in Kerala). For covariates, we focus on pre-determined variables which are unlikely to be outcomes of socioeconomic status, health status, or health care use. Nevertheless, we also consider smoking, body mass index (BMI), and the presence of a self-reported chronic health conditions, as these have been associated with Hb and CRP (Beghé et al., 2004, Carmel, 2001, Daly, 2013, Danesh et al., 2004).

[TABLE 2 HERE]

Figure 2 presents the distribution of CRP in the LASI pilot sample, showing the high risk cut-off at 3 mg/L. We show the combined sample, as we find no gender differences for this outcome. The distribution is skewed to the left, although a substantial proportion has CRP above 3 mg/L (30%). Figures 3 and 4 show the Hb distribution stratified by sex, which in both cases approximates a normal distribution.

[FIGURE 2 HERE]

[FIGURE 3 HERE]

[FIGURE 4 HERE]

We then examine population-level differences in these outcomes, by estimating bivariate associations between biomarkers and explanatory variables. Tables 3 and 4 show this analysis for CRP and Hb, respectively. Each table shows for each subpopulation mean level, along with the associated standard error, the confidence interval, and a T-test for the difference between the mean of that group and the omitted category. For CRP, there are significant bivariate associations (at the 5% level) between the outcome and age of 75 or older (higher), rural location (lower), the states of Kerala and Karnataka (lower), and being obese (higher). The finding on obesity is consistent with existing evidence on the relationship between inflammation and obesity (Danesh et al., 2004).

Table 4 shows that being female and underweight are significantly associated with a lower level of Hb, as is a lack of formal schooling. Having a self-reported chronic health condition and smoking are associated with a higher Hb level. These results are consistent with a socioeconomic gradient in anemia, and with well-documented gender disparities (Rosero-Bixby and Dow, 2009).

[TABLE 3 HERE]

[TABLE 4 HERE]

These bivariate associations do not adjust for other variables, such as age and state, which are likely to be related to both the outcome and the covariates of interest. We seek to establish whether the findings above remain intact after full adjustment for all relevant covariates. For example, the relationship between Hb and education may be explained by the fact that earlier, now older, birth

cohorts have lower levels of educational attainment. Once we establish the significance of key covariates of interest in multivariate analysis, we decompose differences across groups using the Oaxaca-Blinder approach in order to understand the origin of these differences (Blinder, 1973, Liu et al., 2013, Oaxaca, 1973, Powell et al., 2012).

We begin by adopting the following regression model:

$$Biomarker_i = \alpha + \beta_1 Gender + \beta_2 State + \beta_3 Education + \beta_4 Age\ Group + \beta_5 Urban + \beta_6 Caste + \mu_i \quad (1)$$

CRP and Hb are modeled as a function of the covariates of interest in a linear regression model (OLS), which is adjusted for weighting and survey design. Our main coefficients of interest are β_1, β_2, and β_3, reflecting the adjusted association of the outcomes with gender, state, and education. Initially, we do not control for BMI, smoking, or the presence of a self-reported chronic health condition, as these could potentially be outcomes of education or the biomarkers themselves. We then extend the analysis to include these additional variables, and demonstrate that doing so has little effect on our conclusions. We present results for the regression models in the following section. For each outcome, we present both the pooled and gender-stratified analyses.

Given that we observe state-level differences in CRP, we adopt a decomposition approach in order to establish the relevant mechanisms behind these disparities. Conceptually, there are two possible reasons for observing differences in CRP levels between two states (state 1 and state 2): either a difference in the observed covariates ($X_{State\,1}$, $X_{State\,2}$), or a difference in the effects of the covariates ($\beta_{State\,1}$, $\beta_{State\,2}$). Under the standard assumptions for the OLS model where the estimates for β_i are unbiased, we can express this as:

$$E(CRP_{State\ 1}) - E(CRP_{State\ 2}) = E(X_{State\ 1})'\beta_{State\ 1} - E(X_{State\ 2})'\beta_{State\ 2} \quad (2)$$

It follows that (2) can be written in terms of differences in endowments (i.e., establishing what $State\ 2$ outcomes would be if $State\ 2$ had $State\ 1$'s endowments), differences in the relevant coefficients (i.e., establishing what $State\ 2$ outcomes would be if $State\ 2$ had $State\ 1$'s coefficients), and an interaction effect between the two:

$$E(CRP_{State\ 1}) - E(CRP_{State\ 2}) =$$

$$[E(X_{State\ 1}) - E(X_{State\ 2})]'\beta_{State\ 2} +$$

$$E(X_{State\ 2})'(\beta_{State\ 1} - \beta_{State\ 2}) +$$

$$[E(X_{State\ 1}) - E(X_{State\ 2})]'(\beta_{State\ 1} - \beta_{State\ 2}) \quad (3)$$

Specifically, we decompose state level differences in CRP using this approach.

3. Results

Table 5 presents results for the multivariate analysis of CRP. Column 1 shows the results for the pooled sample, controlling for age, gender, state, caste, location, and education. Column 2 adds the potentially endogenous variables (smoking, BMI, and self-reported diagnosis of diabetes, stroke, hypertension, or a heart condition). Columns 3 and 4 replicate these specifications for men, while columns 5 and 6 focus on women only.

The overall results are similar to the bivariate analysis. In the pooled sample, the oldest old (those aged 80 and older) have higher CRP levels, while those in rural areas have lower levels, as do those in Kerala. Adding the additional control variables has little effect on these existing coefficients, although doing so indicates that higher BMI is associated with higher CRP. Prior diagnosis of diabetes, stroke, hypertension, or a heart condition has a marginally significant positive association with CRP. Stratification by gender indicates some differences in the effects of covariates. For example, the age gradient is only present among women. For men, controlling for BMI effectively removes the association with urban/rural residence, suggesting that differences in obesity explain the difference in CRP risk between urban and rural areas. For women, urban residence still has a significant, albeit smaller, positive association with CRP after controlling for BMI.

Table 6 presents the corresponding analysis for Hb. There is clear evidence for gender and education gradients in the pooled sample, with women having around 2g/dL less Hb than men, and with persons with no formal schooling having about 0.6 g/dL less Hb than those with some education. These results are similar to the bivariate estimates. There are also some gender differences in the coefficients in the stratified model. Older men (ages 75+) have reduced levels of Hb, while BMI is positively associated, perhaps reflecting a marker of nutritional status. Smoking is also positively associated with Hb for men, but not for women.

As a robustness check, we implemented binary regression models (logit) for the cut-offs for anemia and cardiovascular risk, and found similar results. Unweighted results are also similar. Finally, we considered a propensity score matching approach for education, and found similar conclusions regarding an SES gradient in Hb: those with no formal schooling have lower levels of Hb, with effects comparable to those shown in Table 6.

[TABLE 5 HERE]

[TABLE 6 HERE]

Finally, we implement the Oaxaca-Blinder approach to analyze state-level differences in CRP, given that we find substantial disparities in table 5. Table 7 shows the relevant means for the base state (Rajasthan), compared to each of the others. Column 1 uses the regression coefficients from the base model in table 5 above, while column 2 controls for BMI, smoking, and prior diagnosis of diabetes, stroke, hypertension, or a heart condition. The decomposition shows the mean difference, the estimated contribution from endowments, the coefficients, and their interaction.

Overall, the difference is only significant between Rajasthan and Kerala, and between Rajasthan and Karnataka. The decomposition indicates that these differences are mainly due to differences in the effect of the explanatory variables, rather than the distribution of endowments across states. The additional controls have little effect on these results.

[TABLE 7 HERE]

4. Conclusions

Using new data from a representative sample of older Indians, we investigate the relationship between respondents' characteristics and objective measures of their health. We demonstrate the feasibility and value of collecting population-based biomarkers among those aged 45 years and older in a developing country. We find that about one- third of Indians have a CRP level considered to be high risk (>3 mg/L), which is comparable to results from the English Longitudinal Study on Ageing (Hamer and Molloy, 2009).

We find that CRP is greater among the oldest old and among urban residents. Although there are substantial state-level differences, there is no evidence of an education gradient for CRP, which is consistent with existing evidence from Costa Rica (Rosero-Bixby and Dow, 2009). These findings suggest that CVD is a health problem affecting individuals of all socioeconomic groups. When we decompose state-level differences, we find that these disparities are mainly due to differences in the effects of risks rather than in their distribution.

Kerala has the lowest CRP levels, raising the question of why individuals in Kerala and Karnataka as opposed to other states like Rajasthan and Punjab carry less risk of CRP as they grow older. As a follow-up to this analysis, we considered policies that might modify the effect of these endowments. As demonstrated in Figure 5, our finding of state differences in CRP is consistent with group-level access to health care; Kerala and Karnataka have the highest proportions of individuals who have visited a formally qualified doctor.

[FIGURE 5 HERE]

We find evidence for an education gradient in Hb, but there is no evidence of state-level differences. Despite economic growth, across all states the risk of anemia associated with malnutrition is higher for women and for those without schooling.

Overall, biomarkers provide a useful complement to other health measures when determining the health status of older individuals in lower and middle income countries, as well as the health disparities between groups within these countries. The use of biomarkers may help to overcome the drawbacks associated with self-reported health measures.

There are important limitations to this study. First, we are only able to document associations. Further data is required to establish, for example, whether the observed relationship between Hb and education is causal. Second, we are presently only able to consider the four pilot states. Future, nationally representative waves of LASI will enable us to examine whether these findings extend to other Indian states, and will provide valuable data for answering other important research questions.

References

Aguayo, V. M., Scott, S., & Ross, J. (2003). "Sierra Leone–investing in nutrition to reduce poverty: a call for action." Public Health Nutrition **6**(07): 653-657.

Arokiasamy, P., Bloom, D., Lee, J., Feeney, K., & Ozolins, M.et al. (2012). Longitudinal aging study in India: vision, design, implementation, and some early results. In J. P. Smith and M. Majmundar (Eds.) Aging in Asia: findings from new and emerging data initiatives (36-74). Washington, D.C.: National Academies Press.

Balarajan, Y., Ramakrishnan, U., Özaltin, E., Shankar, A. H., & Subramanian, S. V. (2011). "Anaemia in low-income and middle-income countries." The Lancet **378**(9809): 2123-2135.

Balarajan, Y. S., Fawzi, W. W. & Subramanian, S. (2013) Changing patterns of social inequalities in anaemia among women in India: cross-sectional study using nationally representative data. *BMJ open,* 3.

Beghé, C., Wilson, A., & Ershler, W. B. (2004). "Prevalence and outcomes of anemia in geriatrics: a systematic review of the literature." The American Journal of Medicine **116**(7, Supplement 1): 3-10.

Bentley, M. & Griffiths, P. (2003). "The burden of anemia among women in India." European Journal of Clinical Nutrition **57**(1): 52-60.

Blinder, A. S. (1973). "Wage discrimination: reduced form and structural estimates." Journal of Human resources: 436-455.

Bloom, D. E., Cafiero, E. T., McGovern, M. E., Prettner, K., Stanciole, A., Weiss, J., Bakkila, S., & Rosenberg, L. (2013). The Economic Impact of Non-communicable Disease in China and India: Estimates, Projections, and Comparisons, National Bureau of Economic Research.

Carmel, R. (2001). "Anemia and aging: an overview of clinical, diagnostic and biological issues." Blood reviews **15**(1): 9-18.

Chaves, P. H., Semba, R. D., Leng, S. X., Woodman, R. C., Ferrucci, L., Guralnik, J. M., & Fried, L. P. (2005). "Impact of anemia and cardiovascular disease on frailty status of community-dwelling older women: the Women's Health and Aging Studies I and II." The Journals of Gerontology Series A: Biological Sciences and Medical Sciences **60**(6): 729-735.

Daly, M. (2013). "The relationship of C-reactive protein to obesity-related depressive symptoms: A longitudinal study." Obesity **21**(2): 248-250.

Danesh, J., Wheeler, J. G., Hirschfield, G. M., Eda, S., Eiriksdottir, G., Rumley, A. . . . & Gudnason, V. (2004). "C-reactive protein and other circulating markers of inflammation in the prediction of coronary heart disease." New England Journal of Medicine **350**(14): 1387-1397.

Datt, G. (1998). Poverty in India and Indian states: an update, FCND Discussion Paper No. 47, Food Consumption and Nutrition Division, International Food Policy Research Institute, Washington, D.C.

Denny, S. D., Kuchibhatla, M. N., & Cohen, H. J. (2006). "Impact of Anemia on Mortality, Cognition, and Function in Community-Dwelling Elderly." The American Journal of Medicine **119**(4): 327-334.

Ghosh, S. (2009) Exploring socioeconomic vulnerability of anaemia among women in eastern Indian States. Journal of biosocial science, 41**,** 763-787.

Government of India, Ministry of Human Resource Development, Department of Education, Planning, Monitoring and Statistics Division (2008). Selected Educational Statistics: MHRD, 2000 – 2001, http://www.educationforallinindia.com/page20.html/ (accessed on March 6, 2012).

Hamer, M. & Molloy, G. J. (2009) Association of C-reactive protein and muscle strength in the English Longitudinal Study of Ageing. Age, 31**,** 171-177.

Johnston, D. W., Propper, C., & Shields, M. (2009). "Comparing subjective and objective measures of health: Evidence from hypertension for the income/health gradient." Journal of health economics **28**(3): 540-552.

Kearney, P. M., Whelton, M., Reynolds, K., Muntner, P., Whelton, P. K., & He, J. (2005). "Global burden of hypertension: analysis of worldwide data." The Lancet **365**(9455): 217-223.

Lee, J., Arokiasamy, P., Chandra, A., Hu, P., Liu, J., & Feeney, K. (2012) "Markers and Drivers: Cardiovascular Health of Middle-Aged and Older Indians," In Aging in Asia: Findings from New and Emerging Data Initiatives, Committee on Policy Research and Data Needs to Meet the Challenge of Aging in Asia, J.P. Smith and M. Majmundar (eds.), Washington, DC: the National Academy Press, 389 – 416.

Lee, J., & Smith, J. P. (2014). "Regional disparities in adult height, educational attainment, and late-life cognition: Findings from the Longitudinal Aging Study in India (LASI)." The Journal of the Economics of Ageing(0).

Liu, H., H. Fang & Zhou, Z. (2013). "Urban–rural disparities of child health and nutritional status in China from 1989 to 2006." Economics & Human Biology **11**(3): 294-309.

Lopez, A. D., Mathers, C. D., Ezzaqti, M., Jamison, D. T., and Murray, C. J. (2006). "Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data." The Lancet **367**(9524): 1747-1757.

McDade, T.W., Burhop, J., & Dohnal, J. (2004). "High-sensitivity enzyme immunoassay for C-reactive protein in dried blood spots." Clinical Chemistry **50**: 652-54.

Murray, C. J., Vos, T., Lozano, R., Naghavi, M., Flaxman, A.D., Michaud, C. … & Brooker, S. (2012). "Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990 – 2010: A systematic analysis of the Global Burden of Disease Study 2010." The Lancet **380**(9859): 2197 – 2223.

Myers, G. L., Rifai, N., Tracy, R. P., Roberts, W. L., Alexander, R. W., Biasucci, L. M. . . . & Waymack, P. P. (2004). "CDC/AHA Workshop on Markers of Inflammation and Cardiovascular Disease Application to Clinical and Public Health Practice: Report From the Laboratory Science Discussion Group." Circulation **110**(25): e545-e549.

Narayan, K. V., Ali, M. K., & Koplan, J. P. (2010). "Global noncommunicable diseases: where worlds meet." New England Journal of Medicine **363**(13): 1196-1198.

Oaxaca, R. (1973). "Male-female wage differentials in urban labor markets." International economic review: 693-709.

O'Broin, S.D., & Gunter, E. W. (1999). Screening of folate status with use of dried blood spots on filter paper. American Journal of Clinical Nutrition **70**(3):359-67.

Powell, L. M., Wada, R., Krauss, R. C., and Wang, Y. (2012). "Ethnic disparities in adolescent body mass index in the United States: the role of parental socioeconomic status and economic contextual factors." Social Science & Medicine **75**(3): 469-476.

Prentice, A. M. (2006). "The emerging epidemic of obesity in developing countries." International journal of epidemiology **35**(1): 93-99.

Rosero-Bixby, L. & Dow, W. H. (2009). "Surprising SES gradients in mortality, health, and biomarkers in a Latin American population of adults." The Journals of Gerontology Series B: Psychological Sciences and Social Sciences **64**(1): 105-117.

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.

Shetty, P. (2012). "Grey matter: ageing in developing countries." The Lancet **379**(9823): 1285-1287.

Vikram, N. K., Misra, A., Dwivedi, M., Sharma, R., Pandey, R., Luthra, K., Chatterjee, A., Dhingra, V., Jailkhani, B. & Talwar, K. (2003) “Correlations of C-reactive protein levels with anthropometric profile, percentage of body fat and lipids in healthy adolescents and young adults in urban North India.” Atherosclerosis*,* 168**,** 305-313.

World Health Organization (2014). India: Country profile, Noncommunicable diseases. Geneva, World Health Organization, http://www.who.int/nmh/countries/ind_en.pdf?ua=1 (Accessed on July 31, 2014).

World Health Organization (2008). Worldwide prevalence of anaemia 1993-2005. WHO Global Database on Anaemia. Geneva, World Health Organization.

World Health Organization (2005). Preventing chronic disease: a vital investment, Geneva, World Health Organization.

Tables

Table 1: Changes in Economic and Social Factors by State

	Punjab	Rajasthan	Karnataka	Kerala
Economic Growth				
1960-61 Mean Per Capita Consumption (Rs/Month)[1]				
Rural	82.06	55.7	59.19	46.64
Urban	83.71	66.6	72.05	53.6
1993-94 Mean Per Capita Consumption (Rs/Month)[1]				
Rural	79.23	58.07	62.52	73.44
Urban	100.34	75.87	79.82	89.32
Male Preference				
Child sex ratio (Girls per 1,000 boys aged 0 – 6)[2]	846	883	943	959
Elementary School Attendance Rates (Ages 5 – 14) Per 1,000[3]				
Boys	897	847	898	968
Girls	882	710	866	985

Source:

1. Datt (1998)
2. The Government of India 2011 Census
3. The Government of India, Ministry of Human Resource Development, Selected Educational Statistics: 2000-01

Table 2: Descriptive Statistics for the Analysis Sample

	Median	Mean	SD	N
Hemoglobin (g/dL)	14.31	14.26	2.50	1,077
Anemic (%)	0	19.8	39.9	1,077
C Reactive Protein (mg/L)	1.68	2.68	3.01	1,106
CRP High Risk (%)	0	29.0	45.4	1,106
Body Mass Index (kg/m2)	21.91	22.35	5.07	1,133

Age Group	**No.**	**%**
44-54	532	46.3
55-64	318	27.7
65-74	202	17.6
75+	97	8.4
Total	1,149	100
Gender		
Male	552	48
Female	598	52
Total	1,150	100
Caste		
Scheduled Caste	188	16.8
Scheduled Tribe	138	12.3
Other Backward Class	395	35.2
None	400	35.7
Total	1,121	100
Education		
Some Schooling	627	54.5
No Schooling	523	45.5
Total	1,150	100

Residency	**No.**	**%**
Urban	298	25.9
Rural	852	74.1
Total	1,150	100
State		
Punjab	288	25
Rajasthan	280	24.3
Kerala	329	28.6
Karnataka	253	22
Total	1,150	100
Diabetes, Stroke, Hypertension, and Heart diseases		
No	862	75
Yes	288	25
Total	1,150	100
Current smokers		
No	971	84.7
Yes	176	15.3
Total	1,147	100

Source: LASI Pilot 2010 biomarker sample. Those under age 45 are excluded. BMI is calculated using measured height and weight. “Diabetes, Stroke, Hypertension, Heart condition” refers to self-reported prior diagnosis of at least one of: diabetes mellitus, stroke, hypertension, and heart disease. The sample is weighted to be nationally representative. “Anemic” is defined as being below 13 g/dL of Hb for men, and below 12 g/dL for women. “High risk” for cardiovascular disease is defined as CRP level above 3 mg/L.

Table 3: Associations between Explanatory Variables and C-Reactive Protein

	Mean (mg/L)	SE	CI	N	T test P Value
Total	2.69	0.1	[2.49,2.90]	1,106	
Gender					
Male	2.78	0.15	[2.48,3.08]	529	
Female	2.61	0.15	[2.32,2.90]	577	(0.432)
Age Group					
44-54	2.47	0.16	[2.15,2.79]	513	
55-64	2.71	0.16	[2.39,3.03]	310	(0.249)
65-74	2.85	0.22	[2.41,3.30]	194	(0.191)
75+	3.61	0.49	[2.61,4.60]	88	(0.030)
Caste					
Scheduled Caste	2.78	0.23	[2.31,3.25]	180	(0.971)
Scheduled Tribe	2.94	0.41	[2.12,3.76]	130	(0.779)
Other Backward Class	2.48	0.16	[2.16,2.80]	390	(0.298)
None	2.79	0.23	[2.33,3.26]	378	
Education					
Some Schooling	2.53	0.11	[2.30,2.75]	499	
No Schooling	2.88	0.18	[2.52,3.24]	275	(0.098)
Residency					
Urban	3.16	0.19	[2.78,3.53]	290	
Rural	2.53	0.12	[2.30,2.77]	816	(0.007)
State					
Punjab	3.06	0.13	[2.79,3.32]	277	
Rajasthan	3.28	0.26	[2.76,3.80]	268	(0.447)
Kerala	1.96	0.17	[1.63,2.30]	314	(0.000)
Karnataka	2.51	0.15	[2.20,2.81]	247	(0.009)
BMI Group					
Underweight (BMI < 18.5)	2.36	0.23	[1.90,2.82]	238	(0.371)
Normal (BMI: 18.5 – 24.9)	2.61	0.17	[2.27,2.95]	569	
Overweight (BMI: 25.0 – 29.9)	2.92	0.18	[2.56,3.29]	215	(0.253)
Obese (BMI: 30.0 or above)	3.77	0.46	[2.85,4.70]	68	(0.023)
Diabetes, Stroke, Hypertension, Heart condition					
No	2.61	0.12	[2.37,2.86]	826	
Yes	2.97	0.19	[2.58,3.36]	280	(0.150)
Current smokers					
No	2.68	0.11	[2.46,2.90]	938	
Yes	2.71	0.27	[2.17,3.26]	166	(0.904)

Note: Sample is weighted to be nationally representative. Confidence intervals and T tests account for survey design. BMI is calculated using measured height and weight. “Diabetes, Stroke, Hypertension, Heart condition” refers to self-reported prior diagnosis of at least one of: diabetes mellitus, stroke, hypertension, and heart disease.

Table 4: Associations between Explanatory Variables and Hemoglobin

	Mean (gram/dL)	SE	CI	N	T test P Value
Total	14.23	0.14	[13.96,14.51]	1,077	
Gender					
Male	15.27	0.16	[14.94,15.60]	517	
Female	13.27	0.14	[12.98,13.55]	560	(0.000)
Age Group					
44-54	14.41	0.14	[14.12,14.70]	506	
55-64	14.22	0.25	[13.72,14.73]	295	(0.510)
65-74	13.96	0.24	[13.48,14.44]	185	(0.060)
75+	13.81	0.32	[13.17,14.46]	90	(0.089)
Caste					
Scheduled Caste	13.79	0.25	[13.29,14.30]	179	(0.159)
Scheduled Tribe	13.99	0.32	[13.34,14.65]	137	(0.573)
Other Backward Class	14.54	0.18	[14.18,14.91]	367	(0.161)
None	14.20	0.19	[13.81,14.59]	366	
Education					
Some Schooling	14.88	0.16	[14.56,15.19]	580	
No Schooling	13.55	0.16	[13.22,13.88]	497	(0.000)
Residency					
Urban	14.69	0.33	[14.02,15.36]	257	
Rural	14.09	0.14	[13.80,14.37]	820	(0.104)
State					
Punjab	14.12	0.36	[13.40,14.84]	264	
Rajasthan	13.73	0.24	[13.25,14.21]	275	(0.368)
Kerala	14.69	0.13	[14.43,14.94]	306	(0.143)
Karnataka	14.46	0.29	[13.87,15.05]	232	(0.472)
BMI Group					
Underweight	13.64	0.24	[13.16,14.12]	240	(0.002)
Normal	14.51	0.15	[14.21,14.81]	551	
Overweight	14.49	0.24	[14.01,14.97]	205	(0.922)
Obese	14	0.41	[13.18,14.82]	66	(0.179)
Diabetes, Stroke, Hypertension, Heart condition					
No	14.11	0.15	[13.81,14.40]	811	
Yes	14.68	0.18	[14.31,15.05]	266	(0.005)
Smokes Now					
No	14.03	0.13	[13.77,14.30]	907	
Yes	15.21	0.21	[14.79,15.64]	167	(0.000)

Note: Sample is weighted to be nationally representative. Confidence intervals and T tests account for survey design. BMI is calculated using measured height and weight. “Diabetes, Stroke, Hypertension, Heart condition” refers to self-reported prior diagnosis of at least one of: diabetes mellitus, stroke, hypertension, and heart disease.

Table 5: OLS Regression Results for C-Reactive Protein

	ALL		MEN		WOMEN	
	(1)	(2)	(1)	(2)	(1)	(2)
Age Group: Reference 45-54						
55-64	0.266	0.135	0.170	0.018	0.397	0.284
	(0.215)	(0.210)	(0.362)	(0.331)	(0.283)	(0.276)
65-74	0.421	0.303	0.468	0.568	0.377	0.107
	(0.289)	(0.268)	(0.467)	(0.436)	(0.303)	(0.267)
75+	1.135**	1.158**	0.753	0.839	1.544***	1.592***
	(0.485)	(0.474)	(0.878)	(0.844)	(0.516)	(0.529)
Female	-0.099	-0.170				
	(0.207)	(0.206)				
Caste: Reference None or other						
Scheduled caste	-0.207	0.011	-0.039	0.182	-0.534	-0.304
	(0.447)	(0.447)	(0.686)	(0.639)	(0.571)	(0.565)
Scheduled tribe	-0.225	-0.324	-0.646	-0.692	0.148	-0.016
	(0.324)	(0.315)	(0.509)	(0.491)	(0.376)	(0.354)
Other backward class	0.025	-0.039	-0.295	-0.327	0.300	0.210
	(0.378)	(0.381)	(0.550)	(0.581)	(0.454)	(0.462)
Some Education	-0.216	0.039	0.020	0.226	-0.558	-0.252
	(0.255)	(0.266)	(0.428)	(0.446)	(0.373)	(0.377)
Rural	-0.772***	-0.670***	-0.617**	-0.493	-0.884**	-0.800**
	(0.209)	(0.219)	(0.292)	(0.298)	(0.379)	(0.357)
State: Reference Punjab						
Rajasthan	0.472	0.742*	0.331	0.571	0.632	0.925
	(0.329)	(0.370)	(0.450)	(0.467)	(0.577)	(0.611)
Kerala	-1.127***	-0.964***	-1.099***	-1.072***	-1.258***	-0.921**
	(0.246)	(0.258)	(0.318)	(0.364)	(0.333)	(0.366)
Karnataka	-0.451*	-0.249	-0.531	-0.458	-0.346	-0.001
	(0.268)	(0.289)	(0.362)	(0.411)	(0.380)	(0.425)
BMI		0.090***		0.115**		0.078***
		(0.024)		(0.045)		(0.023)
Diabetes, Stroke, Hypertension, Heart condition		0.487*		0.453		0.434
		(0.284)		(0.413)		(0.369)
Current Smoker		0.406		0.642		-0.054
		(0.355)		(0.401)		(0.702)
Constant	3.557***	1.098	3.686***	0.640	3.426***	1.154
	(0.398)	(0.748)	(0.607)	(1.085)	(0.518)	(0.792)
Observations	1,077	1,059	517	510	560	549

R-squared	0.052	0.076	0.048	0.072	0.084	0.109

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: BMI is calculated using measured height and weight. “Diabetes, Stroke, Hypertension, Heart condition” refers to self-reported prior diagnosis of at least one of: diabetes mellitus, stroke, hypertension, and heart disease. The sample is weighted to be nationally representative, and standard errors are adjusted for survey design.

Table 6: OLS Regression Results for Heomglobin

	ALL		MEN		WOMEN	
	(1)	(2)	(1)	(2)	(1)	(2)
Age Group: Reference 45-54						
55-64	-0.314	-0.373	-0.416	-0.479	-0.198	-0.258
	(0.237)	(0.233)	(0.343)	(0.335)	(0.239)	(0.240)
65-74	-0.423*	-0.482**	-0.398	-0.354	-0.409	-0.505
	(0.220)	(0.225)	(0.399)	(0.406)	(0.304)	(0.316)
75+	-0.399	-0.361	-0.852**	-0.783**	0.023	0.136
	(0.287)	(0.249)	(0.364)	(0.367)	(0.407)	(0.311)
Female	-1.969***	-1.970***				
	(0.154)	(0.168)				
Caste: Reference None of them						
Scheduled caste	0.583	0.581	0.705	0.775	0.474	0.399
	(0.393)	(0.378)	(0.510)	(0.505)	(0.444)	(0.427)
Scheduled tribe	0.436	0.389	0.696*	0.594	0.163	0.110
	(0.286)	(0.282)	(0.398)	(0.380)	(0.299)	(0.290)
Other backward class	0.030	0.020	0.389	0.243	-0.300	-0.256
	(0.311)	(0.299)	(0.413)	(0.398)	(0.340)	(0.312)
Some Education	-0.707***	-0.620**	-0.709*	-0.573	-0.693**	-0.628*
	(0.237)	(0.246)	(0.387)	(0.385)	(0.315)	(0.313)
Rural	-0.455	-0.431	-0.429	-0.431	-0.478	-0.405
	(0.300)	(0.300)	(0.445)	(0.438)	(0.286)	(0.298)
State: Reference Punjab						
Rajasthan	-0.477	-0.320	-0.356	-0.098	-0.653	-0.473
	(0.417)	(0.422)	(0.463)	(0.455)	(0.509)	(0.504)
Kerala	0.261	0.170	0.385	0.334	0.125	0.072
	(0.385)	(0.411)	(0.476)	(0.462)	(0.443)	(0.485)
Karnataka	-0.088	-0.055	0.108	0.170	-0.271	-0.187
	(0.419)	(0.431)	(0.495)	(0.485)	(0.462)	(0.477)
BMI		0.040***		0.110***		0.015
		(0.012)		(0.031)		(0.011)
Diabetes, Stroke, Hypertension, Heart condition		0.395**		0.328		0.437*
		(0.168)		(0.317)		(0.231)
Current Smoker		0.429*		0.630**		0.129
		(0.215)		(0.266)		(0.317)
Constant	15.979***	14.906***	15.672***	13.056***	14.322***	13.754***
	(0.470)	(0.484)	(0.576)	(0.773)	(0.539)	(0.614)
Observations	1,048	1,030	505	499	543	531
R-squared	0.223	0.239	0.095	0.136	0.073	0.079

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: BMI is calculated from measured height and weight. “Diabetes, Stroke, Hypertension, Heart condition” refers to self-reported prior diagnosis of at least one of: diabetes mellitus, stroke, hypertension, and heart disease. The sample is weighted to be nationally representative, and standard errors are adjusted for survey design.

Table 7: Oaxaca Blinder Decomposition for State Differences in C-Reactive Protein

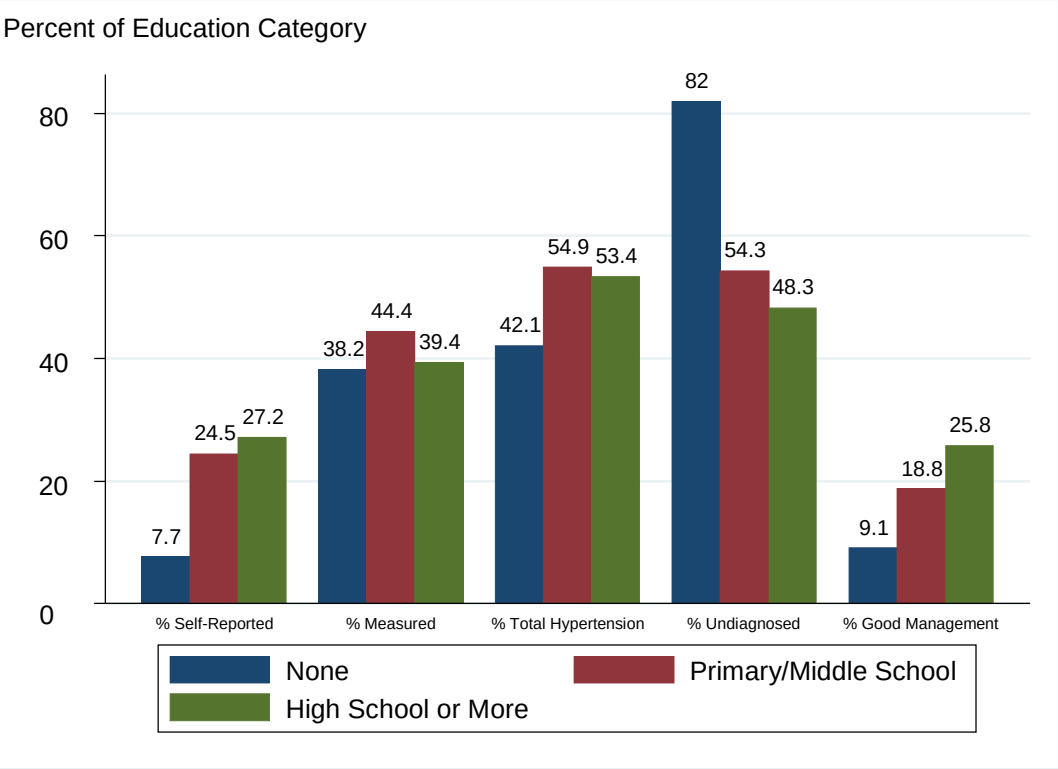
	Punjab (1)	Punjab (2)	Kerala (1)	Kerala (2)	Karnataka (1)	Karnataka (2)
Rajasthan Mean	3.274***	3.220***	3.274***	3.220***	3.274***	3.220***
	(0.257)	(0.264)	(0.278)	(0.285)	(0.257)	(0.264)
Comparison State Mean	3.055***	3.022***	1.926***	1.932***	2.509***	2.459***
	(0.179)	(0.192)	(0.198)	(0.200)	(0.173)	(0.182)
Difference	0.219	0.198	1.348***	1.288***	0.765**	0.761**
	(0.313)	(0.326)	(0.341)	(0.348)	(0.310)	(0.321)
Endowments	-0.054	-0.203	0.323	0.156	-0.353	-0.407
	(0.150)	(0.393)	(0.459)	(0.456)	(0.306)	(0.329)
Coefficients	0.946*	1.251**	1.967**	2.196***	0.799*	0.941**
	(0.467)	(0.545)	(0.717)	(0.582)	(0.430)	(0.415)
Interaction	-0.673	-0.850	-0.942	-1.064	0.318	0.227
	(0.404)	(0.616)	(0.843)	(0.739)	(0.493)	(0.495)
Additional Controls	N	Y	N	Y	N	Y
Observations	532	519	560	550	507	500

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The first column implements the Oaxaca-Blinder decomposition for each state compared to Rajasthan, using the control variables from the specification in column 1 of tables 5 and 6. The second column adds controls for BMI, smoking, and a diabetes/stroke/hypertension/heart condition. The sample is weighted to be representative at the state level, and standard errors are adjusted for survey design.

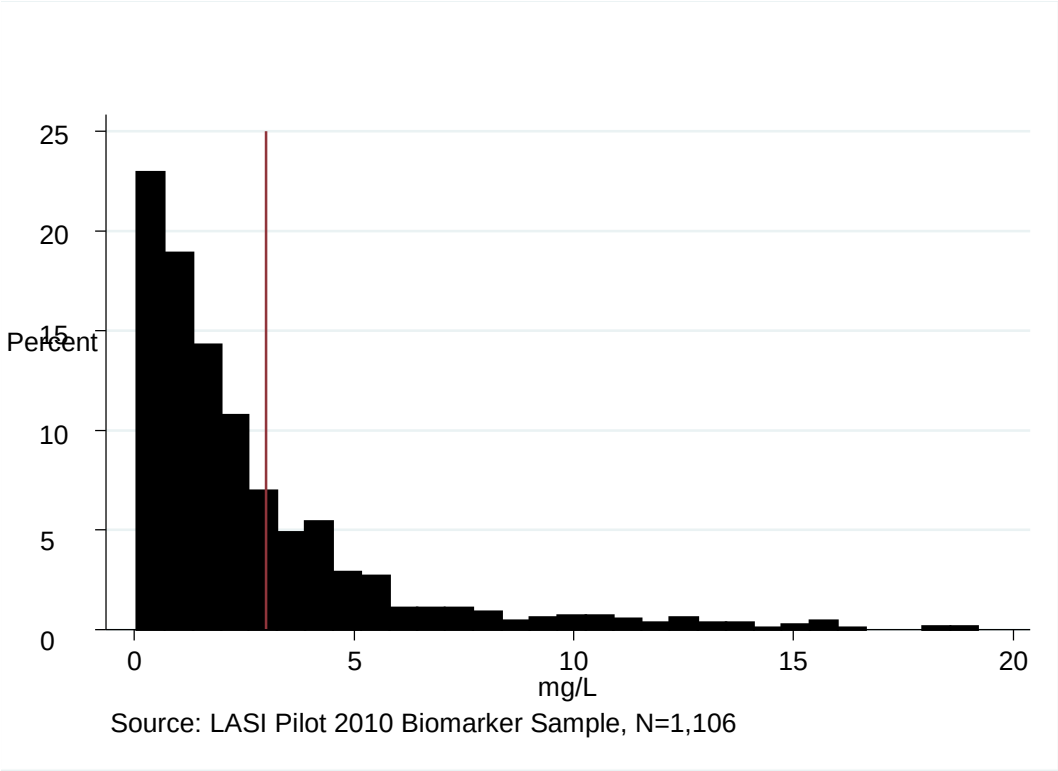
Figures

Figure 1: Educational Disparities in Measured and Reported Hypertension in India



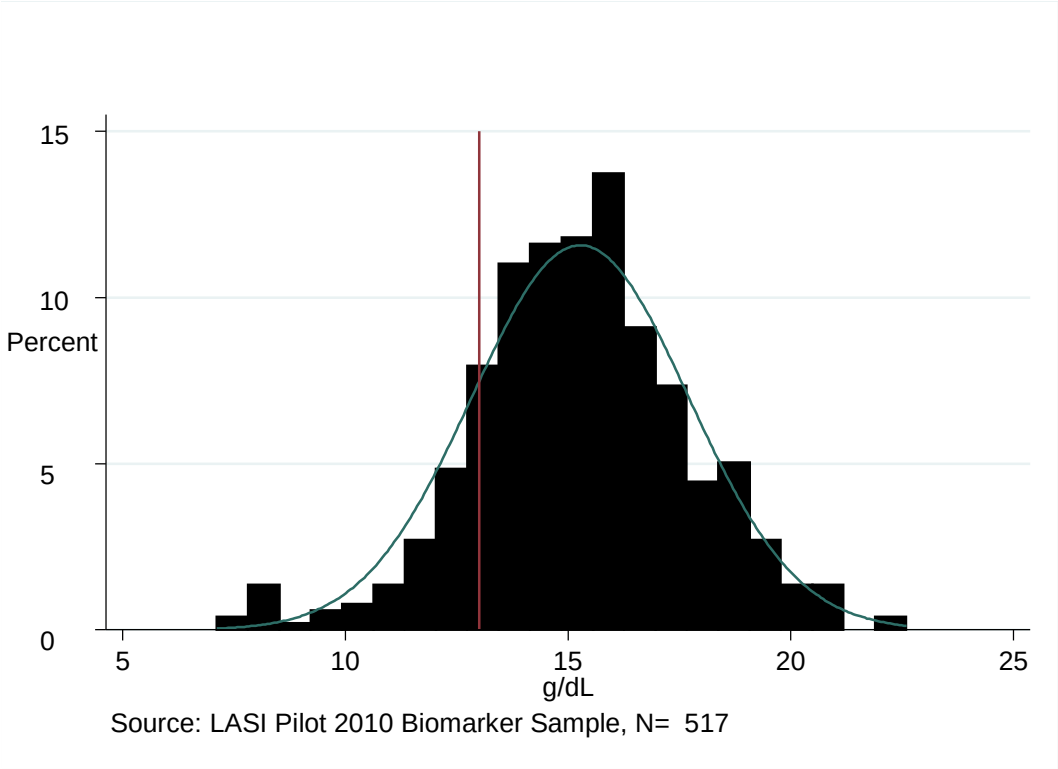
Note for Figure 1: The percent of each education category (no education, primary/middle school, high school or more) with self-reported hypertension, measured hypertension (i.e., blood pressure readings are higher than the following thresholds: systolic > 140 or diastolic > 90), total hypertension (i.e., either self-reported or measured hypertension), undiagnosed hypertension (i.e., not diagnosed according to self-reports but measured hypertension), and good management of their hypertension (i.e., diagnosed to be hypertensive but blood pressure readings are lower than the above thresholds). Source: Lee et al. (2013) National Academy of Science.

Figure 2: C-Reactive Protein Distribution for Men and Women



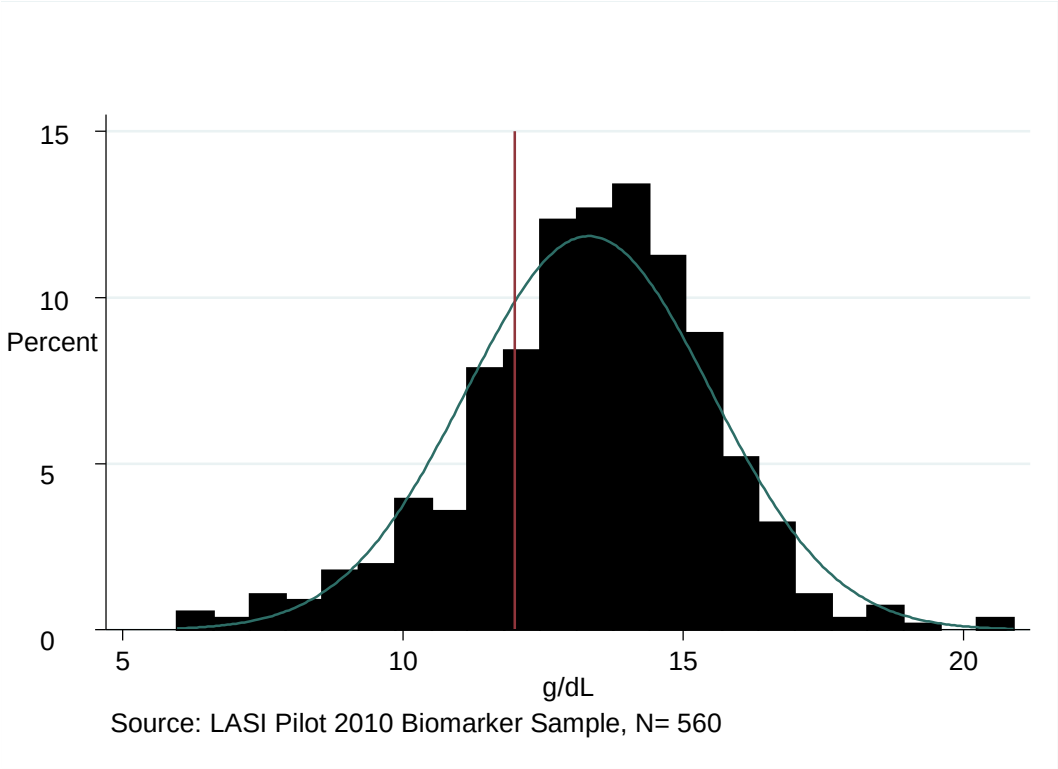
Note: Sample is weighted to be nationally representative. The cut-off for high cardiovascular risk is shown at 3 mg/L (above is at risk).

Figure 3: Hemoglobin Distribution for Men



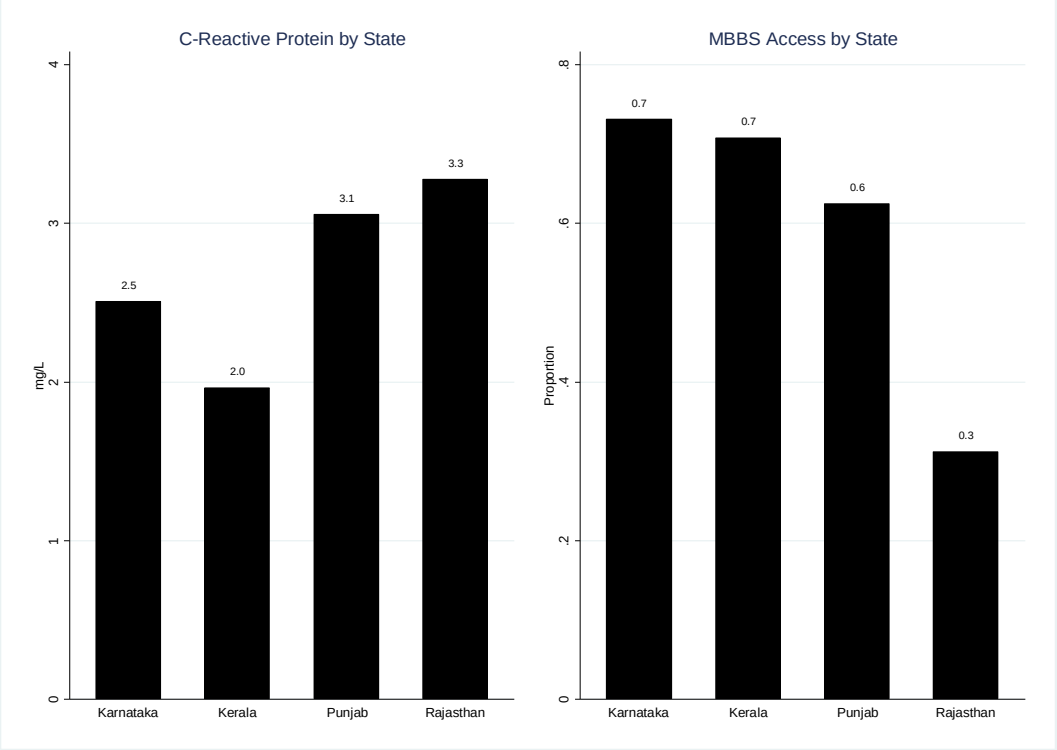
Note: Sample is weighted to be nationally representative. The cut-off for anemia among men is shown at 13 g/dL (below is anemic).

Figure 4: Hemoglobin Distribution for Women



Note: Sample is weighted to be nationally representative. The cutoff for anemia among women is shown at 12 g/dL (below is anemic).

Figure 5: C-Reactive Protein and Health Care Access by State



Note: Sample is weighted to be representative at the state level. The left figure shows the mean CRP level by state, while the right figure shows the proportion of respondents in each state who have ever visited a doctor with a formal qualification (e.g., MMBS degree).