The Impact of Recall Periods on Reported Morbidity and Health Seeking Behavior[1]

Jishnu Das
(World Bank, Washington and Center for Policy Research, New Delhi)
Jeffrey Hammer
(Princeton University)
Carolina Sánchez-Paramo
(World Bank)

Abstract

Between 2000 and 2002, we followed 1621 individuals in Delhi, India using a combination of weekly and monthly-recall health questionnaires. In 2008, we augmented these data with another 8 weeks of surveys during which households were experimentally allocated to surveys with different recall periods in the second half of the survey. We show that the length of the recall period had a large impact on reported morbidity, doctor visits; time spent sick; whether at least one day of work/school was lost due to sickness and; the reported use of self-medication. The effects are more pronounced among the poor than the rich. In one example, differential recall effects across income groups reverse the *sign of the gradient* between doctor visits and per-capita expenditures such that the poor use health care providers more than the rich in the weekly recall surveys but less in monthly recall surveys. We hypothesize that illnesses--especially among the poor--are no longer perceived as "extraordinary events" but have become part of "normal" life. We discuss the implications of these results for health survey methodology, and the economic interpretation of sickness in poor populations.

Introduction

How much does survey design affect survey results and, therefore, our view of the world? Why is it that the rich report more illnesses than the poor in a large number of surveys? These questions seem largely unrelated but are at the core of this paper. Using the length of the recall period as a lens through which to examine these issues, we show that the number used to fill the blank in the standard screening question, "Were you sick in the last ___ weeks?" can change the results. We compare weekly and monthly recall periods using observational and experimental data and find that not only do different recall periods give very different results, but also that these differences are not uniform across income groups. Reported illnesses, use of health care and the reported labor impacts of sickness are lower in monthly reports, with higher declines for the poor. In the case of doctor visits, the recall period used changes the *sign of the gradient* between doctor visits and per-capita expenditures (PCE)—in weekly recall surveys, the poor visit doctors more than the rich.

This study stems from a project in Delhi examining morbidity and health-seeking behavior among 300 urban households.[2] At the beginning of the project, in 2000, we noted that the patterns of morbidity and health care-seeking behavior observed in the field among richer households accorded well with theoretical and cognitive models frequently used in the health economics literature. The (relatively) wealthy became sick, self-medicated as a first resort, and then, sought out treatment from a doctor. For example: "*I fell sick on Monday with a very bad cold; I took aspirin until Wednesday and then, when I was still not better by Friday, I went to the doctor. The doctor gave me antibiotics, which I took until Wednesday, and now I am fine.*"

Yet, among households in poor localities and slums, the underlying models of disease and treatment-seeking seemed to be different. People appeared to be sick all the time and going to a doctor seemed less like an "exceptional" event than a common fact of life, well-integrated into a household's daily routine. Illness narratives among poor households we interviewed were fractured and temporally confusing.[3] The chaos inherent in preliminary interviews with poor households--including numerous doctor visits, uncertainty about when the disease was diagnosed, whether it was "cured," and uncertainty about current health status—is very different from what standard health surveys lead us to believe. The health rounds of India's National Sample Survey, for instance, allow for a fairly well-developed narrative in which people fall sick with a variety of

---

[2] This is a collaborative study between the authors, anthropologists based at Johns Hopkins University, and the Institute of Socio-Economic Research on Development and Democracy (ISERDD) in Delhi.

[3] In one case, an interviewer speaks with a woman from a poor household about the diagnosis and treatment of her tuberculosis. The interview winds back and forth; the path of treatment is never clear, and the respondent keeps contradicting the temporal sequence of events. She cannot give clear answers about what treatment was sought when and what the results from these treatments were. Eventually, it emerges that her tuberculosis was discovered in a hospital when she went for an ante-natal checkup (although we remain unsure, since a little later she tells the interviewer that a private doctor asked for a chest X-ray 6 months before the hospital visit) and that she was then put on a treatment regime. It later emerges that her baby died because, on the advice of her neighbors, she did not breastfeed the child, believing that it would transmit TB to her newborn. But this fact only emerges because the interviewer asks at some stage how the child is now doing.

illnesses and pay high prices for "catastrophic" incidents that often put them in debt. That an individual's health status constantly fluctuated between illness and well-being and that he or she visited different doctors numerous times, incurring potentially large expenditures for seemingly "minor" illnesses, never emerged as a dominant theme.

In an attempt to capture some of the patterns we observed in the field, we designed and implemented a weekly survey module that was eventually interspersed with monthly-recall modules over a 2-year survey period from 2000 to 2002.[4] Given surprisingly large recall effects in the observational data, we returned in 2008 to experimentally evaluate these effects in a sub-sample of households over a 2-month period.

Across the observational and experimental data, both the rich and poor underreport acute health events in monthly-recall surveys, but the poor underreport more and this underreporting seems to arise from the "forgetting" of entire episodes of acute illness. Among the poor, monthly reporting "erases" almost half the morbidity burden of acute illnesses, over a third of doctor visits for both acute and chronic illnesses, and almost half of all self-medication episodes. Monthly reporting also decreases reports that at least one day of work was lost due to an acute illness by close to a third, again with somewhat higher recall impacts among the poor. We explore the possibility that this "forgetting" is driven by a process of normalization of sickness. Consistent with this hypothesis, we are able to show that recall effects are larger for those individuals and households whose initial sickness burden is also higher; arguably, these are households where sickness has become a "normal" event.

Although we are able to present stark differences for reported morbidity and doctor visits, we should note that we are unable to integrate routine and catastrophic bouts of illness into our results and therefore, provide a full accounting of health *expenditures* and the impact of different recall periods on health spending. Our results do not apply to episodes of illness associated with very large expenditures. Consistent with results from Banerjee and others (2006), there does not appear to be systematic recall bias on the amount of total expenditures people incur in a given time period. Somewhat intuitively, everyone remembers large expenditures well.[5] We believe that it would take a very large survey or some other sampling technique to get at the frequency of spending on catastrophic illnesses. Since catastrophic illnesses account for a large fraction

---

[4] Weekly and monthly recall periods were interspersed to avoid surveyor (and respondent) fatigue and to manage the project within the grant amount.

[5] In India, a "large" expenditure is different from the world standard for large expenditures. Nonetheless, such expenditures might still be quite significant to India's poor. We cut off "large" expenditures at a level of about $25 US – just about the co-payment for a doctor's visit in the US. In richer countries, unlike in India, everyone may avail themselves of life-saving but very expensive procedures – the decision usually does not lie with the patient but with doctors. In India, where the vast majority of the population does not have health insurance, life and death decisions are made more frequently by the patients and, due to the out of pocket costs involved, patients frequently don't avail themselves of treatments. So, a full accounting of the cost of the health care system has more endogenous (from the patient's perspective) elements in it than would be the case in richer countries.

of total health expenditures, our results on the impact of recall periods on health expenditures remain tenuous.[6]

The combination of large impacts on reported morbidity, doctor visits and whether a day was lost due to sickness together with smaller impacts on health expenditures suggests that much of what was underreported in the monthly surveys were less "severe" illnesses that required smaller expenditures on the part of the household. Therefore, whether monthly or weekly reports should be the "gold standard" for health surveys depends on the underlying question. If researchers are interested in the ability of households to smooth consumption after a major health shock, monthly (or even annual) reports may be sufficient—our data do not suggest that these episodes are easily forgotten. But if researchers want to look at the process through which sickness evolves from something "minor" to something "severe", monthly reporting may erase a significant fraction of the relevant action. Similarly, little is known about the industrial organization of markets for medical care in low-income countries. Since a large fraction of visits to doctors are for self-limiting minor problems—a fraction often guessed to be 80 percent—the degree of underreporting we find could substantially alter the formulation and tests of hypotheses related to these markets.

The remainder of our paper is organized as follows. Section 1 describes the data. Section 2 presents the evidence of the overall recall effect independent of its interaction with income. Section 3 explores the data further in order to illustrate the differentials across income groups and the robustness of our findings to age, education and respondent effects. Section 4 concludes with a brief discussion of the implications from these findings.

Section I: Data

This paper uses data from two sources. The first is a long-standing panel study of health seeking behavior carried out by the Institute for Socio-Economic Research on Democracy and Development (ISERDD) in Delhi. The study followed 1621 individuals in 300 households from 7 different neighborhoods in Delhi between 2000 and 2002. The localities were chosen through initial contacts developed by the anthropologists on the team. Following these initial contacts, which were used for "entry" into the locality, households within each locality were chosen randomly. In each locality, a street was chosen at random and every fourth household located on the street was asked to participate in the survey. A total of 40-45 households were approached to be part of the survey in each locality, and refusal rates were less than 4 percent in all localities. Considerable time was spent explaining to each household the purpose of the survey and care was taken to

---

[6] Until recently, medical insurance for catastrophic events was limited to a small fraction of the population. Some estimates suggest that 10 percent of the urban population has access to health insurance; most of these are richer households or government employees. India recently launched a health insurance program for the poor, which is in its 3rd year in Delhi.

ensure that households did not believe that participation would render any material benefits. Principal investigators personally conducted recruitment to begin with and households were informed that the purpose of the survey was to "*gather data and information on health and health seeking behavior among households in Delhi*".

Given the lack of data on medical facilities that are not registered (the majority) in Delhi, it is difficult to ascertain whether the neighborhoods chosen had more medical facilities or households were more aware of health problems than others in the city. As a check on the external validity of this sample, Das and Sánchez (2002) compared the ISERDD sample to two representative samples from Delhi—the (Delhi) National Family Health (1993) and National Sample Survey (2000). Their comparison suggests that the ISERDD sample is, on average, slightly younger (with a one year difference in the mean age) with a greater density of households in the upper-income groups. On a host of variables including education levels, marital status, characteristics of the household-head, household size and availability of different types of facilities they find few significant differences. The income differences do not appear to be large (they could mirror the different survey years in a growing economy), and assuming a positive relationship between income and health, should actually bias our results towards less morbidity than what is representative for Delhi as a whole.

Recall periods varied for the same household through the two years of the survey. The 7 localities were divided into two "waves" composed of 4 and 3 localities respectively. Between August 2000 and December 2000, every household in Wave 1 (4 localities) was visited weekly by a member of the ISERDD team and asked about health problems during the previous week, whether treatment was pursued, and how much treatment (if any) cost. Care was taken to ensure visits to the households were always on the same day of the week, which became the focal reference-point for the weekly recall questions. This design was followed to avoid the possibility of "telescoping" (the reporting of events prior to the recall period) in the weekly surveys; it also had costs in terms of missing household observations in the weekly data, which are addressed below. Surveyors also attempted to always interview the same respondent in both weekly and monthly surveys; this person, designated as the primary respondent, was usually the head or the spouse of the head of the household. The primary respondent was chosen to partly reflect the belief that they may be better able to report on sicknesses of their children. These repeated visits followed a larger initial interview that covered a much wider range of household level information that was not likely to change much during the study period—household structure, total expenditures (using the standard "thin round" set of questions for household income and expenditures from the National Sample Survey, which is India's representative household survey), education, and other household characteristics.

Then, between January and April 2001, Wave 1 localities were visited once *a month* and administered a similar survey based on monthly recall. At the same time, households in the three Wave 2 localities were visited weekly and given the same weekly recall survey as had been used with Wave 1 during the previous four month period. Following a break in the summer months, the process was repeated for a second year. Thus, between August and December 2001, Wave 1 households received a weekly survey while Wave 2 households received a monthly survey and then, between January and April 2002, Wave 1 households received a monthly survey while Wave 2 households received a weekly survey.

In both weekly and monthly surveys, information was collected separately on chronic and non-chronic conditions. Chronic conditions were identified through initial interviews, and these illnesses were individually followed up upon each week/month. In the case of chronic illnesses, once identified, individuals would be recorded as sick (with a chronic illness) in every survey (weekly or monthly) and asked about his or her chronic condition during each subsequent visit. Consequently, recall periods, by design, have no effect on the reporting of chronic illnesses or their duration. Therefore, we examine the effects of recall only on doctor visits and self-medication for chronic conditions[7]. Additionally, a general question ("Were you sick last week/month?") was used to screen for acute illnesses, and conditional on a positive response, questions were asked about the duration of the morbidity and the resulting treatment, including visits to doctors and self-medication.

Attrition of the sample was negligible—the survey lost 5 households over 2 years due to migration out of sample localities. Therefore, we have a large number of observations across time asking the same set of participants the same survey questions with varying recall periods (though, of course, with fewer observations for monthly rather than weekly recall periods). These data will be used to compare responses given for weekly versus monthly recall periods. The size of the discrepancies between survey responses in the weekly and monthly data indicates the magnitude of forgetting that takes place with the monthly recall period. Given the structure of the data, we will also be able to account for seasonal effects by including responses from the group of households administered a monthly survey concurrently. Finally, since monthly surveys were administered *after* the completion of weekly surveys, we are able to avoid "telescoping" effects (survey reports capture events preceding the recall period considered) in the monthly data as well.

Our second data source is an experiment carried out in 2008 to verify the large recall effects discovered in the observational data. For this survey, we began with a sub-sample of 205 households who could still be located from the original sample and surveyed them for a total of 2 months. In the first month, the entire sub-sample

---

[7] See (Das and Hammer 2009) for more detailed discussion of how households identify and handle chronic illness.

was interviewed using the weekly-recall module. In the second month, the sub-sample was split into two randomly assigned groups. The "treatment" group continued with weekly recall surveys for another 4 weeks, while the "control" group was administered a single survey collecting data on the basis of monthly recall at the end of the month. The randomization was privately completed using STATA and was stratified by locality. To check the randomization, we compared household characteristics for the control and treatment groups and found that the two sets of households similar on all observable criteria.

The two data sources are complementary. In the "observational" data – that is, the data from families that were tracked for two years – many individual characteristics are held constant over time, including general healthiness and innate ability to recall events. The large number of observations leads to greater precision in the estimation of recall effects. However, a drawback to using this type of data arises from its vulnerability to temporal autocorrelation of illness patterns that differ between individuals and could systematically influence individuals' recollections of events. This temporal autocorrelation would be particularly problematic if the localities in the two different waves differed in their morbidity and treatment seeking patterns. In our analysis, we attempt to tackle this problem by including individual fixed-effects and a full set of monthly dummies. Nonetheless, we cannot entirely rule out the possibility that our results are biased by this autocorrelation.

In the "experimental" data, the autocorrelation of illnesses is entirely controlled for by the randomization process. On the other hand, the experimental sample is not large enough to produce precise heterogeneous treatment effects. In particular, for the observational data, there are a minimum of 36 weekly observations and 5 monthly observations for 300 households, while in the experimental data the sample size is reduced to 4 weekly observations and 1 monthly observation for 205 households. Therefore, we use the experimental data to verify the average treatment effects from the observational estimates, but we rely on the observational data to further examine heterogeneity in the treatment, particularly across income groups.

Summary statistics for all individuals and households in the observational sample are presented in Table 1. The median age of individuals in our sample is 22, 50 percent of all individuals are either illiterate or have less than primary education and are living in nuclear households with an average of 5.4 members per household. About 50 percent of respondents of age greater than 15 are employed, resulting in 1.5 income earners per household, although only a third of income earners are employed in the formal sector (public and private). Table 1 also presents summary statistics by income group. Per capita income among rich households is more than three times that among poor households. Since both household size and the number of employed adults per household are similar across income groups, variation in income levels is mainly due to differences in the earning capacity of employed adults. As expected, average age and education are higher for the rich, who also

enjoy better access to infrastructure facilities and sanitation systems such as flush toilets and piped water. Employment rates are not substantially different across income groups, but the composition of employment does vary with a higher incidence of public sector employment among the middle and upper income groups.

Since households in the experimental sub-sample are those who could still be located eight years later, Appendix Table 1 compares the main outcome variables for the original sample of 300 households to the experimental subsample of 205 households. The means are computed for the first wave of weekly surveys completed in 2000-01. The number of "acute" illnesses refers to the total number of acute illnesses reported in the preceding month, and weekly reports are aggregated up to obtain monthly totals. Since aggregated measures could overstate the recall effect if episodes last longer than a week, we will also examine two other outcome measures—whether or not at least one acute illness was reported and the duration of acute illness— that should not be affected by this aggregation bias. Our measure of the labor and schooling impacts of sickness are based on reports indicating that at least one day of work (or school) was lost due to an acute illness during the relevant recall period; for brevity, we refer to this variable as the "labor impact of acute illness".

For our full sample, based on the weekly data, individuals report just less than 1 acute episode a month, and about 47 percent report at least one episode. The average time spent sick every month is 4 days and individuals report losing at least one day of work/school in 7.8 percent of weeks due to an acute illness. Das and Sánchez (2004) report the types of acute illnesses that were reported during the weekly surveys; the large majority is respiratory illnesses, non-specific myalgias, dermatological disorders and ENT infections. Given that (a) doctors very seldom tell their patients what they are suffering from and (b) even when they do, they is a high probability that they are incorrect, it is unlikely that these reports can be used to deduce what the underlying condition is, including the severity of the illness.[8] Whether or not treatment was sought is measured through doctor visits and self-medication. On average, people visit 0.4 doctors a month (roughly 1 doctor's visit every 10 weeks) and the majority of visits are for acute illnesses. Individuals also self-medicate, and around 30 percent self-medicated at least once in an average month of the survey.

Finally, there appear to be few differences in the initial year of the survey between the larger observational sample and the experimental sub-sample of 205 households, although there is some evidence that more time is spent sick with acute illnesses with higher reported durations in the experimental sub-sample. In addition, the sub-sample households were somewhat more likely to self-medicate and reported somewhat higher labor

---

[8] Das and others (2011) send standardized patients to doctors in urban Delhi and rural Madhya Pradesh. Across all interactions, a diagnosis was articulated in 31.3 percent and the rate of correct diagnosis conditional on a diagnosis being given was 8.8 percent for myocardial infarction, 22.5% for asthma and 9.5% for dysentery.

impacts of acute illness during the experimental survey than the 300 households in the original observational survey. These results are reassuring in terms of the comparability of the two samples and the possibility that repeated surveys may have altered health behavior; at least these broad indicators suggest that households were behaving very similarly when we first visited them and when we followed up eight years later.

Section II: Recall period and health measures

Table 2 shows the differences in means for the main outcome variables between the weekly and monthly recall for the observational and experimental samples. It provides a benchmark from which progressively more detailed specifications deal with, in Tables 3 through 5, respectively: 1) time invariant individual fixed-effects and seasonal effects in a difference in differences specification, 2) variation in recall effects by total number of episodes, and 3) possible biases introduced by modeling alternative patterns of missing values that are more frequent in weekly than monthly data.[9]

As is clear from the table, the response from the monthly-recall group is lower than that of the weekly-recall group for every outcome variable. Furthermore, with only the exception of time spent sick with an acute illness, the estimated differences between the experimental and observational results are virtually identical at similar levels of statistical significance. The tests between the weekly and monthly data are precisely estimated, and the differences between the methods are robust. For time spent suffering from an acute illness, both methods show that the results from monthly recall are significantly different from those from weekly recall, just at slightly different rates.[10]

On every dimension, the declines due to monthly reporting are large and important behaviorally. The number of acute illnesses declines by 0.6 illnesses per person-month or just over 60 percent. This measure may be contaminated if episodes last longer than a week (and are thus double-counted in the weekly but not the monthly data). However, consistent with the total acute illness result, we find that the probability of reporting any acute illness also declines by close to 15 percentage points in the monthly-recall surveys. Given base probabilities in the weekly-recall surveys of 47 (45) percent in the observational (experimental) data, this suggests that close to a third of all acute episodes are "forgotten" once the recall period is increased to a month. Similarly, the total number of doctor visits declines by 0.16 (observational) and 0.18 (experimental) visits per month or 37 percent of the base weekly probabilities. As is clear from decomposition of doctor

---

[9] In the experimental sub-sample, all respondents received weekly surveys in the first month and were allocated to weekly and monthly-recall groups in the second month. Therefore, if morbidity was lower in the second month due to seasonal effects, this would upwardly bias the differences due to the recall period.

[10] In combination with Appendix Table 1, the means also suggest that despite the many surveys that households participated in, there was no change in overall morbidity or the probability of treatment (for evidence of behavior change from surveys, see Zwane and others, 2011). This could be because our survey covered a number of different topics and did not ask respondents about specific behavioral actions.

visits into doctor's visits for acute and chronic illnesses, the majority of these declines are in doctor visits for acute illnesses, which also account for the bulk of visits. Individuals and households report a higher labor impacts due to sickness in the weekly recall surveys. Across the observational and experimental samples, such reports drop from .09 per person-month to .06 (observational) or from .10 to .05 (experimental). Finally, there are large reported declines in the use of self-medication, both measured as the frequency of any self-medication and the total number of medications taken.

Table 2 also shows how recall effects alter our perception of average household morbidity and treatment-seeking patterns. In the weekly data, households (outcome variables are aggregated up to the household level by summing across individuals) report a total of 5 acute illnesses a month. On average, 2.4 individuals in every household report at least one acute illness in a month, and household members spend a total of between 23 and 26 days sick per month. There are between 2 (observational) and 3 (experimental) doctor visits every month, the bulk of which are accounted for by acute illnesses—a result very much in line with our own experiences during the initial field work. Monthly-recall data present a completely different picture. Reported acute illnesses drop and doctor visits are down to 1.28 (observational) to 1.4 (experimental) visits per household per month. With an average of 5.2 members per household, this is just above the average of 3 visits per individual per year reported in the US data (National Ambulatory Medical Survey 2000).

In Table 3, we account for time-invariant individual heterogeneity and seasonality in the observational and experimental data. Specifically, for the observational data, we estimate

$$y_{it} = \alpha + \beta T_{it} + \sum_{j=1}^{2} \gamma_j Timej + \mu_i + \epsilon_{it}$$

where y_{it} is the outcome variable for individual or household i at time t. T_{it} is a dummy variable that takes the value 1 for individuals/households who were surveyed using a weekly recall format at time t; γ_j represent coefficients on 12 seasonal dummies, one for each month of the year, and μ_i represents individual/household effects that are time-invariant but may be correlated with the format of the survey. The estimator β captures the effect of the recall period in the observational data under certain assumptions on the temporal auto-correlation of the outcome measures and the linear separability of recall from seasonal effects. In the case of the experimental data, we estimate a standard difference estimator with the treatment effect estimated as the interaction between the treatment group and the 2nd month of the survey, when the randomization into weekly and monthly recall was implemented. In all specifications, the lagged outcome variable increases the efficiency of the recall estimate and subsumes the individual fixed-effect. Given the nature of the experiment,

the lagged outcome variable is necessarily uncorrelated with treatment status, but serves to increase precision given the smaller sample size.

The estimates from these specifications are presented in Table 3 for both households and individuals. To account for intra-cluster correlations in the error structure, we cluster standard errors at the household/locality level depending on the unit of observation. The results from the means comparisons are largely replicated with some differences. Most notably, the estimated recall estimates are somewhat lower in the experimental data for the number of individuals reporting at least one acute illness during the month, and the effect on doctor visits declines from 18.3 percentage points to 8.7 percentage points. This decline happens primarily because respondents experienced unusually few health events in the second month of the data. Similarly, the labor impact of the recall period decreases to 1 percentage point since there were fewer work/school days lost in the second month. As a fraction of total work/school days lost in the second month, the recall effect is very similar to what we find in the observational data. It is likely that a longer experimental study would have yielded similar results as those using the observational estimates. The household level estimates for both the observational and experimental estimates are again similar to the means comparisons. Again, the estimated effect of recall on doctor visits and the labor impact in the experimental case is lower due to the inclusion of seasonal effects. In the experimental case, the significance is also lower due to the smaller sample size.

Table 4 explores whether "forgetting" is higher among individuals with larger health burdens. Here, we include an additional interaction of the recall effect with the original morbidity burden in the first year of recall surveys. While for the experimental data, these effects retain their causal interpretation, temporal autocorrelation in the observational data limits our ability to move beyond correlations without imposing further assumptions on the nature of the lag structure. The experimental data strongly suggests that the recall effects are stronger among those with a higher initial burden. These effects are significant for the total number of acute illnesses, the labor impact of sickness and the use of self-medication. They are also reasonably large for doctor's visits, but again, the small sample leads to imprecision in the estimated coefficients. In the observational data, the effects are uniformly large and significant and are actually quite close to the experimental estimates. The large size of the coefficients suggests that some portion of the recall effect is driven by those individuals and households with large morbidity burdens. In the observational results, the effect of recall on doctor's visits and the labor impact disappears entirely with the inclusion of the lagged morbidity interaction term, suggesting that when individuals and households have a large number of acute illnesses, they start to systematically forget entire episodes and it is precisely these households that drive the large observed differences between monthly and weekly reporting. This is consistent with our hypothesis

that, in poor settings, individuals have come to treat sickness and doctor's visits as a regular part of their daily lives rather than an extraordinary event.

Figure 1 illustrates this effect in the observational data. The figure plots the number of acute illnesses reported in the second period of data collection on acute illnesses as reported weekly in the first period. The second period data is divided into weekly and monthly recall period. Clearly weekly recall is higher than monthly. Further, the higher the number of illnesses in the initial year, the higher the number of illnesses and the larger the gap between monthly and weekly reports is in the next period. Interpreting this gap as "forgetting," the results suggest that forgetting occurs in strict proportionality with the number of health episodes, a result that even extends into the unreliable area of the data where observations are sparse.

Finally, Table 5 presents a robustness check on our results to account for missing observations. Specifically, in the weekly survey, surveyors were instructed to visit every household on the same day every week to ensure that the reference period was benchmarked to the previous visit. One of the repercussions of this rigid survey schedule was that 15 percent of our weekly observations were missing. In around half of these cases, the observation is missing because the surveyor did not visit the household due to a national holiday and in the remaining half, it is missing because the house was locked or the respondent could not be found despite repeated visits during the day. These missing variables create a particular problem when we aggregate weekly reports into monthly data since in the monthly recall surveys, a week outside the house would not result in a missing observation. In our previous regression specifications, we took the extreme position that the missing variables correspond to a fully healthy report for the individual in the missing week—this ensured that we computed a lower bound for the recall effect. In Table 5, we present estimates from two alternate assumptions—that the data were "missing at random" (which is certainly true for national holidays when surveyors did not visit the household) or that the missing data corresponded to a report of sickness with an associated doctor's visit. The estimates from these two alternate specifications are presented as the "Missing at Random" (MAR) rows and the "Upper Bound" (UB) rows. The "Lower Bound" (LB) rows replicate the estimates from previous regressions for comparability. As seen in the table, the MAR estimates are somewhat higher but fairly close to the LB estimates, while the UB estimates are an order of magnitude higher. The true recall effects are likely somewhere between the MAR and UB estimates, and probably close to the MAR. This is because the simultaneous interviews by anthropologists on the team suggest that, in some cases, households would return to their villages in order to better cope with severe illness. If this is the case, absence may be correlated with illness, leading MAR to underestimate the true recall effect.

Section III: Variation in bias by income

It is possible that these differences in weekly versus monthly data are not a terrible problem. If the underestimation of the frequency of illness and the use of health services were orthogonal to variables that researchers are interested in using to explain these phenomena, then the overall explanatory power of models using monthly data would be lower than with weekly and there would be bias in the scale of all coefficients uniformly – all effects would be about half of their real effect. Relative impacts of different variables would still be recoverable. However, in this section, we use the observational data to show that this recall effect is related to an important determinant of health care demand – household income. In this case, mis-measurement of the dependent variable biases gradients like the income elasticity of doctor visits, and therefore, affects our understanding of how illnesses are reported and how individuals seek treatment.

Specifically, we can decompose the income elasticity of expenditures on health, commonly believed to be quite high, into three parts – the probability of being ill, the probability of seeking treatment conditional on being ill, and the price of services sought. We present four figures that explore each of these components in turn. We then turn to a further aspect of health reporting. It is possible that many of the episodes not reported in the monthly recall surveys are "minor" with few impacts on households, and small expenditures. We therefore also look at the effect of the recall period on labor impacts of sickness, again across poor and non-poor households.

Figure 2 shows the number of family members who reported at least one acute illness in any one month when asked on a weekly versus a monthly basis as a function of the family's log per capita expenditure.[11] In this and other figures, the vertical line represents the income range beyond which the data is very sparse and therefore, less reliable. There is an obvious difference between the recall periods; the outcome variables measured with weekly recall are always significantly larger than those measured with monthly recall. The slope of the weekly recall curve is very steep, and for the highest income values, approaches the monthly recall curve. While more than two episodes of illness are forgotten per month by the poorest, the wealthiest forget less than half an episode per month on average. When measured using monthly data, the difference between the rich and the poor is slightly less than one acute episode per family per month. The apparent recall effect

---

[11] Chronic illnesses are excluded here because the survey methodology guarantees that the chronically ill are recorded as sick (with a chronic illness) once diagnosed regardless of the recall period used. To create monthly measures of reported doctor's visits and total expenditures from weekly data, we can just sum the visits and expenditures. Creating a monthly measure of health events is somewhat more difficult. For example, an event with duration greater than a week would be reported in multiple surveys in the weekly recall group but only in one survey in the monthly recall group, leading us to overestimate the total number of health events per month using the weekly data. As a result, we also variously use the total duration of health events in a month or condition on having at least one health event in a given month to ensure that the results on recall and morbidity are not driven by potential "double-counting".

alone goes a long way in explaining why the poor are less sick than the rich in most self-reported data. In fact, the poor are sicker, a conclusion that is clear in this high frequency data.

Similarly, and possibly more importantly from the policy perspective, this pattern is evident in the number of visits to medical practitioners (Figure 3).[12] Surveys commonly show that doctor's visits are positively correlated with income. In Figure 3, this pattern does, in fact, emerge but with some interesting nuances. In the monthly data – the most common recall period in use – there is a clear increase in doctor visits in the lower half of the sample, which plateaus in the middle of the sample. There may be a further increase at the top of the income spectrum, but this is an area with fewer observations. Except for the leveling off – which, unless a demand function had a careful, non-linear, specification of income effects, rather than the standard log–log specification, the regression would miss – our monthly results are similar to standard results.

However, when we turn to the weekly data, an entirely different story emerges. There may be a tiny increase of doctor visits with income at the lowest level of income, though this is neither noticeably nor significantly different from a flat line. Much more prominent is the dramatic **decline** of doctor visits throughout the reliable range of the data. Again, the number of doctor's visits may increase at the very top of the income spectrum, but the data is too sparse to tell. The decline in doctor's visits as a function of income, when measured in weekly surveys as opposed to the increase when measured monthly, raises suspicions about the conventional wisdom dictating that richer people visit medical practitioners more frequently.

Finally, we examine how recall and PCE affect health care expenditures. One issue that limits our ability to present a full characterization of the impact of recall periods on health expenditures is the sparse data on very high, catastrophic expenditures that can significantly alter the mean and are difficult to capture fully in this kind of micro-study. The impact of recall on expenditures and the limitations arising from the small sample are easily seen in Figures 4 and 5, which show total health expenditures by income. Figure 4 uses our whole sample, while Figure 5 caps health expenditures at Rs.1000 a month.

Figure 4 is difficult to interpret. There is no consistent difference between weekly and monthly recall given that the curves cross four times. However, inconsistencies in the curves are driven almost entirely by a tiny number of large expenditures. We hypothesize that these are particularly memorable events for which the monthly and the weekly recall quality is not likely to vary. When we group households into income terciles, the poorest group, with an average monthly per capita expenditure of Rs. 600, reports a single expenditure

---

[12] The term "doctors" has to be interpreted carefully here since the range of providers of medical services is very wide in India – from MD's and MBBS (standard medical university training) doctors to people knowledgeable of Indian traditional medicine (but prescribe allopathic medicine anyway) to genuine "quacks" with no training whatsoever (Das and Hammer, 2005).

episode of Rs. 20000 and the top 23 expenditures out of 4158 total individual health visits account for over 50 percent of this tercile's total health expenditure. Similarly, for the rich, the top tercile had an average monthly per capita expenditure of about Rs. 3500, and the highest 66 out of 3228 expenditures also account for over 50 percent of the group's health expenditures. Examining the effect of recall on health spending that arises from catastrophic illnesses turns out to be outside the analytical scope of our data.

Figure 5 exclude health expenditures greater than Rs. 1000 in a month (less than 5 percent of the sample). Once again, the monthly and weekly figures are completely different, with the monthly recorded health expenditures both much lower than weekly recorded health expenditures and much more responsive to household expenditures. With an income differential of about a factor of 7, our estimate of the relationship between monthly total expenditure and an income proxy is in line with the literature, which typically reports income elasticity above one. In fact, for the lower half of the sample, the implied income elasticity is well above one; the concavity noticeable in the non-parametric regression moderates the overall effect. For the weekly recorded expenditures, the story is entirely different, with poorer people spending less than others – the slope increases slightly – but by less than a factor of one and by only a half for the most reliable portion of the data. So, contrary to much of the theoretical and empirical literature, controlling for recall bias, it appears that healthcare is a normal but not a luxury good.

One overarching theme from the analysis thus far is that altering the recall period has an impact on reported morbidity and health seeking behavior, but more so for illnesses that are not associated with large expenditures. To the extent that large health expenditures are the ones associated with household debt or inability to smooth consumption, perhaps the illnesses that are forgotten are not particularly important for economic analysis that looks at the effect of illness on household outcomes such as debt, disability or school completion (for other outcomes such as the industrial organization of the medical market where the number of visits is a concern, these minor illnesses will be important). Figure 6 looks at a variable that has received considerable attention from labor economists—the labor impacts of sickness. As before, we aggregate the impacts to the household level and look at the non-parametric relationship between the number of members reporting at least one lost work/school day in the previous month due to an acute illness and (log) per-capita expenditures.

As for reports of acute illnesses and doctor visits, the labor impact of sickness declines with PCE. Among the poorest households, in an average month between 0.7 and 0.8 members say they lost a day of work/school due to sickness in weekly reporting. With monthly reporting, this drops down to between 0.3 and 0.4. Again,

at lower levels of PCE, recall impacts are larger than at higher. The pattern bears more similarity to that of recall of acute events than of expenditures.

Table 6 repeats the basic regressions in Table 3 to show that the response heterogeneity by income visible in the figures is, indeed, born out statistically. In these analyses, we restrict ourselves to the observational data and augment the estimating equation with the inclusion of household PCE and household PCE interacted with recall period. The estimates largely confirm the results from the figures in the observational data. For all our outcome variables, the recall effect is greater for the poor than for the rich. These effects are uniformly significant and suggest, for instance, that at the household level, an increase in (log) PCE by 1 increases the size of the recall effect by close to 0.4 doctor visits per month. The duration estimates at the household level are particularly large. Since log PCE varies between 5 and 10 in our sample, our estimate suggests that monthly surveys may underestimate the total time spent sick by members of a household by as much as 18 days a month. The labor impact of sickness shows a similar pattern with declines in the importance of the recall effect at higher levels of PCE, although in this case the estimate is less precise.

Further Robustness Checks

The recall period affects reports of morbidity and health seeking as well as the labor impacts of sickness. We investigate several ancillary specifications to check whether (a) these results could be driven by specific age populations, (b) specific respondents or, (c) whether the differential impact between the rich and poor could be picking up the effect of attributes correlated with PCE, notably education. For brevity, only the education results are presented in Appendix Table 2.

Age Impacts. One-third of our sample is below the age of 15. To check whether this recall effect is driven by high morbidity and differential recall for sickness among children, we restricted the sample to those above the age of 15 and find no difference in the estimated parameters. The estimates in the observational data for number of acute illnesses, the probability of a single acute illness, the labor impacts of acute illnesses and doctor visits are respectively 0.55 (0.57), 0.13 (0.12), 0.024 (0.029) and 0.13 (0.14) in the full and adult-only samples, the latter in parenthesis. For the labor impacts of acute illnesses where the differences are somewhat larger, the interaction between adults and the recall impact is not significant (p-value = 0.859).

Respondent Impacts. Every effort was made to retain the same primary respondent in all surveys, weekly and monthly. We code an indicator variable for whether the respondent was the designated primary respondent to check for respondent bias. In the case of the weekly surveys, this is aggregated up to the percent of weeks in the relevant month that the respondent was the designated primary respondent. The data suggest that

surveyors were fairly successful at maintaining respondent consistency. In the observational data the primary respondent accounted for 79.8 percent of all reports at the weekly level and the fraction drops somewhat to 77.5 percent in the monthly reports. Despite the qualitatively similar averages, the difference between the two is significant at the 90 percent level of confidence. We therefore estimate recall effects with an additional indicator variable that takes the value 1 if the respondent was the designated primary respondent (for weekly reports aggregated to the monthly level this becomes the fraction of weeks that the primary respondent answered). For the reported number of acute illnesses and total self-medication we find small respondent effects (when the respondent was the primary respondent, reports increase) with significance at the 90 percent level of confidence. In these cases, the estimated recall impact decreases from 0.57 to 0.55 (number of acute illnesses) and from 0.95 to 0.947 (total self-medication) and the declines are not statistically significant. For the other outcome variables considered here, we are unable to detect respondent effects in these specifications (with p-values on respondent identity ranging from 0.32 to 0.76) and there is no impact on the estimated size of the recall effects.

Income or Education? Finally, heterogeneity in the recall impact by PCE could be driven by covariates correlated with PCE rather than PCE itself. Although most naïve estimates of income-elasticity used in the health literature suffer from similar interpretational issues, it is possible that education (which is highly correlated with PCE) rather than PCE drives the greater recall impacts among the poor. We re-estimate the heterogeneous impact of the recall period with PCE at the household-level with an additional interaction between the education of the household head and the recall period. These regression results are reported in Appendix Table 2. Across the outcomes considered here, we find two consistent patterns. First there is a small drop in the size of the recall impact and in its interaction with PCE, suggesting that part of the heterogeneous impacts could indeed arise from differences in education. However, for all our outcome variables, the qualitative (and quantitative) finding that recall impacts are higher for the poor remains robust to the inclusion of the additional interaction. Second, for all specifications, the interaction with education is in the expected direction—the impact of the recall period is lower for more educated households—but the effect sizes are small and generally lack precision (except for the labor impact of acute illness, which is significant at the 90 percent level of confidence).

Conclusion

Common sources of information concerning health conditions—the National Sample Surveys in India and the Living Standards Measurement Surveys of the World Bank—use between 2 weeks and a month as a standard recall period for illnesses. Ordinarily, we assume that decisions about recall length don't bias results. In this paper, we find that they do; the choice of recall period has substantive implications. With shortened

recall periods, poor people are ill for large fractions of the year with ailments that they believe to be short-term and, apparently, easily forgotten. Richer people, who do not forget as easily, do not suffer from acute illnesses nearly as much as the poor. The recall impacts documented here may have implications for our understanding of health reporting across rich and poor households, and measurement and policies regarding access to health care. Take each in turn.

One long-standing question in the literature on health in developing countries is why the rich appear to be "sicker," as assessed by self-reports, than the poor. One explanation is that people have different "thresholds" or intolerance to discomfort and that these thresholds decrease with income (Thomas and Frankenberg, 2002 or Paqueo and Gonzalez, 2003, Strauss and Thomas (1998)). Another is that rich people have more accurate information concerning chronic conditions.[13] Richer people may respond "yes" to the question of whether they've been ill if they have been diagnosed with a chronic illness like diabetes. A poorer person will not know to report the underlying chronic condition unless there has been an acute episode of the ailment in the recall period. This correlation between reported health and information extends beyond chronic illnesses—for instance, Dow and others (2008) show in an experimental context that increasing the price of health care leads to a drop in physician visits but an improvement in reported health status. They interpret this result as consistent with a positive correlation between information about health and the probability of reporting sick.[14]

This paper suggests that in addition to the problems of reporting thresholds and information about chronic illnesses, survey design may drive some of the results relating self-reported illness and income. Since the effect of the recall period declines with PCE, it influences estimates of the income elasticity of doctor visits, use of medicines, number of illnesses, and amount spent on health care. We cannot decompose the differential reporting of rich and poor into shares contributed by the competing hypotheses, but we do offer evidence of the substantial role played by recall effects. Thus, while differential sensitivity to severity across income groups is an important factor in self-reports, recall bias, awareness of chronic illness and the interaction between the two adds another set of considerations to explain the puzzling empirical regularity of the "sicker rich."

Patterns of health usage reported here are also problematic for measurements of inequality in access. Standard theories about inequality in health access or the links between health expenditures and per-capita expenditures are consistent with our monthly recall data but harder to reconcile with observed patterns once

---

[13] See, for example, the interchange between Murray and Chen (1992,1993) and Riley (1993).

[14] In the Delhi sample, we examined this possibility and also find substantial evidence of a link between information and the reporting of chronic illnesses (Das and Hammer, 2009).

the recall period is shortened to a week. For instance, in the health inequality literature, researchers have suggested that inequality in access explains the difference in the number of doctor's visits between the rich and the poor (Wagstaff and Van Doorslaer, 2000). This hypothesis makes sense in our monthly recall data where the poor visit doctors less than the rich, but is not consistent with the patterns observed in the weekly data where the poor visit doctors more than the rich. It might be possible to reconcile the results with inequality in access to *high quality health care*, but what a measure of health inequality that accounts for quality may look like (and what measurements are required to construct it) is an open question.

The relative frequency of visits to medical providers in our data by both the rich and the poor also requires us to reconsider some commonly prescribed solutions to the health care problem in poor countries. Since the Alma Ata declaration in 1978, policymakers have assumed that they should focus on providing poor people with better "access" to primary medical care.[15] Because the main obstacles to access are the price of care and distance, the obvious solution was to open more facilities close to poor people and heavily subsidize (or, preferably, make free) primary care. Our results show that the use of health care providers by the very poor was *higher* than that of the somewhat better-off. Furthermore, the out-of-pocket expenditures of the very poor correspond to the higher-than-previously-thought use of health services. Again, with shorter recall periods the degree to which very poor people choose to spend on health rises accordingly. Yet, the health of the poor in Delhi—as measured through outcomes data in the National Family Health Survey or the health rounds of the National Sample Survey—is much worse than that of the rich. If the use of health care providers and the extent of health spending do not decrease gaps in health outcomes between the poor and the rich, what is the missing link? Poor health outcomes may be driven by lack of knowledge, particularly about appropriate preventive measures, the inability of the poor to distinguish between bad and good quality health care, or poor diagnosis of underlying conditions.[16] But simple health care access is not the root of the problem, at least in this context.

---

[15] This position has been reiterated in the World Health Report 2008 by the WHO on the thirtieth anniversary of Alma Ata.
[16] Jalan and Somnathan (2008), Das and Hammer (2009)

References

Banerjee, Abhijit, Deaton, Angus and Esther Duflo. 2004. "Wealth, Health and Health Services in Rural Rajasthan." *American Economic Review*, Papers and Proceedings 94(2): 326-330

Chaudhury, Nazmul., Hammer, Jeffrey, Kremer, Michael, Muralidharan, K. and F. Halsey Rogers. 2006. "Missing in Action: Teachers and Health Workers in Developing Countries." *Journal of Economic Perspectives* 20(1) Spring.

Das, Jishnu and Carolina Sanchez. 2002. "Short but not Sweet: New Evidence on Short Duration Morbidities From India." *World Bank Policy Research Working Paper 2971.*

Das, Jishnu and Jeffrey Hammer. 2005. "Which Doctor? Combining Vignettes and Item-Response to Measure Doctor Quality," *Journal of Development Economics*, 78: 348-383.

Das, Jishnu and Jeffrey Hammer. 2007. "Money for Nothing: The Dire Straits of Medical Practice in Delhi, India," *Journal of Development Economics*, 83(1): 1-36.

Das, Jishnu and Jeffrey Hammer. 2009. "Chronically Misinformed: Chronic Illness and Information in a Low-Income Country." The World Bank, Processed.

Das, Jishnu, Alaka Holla, Veena Das, Manoj Mohanan, Diana Tabak and Brian Chan. 2011. "The Quality of Medical Care in India: Evidence from a Standardized Patient Study." Processed.

Dow, Will, Gertler, Paul, Schoeni, Robert, Strauss, John, and Duncan Thomas. 1997. "Health Care Prices, Health and Labor Outcomes: Experimental Evidence." *RAND Labor and Population Program Working Paper Series 97-01.*

Jalan, Jyotsna and E. Somnathan. 2008. "The Importance of Being Informed: Experimental Evidence on Demand for Environmental Quality." *Journal of Development Economics,* 87(1): 14-28.

Leonard, Kenneth L., Masatu, Melkiory C. and Alex Vialou. 2007. "Getting Doctors to Do Their Best: The Roles of Ability and Motivation in Health Care." *Journal of Human Resources*, 42 (3): 682-700.

Kakwani, Nanak & Wagstaff, Adam & Eddy van Doorslaer 1997. "Socioeconomic Inequalities in Health: Measurement, Computation, and Statistical Inference." *Journal of Econometrics*, 77(1): 87-103.

Murray, Christopher J. L. and Lincoln C. Chen. 1992. "Understanding Morbidity Change." *Population and Development Review*, 18(3): 481-503

"National Hospital Ambulatory Medical Care Survey (NHAMCS): Emergency Department." 2000. National Center for Health Statistics and US Census Bureau [Distributor]

Paqueo, Vincente and Christian Gonzalez. 2003. "Economic Analysis of Health Care Utilization and Perceived Illness: Ethnicity and Other Factors." *World Bank Policy Research Working Paper 3125*.

Riley, James. 1993. "Understanding Morbidity Change: Comment on an Article by Murray and Chen." *Population and Development Review*, 19(4): 807-811.

Strauss, John and Duncan Thomas. 1998. "Health, Nutrition, and Economic Development." *Journal of Economic Literature*, 36(2): 766-817.

Thomas, Duncan and Elizabeth Frankenberg. 2002. "Health, Nutrition, and Prosperity: A Microeconomic Perspective." *Bulletin of the World Health Organization*, 80(2): 106-113.

Wagstaff, A. and E. van Dorrslaer. 2000. "Measuring and Testing for the Inequity in the Delivery of Health Care." *Journal of Human Resources*, 35(4): 716-733.

World Health Organization. 2008. "The World Health Report 2008: Now More Than Ever." Geneva, Switzerland: WHO Press.

Alix Peterson Zwane, Jonathan Zinman, Eric Van Dusen, William Pariente, Clair Null, Edward Miguel, Michael Kremer, Dean S. Karlan, Richard Hornbeck, Xavier Giné, Esther Duflo, Florencia Devoto, Bruno Crepon, and Abhijit Banerjee. 2011. "Being surveyed can change later behavior and related parameter estimates." PNAS 2011 108 (5) 1821-1826.

Table 1.A: Sample Proportions - Individuals					
Individual and Household Characteristics	Categories [A]	Poor (1)	Middle (2)	Rich (3)	Total (4)
Income	Average Monthly Income – Rs. (per capita in 2000)	1,666	2,384	5,493	3,181
Morbidity	Reported ever sick	82.74	77.4	78.79	80.49
	(Diagnosed) Chronic Condition	9.33	15.63	31.81	18.14
Gender	Female	45.07	48.77	50.72	48.05
Age Distribution	0-4	10.85	6.74	5.29	8.05
	5-14	32.9	22.16	15.64	24.03
	15-29	25.18	32.95	27.27	28.34
	30-49	25	26.4	31.92	27.52
	>50	6.07	11.75	19.87	12.05
Education of household head	Illiterate	44.8	24.47	5.1	22.67
	Primary-Middle School	36.78	39.36	16.32	30.33
	Middle-High School	16.09	29.79	21.42	23.00
	> High School	2.30	6.38	57.14	24.00
Marital Status	Now Married	38.73	41.05	50.73	43.16
	Never Married	58.27	55.18	42.62	52.47
	Widowed	2.99	3.58	6.44	4.24
Migrant Status	First Generation Migrant	34.86	34.46	30.56	33.42
Labor Force Status	Public Sector	4.66	8.8	8.38	7.46
	Private Sector	16.85	10.26	9.22	11.76
Employment Status (> 15 years old)	Informal Sector	29.75	25.22	28.21	27.61
	Unemployed	3.94	3.52	1.68	2.97
	Retired	1.08	2.64	4.19	2.76
	Out of Labor Force	43.73	49.56	48.32	47.44
Sanitation	Flush System	0	75.26	92.78	56.01
	Septic Tank	25.77	9.28	1.03	12.03
	Service Latrine	52.58	12.37	1.03	21.99
	Other Latrine	21.65	3.09	5.15	9.96
Light Source	% Using Electricity as Main Light Source	88.66	98.97	95.88	94.5
Drinking Water Source	Tap Water	77.32	96.91	100	91.41
	Tank/Pond	22.68	2.06	0	8.25
Cooking Fuel	LPG (Petroleum Gas Cylinders)	45.36	83.33	94.85	74.48
	Firewood or Kerosene	51.54	12.50	0	8.28
Number of individuals [C]		568	531	481	1,621
Number of households		97	97	97	300
Household size		5.85	5.47	4.96	5.4

Notes: All numbers in percentages. Income information is in Rupees for the year 2000. The income information is not available for 9 households (41 individuals

	Observation	Observational data		Experimental data		Difference (Mean Monthly - Mean Weekly)	
		Monthly (1)	Weekly (2)	Monthly (3)	Weekly (4)	Observational (5)	Experimental (6)

Table 2: Mean Comparison for Weekly and Monthly Recall Period

Variable	Observation	Monthly (1)	Weekly (2)	Monthly (3)	Weekly (4)	Observational (5)	Experimental (6)
Number of acute illnesses	Individual	0.375	0.962	0.360	0.943	-0.587***	-0.582***
		(0.571)	(1.251)	(0.570)	(1.286)	(0.013)	(0.056)
	Household	1.945	5.064	1.942	5.150	-3.119***	-3.208***
		(1.673)	(4.087)	(1.656)	(3.786)	(0.094)	(0.385)
At least one acute was reported	Individual	0.329	0.475	0.314	0.454	-0.146***	-0.140***
		(0.470)	(0.499)	(0.464)	(0.498)	(0.006)	(0.024)
	Household	1.533	2.377	1.673	2.479	-0.844***	-0.806***
		(1.414)	(1.781)	(1.424)	(1.634)	(0.046)	(0.180)
Duration of acute illnesses	Individual	4.191	4.492	4.048	4.818	-0.300***	-0.770*
		(8.629)	(7.182)	(8.435)	(7.992)	(0.103)	(0.397)
	Household	21.759	23.653	21.811	26.319	-1.895***	-4.508*
		(23.089)	(21.857)	(24.377)	(22.527)	(0.668)	(2.619)
Total number of doctor visits	Individual	0.247	0.403	0.256	0.439	-0.156***	-0.183***
		(0.560)	(0.817)	(0.523)	(0.849)	(0.009)	(0.038)
	Household	1.279	2.120	1.379	2.397	-0.841***	-1.019***
		(1.512)	(2.306)	(1.394)	(2.346)	(0.058)	(0.245)
Number of doctor visits for acute illnesses	Individual	0.196	0.334	0.209	0.383	-0.138***	-0.174***
		(0.483)	(0.714)	(0.476)	(0.772)	(0.008)	(0.035)
	Household	1.016	1.758	1.126	2.091	-0.742***	-0.965***
		(1.325)	(2.070)	(1.326)	(2.175)	(0.052)	(0.227)
Number of doctor visits for chronic illnesses	Individual	0.051	0.069	0.047	0.056	-0.018***	-0.009
		(0.263)	(0.380)	(0.251)	(0.300)	(0.004)	(0.014)
	Household	0.264	0.363	0.252	0.306	-0.099***	-0.054
		(0.605)	(0.898)	(0.606)	(0.744)	(0.023)	(0.081)
At least one day of work/school lost due to sickness	Individual	0.063	0.096	.050	0.10	-0.033***	-0.053
		(0.024)	(0.295)	(0.219)	(0.305)	(.004)	(0.014)
	Household	0.293	0.481	0.272	0.570	-0.187***	-0.298***
		(0.616)	(0.805)	(0.563)	(0.923)	(0.020)	(0.096)
Total number of self-medication taken	Individual	0.390	1.330	0.423	1.240	-0.941***	-0.816***
		(0.958)	(2.790)	(1.028)	(2.556)	(0.027)	(0.111)
	Household	1.818	6.656	2.260	6.772	-4.838***	-4.512***
		(2.694)	(7.845)	(3.069)	(7.022)	(0.168)	(0.711)
At least one self medication was taken	Individual	0.205	0.348	0.218	0.340	-0.143***	-0.122***
		(0.404)	(0.476)	(0.413)	(0.474)	(0.006)	(0.023)
	Household	0.956	1.742	1.163	1.857	-0.786***	-0.693***
		(1.157)	(1.479)	(1.158)	(1.372)	(0.038)	(0.150)

Notes: * Significant at the 10%; **significant at the 5%; *** significant at 1%. Columns 1-4 contain means with standard deviations in parentheses. Columns 5 and 6 contain mean difference between monthly and weekly data, and parentheses contain standard errors.

Table 3: Effect of Recall Controlling for Time-Invariant Individual Heterogeneity and Seasonality							
	Observational Data				Experimental Data		
	Individual (1)	R^2 (2)	Household (3)	R^2 (4)	Individual (5)	R^2 (6)	Household (7)	R^2 (6)
Number of Acute Illnesses	0.557***	0.134	2.971***	0.315	0.507***	0.336	2.780***	0.503
	(0.025)		(0.483)		(0.057)		(0.560)	
At Least One Acute Illness	0.132***	0.044	0.760***	0.146	0.092***	0.154	0.532**	0.397
	(0.009)		(0.129)		(0.030)		(0.189)	
Doctor's Visits	0.134***	0.027	0.734***	0.108	0.087**	0.107	0.465	0.207
	(0.012)		(0.186)		(0.040)		(0.300)	
Duration of Acute Illnesses	0.380**	0.004	2.267	0.397	0.553	0.017	3.234	0.506
	(0.164)		(1.623)		(0.432)		(2.689)	
At least one lost workday due to sickness	0.024***	0.017	0.144***	0.072	.009	0.033	0.038	0.122
	(.005)		(0.053)		(0.57)		(0.083)	
Number of Self-Medications Taken	0.950***	0.103	4.750***	0.401	0.927***	0.292	5.018***	0.437
	(0.062)		(0.802)		(0.116)		(1.099)	
At Least One Self-Medication Taken	0.133***	0.049	0.716***	0.186	0.122***	0.169	0.673**	0.227
	(0.008)		(0.109)		(0.029)		(0.209)	

Notes: * Significant at the 10%; **significant at the 5%; *** significant at 1%. *Weekly:* Sample size for individual observational regressions is 23669 and for household observational regressions it is 1094. Sample size for experimental individual regressions is 4528 and for household experimental household regressions is 205. All regressions in observational data are with individual fixed effects and calendar month dummies with clustering at the household level or with household fixed effects and clustering at the locality. All Experimental results are with one lag, month dummies and clustering at the household level. Robust standard errors in parentheses.

Table 4: Effect of Disease Burden on Recall

	Observational Data			Experimental Data		
	Weekly	Weekly*Acute in Period One for Weekly Data	R^2	Weekly	Weekly*Acute in Period One for Weekly Data	R^2
	(1)	(2)	(3)	(4)	(5)	(6)
Number of Acute Illnesses	0.088***	0.137***	0.224	0.130**	0.379***	0.374
	(0.019)	(0.004)		(0.056)	(0.047)	
At Least One Acute Illness	0.067***	0.019***	0.051	0.087**	0.012	0.216
	(0.011)	(0.002)		(0.035)	(0.018)	
Doctor's Visits	0.004	0.038***	0.041	0.043	0.048	0.123
	(0.013)	(0.004)		(0.041)	(0.033)	
Duration of Acute Illnesses	-0.475***	0.249***	0.01	0.397	0.153	0.397
	(0.178)	(0.038)		(0.372)	(0.412)	
At least one workday lost due to sickness	0.002	0.006***	0.025	0.129**	0.378***	0.374
	(0.005)	(0.0013)		(0.055)	(0.047)	
Number of Self-Medications Taken	0.523***	0.124***	0.128	0.348***	0.581***	0.436
	(0.067)	(0.016)		(0.100)	(0.108)	
At Least One Self-Medication Taken	0.056***	0.022***	0.064	0.062**	0.059***	0.213
	(0.008)	(0.002)		(0.031)	(0.019)	

Note: * Significant at the 10%; **significant at the 5%; *** significant at 1%.Sample size for individual observational regressions is 23669 and for household observational regressions it is 1094. Sample size for experimental individual regressions is 4528 and for household experimental household regressions is 205.All regressions in observational data are with individual fixed effects and calendar month dummies with clustering at the household level. All Experimental results are with one lag, month dummies and clustering at the household level. Robust standard errors in parentheses

Table 5: Robustness Check Using MAR and UB Assumptions For Missing Observations
Household Level

	Observational Data		Experimental Data	
	Weekly	R^2	Weekly	R^2
	(1)	(2)	(3)	(4)
Number of acute illnesses (LB)	2.971***	0.315	2.780***	0.503
	(0.483)		(0.560)	
Doctor visits (LB)	0.734***	0.108	0.465	0.207
	(0.186)		(0.300)	
Number of acute illnesses (MAR)	3.479***	0.386	2.901***	0.520
	(0.528)		(0.537)	
Doctor visits (MAR)	0.935***	0.123	0.506	0.213
	(0.215)		(0.286)	
Number of acute illnesses (UB)	4.836***	0.474	3.742***	0.536
	(0.535)		(0.483)	
Doctor visits (UB)	2.575***	0.274	1.425***	0.257
	(0.246)		(0.210)	

* Significant at the 10%; **significant at the 5%; *** significant at 1%. Weekly: This variable equals 1 if the data were collected weekly, and is zero otherwise. Lower Bound (LB): Assigns the value 0 for any observation when the weekly observation is missing but there is at least one observation for this month. Upper Bound (UB): Assigns the value 1 for any observation when the weekly observation is missing but there is at least one observation for this month. Missing at random (MAR): Scale up the outcome variable by the missing weeks for the individual (assigns the average value observed for this person in this month when the weekly observation is missing). Note: All regressions in observational data are with household fixed effects and calendar month dummies with clustering at locality. All Experimental results are with one lag, month and clustering at locality. Robust standard errors in parentheses

Table 6: Effect of Recall Period and Heterogeneity by PCE (Observational Only)

	Individual			Household		
	Weekly	Weekly*Log pc expenditures	R^2	Weekly	Weekly*Log pc expenditures	R^2
	(1)	(3)	(4)	(5)	(7)	(8)
Number of acute illnesses	0.890***	-0.048	0.133	9.459***	-0.906***	0.360
	(0.217)	(0.030)		(2.595)	(0.320)	
At least one acute	0.290***	-0.023**	0.044	2.804***	-0.285***	0.168
	(0.078)	(0.011)		(0.823)	(0.106)	
Doctor's visits	0.482***	-0.050***	0.028	3.560***	-0.397***	0.118
	(0.116)	(0.016)		(1.041)	(0.134)	
Duration of acute illnesses	4.846***	-0.640***	0.006	28.255***	-3.653***	0.025
	(1.293)	(0.181)		(8.457)	(1.055)	
At least one workday lost due to illness	0.048	-0.004	0.0172	0.626*	-0.067	0.077
	(0.039)	(0.005)		(0.346)	(0.042)	
Number of self-medication taken	-3.896***	0.696***	0.132	-10.222**	2.127***	0.334
	(0.637)	(0.094)		(4.699)	(0.666)	
At least one self medication was taken	0.294***	-0.023**	0.050	2.371***	-0.228***	0.184
	(0.071)	(0.010)		(0.715)	(0.092)	

* Significant at the 10%; **significant at the 5%; *** significant at 1%.Sample size for individual observational regressions is 23669 and for household observational regressions it is 1094. Sample size for experimental individual regressions is 4528 and for household experimental household regressions is 205. All regressions in observational data are with individual fixed effects and calendar month dummies with clustering at the household level. All Experimental results are with one lag, month dummies and clustering at the household level. Robust standard errors in parentheses

Figure 1: Number of Acute Illnesses by Health Burden From Previous Study and Recall Period

Recall and Morbidity Burden

Sparse data beyond here

Acute Illnesses Reported: Year 1
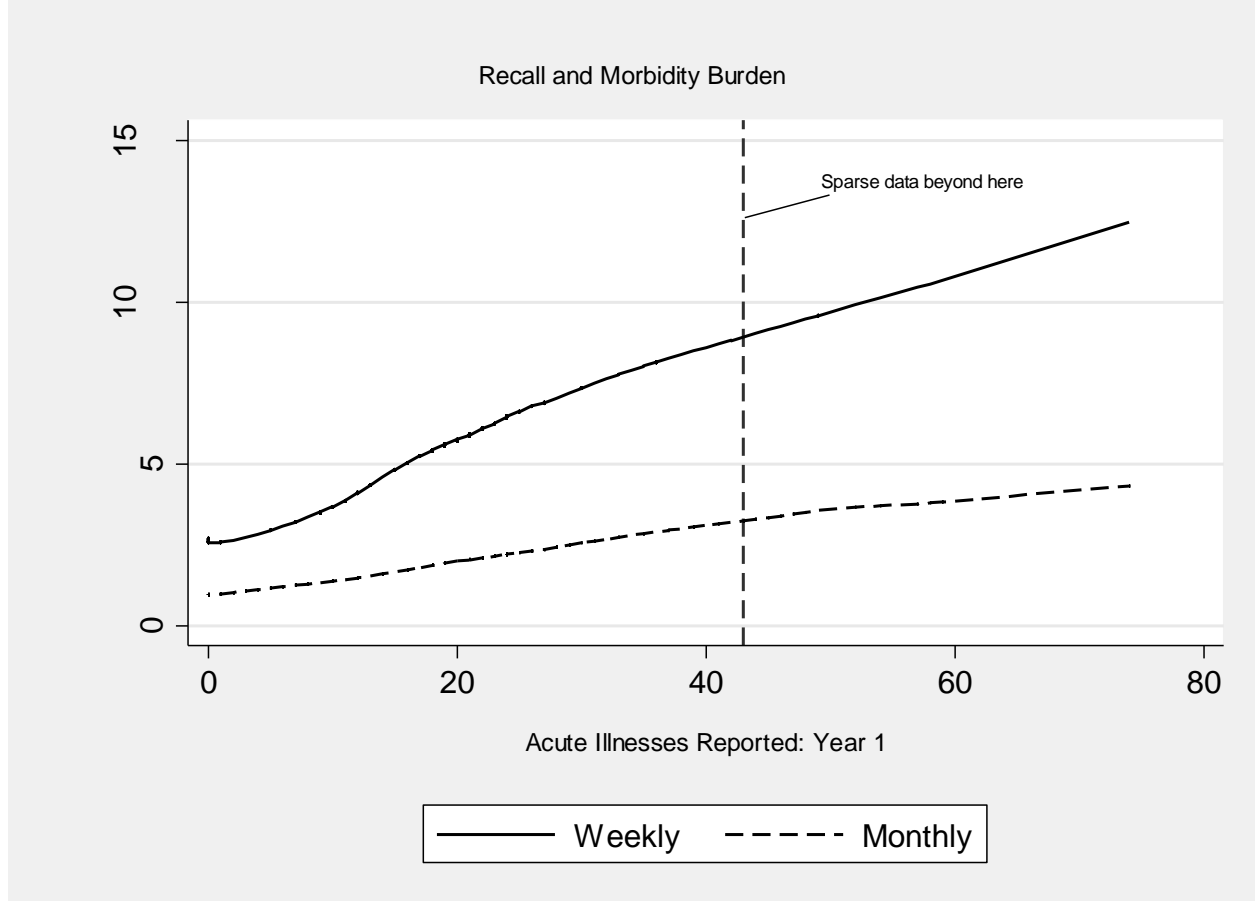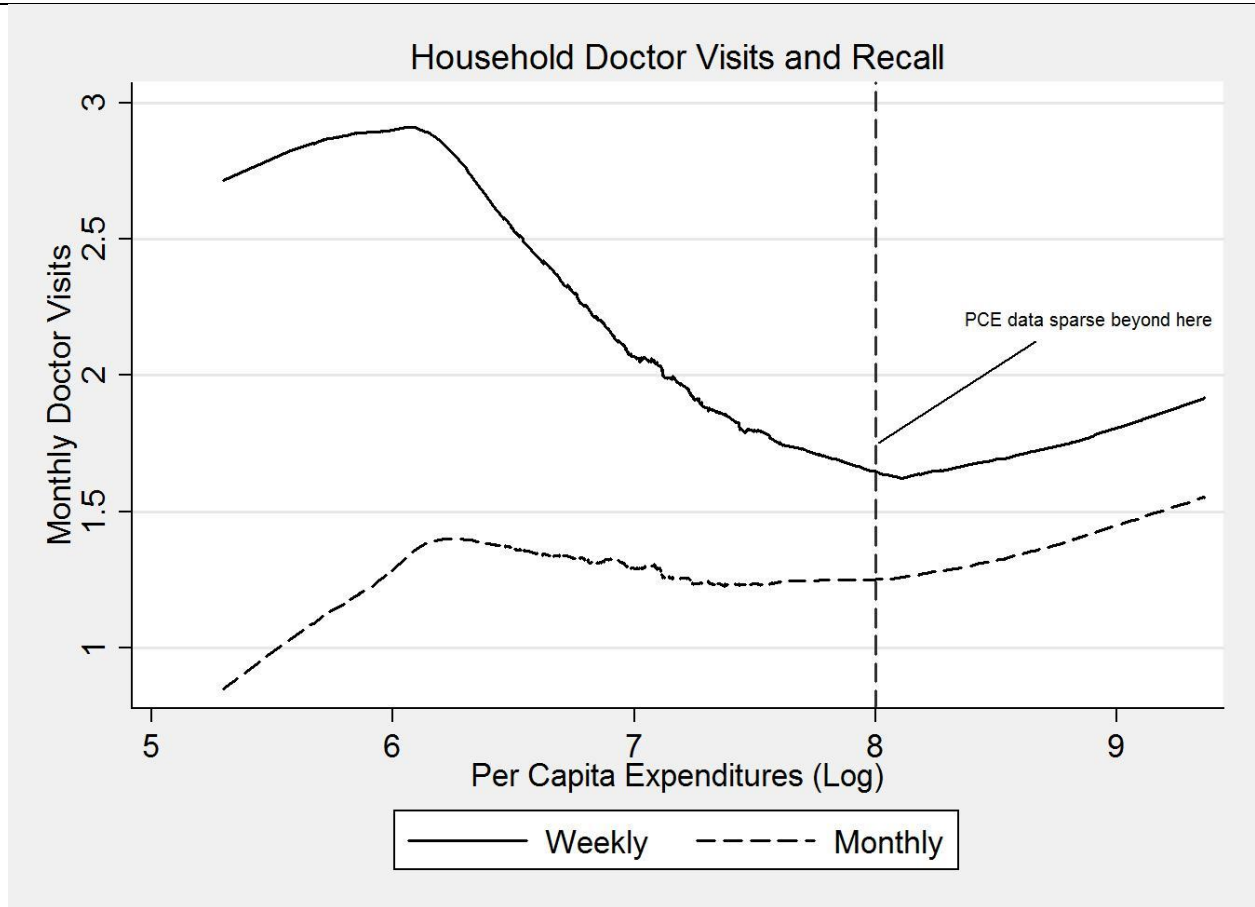
Weekly ----- Monthly

Note: The figure shows the non-parametric (lowess) relationship between acute illnesses reported in the first year of observational data collected weekly with acute illnesses reported in the second year collected weekly/monthly. Data are sparse beyond the indicated vertical line.

**Figure 2: Number of Acute Illnesses By Recall Period and Log PCE**
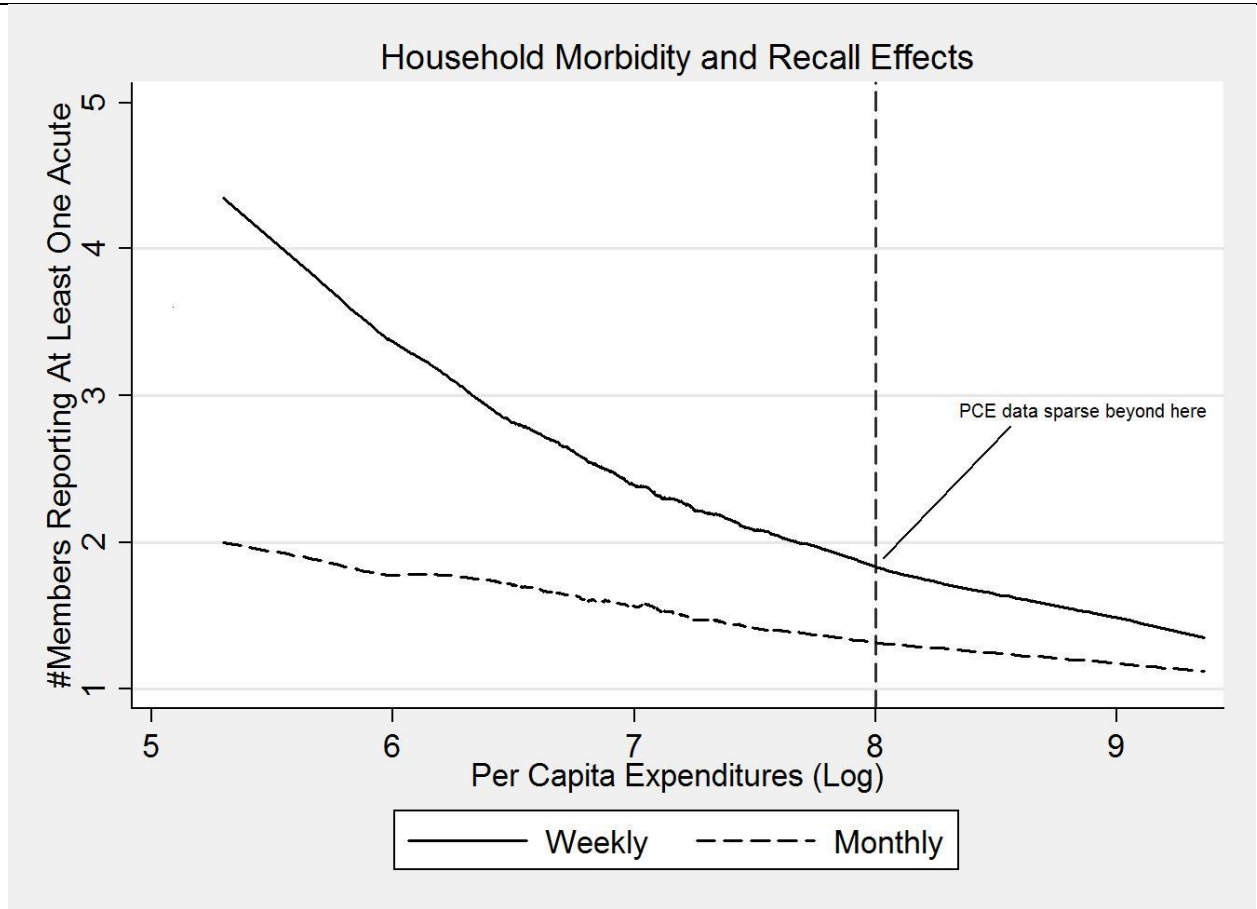
Note: This figure shows the non-parametric (lowess) relationship between the number of members reporting at least one acute illness in the previous month and (log) per-capita expenditures using monthly and weekly recall periods. The data on PCE are sparse beyond the indicated vertical line.

**Figure 3: Number of Visits to Medical Practitioners by Recall Period and Log PCE**

Note: This figure shows the non-parametric (lowess) relationship between the total number of doctor visits in the previous month and (log) per-capita expenditures using monthly and weekly recall periods. The data on PCE are sparse beyond the indicated vertical line.

Figure 4: Total Health Expenditures by Recall Period and Log PCE
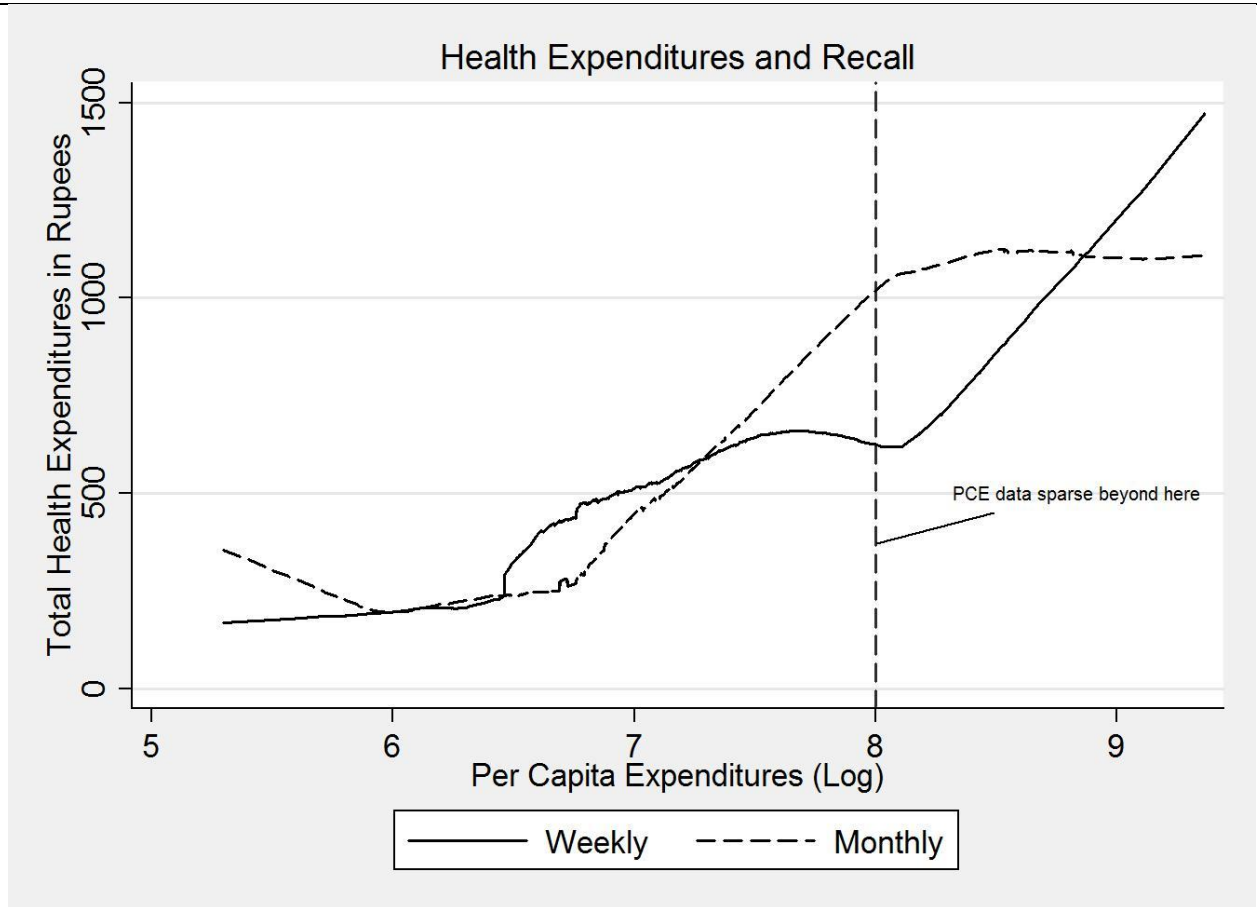
Note: This figure shows the non-parametric (lowess) relationship between health expenditure in the previous month and (log) per-capita expenditures using monthly and weekly recall periods. The data on PCE are sparse beyond the indicated vertical line.

Figure 5: Total Health Expenditures by Recall Period and Log PCE, Capped at 1000R
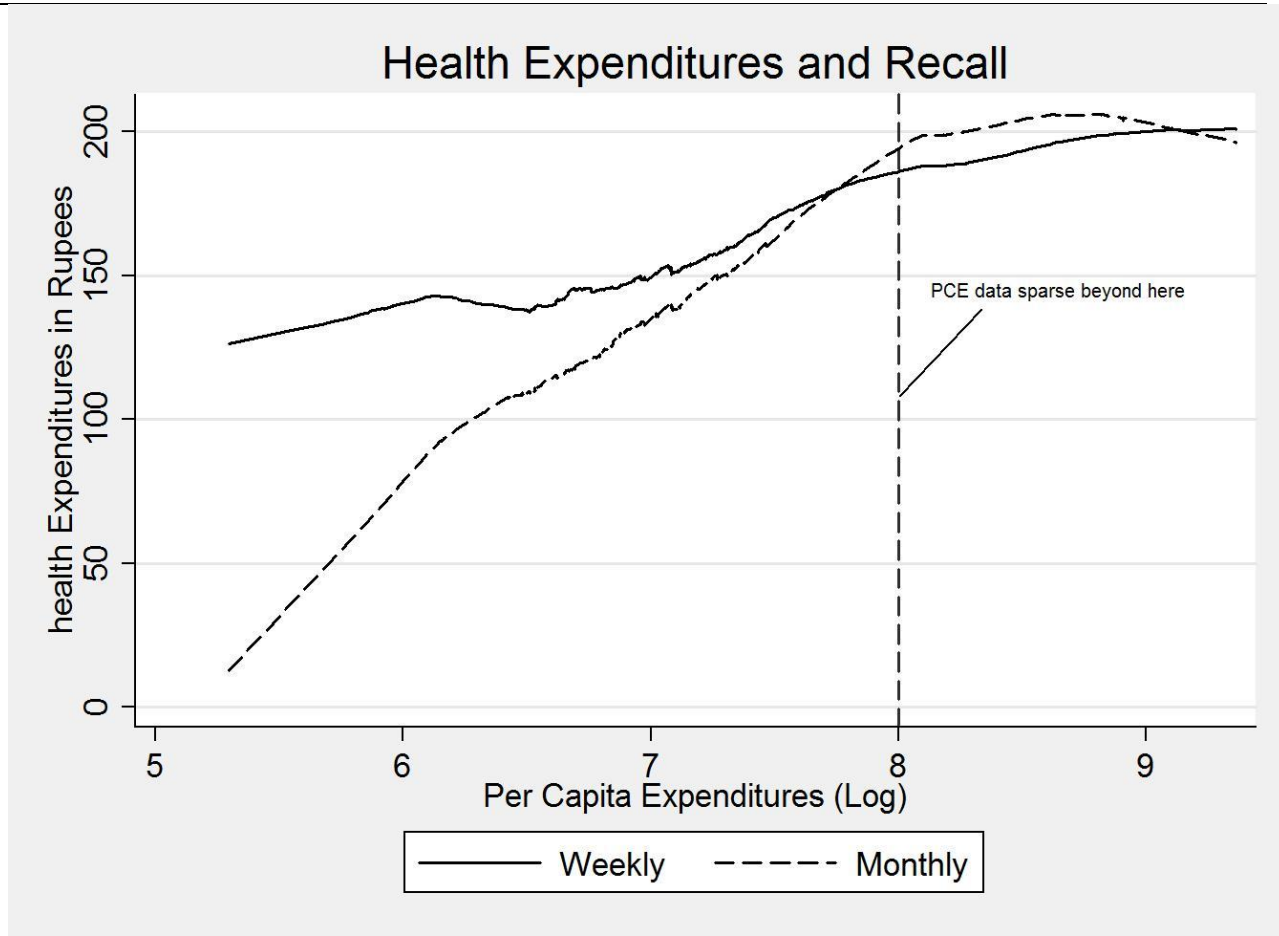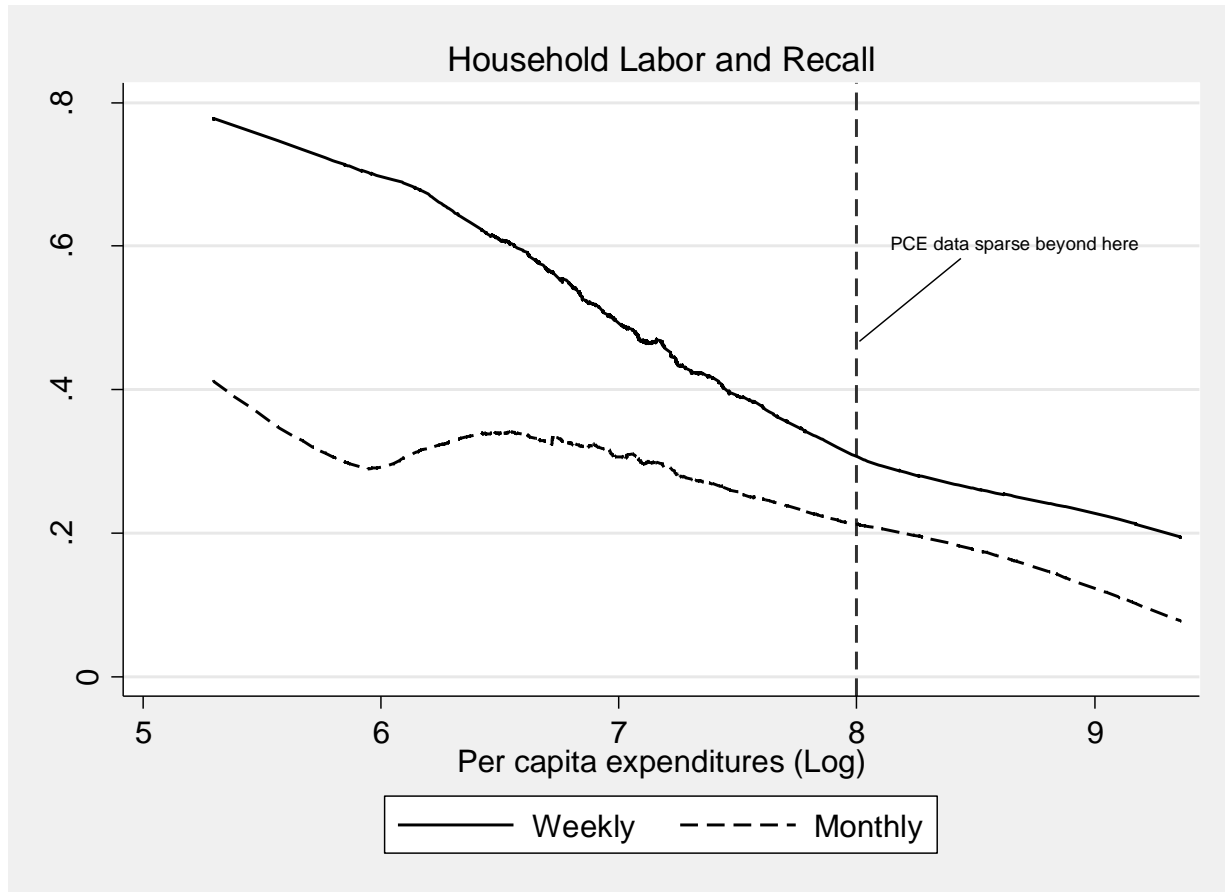

Note: This figure shows the non-parametric (lowess) relationship between health expenditure in the previous month and (log) per-capita expenditures using monthly and weekly recall periods. We exclude the 5 percent of expenditure episodes exceeding Rs.1000 from the sample. The data on PCE are sparse beyond the indicated vertical line.

Household Labor and Recall

Note: This figure shows the non-parametric (lowess) relationship between the likelihood of reporting at least one day of work lost in the previous month due to an acute illness and (log) per-capita expenditures using monthly and weekly recall periods. The data on PCE are sparse beyond the indicated vertical line.

Appendix

Table 1: Outcome Comparisons for Observational and Experimental Sample

Outcome Variable	Weekly report for individuals in both databases		Difference in weekly reports (Year "0" observational data - Year 2008 experimental data)
	Individuals surveyed 2000 (observational data)	Individuals surveyed 2008 (experimental data)	
	(1)	(2)	(3)
# Acute Illnesses	0.911	0.936	-0.024
	(1.200)	(1.286)	(0.038)
At least one acute illness	0.469	0.450	0.020
	(0.499)	(0.498)	(0.016)
Duration of Acute Illness	4.064	4.767	-0.703***
	(6.618)	(7.984)	(0.220)
Total Doctor Visits	0.413	0.430	-0.017
	(0.799)	(0.848)	(0.025)
Doctor Visits for Acute	0.351	0.364	-0.013
	(0.713)	(0.755)	(0.023)
Doctor Visits for Chronic	0.061	0.066	-0.004
	(0.349)	(0.324)	(0.011)
At least one workday lost due to illness	0.078	0.125	-.047***
	(0.268)	(0.33)	(0.101)
Total Self-Medication	1.059	1.321	-0.262***
	(2.465)	(2.651)	(0.079)
At least one self-medication	0.305	0.353	-0.048***
	(0.461)	(0.478)	(0.015)

Note: This table compares the main outcome variables used in the paper for the individuals who comprised the original ISERDD sample with those in the 205 households chosen for follow-up in the experimental study. The means of the outcome variables are presented for the individuals of the original ISERDD sample in the first year of observation. For the 205 households in the experimental study, we provide the means of the outcome variables *in 2000*, when they were first observed. Standard Errors are in parentheses. *** is significant at the 1% level of confidence.

Table 2: Heterogeneity in the Recall Effect by PCE and Education

	Interaction with PCE only			Interaction with PCE and Education			
	Weekly	Weekly*Log pc expenditures	R²	Weekly	Weekly*Log pc expenditures	Weekly*Education of head	R²
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Number of acute illnesses	9.459***	-0.906***	0.360	8.860***	-0.793***	-0.441	0.357
	(2.595)	(0.320)		(2.45)	(0.292)	(0.287)	
At least one acute	2.804***	-0.285***	0.168	2.655***	-0.257***	-0.103	0.167
	(0.823)	(0.106)		(0.783)	(0.099)	(.080)	
Doctor's visits	3.560***	-0.397***	0.118	3.027***	-0.317***	-0.215	0.112
	(1.041)	(0.134)		(0.701)	(0.076)	(0.155)	
Duration of acute illnesses	28.255***	-3.653***	0.025	27.697***	-3.535***	-0.636	0.025
	(8.457)	(1.055)		(7.177)	(0.819)	(1.545)	
At least one workday lost due to illness	0.626*	-0.067	0.077	0.462	-0.036	-0.121*	0.079
	(0.346)	(0.042)		(0.309)	(0.037)	(0.062)	
Number of self-medication taken	-10.222**	2.127***	0.334	-9.49**	1.993***	0.452	0.334
	(4.699)	(0.666)		(4.716)	(0.699)	(0.745)	
At least one self medication was taken	2.371***	-0.228***	0.184	2.471***	-0.246***	0.066	0.185
	(0.715)	(0.092)		(0.660)	(0.084)	(0.094)	

* Significant at the 10%; **significant at the 5%; *** significant at 1%. Estimates are reported from two specifications investigating the heterogeneity of the recall impact with PCE and education of the household head, the latter coded as 0 if the head reported middle school or lower education and 1 for education above middle school. Regressions are at the household level using observational data and sample size varies from 4285 to 4622 observations. Columns 1-3 report estimates from a specification that only includes the interaction with PCE; Columns 5-7 include interactions of the weekly dummy with education as well. *Weekly*: This variable is equal to one if the data was collected weekly, and is zero otherwise. All specifications include household fixed-effects with clustering at the locality level. Robust standard errors in parentheses