Appendix on Cost Calculations

This appendix presents the cost calculations. We did not have access to individual-level NAAL data, which would have enabled us to undertake the necessary econometric analyses to generate more precise and reliable estimates. Please refer to the caveats and limitations of these calculations, which are described in the brief. These calculations are intended only to be suggestive, and to motivate future research: they should not be taken out of context.

As noted in the brief, we employ many of the same assumptions used by Friedland (2002), who estimated the direct medical cost of low functional adult literacy; however, we employ the newly released health literacy survey data results from NAAL in our calculations. We also use contemporary cost and census population data from MEPS and the U.S. Census Bureau.

It is important to emphasize that there are other sets of assumptions and approaches that could be used to arrive at these top-line calculations. Such alternatives could also be very reasonable in approximating the order of magnitude of the direct healthcare cost burden of low health literacy levels. We describe only our approach.

Even in such rudimentary approximations, such as the ones described here, it is necessary to undertake sensitivity analyses, in order to better understand the key drivers of the economic cost of low health literacy to the U.S.

We outline our base case calculation below. We first present the data and key assumptions (along with variable designations for demonstrating our calculations), and then document our methods.

Key Data and Estimates:

1. According to the most recent MEPS data, 2006 per capita medical expenditures (C_A) in the U.S. were approximately $3,905.

2. U.S. Census Bureau data reports the U.S. adult population (Q) in 2006 was 225.7 million.

3. The 2003 NAAL health literacy survey reports that the proportion of adults at *Below Basic* health literacy levels was 14 percent, with another 22 percent classified as having only *Basic* health literacy levels. We define the proportion of adults with low health literacy (our term) as *p*.

4. Friedland (2002) reports that adults in the bottom 20 percent of predicted *functional literacy* scores (low literacy) have average per capita medical expenditures that are approximately twice (196 percent) as much as the per average cost for the entire population. We define this ratio to be λ. We define average healthcare costs in adults with low health and not health literacy as C_L and C_{NL}.

5. Friedland (2002) dichotomizes incremental direct medical costs incurred by low literacy adults into the following: the proportion attributable to lowfunctional literacy and the proportion attributable to other factors (covariates). He models the former using a range from 1/3 to 2/3. We define this proportion of the incremental costs to be α.

Methods—Annual Estimates:

To simplify the exposition, we define the ratio of average direct medical expenditures for adults with *low* health literacy to the average direct medical expenditures for the entire adult population as follows:

$$C_L = \lambda C_A \quad (1)$$

The following equations will also be useful:

$$C_A = pC_L + (1-p)C_{NL} \quad (2)$$

$$\Delta C = C_L - C_{NL} \quad (3)$$

Obviously, equation (3) is the incremental, or marginal, direct medical cost associated with having low health literacy, relative to not having low health literacy. The proportion of this marginal cost, α, that is attributable to low health literacy is unclear and cannot be answered rigorously without adequate data and appropriate econometric techniques. We were unsuccessful in obtaining the necessary individual level data from the Department of Education's National Center for Education Statistics despite repeated efforts and requests for help. In fact, we could not get a single phone call or email returned by the survey's lead researcher, Shieda White. For this reason we rely on Friedland's analysis and model of the cost of low adult literacy to generate our estimates of the cost of low health literacy in the U.S.

Combining and re-arranging terms in (1)-(3) yields the following useful algebraic representation of incremental costs (conditional on $0 > \lambda p > 0$):

$$\Delta C = C_A\left(\lambda - \frac{1-\lambda p}{1-p}\right) \quad (4)$$

It is also obviously the case that λ is necessarily greater than unity by definition and *p* lies on the interval [0, 1]. Multiplying equation (4) by α , which also lies on the interval [0, 1], generates a measure the proportion of the incremental cost between low health literacy individuals and not low health literacy individuals that is attributable (independently caused by) low health literacy. Therefore, the fraction of national health expenditures attributable to low health literacy may be expressed as follows:

$$\frac{pQ\left(\alpha C_A\left(\lambda - \frac{1-\lambda p}{1-p}\right)\right)}{Q\left(pC_L + (1-p)C_{NL}\right)} = p\alpha\left(\lambda - \frac{1-\lambda p}{1-p}\right) \quad (5)$$

This expression is, of course, subject to the same parametric constraints mentioned previously. A key consideration in generating a cost estimate is how to map the estimate of incremental costs (based on predicted adult literacy from MEPS) associated with low functional literacy into our calculations using the new NAAL health literacy data. Any mapping will necessarily be speculative, as is the case with several other aspects of our estimate.

Our approach was the following. It seems plausible, if not probable, that individuals with low health literacy scores are more likely to come from the tail of the distribution associated with high healthcare expenditures than individuals with low functional literacy scores. Health literacy is obviously a more direct and precise measure of an individual's ability to obtain, process and make appropriate health decisions than functional literacy; it may capture additional elements (that functional literacy does not) of the challenges faced by some individuals in navigating the U.S. healthcare system and managing their own healthcare needs and requirements. Thus, Friedland's estimate of 196 percent higher costs (relative to the population average) for individuals below the first quintile of predicted functional literacy scores will underestimate this ratio (of costs) for individuals below the first quintile of health literacy scores. [We hope to test this empirically, and also address a critical endogeneity issue, if we are eventually able to obtain the individual-level data from the 2003 NAAL.] For this reason, we believe a 20 percent threshold for our own calculations will be a lower bound. It seems plausible, therefore, to model the range from 20 percent to 36 percent (the latter represents, of course, the percentage of individuals at *Below Basic* and *Basic* health literacy from NAAL). We acknowledge, as Friedland did per his cutoff point, that our upper bound is arbitrary.

In sum, our calculations are simple and based on a number of assumptions. The parameter values and ranges used in our calculations are summarized below:

C_A= $3,905

Q = 225,700,000

Lower bound p = 0.20

Upper bound p = 0.36

Lower bound $\alpha = 1/3$

Upper bound $\alpha = 2/3$

These values were used to generate the follow tableau of estimated annual costs.

Table A1: Annual Healthcare Cost Estimates Attributable to Low Health Literacy Levels in the U.S.

Parameter Values	$\alpha = 1/3$	$\alpha = 1/2$	$\alpha = 2/3$
p = 0.20	$70,508,680,000	$105,763,020,000	$141,017,360,000
p = 0.28	$109,680,168,889	$164,520,253,333	$219,360,337,778
p = 0.36	$158,644,530,000	$237,966,795,000	$317,289,060,000

Obviously the range of cost estimates is very large. In the brief we report the range from $106 billion to $238 billion, i.e., when α = 1/2.

Methods—Present Value Long Run Estimates:

Calculating present value, long run costs over a horizon of t years is straightforward. We first consider the case of an infinite time horizon, as $t \to \infty$. From a social welfare perspective this is the appropriate horizon.

For simplicity, we assume that annual healthcare costs attributable to low health literacy levels (as described in this appendix) remain constant over time. If r is social discount rate, then present value healthcare costs from low health literacy levels is represented as follows:

$$\Omega = \sum_{t=0}^{\infty} \frac{\omega}{(1+r)^t} = \omega + \frac{\omega}{r} = \omega\left(1 + \frac{1}{r}\right) \tag{6}$$

It is straightforward to show this infinite geometric series converges; it is a basic perpetuity and inclusive of current year's cost. The annual cost due to low health literacy levels is measured, as has been shown before, as follows:

$$\omega = pQ\alpha C_A\left(\lambda - \frac{1-\lambda p}{1-p}\right) \tag{7}$$

Alternatively, shorter time horizons may be considered when calculating these costs. Equation (8) is the present value cost of low health literacy levels over the finite time horizon of n years (and inclusive the current year's cost).

$$\Omega = \sum_{t=0}^{n} \frac{\omega}{(1+r)^t} = \omega + \frac{\omega\left(1 - \frac{1}{(1+r)^n}\right)}{r} \tag{8}$$

It is easy to see by inspection that the ratio on the right-hand side in (8) is simply the present value difference between to perpetuities: one that begins in t = 1 and the other that begins in year n.

Table A2 summarizes the present value cost estimates over 5, 10, 25, and 50 years—inclusive of the base, or current, year. Thus, we are considering n future years plus the current year—a total of n+1 years of costs. An infinite time horizon calculation is also shown. We consider the same p values as used in Table A1, but use the base case (midpoint) value of α = ½.

Table A2: Finite Time Horizon Present Value Healthcare Cost Estimates Attributable to Low Health Literacy Levels in the U.S.

Parameter Values	$p = 0.20$	$p = 0.28$	$p = 0.36$
5 Years	$539,412,283,422	$839,085,774,212	$1,213,677,637,700
10 Years	$848,598,214,985	$1,320,041,667,755	$1,909,345,983,717
25 Years	$1,338,281,170,753	$2,081,770,710,061	$3,011,132,634,195
50 Years	$1,565,371,626,311	$2,435,022,529,817	$3,522,086,159,199
∞	$1,616,663,305,714	$2,514,809,586,667	$3,637,492,437,857

The values in Table A2 demonstrate the sensitivity of the present value cost estimates to both the time horizon considered and, of course, the assumed proportion of incremental costs between groups attributable to low health literacy levels. Sensitivity analyses across the other model parameters are also easily performed using the interactive model we have developed.

This cost calculation exercise is a good faith effort to gain insight into the order of magnitude of the economic costs of low health literacy levels in the U.S., but it is only within this context that our results should be considered. Only rigorous econometric analyses using individual level data from the 2003 NAAL survey has the potential to generate sufficiently precise estimates of these costs. Our first approximation calculations have only endeavored to better understand the potential order of magnitude of these costs.