Disease and Development: A Reply to Bloom, Canning, and Fink

Daron Acemoglu

Massachusetts Institute of Technology

Simon Johnson

Massachusetts Institute of Technology

I. Introduction

Beginning in the 1940s, a wave of health innovations and more effective international public health measures led to a rapid and large improvement in health; for example, in some relatively poor countries, life expectancy at birth quickly rose from around 40 years to over 60 years. In Acemoglu and Johnson (2006, 2007), we constructed an instrument for these changes in life expectancy: "predicted mortality," which is calculated from initial mortality by disease and the timing of global disease interventions. Across a wide range of specifications, our work suggests no positive effects—over 40- or 60-year horizons—of life expectancy on GDP per capita (or GDP per working-age population).

Bloom, Canning, and Fink (2014, in this issue) argue that the level of life expectancy in 1940 affected subsequent growth rates and should be included in our long-difference specifications; that is, the level of life expectancy in 1940 should be included on the right-hand side when 1940–80 or 1940–2000 changes in GDP per capita are the dependent variables. In a linear regression framework, their specification introduces a great deal of multicollinearity, and the standard errors become very large.

The specifications in Acemoglu and Johnson (2006, 2007) allowed for potential long-run effects of health improvements and supported our empirical strategy by showing that changes in the predicted mortality instrument were uncorrelated with its own past changes and past changes in population, GDP, and GDP per capita. There are three further ways to assess Bloom et al.'s concerns. First, we include the initial level of life expectancy from 1900, interacted with time dummies in our decadal panel data set (which runs from 1940). Second, we use a nonlinear estimator suggested by Bloom et al.'s framework to estimate directly their proposed equation with reasonable precision. Third, from microeconomic estimates in Ashraf, Lester, and Weil (2008), we calculate potential macroeconomic effects of current life expectancy on future growth and examine the implications for our baseline findings. Our results remain robust throughout.

II.   The Estimating Framework

Our estimating equation in Acemoglu and Johnson (2006, 2007) was

$$y_{it+k} = \pi x_{it} + \zeta_i + \mu_t + \mathbf{Z}'_{it}\boldsymbol{\theta} + \varepsilon_{it+k}. \tag{1}$$

Here i denotes country and t is time period; y is log GDP per capita; x is log life expectancy (at birth or at other ages); the ζ_i's denote a full set of fixed effects to capture cross-country differences in time-invariant characteristics; the μ_t's incorporate time-varying factors common across all countries; and $\mathbf{Z}_{it}$ denotes a vector of other controls. (We use the subscript $it + k$ as shorthand for $i, t + k$.) The case in which $k > 0$ allows for lagged effects of life expectancy.

We instrumented life expectancy with *predicted mortality*, constructed as

$$M^I_{it} = \sum_{d \in \mathcal{D}} [(1 - I_{dt})M_{di40} + I_{dt}M_{dFt}], \tag{2}$$

where M_{dit} denotes mortality in country i from disease d at time t, I_{dt} is a dummy for intervention on disease d at time t (equal to one for all dates after the intervention), and $\mathcal{D}$ denotes a set of 15 infectious diseases for which we have data, including most major communicable causes of death around the world in 1940, as well as some less common killers. The variable M_{di40} refers to the pre-intervention mortality from disease d in the same units, while M_{dFt} is the mortality rate from disease d at the health frontier of the world at time t. For our baseline instrument, M_{dFt} is set equal to zero.

Any change in life expectancy is unlikely to have its full effect on any demographic or economic variables instantaneously—or even in the same decade. For this reason, in Acemoglu and Johnson (2007), we estimated

equation (1) in long differences, that is, regressing change on change in a panel including only two years, t_0 and t_1 (in practice 1940 and 1980 or 1940 and 2000). In Acemoglu and Johnson (2006), we also presented a range of panel specifications using decadal observations; these results were very similar to those from the long-difference specifications that were emphasized in Acemoglu and Johnson (2007). We explicitly discussed the adjustment dynamics of population and GDP and allowed for potential health effects to show up after a long lag: after 40 or 60 years in the long-difference specifications and with 10-, 20-, 30-, or 40-year horizons in panel specifications.

Bloom et al. propose a "partial adjustment model" that takes the following form:[1]

$$\Delta y_{it} = \pi \Delta x_{it} + \lambda \pi x_{it-1} - \lambda y_{it-1} + \alpha_t + \xi_{it}, \tag{3}$$

where $\Delta y_{it} \equiv y_{it} - y_{it-1}$, and Δx_{it} is defined similarly. They derive this from our equation (1), assuming an AR(1) specification for the error term ($\varepsilon_{it} = \lambda \varepsilon_{it-1} + \xi_{it}$). This equation allows for convergence dynamics (through the λ term) and a potential impact of the lagged level of log life expectancy, x_{it-1}, on subsequent changes in GDP per capita.

III.   The Impact of Initial Life Expectancy

Bloom et al.'s instrumental variable regressions generate very imprecise estimates for the effect of life expectancy on GDP per capita. This simply reflects the fact that it is impossible to distinguish the impact of the level of life expectancy in 1940 (x_{it-1}) and of the subsequent change in life expectancy (Δx_{it}) in long difference using only the variation in predicted mortality (M_{it}^I).

If the true effect of life expectancy on GDP per capita were positive—for example, because the level of life expectancy affects subsequent changes in GDP per capita—then estimates of the relationship between changes in life expectancy and changes in GDP per capita over a 60-year horizon should capture much of these positive effects even if there are reasonable lags.[2] Our long-difference specifications should thus reveal any long-run, positive relationship between life expectancy and GDP per capita. Our estimates in Acemoglu and Johnson (2006, 2007) using 60-year changes show no such positive effect.

---

[1] This is their eq. (3), using their notation, except that we denote the error term by ξ_{it} to distinguish it from our error term, ε_{it}, in (1), and we use π instead of their β for consistency.

[2] For example, even if only a third of the impact of lagged life expectancy on GDP per capita materializes over a generation (i.e., over 20 years), the bulk of these effects should be evident in our specification using 60-year changes (1940–2000).

There are three further ways to check if potential long-term effects from lagged life expectancy modify any of our conclusions: (*a*) run panel regressions including initial life expectancy in 1900, interacted with time dummies; (*b*) employ a nonlinear estimator implied by Bloom et al.'s equation (3); and (*c*) use reasonable estimates for direct effects of health improvements based on microeconomic evidence.

A. *Controlling for Initial Life Expectancy*

To facilitate comparison with models that control for the effect of initial life expectancy, column 1 of table 1 reports baseline estimates of (1) using decadal observations as in the panel data models of Acemoglu and Johnson (2006). Panel A is for 1940–80 and panel B is for 1940–2000.[3] The standard errors in this and subsequent models are robust and allow for arbitrary serial correlation at the country level. In column 1 of panel A, $\hat{\pi} = -1.307$, with a standard error (SE) of 0.455, indicating a negative impact of life expectancy on GDP per capita. Column 1 of panel B shows a similar estimate that is larger in absolute value (i.e., more negative) for the period 1940–2000, $\hat{\pi} = -1.394$.[4]

The remaining columns in table 1 include a full set of time interactions with log life expectancy in 1900—allowing initial life expectancy to flexibly affect future GDP per capita. Using the 1900 value for life expectancy rather than the 1940 level alleviates the mechanical correlation between 1940 life expectancy and predicted mortality. It is equally valid if there is an impact from the level of initial life expectancy on future growth as proposed by Bloom et al.[5]

Column 2 shows results from including these interactions without controlling for lagged GDP per capita. In panel A, the estimate is $\hat{\pi} = -0.100$ (SE = 0.421). Thus there is a negative (and far from significant)

---

[3] We have data on GDP, life expectancy, and other variables of interest every 10 years from 1940 to 2000. We also look at the period 1940–80 to avoid the potential effects of the onset of HIV-AIDS as a global disease.

[4] These balanced panel estimates are very close to those reported in cols. 1 and 2 of the unbalanced panel of table 11 in Acemoglu and Johnson (2006) and to the long differences in cols. 1 and 2 of table 7, panel B, in Acemoglu and Johnson (2007).

[5] Econometrically, we are controlling for the effects of initial life expectancy by including a full set of time dummies interacted with initial life expectancy, i.e., terms of the form $\omega_t \times x_{i1900}$ (one for each t). This strategy potentially controls for two types of effects. The first is that life expectancy in 1900, x_{i1900}, directly affects outcomes in subsequent years. The second is that the year t equation contains the term $\omega_t \times x_{it-1}$ (thus allowing for a general impact of lagged life expectancy). In this latter case, we can substitute for x_{it-1} in terms of log life expectancy in 1900, x_{i1900}. For example, following the model for the dynamics of life expectancy estimated in table 6 of Acemoglu and Johnson (2007, 957, eq. [12]), suppose that $x_{it} = \nu x_{it-1} + \eta_{it}$, with decadal observations and η_{it} being serially uncorrelated and orthogonal to other variables. Then substitute for x_{it-1} and its lags successively to obtain $x_{it-1} = \nu^{t-1} x_{i1900} + \eta_{it-1} + \nu \eta_{it-2} + \nu^2 \eta_{it-3} + \cdots$, with $x_{i0} = x_{i1900}$. Then the coefficient on x_{it-1} in the year t equation would be $\omega_t \times \nu^{t-1}$, and all other coefficients can be estimated consistently.

TABLE 1
EFFECT OF LIFE EXPECTANCY ON GDP PER CAPITA, CONTROLLING FLEXIBLY FOR THE
IMPACT OF LIFE EXPECTANCY IN 1900 AND CONVERGENCE DYNAMICS, USING PANEL DATA

	DEPENDENT VARIABLE: LOG GDP PER CAPITA				
	Baseline Specification	Including Life Expectancy in 1900, Interacted with Time Dummies			
	2SLS (1)	2SLS (2)	2SLS (3)	2SLS (4)	GMM (5)
	A. 1940–80 Balanced Panel				
Lagged GDP per capita				.552 (.080)	.414 (.296)
Life expectancy	−1.307 (.455)	−.100 (.421)	−.270 (.522)	−.478 (.443)	−.171 (.393)
Countries	47	47	47	47	47
Periods	5	5	5	5	4
Moments					25
Hansen p-value					.29
AR2 p-value					.22
	B. 1940–2000 Balanced Panel				
Lagged GDP per capita				.817 (.047)	.821 (.144)
Life expectancy	−1.394 (.362)	−.928 (.486)	−1.317 (.627)	−.965 (.425)	−.598 (.234)
Countries	47	47	47	47	47
Periods	7	7	7	7	6
Moments					44
Hansen p-value					.31
AR2 p-value					.43

NOTE.—The 2SLS specifications (cols. 1, 2, 3, and 4) include a full set of country and year fixed effects. Columns 2, 3, 4, and 5 include a full set of year dummies interacted with life expectancy in 1900. Regressions in col. 3 also include a full set of year dummies interacted with the log of GDP per capita in 1940. Arellano and Bond's GMM estimator (col. 5) removes country fixed effects by taking first differences (hence the lower number of time periods) and then constructs moment conditions using all predetermined lags of GDP per capita and predicted mortality as instruments. It is estimated in two steps and thus is optimally weighted. Robust standard errors corrected for arbitrary serial correlation (clustered) at the country level are reported in cols. 1, 2, 3, and 4, and robust standard errors are reported in col. 5. Panel A contains estimates using a balanced panel of 47 countries from 1940 to 1980. Panel B contains estimates using a balanced panel of the same 47 countries from 1940 to 2000. See Acemoglu and Johnson (2007) for the construction of the predicted mortality instrument, definitions, and data sources.

impact of life expectancy on GDP per capita, which is much smaller than the estimate in column 1. The estimate in panel B ($\hat{\pi} = -0.928$, SE = 0.486) is larger in absolute value (i.e., more negative), much closer to the estimate in column 1, and marginally statistically significant.

In addition to year dummies interacted with initial life expectancy, column 3 adds a full set of time interactions with log GDP per capita in 1940. These interactions are useful since any correlation with initial GDP

per capita might otherwise load onto initial life expectancy. In panel A, we now estimate $\hat{\pi} = -0.270$ (SE = 0.522). The coefficient on life expectancy in panel B is larger, -1.317, very similar to our baseline estimate in column 1, and statistically significant at 5 percent (SE = 0.627).

Columns 4 and 5 add lagged log GDP per capita to the right-hand side, allowing for convergence effects. These two columns, respectively, use the standard two-stage least-squares (2SLS) estimator and Arellano and Bond's (1991) optimally weighted two-step generalized method of moments (GMM) estimator, with predicted mortality as the external instrument. The results are again broadly consistent with our baseline results. The GMM estimate in column 5 is $\hat{\pi} = -0.171$ (SE = 0.393) in the 1940–80 panel and a larger (in absolute value), more precise, and statistically significant $\hat{\pi} = -0.598$ (SE = 0.234) in the 1940–2000 panel.

Overall, controlling for the effects of initial life expectancy changes our point estimates, especially for the 1940–80 period. However, in no case is there any evidence for a positive effect of life expectancy on GDP per capita, and the estimates in table 1 for 1940–2000 show statistically significant negative effects of life expectancy on GDP per capita that are close in magnitude to the baseline results of Acemoglu and Johnson (2006, 2007).

B. Nonlinear Generalized Method of Moments

We can directly estimate Bloom et al.'s proposed equation (3) using a nonlinear GMM approach (with nonlinear equivalents of the moment conditions used in col. 5 of table 1). Estimates for π and λ obtained in this fashion are shown in table A1 of the online appendix. These imply long-run negative effects of life expectancy on GDP per capita that are very similar to our baseline results for both 1940–80 and 1940–2000. For example, $\hat{\pi}$ is estimated as -1.261 for 1940–80 and -1.548 for 1940–2000, while our original estimates ranged from -1.21 to -2.70.[6]

C. Directly Incorporating Lagged Effects of Life Expectancy

An alternative strategy is to directly incorporate the potential effect of initial life expectancy in the long-difference specification from Acemoglu and Johnson (2007). Rewriting Bloom et al.'s estimating equation gives

$$\Delta \tilde{y}_{it} \equiv \Delta y_{it} - \kappa x_{it-1} = \pi \Delta x_{it} - \lambda y_{it-1} + \alpha_t + \xi_{it}. \qquad (4)$$

---

[6] When we set $\Delta y_{it} = \Delta x_{it} = 0$ in eq. (3), it can be verified that their π measures the long-run (e.g., 40 or 60 years) impact of life expectancy on GDP per capita, exactly as does the parameter π in our eq. (1).

Although we do not know the precise value of κ ($= \lambda\pi$), the micro-economic literature—surveyed by Ashraf et al. (2008)—provides guidance on how large this could be. Specifically, we use their estimates to obtain an upper bound for plausible values of κ by supposing that all the potential effects of initial life expectancy are captured by κ.

In our sample, life expectancy among the countries with high initial mortality increased from about 40 to over 60 between 1940 and 1980. Increasing median life expectancy from 40 to 60 years would, according to Ashraf et al.'s base estimate, raise GDP per capita by 15 percent in the long run (over 60 years). When their high estimate is used—which assumes that all impacts of health are as positive as any microeconomic study could suggest—the increase in GDP per capita is 25 percent and the full long-run effect is achieved within 40 years.

In terms of equation (4), supposing that the 15 percent long-run effect is all captured by κ, this would imply a value of κ equal to 0.343, while a 25 percent long-run effect implies that $\kappa = 0.54$.[7] We use $\kappa = 0.3$, 0.4, 0.5, and 0.6 to span a range for the upper bound for the effects of increased life expectancy on future growth. The estimate of 0.6, in particular, represents the strongest possible case for the Bloom et al. hypothesis.

We estimate equation (4) using 2SLS in two sets of specifications. First, we estimate (4) assuming no mean reversion, that is, setting $\lambda = 0$ (odd-numbered columns in table 2). Second, we estimate (4) including log GDP per capita in 1940 on the right-hand side to control for potential convergence effects in GDP per capita (even-numbered columns in table 2). In either case, there is no evidence of a positive coefficient for π.

For example, for 1940–80, with $\kappa = 0.6$ and log GDP per capita in 1940 included, the coefficient on change in life expectancy is -0.551 (panel A, col. 8). For 1940–2000, in column 8 of panel B, there is a significant negative coefficient on change in life expectancy: -2.534 (SE 1.042).

As shown in table 2, every 0.1 increase in κ reduces the negative effect of life expectancy by about 0.15 in absolute terms. This implies that to reach even a zero coefficient on change in life expectancy for the odd-numbered columns of panel A (for 1940–80 and without controlling for GDP per capita in 1940) would require a κ of around 0.9. This is far larger than anything that can reasonably be supported using the available microeconomic evidence. To imagine a positive effect for life

---

[7] We translate between Ashraf et al.'s simulation parameters and our regression coefficients as follows. A 15 percent increase in GDP per capita means that the level of GDP per capita ends up at 1.15 (i.e., if it starts at 1), so the impact measured in natural logarithms is $\ln(1.15) = 0.139$. Initial life expectancy is 40 years and $\ln(40) = 3.69$. Final life expectancy is 60 years and $\ln(60) = 4.09$. The change in log life expectancy is 0.405. Assuming that all of this is accounted for by κ gives an upper bound for $\kappa = (0.139/0.405) = 0.343$ in the base case.

TABLE 2

EFFECT OF LIFE EXPECTANCY ON GDP PER CAPITA CONTROLLING DIRECTLY FOR THE IMPACT OF LIFE EXPECTANCY IN 1940 AND CONVERGENCE DYNAMICS, USING LONG DIFFERENCES

DEPENDENT VARIABLE: CHANGE IN LOG GDP PER CAPITA

	$\kappa = .3$		$\kappa = .4$		$\kappa = .5$		$\kappa = .6$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Long Differences between 1940 and 1980								
Change in life expectancy	−.820	−1.008	−.655	−.855	−.489	−.703	−.324	−.551
	(.331)	(.757)	(.325)	(.746)	(.319)	(.736)	(.313)	(.726)
GDP per capita 1940		−.051		−.054		−.058		−.061
		(.169)		(.167)		(.164)		(.162)
B. Long Differences between 1940 and 2000								
Change in life expectancy	−1.111	−2.894	−.979	−2.774	−.848	−2.654	−.716	−2.534
	(.395)	(1.058)	(.394)	(1.053)	(.392)	(1.048)	(.391)	(1.042)
GDP per capita 1940		−.604		−.608		−.612		−.616
		(.266)		(.264)		(.263)		(.261)

NOTE.—2SLS estimates of eq. (8), from the text, assuming different values for κ. Columns 1 and 2 show estimates with $\kappa = 0.3$, cols. 3 and 4 show estimates with $\kappa = 0.4$, cols. 5 and 6 show estimates with $\kappa = 0.5$, and cols. 7 and 8 show estimates with $\kappa = 0.6$. In all models, the change in life expectancy is instrumented using the change in predicted mortality during the corresponding time period. Robust standard errors are in parentheses. Panel A contains estimates for a cross-sectional regression with data for 47 countries in 1940 and 1980. Panel B contains estimates for a cross-sectional regression with data for the same 47 countries in 1940 and 2000. See Acemoglu and Johnson (2007) for the construction of the predicted mortality instrument, definitions, and data sources.

expectancy on GDP per capita in the other specifications in panel A or in panel B (for 1940–2000) is even more far-fetched.

IV. Conclusion

Estimates using 40-year or 60-year differences in Acemoglu and Johnson (2006, 2007), which should capture any slow-acting effects of health improvements, did not show any evidence for a positive impact of life expectancy on GDP per capita. In this note, we report three additional approaches for assessing the potential effects of initial life expectancy on subsequent changes in GDP per capita. All these approaches confirm that our main results are robust: there is no evidence that increases in life expectancy after 1940 had a positive effect on GDP per capita growth.

References

Acemoglu, Daron, and Simon Johnson. 2006. "Disease and Development: The Effect of Life Expectancy on Economic Growth." Working paper no. 12269 (June), NBER, Cambridge, MA.
———. 2007. "Disease and Development: The Effect of Life Expectancy on Economic Growth." *J.P.E.* 115 (6): 925–85.
Arellano, Manuel, and Stephen R. Bond. 1991. "Some Specification Tests for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *Rev. Econ. Studies* 58 (2): 277–98.
Ashraf, Quamrul H., Ashley Lester, and David N. Weil. 2008. "When Does Improving Health Raise GDP?" *NBER Macroeconomics Ann.* 23:157–204.
Bloom, David E., David Canning, and Gunther Fink. 2014. "Disease and Development Revisited." *J.P.E.* 122 (6): 1355–66.