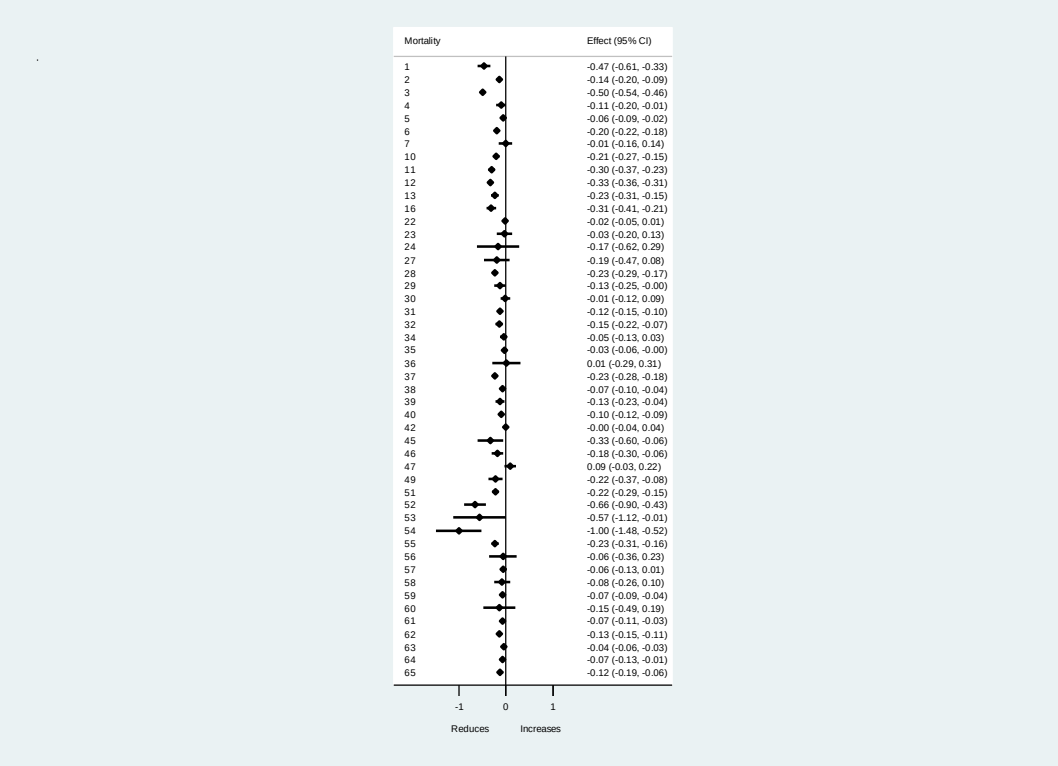
Online Appendix

The impact of healthcare spending on health outcomes: A meta-regression analysis

1. Data

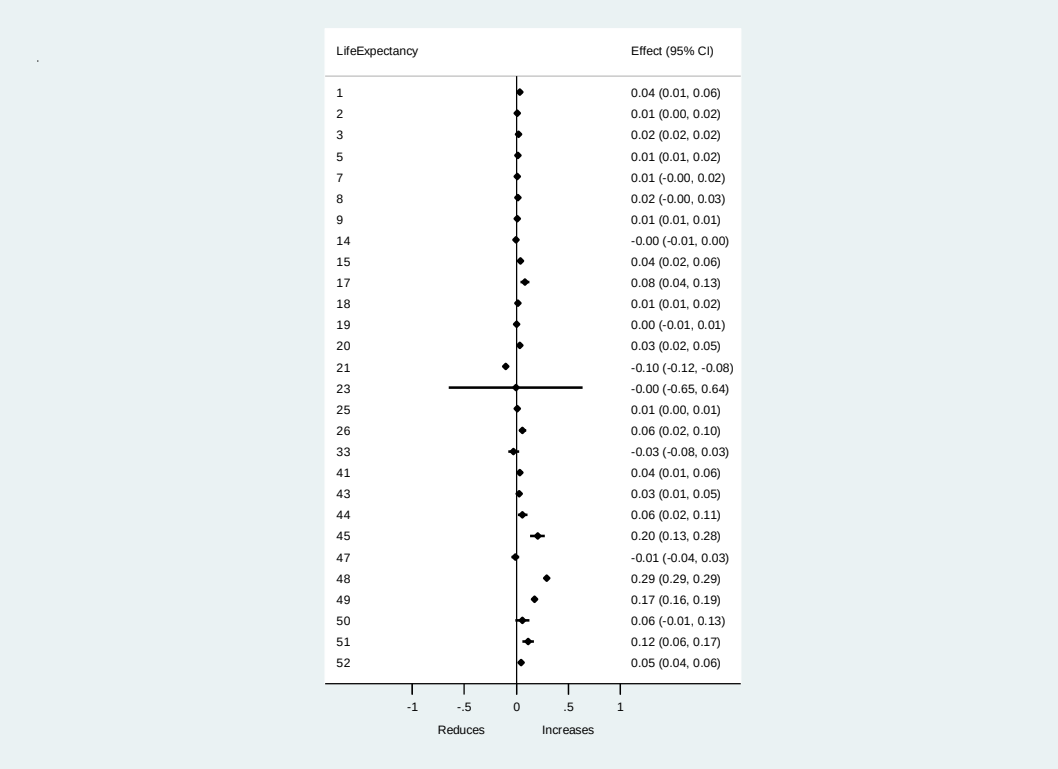
The meta-data can be downloaded at [*address suppressed to maintain anonymity*]. As mentioned in the paper, funnel plots are a useful means of illustrating heterogeneity and truncation in the distribution of spending elasticity estimates, and thus act as a visual signal of publication bias. As an alternative, we provide in Figures A1 and A2 forest plots pertaining to the spending elasticity for the mortality rate and life expectancy, respectively. Differences in effect sizes and associated confidence intervals indicate heterogeneity across studies.

Figure A1: Forest plot (Healthcare spending elasticity, mortality rate)



Note: Effect sizes are weighted average for each study.

Figure A2: Forest plot (Healthcare spending elasticity, life expectancy)



Note: Effect sizes are weighted average for each study.

2. Treatment of outliers

Figures A3 and A4 are funnel plots illustrating outliers in estimates of healthcare spending elasticities for the mortality rate and life expectancy, respectively. The outliers are highlighted with a square. Figure A3 illustrates the two outliers, while Figure A4 illustrates the one outlier, as discussed in the text. The two outliers in Figure A3 come from Nixon and Ulmann (2006) and Gupta et al. (2003), while the one outlier in Figure A4 comes from Hall et al (2012). Interestingly, Hall et al. (2012) estimate a generalized cointegration model, which by focusing on the time series properties of healthcare spending and life expectancy is different from the typical study encountered in literature.

Figure A3: Outliers in the mortality rate data

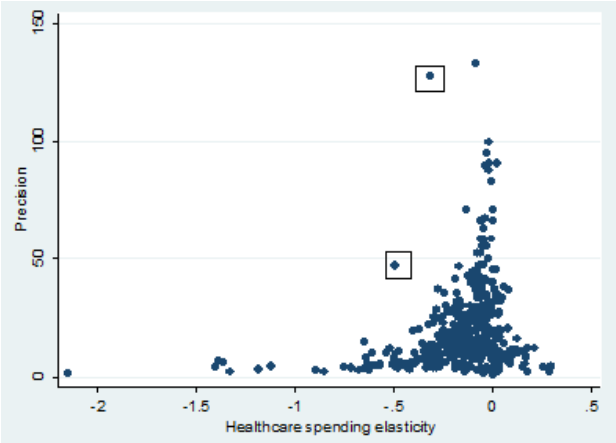
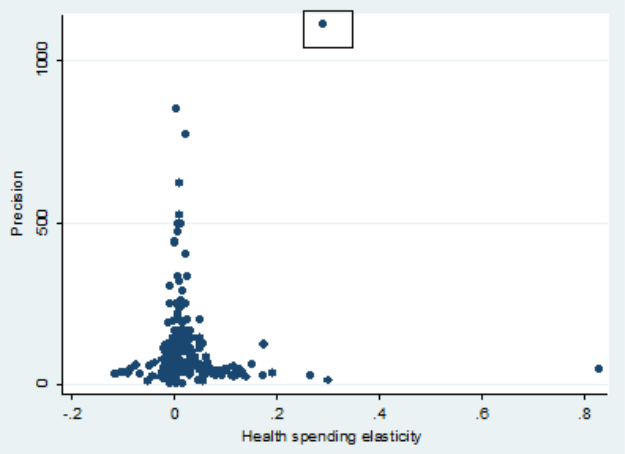


Figure A4: Outliers in the life expectancy data



As noted in the text, our preferred approach is to remove these outliers from the meta-regression analysis (MRA) and run weighted least squares (*WLS*). Table A1 reports the robustness of our estimates to alternate approaches. For comparison purposes, Column (1) reports the results from the general model for the mortality rate estimated using *WLS* without the 2 outliers in the mortality data. These are the same estimates as provided in Column (1) of Table 3. In Column (2) we again use WLS, but this time with the two mortality rate outliers added. In Column (3) we also include the mortality rate outliers, but apply iteratively reweighted least squares (*Huber's M-estimator*) to control for the influence of outliers. Columns (4) – (6) are respective counterparts for life expectancy.

Table A1
Estimation results for alternate treatment of outliers.

Variable	Mortality rate, without outliers (1)	Mortality rate, with outliers (2)	Mortality rate, iteratively reweighted least squares (3)	Life expectancy, without outliers (4)	Life expectancy, with outliers (5)	Life expectancy, iteratively reweighted least squares (6)
Constant	-0.089**	-0.230***	-0.140***	0.004	0.195***	0.041
	(2.64)	(3.31)	(3.67)	(0.21)	(4.61)	(1.23)
SE	-1.876***	-1.156**	-1.387***	1.222	-1.952*	0.091
	(6.24)	(2.52)	(5.16)	(1.08)	(1.97)	(0.10)
Child	0.004	-0.042	-0.006			
	(0.18)	(-1.29)	(0.21)			
Older				0.012**	-0.037**	0.002
				(2.15)	(2.39)	(0.27)
Female	-0.005	0.019	-0.001	-0.000	-0.022***	-0.006
	(0.57)	(0.75)	(0.05)	(0.06)	(3.27)	(0.96)
OECD	-0.018	0.012	-0.029**	-0.012	0.098***	0.015
	(1.35)	(0.49)	(2.31)	(1.24)	(5.24)	(0.62)
Pharm	0.028	-0.035	0.020	0.010	-0.083***	-0.007
	(1.13)	(-0.80)	(0.95)	(0.71)	(3.23)	(0.39)
Public	-0.025	-0.0153	-0.024**	-0.007*	-0.010**	-0.008**
	(1.64)	(-0.88)	(2.12)	(1.95)	(2.13)	(2.56)
Share	0.016	0.063*	0.056**	0.023	-0.211	-0.027
	(0.66)	(1.74)	(2.44)	(0.91)	(4.38)	(0.72)
Panel	0.014	0.099*	0.049**	0.001	-0.052	-0.005
	(0.67)	(1.91)	(1.98)	(0.12)	(1.42)	(0.48)
Sub-national	0.074***	0.133***	0.107***	-0.004	0.011	0.001
	(4.42)	(3.77)	(5.88)	(0.98)	(0.48)	(0.20)
Income	0.062***	0.086	0.048*	-0.000	-0.087**	-0.025
	(2.89)	(0.99)	(1.81)	(0.04)	(2.43)	(1.13)
Lag spending	-0.079***	-0.074**	-0.070***	-0.025	-0.027	-0.007
	(5.72)	(2.59)	(5.11)	(1.62)	(0.58)	(1.21)
Lag outcome	0.020*	0.025*	0.014	0.011	-0.023	-0.013
	(1.89)	(1.83)	(1.36)	(0.77)	(0.51)	(0.79)
Working paper	-0.011	-0.015	-0.009	0.028**	-0.059	0.018
	(0.61)	(-0.63)	(0.75)	(2.34)	(1.57)	(1.30)
Average year	-0.001	0.001	-0.001	-0.003	0.019***	0.002
	(0.91)	(0.06)	(1.24)	(1.53)	(6.85)	(0.47)
Endogeneity	-0.007	-0.001	-0.010	0.045*	-0.142***	0.001
	(0.59)	(-0.05)	(1.09)	(1.80)	(3.26)	(0.02)
n (k)	629 (47)	631 (48)	631 (48)	256 (27)	257 (28)	257 (28)
Adjusted R^2	0.40	0.41	-	0.13	0.85	-
Predicted E	-0.13	-0.23	-0.19	0.04	-0.01	0.04
(95% CI)	(-0.20 to -0.06)	(-0.37 to -0.10)	(-0.26 to -0.12)	(0 to 0.08)	(-0.13 to 0.10)	(0.01 to 0.06)

Notes: For Columns (1) – (3), *Predicted E* is the predicted spending elasticity for public spending, conditional on income and lagged spending being included in the health production function. For Columns (4) – (6), *Predicted E* is the predicted spending elasticity for life expectancy at an older age, conditional on endogeneity correction and lag spending being included in the health production function.

The results for the mortality rate in Columns (2) and (3) produce larger absolute point estimates for the spending elasticity compared to Table 3 of the paper. However, there is significant overlap in the confidence intervals. Moreover, the covariates that are statistically significant in the

MRA are quite similar with and without the two outliers. The results for life expectancy in Columns (5) and (6) offer very different coefficient estimates when the single outlier/leverage observation is included. This suggests caution with including this observation in the meta-data. However, when the *M-estimator* is applied in Column (6) the results are qualitatively similar to Column (4).

3. Robustness checks

Tables A2 and A3 report results of various robustness checks for the mortality rate and life expectancy meta-regressions, respectively. Regarding Table A2, Column (1) reports the specific model, reproducing Column (5) of Table 3. The figures in parentheses are absolute t-statistics constructed using clustered adjusted standard errors. However, these may be biased because of the relatively small number of studies (less than the rule of thumb of 42), and hence we also applied the wild bootstrap method to correct standard errors for uneven number of estimates within studies. The associated p-values are reported in brackets in Column (1). In Column (2) we replace inverse variance weights with sample size weights. Random effects weights are used in Column (3), which is the specific model counterpart to the general model results reported in Column (4) of Table 3. Column (4) reports the results of a multilevel model estimated using *REML*. Instead of correcting standard errors, this model directly incorporates the multilevel structure of the data, i.e. estimates clustered within studies. Finally, Column (5) reports results using *WAAP* to estimate the specific model, with absolute t-statistic and p-values provided similar to those in Column (1). Similar procedures are used for the life expectancy meta-regressions reported in Table A3.

Perusing Table A2, although there is variation across the models, the coefficient estimates and their significance are similar to those reported in the specific model of Column (1). For instance, we continue to find evidence in favor of publication selection bias, as well as spending elasticities being sensitive to the inclusion of income and lag spending in the production function. Further, spending at the sub-national level affects the spending elasticity, while public spending on

healthcare has a greater impact on mortality rates, *ceteris paribus*. Also, across the columns of Table A2, the predicted spending elasticities are similar to those values reported in Table 3.

Table A2
Estimation results for alternate weights and treatment of data dependence, mortality rate.

Variable	Specific model (1)	Sample size weights (2)	RE weights, *REML* (3)	Multilevel, *REML* (4)	Adequate Power (5)
Constant	-0.072**	-0.197*	-0.106**	-0.211***	-0.035
	(2.66) [0.06]	(1.84)	(2.46)	(4.14)	(1.59) [0.19]
SE	-1.948***	-1.589***	-1.583***	-0.549***	-2.691***
	(8.22) [0.01]	(3.60)	(6.12)	(5.73)	(7.71) [0.01]
OECD	-0.008	-0.018	-0.005	-0.040	-0.007
	(0.89) [0.39]	(0.72)	(0.28)	(1.31)	(0.18) [0.49]
Pharm	0.029	0.050*	0.056**	0.123***	0.020
	(1.35) [0.20]	(1.73)	(2.45)	(3.46)	(0.98) [0.49]
Public	-0.031**	-0.071***	-0.055***	-0.020	-0.024
	(2.14) [0.08]	(3.67)	(3.53)	(1.03)	(1.36) [0.26]
Sub-national	0.065***	0.128***	0.067***	-0.072	0.059***
	(4.50) [0.00]	(3.51)	(2.66)	(1.26)	(4.37) [0.00]
Income	0.067***	0.178*	0.083**	0.131***	0.045***
	(3.10) [0.06]	(1.65)	(2.19)	(3.27)	(3.26) [0.04]
Lag spending	-0.079***	-0.031	-0.043**	-0.016	-0.083***
	(6.11) [0.01]	(1.01)	(2.14)	(0.29)	(6.25) [0.01]
Lag outcome	0.026*	0.032	0.055**	0.010	0.020
	(1.69) [0.24]	(1.48)	(2.43)	(0.27)	(1.48) [0.33]
n (k)	629 (47)	577 (42)	629 (47)	629 (47)	284 (34)
Adjusted R^2	0.39	0.38	0.57	--	0.33
Predicted E	-0.11	-0.12	-0.12	-0.12	-0.10
95% CI	-0.16 to -0.07	-0.24 to -0.01	-0.18 to -0.06	-0.23 to -0.01	-0.14 to -0.06

Notes: Columns (1) and (2) report unrestricted *FE-WLS* results, using inverse variance weights. Column (1) reports results of the specific model from Column (5) of Table 3, while Column (2) reports the results of the specific model using sample size weights. Column (3) reports random-effects estimates, using *REML*. Column (4) reports multi-level or hierarchical linear model results, also estimated using *REML*. Column (5) reports *WAAP* estimator results. n (k) denotes number of observations (studies). *Predicted E* is the predicted spending elasticity for public spending, conditional on income and lagged spending being included in the health production function. Figures in parentheses are absolute *t*-statistics using standard errors adjusted for clustering of observations within studies; except for column (5). Figures in brackets in columns (1) and (5) are *p*-values constructed by re-estimating these models using the wild bootstrap to correct for uneven number of observations within studies, as well as study dependence.
*** p<0.01; ** p<0.05; * p<0.10.

Concerning life expectancy, similar to Table 4 in the paper, significance is sparser. Across Columns (1) – (5) of Table A3, evidence tends not to favor publication selection bias. Also, to varying degrees of significance, the spending elasticity for older populations is slightly larger than the population in general, as are spending elasticities reported in working papers, as well as those based on models correcting for endogeneity of healthcare spending. Further, predicted spending

elasticities in Table A3 are similar to those reported in Table 4 of the paper. Thus, comparing the spending elasticity for the mortality rate to the spending elasticity for life expectancy, the effect of spending on mortality remains much larger compared to life expectancy.

Table A3
Estimation results for alternate weights and treatment of data dependence, life expectancy.

Variable	Specific model (1)	Sample size weights (2)	RE weights, *REML* (3)	Multilevel, *REML* (4)	Adequate Power (5)
Constant	0.013***	-0.036	-0.027	0.046**	0.013***
	(8.08) [0.00]	(1.04)	(0.77)	(2.06)	(7.68) [0.00]
SE	1.051	0.037	0.462	-0.128*	1.147
	(1.32) [0.28]	(0.34)	(1.10)	(1.68)	(1.26) [0.29]
Older	0.007*	0.027***	0.018	0.051***	0.006
	(1.87) [0.04]	(3.35)	(1.63)	(4.12)	(1.58) [0.09]
OECD	-0.009**	0.033	0.031	-0.039	-0.009**
	(2.24) [0.06]	(0.99)	(0.91)	(1.40)	(2.17) [0.06]
Public	-0.004	-0.002	0.004	-0.002	-0.004
	(0.84) [0.52]	(0.26)	(0.42)	(0.17)	(0.77) [0.56]
Lag spending	-0.012***	0.009	-0.001	-0.043*	-0.012***
	(2.73) [0.01]	(0.87)	(0.08)	(1.91)	(2.72) [0.01]
Working paper	0.022**	0.061***	0.041***	0.013	0.021*
	(2.10) [0.07]	(8.05)	(4.50)	(0.34)	(1.76) [0.14]
Average year	-0.002***	0.000	0.000	-0.005***	-0.002***
	(9.48) [0.01]	(0.20)	(0.04)	(2.63)	(9.48) [0.01]
Endogeneity	0.022***	0.061**	0.059*	0.011	0.021***
	(3.83) [0.00]	(2.16)	(1.71)	(0.32)	(3.71) [0.00]
n (k)	256 (27)	256 (27)	256 (27)	256 (27)	246 (26)
Adjusted R^2	0.13	0.17	0.07	--	0.13
Predicted E	0.03	0.03	0.03	0.01	0.02
95% CI	0.01 to 0.05	-0.01 to 0.07	-0.01 to 0.08	-0.06 to 0.08	0.01 to 0.04

Notes: Columns (1) and (2) report unrestricted *FE-WLS* results, using inverse variance weights. Column (1) reports results of the specific model from Column (5) of Table 4, while Column (2) reports the results of the specific model using sample size weights. Column (3) reports random-effects estimates, using *REML*. Column (4) reports multi-level or hierarchical linear model results, also estimated using *REML*. Column (5) reports *WAAP* estimator results. n (k) denotes number of observations (studies). *Predicted E* is the predicted spending elasticity for life expectancy at an older age, conditional on endogeneity correction and lag spending being included in the production function. Figures in parentheses are absolute *t*-statistics using standard errors adjusted for clustering of observations within studies; except for column (5). Figures in brackets in columns (1) and (5) are *p*-values constructed by re-estimating these models using the wild bootstrap to correct for uneven number of observations within studies, as well as study dependence.
*** p<0.01; ** p<0.05; * p<0.10.

As a further check on the results, we removed each study, one at a time, and re-estimated the MRA models. For this exercise we estimate the general MRA model (Column (1) of Tables 3 and 4). Figures A5 and A6 report the results in the form of a forest plot. Each row reports the predicted spending elasticity and its 95% CI after removing the listed study. Since the predicted elasticities

are similar across the rows in each plot, this indicates they are robust to the exclusion of individual studies.

Figure A5: Sensitivity of predicted spending elasticity to individual studies, mortality rate

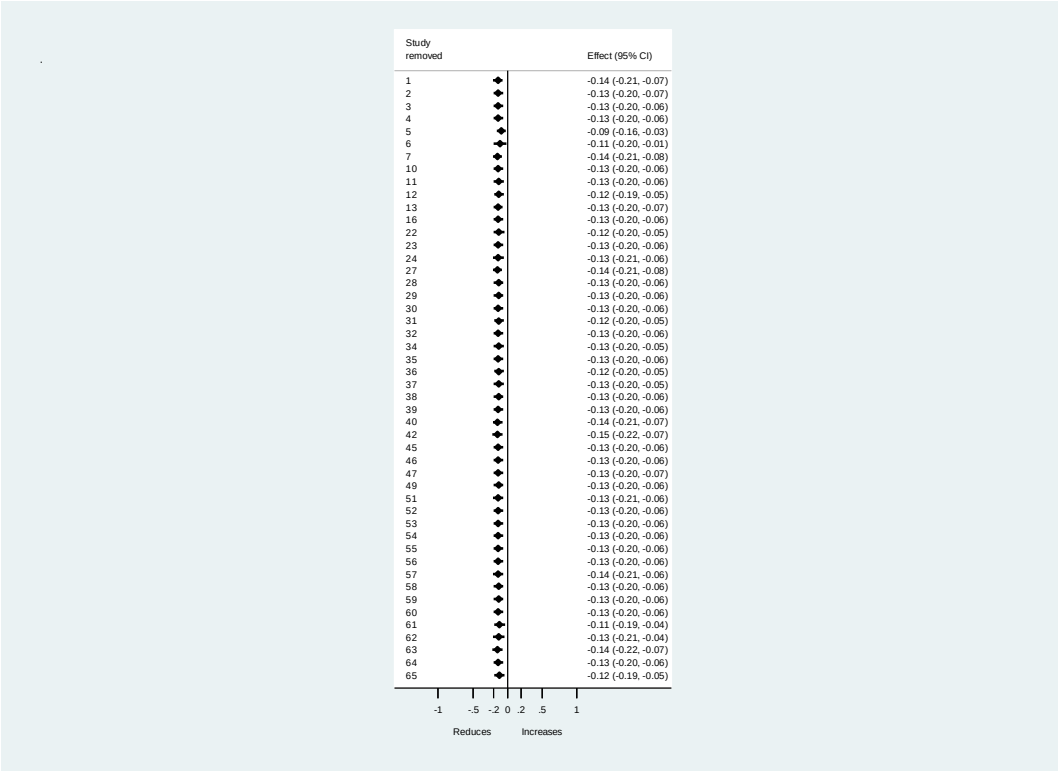


Figure A6: Sensitivity of predicted spending elasticity to individual studies, life expectancy