

The Impact of New Drugs on US Longevity and Medical Expenditure, 1990–2003: Evidence from Longitudinal, Disease-Level Data

By FRANK R. LICHTENBERG*

A number of econometric studies (Charles R. Hulten 1992; Byong-Hong Bahk and Michael Gort 1993; Plutarchos Sakellaris and Daniel J. Wilson 2004) have investigated the hypothesis that capital equipment employed by US manufacturing firms embodies technological change, i.e., that each successive vintage of investment is more productive than the last. Equipment is expected to embody significant technical progress due to the relatively high research and development (R&D) intensity of equipment manufacturers. The method that has been used to test the equipment-embodied technical change hypothesis is to estimate manufacturing production functions, including (mean) vintage of equipment as well as quantities of capital and labor. These studies have concluded that technical progress embodied in equipment is a major source of manufacturing productivity growth.

Embodied technical progress may also be an important source of economic growth in healthcare. One important input in the production of health—pharmaceuticals—is even more R&D intensive than equipment. According to the National Science Foundation (NSF), the R&D intensity of drugs and medicine manufacturing is 74 percent higher than the R&D intensity of machinery and equipment manufacturing. Therefore, it is quite plausible that there is also a high rate of pharmaceutical-embodied technical progress.

This study examines the effect of changes in the vintage distribution of prescription drugs on US longevity and medical expenditure during the 1990–2003 period. We will estimate the following model, using longitudinal disease-level data:

$$(1) \quad \ln Y_{it} = \beta X_{it} + \alpha_i + \delta_t + \varepsilon_{it},$$

where Y_{it} is a measure of mortality or healthcare utilization, and X_{it} is a measure of prescription drug vintage for medical condition (disease) i in

year t . Since the model includes condition and year fixed effects, it is a difference-in-differences model. Negative and significant estimates of β would indicate that conditions with above average increases in prescription drug vintage had above average declines (or below average increases) in mortality and hospitalization.

Equation (1) will be estimated using weighted least squares (WLS), where the weight is equal to $Y_i = (1/T) \sum_t Y_{it}$. Since the dependent variable is $\ln Y_{it}$, and the model includes fixed condition effects, we are in effect analyzing percentage deviations from condition means. Low-mean conditions exhibit much more volatility (noise) than high-mean conditions, so it is appropriate to give more weight to the percentage deviations from high-mean conditions.

We will use data on mortality for the 1990–2002 period, and data on healthcare utilization (hospital discharges) for the 1993–2003 period. We have prescription drug data for the years 1996 to 2003, but can impute values of prescription drug vintage in earlier years (1990 and 1993).

There are two different ways to estimate equation (1). One is to use data for all available years (1990 and 1996–2002 for mortality; 1993 and 1996–2003 for healthcare utilization). However, since the disturbances of equation (1) are likely to exhibit serial correlation, the standard errors of WLS estimates based on (nearly) annual data are likely to be underestimated.

The second approach is to estimate equation (1) using data only for the first and last years of the sample period ($t = 1$ and $t = T$, respectively). This is equivalent to estimating the “long-difference” model

$$(2) \quad \ln Y_{iT} - \ln Y_{i1} = \beta (X_{iT} - X_{i1}) + (\delta_T - \delta_1) + (\varepsilon_{iT} - \varepsilon_{i1}).$$

Since there is only one observation per condition, there can be no serial correlation.¹

¹ By construction, $X_{i1} = 0$ for all i .

* Graduate School of Business, Columbia University, 614 Uris Hall, 3022 Broadway, New York, NY 10027 (e-mail: frl1@columbia.edu).

Estimation of equation (2), as opposed to equation (1), may affect the *point estimate*, as well as the estimated standard error of β . The estimate of β from equation (1) is based on deviations from condition means; these are higher-frequency fluctuations than the long differences on which equation (2) is based. It is often the case that the response to a stimulus is larger in the long run than it is in the short run. For example, the permanent income hypothesis was developed to explain why consumption expenditure is more responsive to long-run changes in disposable income than it is to short-run changes. Similarly, mortality and healthcare utilization may be more responsive to long-run changes in the vintage distribution of prescription drugs than to short-run changes. A person taking a new drug for a chronic condition may have to take it for several years before its full health impact is realized.

I. Measures of Mortality, Healthcare Utilization, and Prescription Drug Vintage

Y_{it} is one of the following measures of mortality or healthcare utilization:

- $LYL65_{it}$: years of potential life lost before age 65 due to condition i in year t ;
- $LYL75_{it}$: years of potential life lost before age 75 due to condition i in year t ;
- $HOSP_TOT_{it}$: number of hospital admissions (or discharges) due to condition i in year t ;
- $HOSP_LTC_{it}$: number of hospital discharges to other institutions (nursing homes, rehabilitation facilities) due to condition i in year t ;
- $HOSP_DEAD_{it}$: number of hospital stays in which the patient died due to condition i in year t .

X_{it} is one of the following measures of prescription drug vintage:

- $POST1990_{it}$: the percent of prescriptions used to treat condition i in year t that contained active ingredients approved by the Food and Drug Administration (FDA) after 1990.
- $POST1993_{it}$: the percent of prescriptions used to treat condition i in year t that contained active ingredients approved by the FDA after 1993.

The disease classification we use is the International Classification of Diseases Tenth Revision 113 Cause-of-Death classification (<http://www.nber.org/mortality/2002/docs/113cause.txt>). Some of the 113 causes are subtotals of other causes. We excluded these subtotals. The classification we used contained 92 nonoverlapping diseases.

II. Data

Mortality data.—Data on LYL65 and LYL75 were computed from the Multiple Cause-of-Death Mortality Data from the National Vital Statistics System of the National Center for Health Statistics. Each record in the microdata is based on information abstracted from death certificates filed in vital statistics offices in each state and the District of Columbia. Causes of death were coded according to the International Classification of Diseases Ninth Revision, which includes data from the 1991–1998 period and the Tenth Revision, for which data begin in 1999. The average number of records (deaths) per year is about 2.3 million.

Hospital discharge data.—Annual data on hospital discharges, by cause, for the 1993–2002 period were obtained from HCUPnet (<http://hcupnet.ahrq.gov/>). HCUPnet reports the total number of discharges, the number of patients discharged to nursing homes, and the number of deaths. We obtained data on the number of discharges by *principal* diagnosis (the condition that is the chief reason for the hospital stay as determined after evaluation) rather than by *all-listed* diagnoses, which include all diagnoses at the time of admission and those that develop during the stay. Discharge status (e.g., whether the patient was discharged to a nursing home) cannot be determined for all listed diagnoses from HCUPnet.

Prescription drug data.—Data on prescribed medicines used to treat condition i in year t ($t = 1996$ – 2003) were obtained from the Medical Expenditure Panel Survey (<http://www.meps.ahrq.gov/mepsweb/>) prescribed medicines files. Each record in these files contains a National Drug Code (NDC) and the total amount paid for the prescription by all payers. It also contains up to three (self-reported) diagnosis (ICD9) codes.

The second and third are usually blank, so we used only the first diagnosis code. We used the Veterans Administration Pharmacy Benefits Management National Formulary to determine the active ingredient(s) associated with each NDC. We used data from Drugs@FDA to determine the year in which the FDA first approved each active ingredient.

Table 1 shows that the share of post-1990 prescriptions increased from 18 percent in 1996 to 50 percent in 2003.

III. Estimates

Estimates of equation (1) are presented in Table 2. Panel A shows estimates based on data for all available years. Panel B shows estimates based on data for only the first and last years. We will discuss the panel A estimates first.

Column 1 shows estimates from the regression of $\ln \text{LYL65}$ on POST1990 . There appears to be a highly significant inverse relationship between these variables, which suggests that conditions experiencing greater pharmaceutical innovation tend to have greater declines in mortality before age 65. As noted above, however, the standard error of β is likely to be underestimated due to serial correlation, so the statistical significance of this estimate is not guaranteed. We should reserve judgment on this until we examine the long-difference estimates in panel B.

Column 2 shows estimates from the regression of $\ln \text{LYL75}$ on POST1990 . This estimate also appears to be highly significant, but the point estimate here is 40 percent smaller than the point estimate in column 1.

Column 3 shows estimates from the regression of $\ln \text{DISCH_TOT}$ on POST1993 . Conditions with larger increases in the fraction of prescriptions for post-1993 drugs had smaller increases, or larger declines, in the total number of hospital admissions. Column 4 shows estimates from the regression of the log of the number of hospital discharges to nursing homes on POST1993 . Once again, the estimate of β is negative and appears to be highly significant. Conditions with larger increases in the fraction of prescriptions for post-1993 drugs had smaller increases, or larger declines, in the number of hospital discharges to nursing homes. But the estimates in column 5 show that more rapid pharmaceutical innovation was *not* associated with a decline in

TABLE 1—TRENDS IN PRESCRIPTION DRUG VINTAGE DURING THE 1996–2003 PERIOD

Year	Number of sample Rx's	POST1990
1996	147,222	18 %
1997	234,405	22 %
1998	171,906	28 %
1999	173,831	33 %
2000	182,524	37 %
2001	277,744	42 %
2002	339,097	46 %
2003	304,149	50 %

the number of in-hospital deaths. Indeed, more rapid pharmaceutical innovation was associated with an *increase* in the probability of dying in the hospital, conditional on being admitted. The estimates in column 1 and 2, however, suggest that pharmaceutical innovation reduced overall mortality before age 75 and especially before age 65.²

Now we will discuss estimates in panel B of Table 2. As expected, in column 1 the standard error of the panel B estimate is about three times as large as the standard error of the panel A estimate. However, the panel B *point estimate* of β is also about three times as large as the panel A point estimate. Therefore, the panel B point estimate is also highly statistically significant (p -value < 0.0001). As discussed above, the difference in point estimates may be attributable to the fact that mortality is more responsive to long-run changes in the vintage distribution of prescription drugs than it is to short-run changes.

In columns 2 through 4, the panel B point estimates are also much larger (two to four times larger) than the corresponding panel A estimates, and are statistically significant at the 5-percent level or lower. In column 5, the panel B estimate, like the panel A estimate, is not statistically significant.

IV. Discussion

These estimates suggest that there is a strong inverse relationship across conditions between pharmaceutical innovation and changes in mortality, hospital admissions, and hospital discharges to nursing homes. We will use the estimates to

² About one in three deaths occur in hospitals.

TABLE 2—ESTIMATES OF EQUATION (1)

	(1)	(2)	(3)	(4)	(5)
Dependent variable	ln LYL65	ln LYL75	ln DISCH_TOT	ln DISCH_LTC	ln DISCH_DEAD
Vintage variable	POST1990	POST1990	POST1993	POST1993	POST1993
<i>Panel A: All available years</i>					
	1990, 1996–2002	1990, 1996–2002	1993, 1996–2003	1993, 1996–2003	1993, 1996–2003
Number of observations	550	552	595	594	592
β	-0.349	-0.210	-0.310	-0.314	-0.051
std. err.	0.057	0.050	0.084	0.081	0.061
<i>t</i> -stat	-6.13	-4.24	-3.67	-3.88	-0.83
<i>p</i> -value	<0.0001	<.0001	0.0003	0.0001	0.4078
<i>Panel B: First and last years only</i>					
	1990, 2002	1990, 2002	1993, 2003	1993, 2003	1993, 2003
Number of observations	149	150	146	145	144
β	-1.125	-0.788	-0.980	-0.652	-0.060
std. err.	0.189	0.196	0.300	0.315	0.245
<i>t</i> -stat	-5.96	-4.01	-3.26	-2.07	-0.25
<i>p</i> -value	<0.0001	0.0002	0.0019	0.0431	0.8069

compare the costs and benefits of increasing use of new drugs during the sample period.

The estimates imply that, in the absence of any pharmaceutical innovation during the sample period, mortality, hospital admissions, and hospital discharges to nursing homes would have been higher at the end of the period (2002 or 2003) than they were. The implied (absolute) increases in these variables that would have occurred, absent any innovation, may be calculated as follows:

$$Y_PRED_T - Y_T = Y_T \exp(\beta(X_T - X_1) - 1),$$

where

Y_PRED_T = the predicted (counterfactual) value of Y in period T ;

Y_T = the actual value of Y in period T ;

X_T = the value of X in period T ;

X_1 = the value of X in period 1.

These calculations are shown in Table 3. First, we will use the estimates of β , based on data for all available years (panel A), to compare the costs and benefits of increasing the use of new drugs during the sample period. Then we will conduct similar calculations using the estimates of β based on data for first and last years (panel B).

Table 3, column 1, shows the LYL65 calculations. The Centers for Disease Control and

Prevention (CDC) reports 8,229,329 years of potential life were lost before the age of 65 in 2002.³ Our estimates imply that, if the 2002 sample mean value of POST1990 had been equal to its 1990 value (zero) instead of its actual value (50 percent), LYL65 would have been 19 percent higher, and an additional 1.57 million years of potential life would have been lost before the age of 65.

Column 2 shows the LYL75 calculations. The CDC reports 15,418,775 years of potential life were lost before the age of 75 in 2002. Our estimates imply that, if the 2002 sample mean value of POST1990 had been equal to its 1990 value instead of its actual value, LYL75 would have been 11 percent higher and an additional 1.70 million years of potential life would have been lost before the age of 75. The fact that the LYL75 estimate is not much larger than the LYL65 estimate could signify that many of the people who previously would have died before age 65 died between the ages of 65 and 75.

Column 3 shows hospital expenditure calculations. CMS reports that 2003 US hospital care expenditure was \$515.9 billion. Our estimates imply that, if the 2003 sample mean value of POST1993 had been equal to its 1993 value (zero) instead of its actual value (34 percent),

³ We exclude years of potential life lost to unintentional injury, suicide, and homicide, since our analysis examines mortality and hospitalization from natural causes only.

TABLE 3—ESTIMATED EFFECTS OF NEW DRUG USE ON MORTALITY IN 2002 AND MEDICAL EXPENDITURE IN 2003

Variable	(1) LYL65 in 2002	(2) LYL75 in 2002	(3) Hospital Care expenditure in 2003	(4) Nursing Home expenditure in 2003
Y_T	8,229,329	15,418,775	\$515,900,000,000	\$110,800,000,000
$(X_T - X_1)$	0.499	0.499	0.344	0.344
<i>Panel A: Estimates based on data for all available years</i>				
$-\beta(X_T - X_1)$	0.174	0.105	0.107	0.108
$Y_PRED_T - Y_T$	1,566,673	1,701,974	\$58,050,711,511	\$12,629,163,198
<i>Panel B: Estimates based on data for first and last years only</i>				
$-\beta(X_T - X_1)$	0.561	0.393	0.337	0.224
$Y_PRED_T - Y_T$	6,198,015	7,432,321	\$206,577,401,177	\$27,842,189,879

2003 hospital expenditure would have been \$58 billion (11 percent) higher.⁴

Column 4 shows nursing home expenditure calculations. CMS reports that 2003 US nursing home expenditure was \$110.8 billion. Our estimates imply that, if the 2003 sample mean value of POST1993 had been equal to its 1993 value instead of its actual value, 2003 nursing home expenditure would have been \$12.6 billion (11 percent) higher. This assumes that the percentage reduction in nursing home expenditure is equal to the percentage reduction in the number of hospital discharges to nursing homes. According to the National Nursing Home Survey, about half of nursing home residents were admitted directly from a hospital. Therefore, the percentage reduction in nursing home expenditure could be smaller than the percentage reduction in the number of hospital discharges to nursing homes. If utilization of newer drugs had *no* effect on admissions to nursing homes from places other than hospitals, the reduction in 2003 nursing home expenditure attributable to the use of post-1993 drugs was about \$6.3 billion. I will use the midpoint of these two figures (\$9.5 billion) as my estimate of the reduction in 2003 nursing home expenditure attributable to the use of post-1993 drugs.

To summarize, our more conservative estimates imply that use of post-1990 drugs in 2002 reduced the number of life years lost before age 65 by 1.56 million and the number of life years lost before age 75 by about 1.70 million. The

use of post-1993 drugs in 2003 reduced hospital expenditure by \$58 billion and nursing home expenditure by \$9.5 billion.

CMS also reports that US expenditure on prescription drugs was \$162 billion in 2002 and \$179 billion in 2003. I estimate that just over half the money spent on prescriptions in 2003 was on post-1993 drugs. Medicaid data for 2003 indicate that between 46 percent and 56 percent of Medicaid drug expenditure in that year was on post-1993 drugs.⁵ Data reported by NDCHealth (NDCHealth Pharmaceutical Audit Suite, <http://www.rxlist.com>) indicate that in 2004 53 percent of expenditure on the top 100 drugs went to drugs approved after 1994. It is therefore reasonable to assume that 53 percent of 2003 drug expenditure went to post-1993 drugs. This implies that \$95 billion (= 53 percent * \$179 billion) was spent on post-1993 drugs in 2003.

Since the more conservative estimates imply that the use of post-1993 drugs reduced hospital expenditure by \$58 billion and nursing home expenditure by \$9.5 billion in 2003, they imply that the net cost of these drugs was \$27 billion (= \$95 billion - \$58 billion - \$9.5 billion). The net cost per life year saved before age 75 was \$15,974 (= \$27 billion/1.7 million life years). Treatments of this cost level are generally considered to be quite cost-effective. Moreover, the

⁴ This assumes that the percentage reduction in hospital expenditure is equal to the percentage reduction in the number of admissions.

⁵ The FDA approval year of some drugs is unknown. These drugs accounted for 19 percent of 2003 Medicaid drug expenditure. Drugs with unknown approval years are likely to be older than drugs with known approval years, and their average prices are lower. Post-1993 drugs accounted for 56 percent of spending on drugs with known approval years and 46 percent of spending on all drugs.

true net cost per life year gained could be much lower for several reasons.

First, this calculation does not account for the likely reduction in years of potential life lost after age 75. Second, it does not account for other benefits of pharmaceutical innovation, such as the reduction in work-loss days and illness-induced withdrawal from the labor market. Third, since this estimate is based on data for all available years, it may reflect only the short-run benefits of new drugs. Estimates based on data for the first and last years only, shown in panel B of Table 2, may capture the long-run benefits. These estimates indicate that the net cost of new drugs was *negative*: the reduction in the sum of hospital and nursing home costs was 2.4 times as great as the cost of the drugs. They also indicate that in 2003 new drugs reduced the number of years of potential life lost before age 75 by 7.4 million.

Use of newer drugs may have cross-disease spillover effects—using newer drugs for one disease may either increase or decrease mortality from other diseases (in part due to “competing risks”). Such spillovers could be either negative or positive. For example, using a newer drug to treat cardiovascular disease might reduce cardiovascular mortality but increase life years lost due to cancer. On the other hand, using a newer

drug to treat depression and other mental disorders might lead to better management of cardiovascular disease.

The models we have estimated control (via year fixed effects) for nondrug (device/procedure) innovation common to all diseases, but not for disease-specific device/procedure innovation, which is difficult to measure. Since device/procedure innovation may either substitute for or complement drug innovation, controlling for disease-specific device/procedure innovation could either decrease or increase our estimate of the cost per life year gained from using newer drugs. This issue merits further research.

REFERENCES

- ▶ **Bahk, Byong-Hong, and Michael Gort.** 1993. “Decomposing Learning by Doing in New Plants.” *Journal of Political Economy*, 101 (4): 561-83.
- Hulten, Charles R.** 1992. “Growth Accounting When Technical Change Is Embodied in Capital.” *American Economic Review*, 82(4): 964-80.
- Sakellaris, Plutarchos, and Daniel J. Wilson.** 2004. “Quantifying Embodied Technological Change.” *Review of Economic Dynamics*, 7(1): 1-26.