Online Appendix for:
Patents and Research Investments:
Assessing the Empirical Evidence

By ERIC BUDISH, BENJAMIN N. ROIN, AND HEIDI WILLIAMS

ONLINE APPENDIX

A1. Elasticity calculation from Budish, Roin and Williams (2015)

This section re-prints (with minimal edits) Online Appendix Section A.10 of Budish, Roin and Williams (2015), as a point of comparison for the other elasticity estimates we discuss in this paper.

As described in Section II, the empirical estimates in Budish, Roin and Williams (2015) provide an estimate of how R&D investment changes with the 5-year survival rate (our proxy for clinical trial length), $\frac{\partial(\text{R\&D investment})}{\partial(\text{5-year survival rate})}$. To translate this estimate into an estimate of how R&D investment would respond to an increase in the patent term, we would like to scale $\frac{\partial(\text{R\&D investment})}{\partial(\text{5-year survival rate})}$ by an estimate of how the patent term varies with the 5-year survival rate. In practice, we do this scaling — under a strong assumption, as detailed below — using an estimate of how a drug's commercialization lag varies with the 5-year survival rate, since one less year of commercialization lag is equivalent (under this assumption) to one additional year of patent life. By combining these estimates, we can then estimate the elasticity of interest:

$$\frac{\frac{\partial(\text{R\&D investment})}{\partial(\text{5-year survival rate})}}{\frac{\partial(\text{commercialization lag})}{\partial(\text{5-year survival rate})}} = \frac{\partial\left(\text{R\&D investment}\right)}{\partial\left(\text{commercialization lag}\right)} \approx \frac{\partial\left(\text{R\&D investment}\right)}{\partial\left(\text{patent term}\right)}$$

The conceptual problem with estimating $\frac{\partial(\text{commercialization lag})}{\partial(\text{5-year survival rate})}$ is that - by construction - we only observe clinical trial length *conditional* on a drug compound being placed in clinical trials. Because - consistent with the theoretical model presented in Budish, Roin and Williams (2015) - we document that fewer drug compounds are placed in clinical trials for patients with longer survival times, we expect selection into clinical trials to bias the relationship between patient survival and clinical trial length in the set of observed clinical trials.[1] Given this selection bias in which trials are observed in our data, we cannot obtain an unbiased empirical estimate of $\frac{\partial(\text{commercialization lag})}{\partial(\text{5-year survival rate})}$. To overcome this selection problem, we instead calibrate the relationship between commercialization lag and the 5-year survival rate using the power calculation outlined in Online Appendix Section A.9 of Budish, Roin and Williams (2015).

We can approximate our estimate of $\frac{\partial(\text{commercialization lag})}{\partial(\text{5-year survival rate})}$ with an estimate of $\frac{\partial(\text{clinical trial length})}{\partial(\text{5-year survival rate})}$, given that we expect commercialization lag to scale one-for-one with clinical trial length. In the language of the power calculation outlined in Online Appendix Section A.9 of Budish, Roin and Williams (2015), we can re-write this elasticity as:

$$\frac{\partial\left(\text{clinical trial length}\right)}{\partial\left(\text{5-year survival rate}\right)} = \frac{\partial k}{\partial \mu} = \frac{k\mu^{k-1} + Rk(1 - R(1-\mu))^{k-1}}{-[\mu^k \ln \mu + (1 - R(1-\mu))^k \ln(1 - (R(1-\mu)))]}$$

where μ is the per-period survival rate of untreated individuals, k is the number of periods of patient follow-up, and R is a constant per-period multiplicative treatment effect such that in a given period $1 - \mu$ individuals die in the control group and $R(1 - \mu)$ individuals

---

[1] As discussed in detail in Budish, Roin and Williams (2015), perhaps the most natural selection story is that firms are only willing to place a drug compound in clinical trials for patients with long expected survival times if they receive permission to use a surrogate endpoint in place of survival as an endpoint; in this case, the relationship between patient survival and clinical trial length would be biased towards zero. If we estimate this relationship in our data, we do estimate a statistically significant relationship; however, the magnitude is implausibly small, consistent with our prior that this relationship would be biased towards zero (a ten percentage point increase in the five-year survival rate is associated with a 1.5 percent increase in average clinical trial length - an increase on the order of one month).

die in the treatment group, where R is constrained such that $R(1 - \mu)$ is bounded by 0 and 1.

Intuitively, μ and k come in pairs - not all μ and k will generate sufficient statistical power conditional on a given technology (R). Here, we take the two (μ, k) pairs from the examples in the introduction of Budish, Roin and Williams (2015) given that by construction these are feasible pairs (given that the trials were completed), that we know these trials looked at survival outcomes (rather than some alternative surrogate endpoints), and that these examples span different ends of the spectrum of available technologies. We assume a technology of $R = 0.8$, which translates to a 20 percent improvement in the five-year survival rate; this choice of R is arbitrary but we explore robustness to alternative values of R below. Given the assumed value of R, the two examples in the introduction of Budish, Roin and Williams (2015) can be written as:

1) Metastatic prostate cancer: 5-year survival rate of 20 percent ($\mu = 0.2$)

   - Follow-up time of 12.8 months (($12.8/12$)/5 implies $k = 0.213$ units in 5-year increments)
   - Total trial length of 3 years (3/5 implies $k = 0.6$ in 5-year increments)

2) Localized prostate cancer: 5-year survival rate of 80 percent ($\mu = 0.8$)

   - Follow-up time of 9.1 years (9.1/5 implies $k = 1.82$ units in 5-year increments)
   - Total trial length of 18 years (18/5 implies $k = 3.6$ units in 5-year increments)

Plugging in these values for μ, k, and R into the above formula for $\frac{\partial k}{\partial \mu}$ gives estimates of 2.234 for metastatic prostate cancer, and 0.766 for localized prostate cancer. Those estimates are in units of 5-year increments, and multiplying them by 5 to translate them into a 1-year unit gives 11.170 and 3.827. In words, a change from 0 to 1 in the 5-year survival rate translates to between a 3.827-11.170 year increase in patient follow-up time. Therefore, we use this 3.827-11.170 range as our estimate of $\frac{\partial(\text{commercialization lag})}{\partial(\text{5-year survival rate})}$.

Our estimate from Table 2 Column (1) in Budish, Roin and Williams (2015) implies that a change from 0 to 1 in the 5-year survival rate translates into an 86.9% reduction in R&D investment. Dividing this estimate of $\frac{\partial(\text{R\&D investment})}{\partial(\text{5-year survival rate})}$ by our estimates of $\frac{\partial(\text{commercialization lag})}{\partial(\text{5-year survival rate})}$ (following the formula on the previous page) implies an estimated semi-elasticity of R&D investment with respect to a one-year change in commercialization lag of between 7.779% (based on metastatic prostate cancer; 86.9/11.170 = 7.779) and 22.707% (based on localized prostate cancer; 86.9/3.827 = 22.707).

Alternatively, we can do the same calculation using total trial length (3 and 18 years) rather than follow-up times (12.8 months and 9.1 years). Reassuringly, we obtain nearly identical estimates: 7.993% (based on metastatic prostate cancer; 86.9/10.872 = 7.993) and 23.416% (based on localized prostate cancer; 86.9/3.711 = 23.416).

We can investigate sensitivity of our estimates to different assumed values of R, the quality of the technology. A 'reasonable' range of R might be between 0.15-0.95, in which case our estimated elasticities fall between 6-54%.[2]

The above back-of-the-envelope calculation requires several assumptions, but gives some sense of magnitudes. A more important and substantive assumption is needed to conclude that our estimate of $\frac{\partial(\text{R\&D investment})}{\partial(\text{commercialization lag})}$ provides a valid estimate of $\frac{\partial(\text{R\&D investment})}{\partial(\text{patent term})}$.

---

[2]Our metastatic prostate cancer example - where the treatment resulted in a gain of 3.9 months on average - corresponds to $R = 0.961$, which implies elasticity estimates between 6-20%. On the other extreme Gleevec, often referenced as a "miracle" drug, is estimated to have increased the five-year survival rate from 30% to 89% - implying $R = 0.157$, and elasticity estimates between 17-54%.

Specifically, these two estimates are equivalent only if a one year reduction in commercialization lag matters only through inducing a one year increase in effective patent life, and not through other channels. This would not be the case if, for example, cost differences between short and long clinical trials are large enough to be a quantitatively important driver of R&D investment decisions.

*

REFERENCES

Budish, Eric, Benjamin N. Roin, and Heidi Williams. 2015. "Do firms underinvest in long-term research? Evidence from cancer clinical trials." *American Economic Review*, 105(7): 2044–2085.