Reanalysis of "Competition and Innovation: An Inverted-U Relationship"[1]

Andrew Gelman[2]

29 Oct 2025

Two of the winners of the 2025 Nobel prize in economics, credited "for having explained innovation-driven economic growth," were coauthors of an influential (7700 Google scholar citations) paper, Aghion et al. (2005) which claimed to "find strong evidence of an inverted U relationship" between "product market competition and innovation" and presented the visualization shown in Figure 1 with raw data and fitted quadratic regression (the coefficient for the quadratic term was reported as statistically significantly different from zero), and Figure 2 showing the fitted curve and alternative nonparametric fit.

Reading the paper left me with several concerns. In roughly increasing order of importance:

- They use Poisson regression, but the data are not counts, and it was not clear if their analysis corrected for overdispersion.
- The paper's all about an inverted U relationship, but this is driven by fitting a quadratic curve rather than, say, a curve with diminishing returns.
- The line in Figure 1 does not seem to go through the data points. In particular, the curve seems to be too low at the rightmost part of the graph, and the nonparametric curve seems to be an even worse fit.
- The model predicts patents from profit margin in the same year, but to the extent the model is appropriate I think you'd expect a lag.
- The outcome variable is some weighted count of patents but it's being used as a measure of the more abstract concept of "innovation," and the predictor is an average of profit margins but it's being used as labeled a measure of the more abstract concept of "competition."

I downloaded the data (available at Bloom, 2025) and performed a series of reanalyses, culminating in a model that addressed the first three of the above concerns:

- Rather than using a Poisson or quasipoisson regression for weighted patents, I fit a linear regression to log(weighted patents + 1).
- Rather than a quadratic regression, I fit from a family of hinge functions, $E(y|x) = a + b_0*(x-x_0) + (b_1-b_0)*\delta*\log(1+\exp((x-x_0)/\delta))$, as pictured in Figure 3.
- Following the original paper, I also included industry and year effects, but using a multilevel model, as is appropriate given the sparsity of the data in some industries.

Justifications of these choices and details of various intermediate models are at Gelman (2025).

---

[1] Submitted to *Journal of Robustness Reports.*

[2] Department of Statistics and Department of Political Science, Columbia University, New York.

Figure 4 graphs the data and fitted curves by industry. The posterior probability is over 99% that the slope b_0 is positive and the slope b_1 is negative, so our analysis supports the inverted U.

However, we still don't buy the paper's substantive claims. We doubt the use of patent counts to proxy for innovation; also there are concerns about the selection of data used in the study. Also, Figure 4 shows that fitted curves, they explain very little of the variation in the outcome. In addition, the U shape is highly sensitive to the choice of fitting the same curve (except for additive up and down shifts) to all the industries.

References

Philippe Aghion, Nick Bloom, Richard Blundell, Rachel Griffith, Peter Howitt (2005). Competition and Innovation: An Inverted-U Relationship. *Quarterly Journal of Economics* 120, 701-728.

Nicholas Bloom (2025). Data for "Competition and Innovation: An Inverted U Relationship." https://nbloom.people.stanford.edu/research

Andrew Gelman (2025). Reanalysis of that Nobel prizewinning study of patents (with R and Stan code). *Statistical Modeling, Causal Inference, and Social Science*, 21 Oct. https://statmodeling.stat.columbia.edu/2025/10/21/reanalysis-of-that-nobel-prizewinning-study-of-patents-and-innovation/

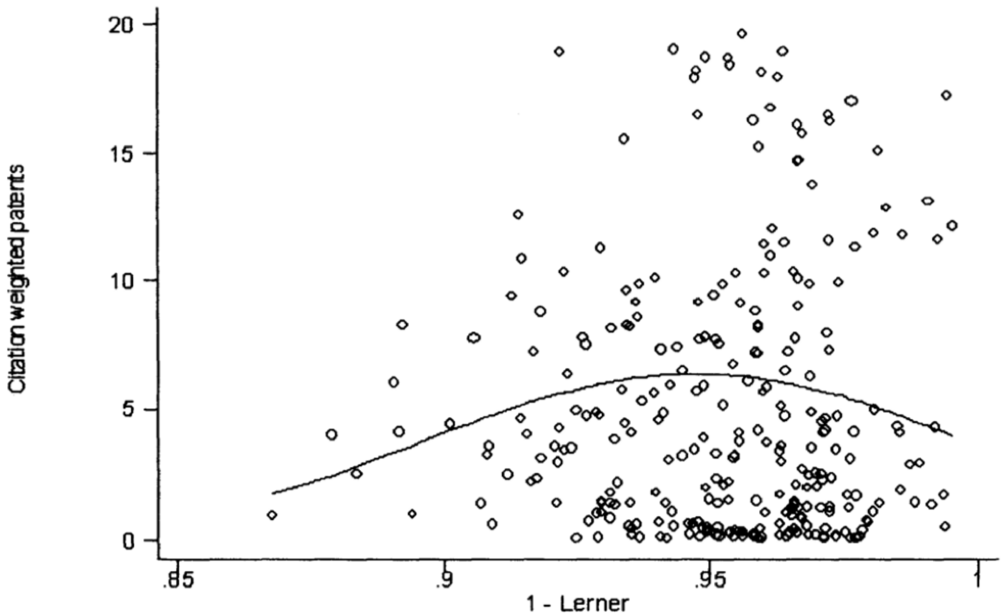

Figure 1, from Aghion et al. (2005). Original figure is labeled, “Scatter Plot of Innovation on Competition,” with caption, “The figure plots a measure of competition on the x-axis against citation-weighted patents on the y-axis. Each point represents an industry-year. The scatter shows all data points that lie in between the tenth and ninetieth deciles in the citation-weighted patents distribution. The exponential quadratic curve that is overlaid is reported in column (2) of Table I.”

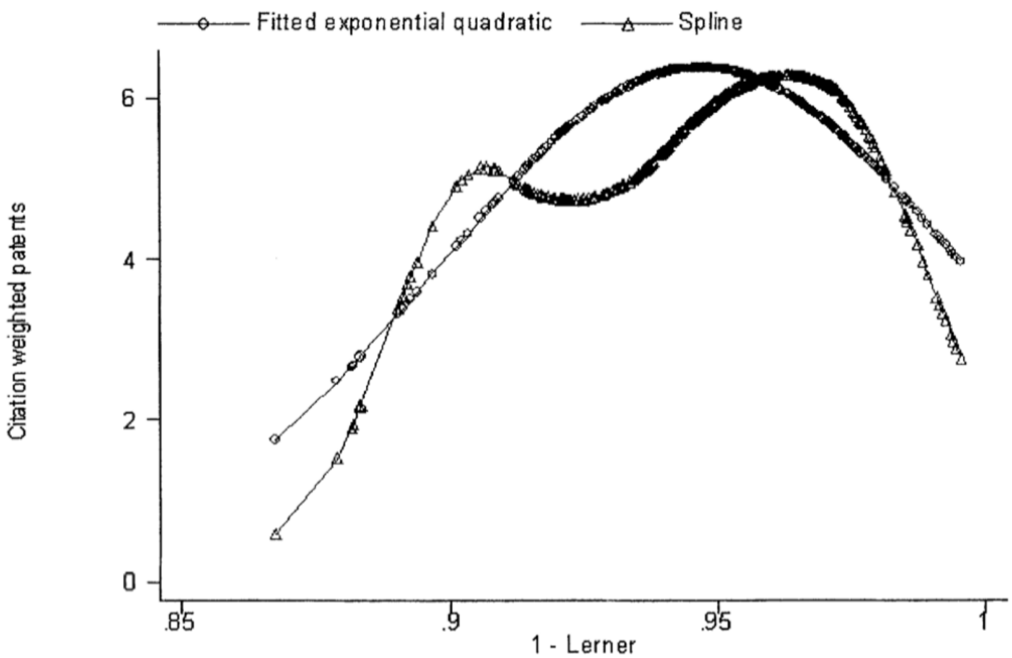

Figure 2, from Aghion et al. (2005). Original figure is labeled, "Innovation and Competition: Exponential Quadratic and the Semiparametric Specifications with Year and Industry Effects," with caption, "The figure plots a measure of competition on the x-axis against citation-weighted patents on the y-axis. Each point represents an industry-year. The circles show the exponential quadratic curve that is reported in column (2) of Table I. The triangles show a nonparametric spline."

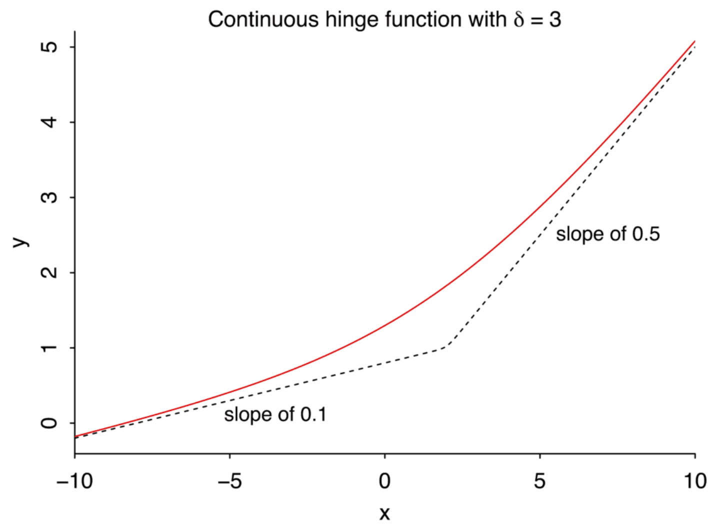

Figure 3. Example of a hinge function, a flexible family of curves that allows an inverse U shape (if the initial slope b_0 is positive and the later slope b_1 is negative) but, unlike the quadratic, also allows for diminishing returns without necessarily ever curving back down.

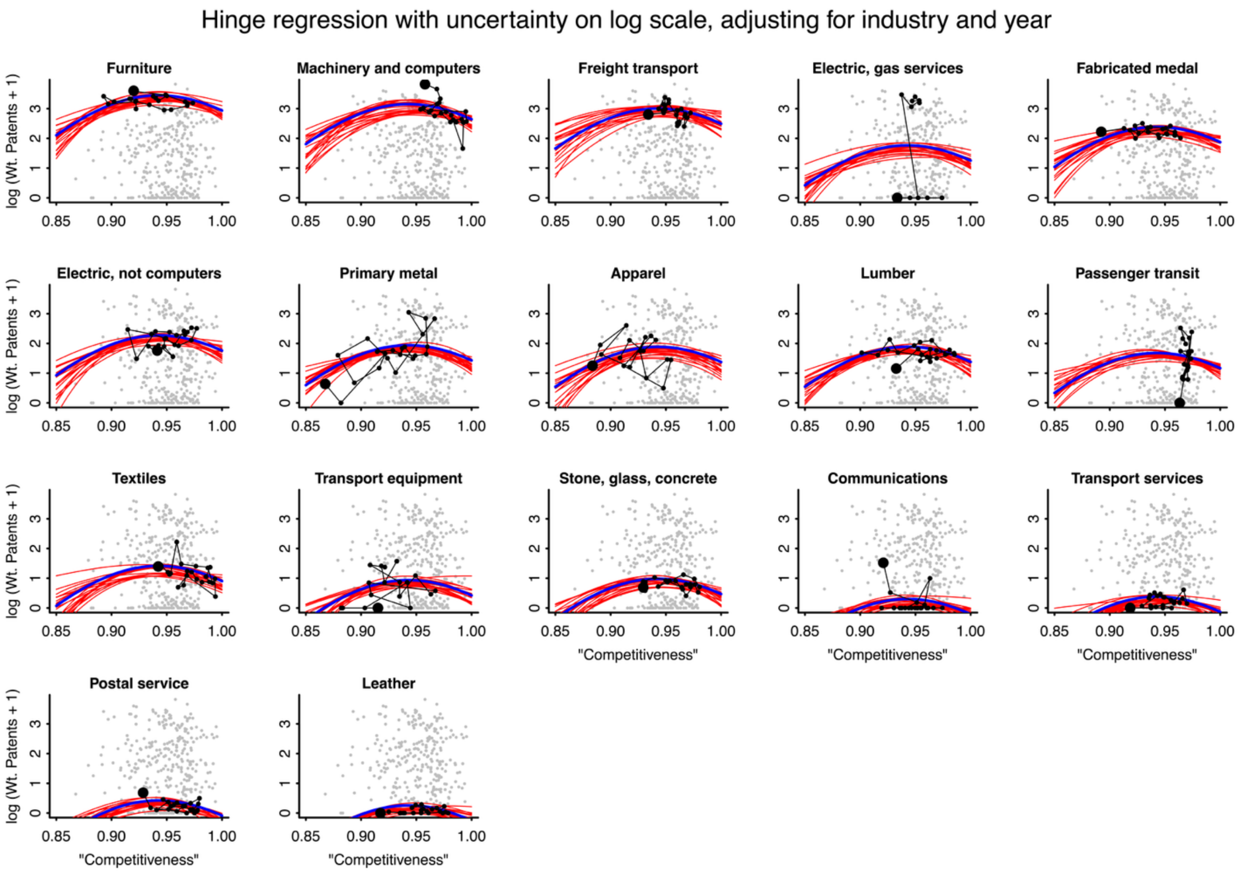

Figure 4. Data and fitted hinge model. For each plot, black dots show the data for that industry, with a large dot showing the data from the first year of the data and thin black lines showing the time sequence. The model had additive effects for industry and year; for each industry, the curves show the fitted model corresponding to an average year, with the blue curve representing the posterior median of the parameters and the red curves corresponding to 20 draws from the posterior distribution.