Reconsidering of "Competition and Innovation: An Inverted-U Relationship"

Andrew Gelman[1]⋆

1 Department of Statistics, and Department of Political Science, Columbia University, New York, USA

⋆ ag389@columbia.edu

JOURNAL OF ROBUSTNESS REPORTS

Abstract

The inverted U held up in reanalysis, but the result is sensitive to the assumption of a common curve for all industries.

Publication information to appear upon publication.

Received Date
Accepted Date
Published Date

Target article

P. Aghion, N. Bloom, R. Blundell, R. Griffith and P. Howitt, *Competition and innovation: An inverted-U relationship*, Quarterly Journal of Economics **120**, 701 (2005), doi:10.1093/qje/120.2.701

1 Goal

To assess dependence on modeling choices of a prominent claim regarding economic competition and corporate innovation. The reanalysis does not address concerns about data selection and quality.

2 Methods

[1] reported "strong evidence of an inverted U relationship" between "product market competition and innovation," with the data graph in Figure 1A and a fitted quadratic regression and alternative nonparametric fit in Figure 1B. This is a much-cited paper that gained additional notice after two of its authors won the Nobel prize.

Several concerns arose:

- Poisson regression was used for non-count data and there is a concern about overdispersion.
- The paper reported an inverted U relationship, but this was based on fitting a quadratic, a curve that does not allow for diminishing but nonnegative returns.
- The quadratic and the nonparametric curve in Figure 1B fit the data poorly in the crucial declining phase of the curves.

- The model predicts patents from profit margin in the same year, with no lag.
- The outcome is some weighted count of patents but represents the more abstract concept of "innovation," and the predictor is an average of profit margins but represents the more abstract concept of "competition."

I performed a series of reanalyses, culminating in a model that addressed the first three of the above issues:

- Rather than using a Poisson or quasipoisson regression, I fit linear models on $\sqrt{y}$, a standard variance-stabilizing transformation for count data.
- Rather than a quadratic regression, I used the hinge function, $E(\sqrt{y}|x) = a + b_0(x - x_0)) + (b_1 - b_0)\delta \log(1 + \exp((x - x_0)/\delta)$, which has an inverted U shape if $b_1 < 0 < b_0$ [2]. I fit using Bayesian inference with δ set to 0.05 (a potential sharp hinge to allow for a strong U curve), x_0 given a normal$(1, 1)$ prior (so that the hinge could fall within or outside the range of the data), and with the other parameters given weak normal$(0, 100)$ priors.
- 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 intermediate models are at [3].

3 Results

Figure 1C graphs the data and fitted curves by industry, which follow the inverted U, with a posterior probability of over 99% that the slope is positive and then negative. I obtained very similar results when modeling $\log(y + 1)$. In both cases, the U shape is robust but only because these models enforce the curve to have a constant same shape for all industries. I did not perform any reanalyses to address the concerns about lagging and measurement issues.

4 Conclusion

My reanalysis differed from the published paper in the use of a transformation, the use of multilevel modeling, and a different functional form for the curve, fit using Bayesian inference. The result was qualitatively similar to that in the published paper. I remain concerned about measurement issues and because the quantitative result is sensitive to the assumption that the curve is the same is across industries, a model assumption that remained in my reanalysis.

Acknowledgments and Disclosures

Reproducibility We were able to computationally reproduce the original analysis and results.

Code and Data Availability Data and code are at https://osf.io/x6h23/.

Funding The author(s) received no financial support for the research, authorship, and/or publication of this article.

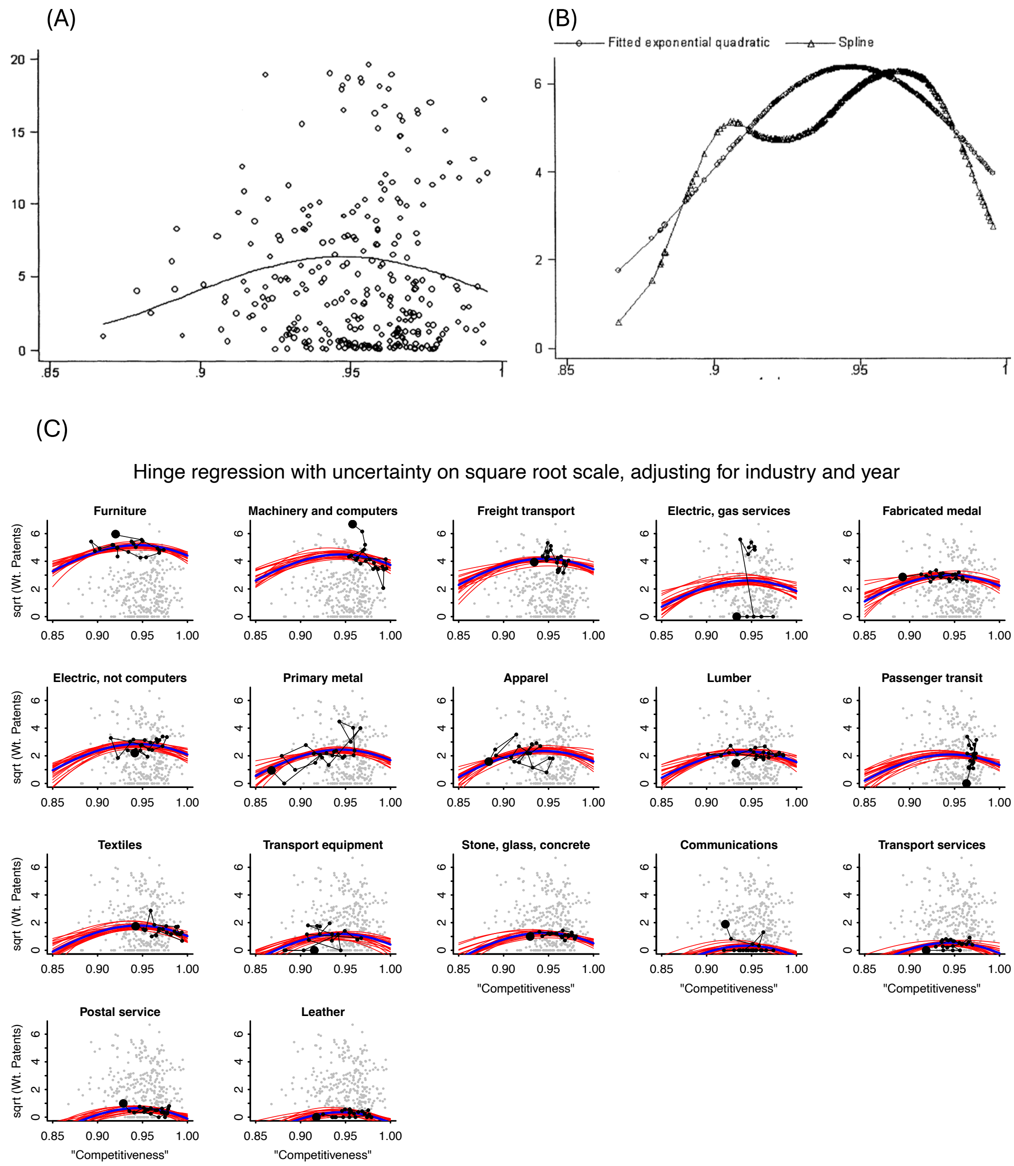

Figure 1: (A) from [1]. Original figure is labeled, "Scatter Plot of Innovation on Competition," with caption, "The figure plots a measure of competition on the x-axis against citation-weighted patents on the y-axis. Each point represents an industry-year. The scatter shows all data points that lie in 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." The quadratic coefficient was statistically significantly different from zero. (B) from [1]. 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." (C) Data by industry 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.

References

[1] P. Aghion, N. Bloom, R. Blundell, R. Griffith and P. Howitt, *Competition and innovation: An inverted-U relationship*, Quarterly Journal of Economics **120**, 701 (2005), doi:10.1093/qje/120.2.701.

[2] A. Gelman, *A continuous hinge function for statistical modeling*, Statistical Modeling, Causal Inference, and Social Science, 19 May (2017), https://statmodeling.stat.columbia.edu/2017/05/19/continuous-hinge-function-bayesian-modeling/.

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