Patent Citations and the Geography of Knowledge Spillovers: A Reassessment: Reply

By PETER THOMPSON AND MELANIE FOX-KEAN*

It is perhaps useful to summarize the main points of our reassessment of Adam B. Jaffe, Manuel Trajtenberg, and Rebecca Henderson's (1993, henceforth JTH) pioneering work on the geography of knowledge spillovers in the language of the econometrics of matching. Let D_i be a dichotomous variable indicating whether or not a pair of patents has the same inventor location. Let T_i be the treatment, taking the value 1 if the patent pair are linked by citation and zero if not, and let X_i be a vector of covariates. JTH's evidence consists of showing that the average treatment effect (ATE), $E(D_i T_i = 1) - E(D_i T_i = 0)$, is positive. The ATE is estimated nonparametrically by matching individuals with common values of X_i. In JTH, X_i consists of two variables: the filing date of the patent and the three-digit primary USPTO technology class. The technology class is a proxy for numerous unobserved factors that would cause inventors to locate near inventors they cite, even in the absence of localized knowledge spillovers. In our reassessment we present some examples to argue that the three-digit patent class serves its intended function only imperfectly. Consequently, unobserved differences between the treatment and non-treatment groups persist even after conditioning on X_i, and the estimator of the ATE must be inconsistent (e.g., Guido W. Imbens, 2004). We present some simple calculations showing that this inconsistency can easily produce evidence that spillovers are strongly localized even when they are not. Finally, to show that these issues matter empirically in the data, we construct a sample that replaces the three-digit technology class with alternative covariates based on the patent subclass. Doing so eliminates much of the evidence that intra-national knowledge spillovers are localized.

We do not think our claims that the three-digit patent classification system is a poor proxy, and that this can produce significant spurious localization effects, are contentious.[1] Indeed, Henderson et al. (2005) do not take issue with these claims. Instead, they focus on the limitations of our use of the patent subclass to construct alternative covariates for matching. For a number of reasons, which Henderson et al. explain clearly and which do not need repeating here, our estimates of the ATEs using the patent subclass are unreliable. We had tried in our paper to be explicit about these limitations and, although Henderson et al. greatly expand on the theme, we are in considerable agreement. There are perhaps just two areas where we disagree. First, while Henderson et al. see only sources of bias that cause us to underestimate true localization effects, we see sources of bias that operate in both directions. Second, while Henderson et al. conclude that matching on patent subclass is too unreliable to be worthwhile, we have concluded that matching on *any* patent classification is too unreliable.

Perhaps because of our pessimistic conclusions, Henderson et al. claim that we interpret our findings as casting doubt on any attempt to assess the localization of spillovers from patent citation data. This is going further than we would have liked. Patent citations are our only direct indicator of knowledge flows, and we did not intend to throw the baby out with the bathwater. The challenge is simply to clean the

* Thompson: Department of Economics, Florida International University, Miami, FL 33199 (e-mail: peter. thompson2@fiu.edu); Fox-Kean: Department of Economics, University of Houston, Houston, TX, 77204-5882 (e-mail: mefox@mail.uh.edu).

[1] Readers who remain undecided can undertake a simple smell test for themselves using the U.S. Patent and Trademark Office (USPTO) Web site: choose a handful of patents at random and then follow the JTH recipe for selecting its matched control. Our experience is that a great many of these matched controls do not appear to be doing the job expected of them.

bathwater. Jaffe and Trajtenberg (2002) have made considerable efforts in this regard, and Henderson et al. suggest combining the three-digit class and subclass data. There are other possibilities. For example, the examiner's *Field of Search*, which contains multiple technology classes derived from an assessment of the general thrust of a patent, might serve as a better proxy.[2] But these alternatives represent attempts to ameliorate, not solve, the underlying problem. We remain convinced that the best way forward is to devise identification systems that avoid entirely technology classification systems that were devised for the sole purpose of helping examiners locate prior art. Thompson (2004) proposed one identification scheme and found new evidence that knowledge spillovers are indeed localized.

---

[2] Unfortunately, *Field of Search* is not a searchable field in the USPTO Web site.

REFERENCES

Henderson, Rebecca; Jaffe, Adam B. and Trajtenberg, Manuel. "Patent Citations and the Geography of Knowledge Spillovers. A Reassessment: Comment." *American Economic Review*, 2005, *95*(1) [add page numbers].

Imbens, Guido W. "Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review." *The Review of Economics and Statistics*, 2004, *86*(1), pp. 4–29.

Jaffe, Adam B. and Trajtenberg, Manuel. *Patents, Citations and Innovations: A Window on the Knowledge Economy.* Cambridge, MA: MIT Press, 2002.

Jaffe, Adam B.; Trajtenberg, Manuel and Henderson, Rebecca. "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." *Quarterly Journal of Economics*, 1993, *108*(3), pp. 577–98.

Thompson, Peter. "Patent Citations and the Geography of Knowledge Spillovers. What do Patent Examiners Know?" Unpublished Paper, 2004.