NBER call for proposals on: Economics of Research and Innovation in Agriculture

Proposal title: **"The Roots of Agricultural Innovation: Evidence from Patents"**

Team members: Matthew S. Clancy, Economic Research Service, U.S.D.A.
Email: matthew.clancy@ers.usda.gov

GianCarlo Moschini, Department of Economics, Iowa State University.
Email: moschini@iastate.edu

Paul Heisey, Economic Research Service, U.S.D.A.
Email: pheisey@ers.usda.gov

Motivation

Considerable research has documented sizeable rates of return from past investments in agricultural research and development (R&D), from both public and private sources (e.g., Alston, Norton and Pardey 1995). Despite widespread consensus on this general conclusion, empirical measurement of the link between R&D and agricultural productivity faces serious conceptual challenges and data limitations. One of these has to do with spillovers. The immediate output of R&D is new knowledge—ideas, for short. As noted by Griliches (1992), "It is quite possible for an idea to have its entire effect elsewhere than where it was originated." In our context, new agricultural technologies may depend on ideas sourced from non-agricultural industries, as much as agricultural R&D. A second challenge has to do with the links between basic and applied R&D, and between science and technology. In Vannevar Bush's classical linear model of R&D, fundamental research is later commercialized: "New products, new industries, and more jobs require continuous additions to knowledge of the laws of nature, and the application of that knowledge to practical purposes" (Bush 1945). Yet, the linear model of R&D is frequently criticized as overly simplistic (Brooks 1994). Evidence on the ultimate roots of the knowledge underlying agricultural innovations can improve the positive undertaking of assessing the actual impact of the agricultural sector's own R&D efforts, and also may help with the normative task of designing science policy for agriculture.

Research Objectives

The overarching objective of this proposal is to shed light on the flow of ideas underpinning agricultural innovations, and to characterize the nature of knowledge spillovers in this context. We propose to do so by a novel use of patent data. A patent can be interpreted as the realization of a successful innovation. Information disclosed in the patent can be quite helpful in tracing the ancestry of the ideas that spawned it.

To the best of our knowledge, a large-scale quantitative effort to characterize the sources of the ideas that are inputs to new agricultural technologies has not yet been attempted. Rich new patent data, coupled with data processing technology, makes this objective possible. We propose to use US patent data on 6 different agricultural sub-sectors to characterize the sources of agricultural ideas. Specifically, we will examine the extent to which patents draw on ideas developed in other sectors of the economy, and the broader academic community. We will determine the extent to which agricultural technologies rely on ideas developed by primarily agricultural firms, by firms with some agricultural R&D, and by non-agricultural firms. We will further describe the extent to which agricultural technologies rely on ideas from scientific publication, differentiating agricultural science from other scientific disciplines. Finally, because patents are quite heterogenous in their value (there

are a few very valuable patents and many not-very-valuable patents) we use a number of proxies for the value of patents to link the value of patents and the sources of ideas.

The use of patents as an indicator of innovation has a long history in empirical economics (Nagaoka, Motohashi, and Goto 2010), but has so far found limited use to study agricultural innovations, with a few exceptions (Albers et al. 2016, Pardey et al. 2016). This is partly due to the fact that, historically, most R&D in agriculture was conducted by the public sector (Huffman and Evenson 2006), and the nature of this R&D in the past tended to leave few traces of research outputs in the form of patents. This has changed in the last few decades, due to a combination of factors: the general broadening of the scope of patents to encompassed biological innovations; the 1980 Bayh-Dole act promoting the use of patents by public institutions; and, a drastically increased role for private industry in agricultural R&D (Pray and Fuglie 2015; Clancy and Moschini 2017). As a result, patents now cover an increasingly large share of US agricultural innovations.

The work we propose is pertinent to several topics listed in the NBER call for proposals. We intend to document the differential value of patents and innovations in different agricultural subsectors. We examine the strength and direction of knowledge spillovers across firms, and from sector to sector. Our work also sheds light on the flow of ideas from universities and academia to the private sector, and establishes the extent to which private agricultural R&D relies on (typically publicly funded) agricultural science, as well as other scientific disciplines. In our view, these are fundamental questions for agricultural R&D whose answers have remained elusive.

Data

The proposed analysis relies on seven sets of US patents. We focus particularly on six different groups of agricultural technologies: agricultural machinery, animal health products, biofuel inputs and processes, biocides, plant cultivars and other agricultural technologies. Our seventh category is the set of all US utility patents, which we use as a benchmark.

The biofuels inputs and processes (hereafter, "biofuel") dataset was extensively described by Albers et al. (2016), who generously shared it with us. This dataset is global in nature, but we use the subset of US patents. The plant cultivars (hereafter, "plants") dataset was developed and shared with us by Paul Heisey at USDA Economic Research Service. The animal health product (hereafter, "animal health") dataset was built by Matthew Clancy, also at USDA Economic Research Service, by using archival FDA data on patents associated with veterinary drugs.

The remaining datasets were assembled by using the US Patent and Trademark Office (USPTO)'s cooperative patent classification (CPC) system. The agricultural machinery (hereafter, "ag machines") dataset includes all patents with a primary CPC code associated with agricultural machinery (excluding hand and animal-driven tools). The biocides dataset includes insecticides, pesticides, herbicides, and other biological pest management chemicals (and excludes chemicals used for the preservation of biological material), again, as determined by the primary CPC classification of patents. Finally, the other agricultural technology (hereafter "other ag") dataset is composed of all remaining patents with primary classification A01, but excluding CPC codes associated primarily with hunting, recreational fishing, and forestry.

Some summary statistics are provided in **Table 1**. The first row (Patents) gives the number of patent grants included in each dataset. Among the agricultural datasets, we have the most data for Other Ag and Ag Machines, and the least for Animal Health. Nonetheless, in all cases we have hundreds of patents, and in most cases we have thousands. We focus on the forty years between 1976 and 2016

(the plants database only starts in 1983, however, when patents for plant cultivars became available in the aftermath of the 1980 US Supreme Court decision in Diamond v. Chakrabarty.) Moreover, some datasets end in 2014 and 2015, because this is as far as they had been developed when we acquired them for this project. The date range is for granted patents, but we also possess information on the application year of each patent.

Patents are assigned to a unique person or organization – the assignee – and the Assignees row of Table 1 gives the number of distinct assignees named. These range from 141 for Animal Health, to 6,390 for Other Ag. The size of the collective universe of assignees with at least one patent in our ag datasets is 12,080 (some assignees operate in multiple subsectors).

Table 1. Summary Statistics

	Ag Machine	Animal Health	Biocides	Biofuel	Plants	Other Ag
Patents	19,362	468	12,774	4,087	8,206	26,844
Start Year	1976	1976	1976	1976	1983	1976
End Year	2016	2016	2016	2014	2015	2016
Assignees	2,854	141	2,382	1,496	219	6,390
Fully renewed	38.6%	84.6%	40.7%	36.0%	88.2%	26.2%
KPSS (mean)	8.5	34.5	8.6	15.5	19.6	12.2
Citations received (mean)	1.6	2.3	1.0	3.0	2.5	1.4
Citations made (mean)	13.1	8.2	8.3	16.5	6.9	11.9
Non-patent references (mean)	0.8	6.5	5.6	17.0	8.4	3.5

Next, we report three metrics for the value of patents. Valuing patents is challenging, given the absence of market prices for most patents and each of the three proxies we use has different potential biases. The row "Fully Renewed" gives the share of patents for which patent-holders paid all renewal fees. This is a proxy for the private value of continued patent protection, but it is quite coarse and we can only observe it for patents granted prior to 2005 (because the last payment is due 12 years after the patent issue date). The row "KPSS (mean)" gives the mean Kogan et al. (2017) estimate of the patent value, made from stock market fluctuations around the dates of a patents grant. This second (noisy) estimate of the private value of patents is available in all years, but only for patents belonging to publicly traded firms at the time of their grant. Finally, the row "Citations received (mean)" gives the average number of citations received by patents in the five years after they were granted, for patents granted prior to 2012. Citation receipt has been shown to be correlated with both the private and social value of patents. It is apparent that there is significant heterogeneity in the value of patents across different agricultural subsectors.

The last two rows of Table 1 relate to our primary means of analyzing the source of ideas in agriculture. We use citations to other patents to track knowledge flows across the world of patented technology, and non-patent references to track knowledge flows from academia into the world of patented technology. Because patents make multiple references and citations, we typically have tens of thousands of references to work with. Note again the heterogeneity across subsectors, with the more biologically and chemically based subsectors citing far more non-patent references than the ag machinery and other ag subsectors.

Methods

The proposed work envisions a component based on analyzing citations to patents, and another component analyzing non-patent references. In our patent citation analysis, we are mostly concerned with the nature of the cited patent's assignee. In particular, we would like to see how often patents cite (i) patents owned by the public sector, (ii) patents from firms operating in their sector, (iii) patents from firms operating in agriculture, and (iv) patents from non-agricultural firms.

We use assignee data to define the boundaries of the agricultural R&D world. Organizations with >50% of their total patent stock (1976-2016) in our ag datasets will be designated as primary Ag R&D firms, while those with >0 patents in our ag datasets will be designated as secondary Ag R&D firms. Organizations with zero patents in our ag datasets are designated as non-ag firms.

Turning to non-patent references, we intend to use references to journal articles as a way to track the use of knowledge from academia and, more specifically, distinct disciplines, based on the discipline of the journal. While an audit of several hundred randomly selected references verifies that, in fact, the majority of non-patent references are academic journal articles, it will be desirable to separate journal article citations from other types of reference, and to sort journals into disciplines.

For both components of our analysis we will establish summary statistics about the overall subsector's use of ideas, and econometrically link the use of ideas and the value of patents. For the latter, we will run regressions against each of our three proxies for patent value. Explanatory variables we would like to investigate include the share of all references to academic journals, the share to non-agricultural academic journals, the share of citations to the patents of firms with a secondary agriculture focus, the share of citations to the patents of non-agricultural firms, and the share of citations to publicly owned patents. Control variables (some of which may also be of first-order interest to NBER participants) could include the R&D size of the assignee (measured by number of patents), the agricultural focus of the assignee (measured by the share of its patents in agriculture), the number of citations the patent makes, and the grant or application year. We will run both pooled regressions (which will include dummies for subsectors), and subsector-specific regressions, which will allow for parameter heterogeneity across subsectors.

Preliminary Results

Table 2 presents some preliminary results about the extent to which different agricultural subsectors draw on ideas from outside their subsector and from outside the agricultural sector.

Table 2. Citations to patents granted in 1976-2015 belonging to agricultural assignees

	Total Citations (1976-2015)	Own subector Citation Share	Other Ag Citation Share	Non-Ag Citation Share
Ag Machinery	169,932	63.5%	2.7%	33.9%
Animal Health	3,230	51.2%	22.0%	26.7%
Biocides	88,093	80.6%	4.4%	19.4%
Biofuels	60,420	50.5%	12.8%	36.7%
Plants	56,028	92.6%	5.9%	1.5%
Other Ag	214,915	40.4%	2.5%	57.1%
All Patents	65,367,409	-	-	66.1%

Note: Citations are classified as belonging to: (i) own subsector if the patent assignee has at least one patent belonging to the subsector; (ii) other ag if the patent assignee has at least one patent in another agricultural subsector and the assignee is not already classified as being in the same subsector; and, (iii) non-ag if the patent assignee has no agricultural patents.

Note that our definition of "non-agricultural" patents is extremely coarse: it excludes any patent belonging to a firm with a single patent that we characterize as agricultural (even if this assignee has thousands more non-agricultural patents). Yet, even by this extremely restrictive definition of "non-agricultural", we see a substantial share of patent citations are drawn from non-agricultural firms. A notable exception is plant cultivar patents, for which just 1.5% of citations flow to non-ag firms. For the other firms though, 19-57% of citations come from firms outside the agricultural universe. That said, for every subsector besides "other ag", the largest share of citations comes from firms with at least one patent in the *same* subsector. For biocides and plants, own-subsector firms account for the bulk of all patent citations. These preliminary results suggest non-agricultural R&D has substantial spillovers to non-plant agriculture.

References

Albers, S.C., A.M. Berklund and G.D. Graff. 2016. The Rise and Fall of Innovation in Biofuels. *Nature Biotechnology* 34: 814-821.

Alston, J.M., Norton, G.W., and Pardey, P.G. 1995. *Science under scarcity: principles and practice for agricultural research evaluation and priority setting*. Cornell University Press.

Brooks, H. 1994. The Relationship Between Science and Technology. *Research Policy* 23: 477-486.

Bush, V. 1945. *Science: The Endless Frontier*. Washington, DC: United States Government Printing Office.

Clancy, M.S., and G. Moschini. 2017. Intellectual Property Rights and the Ascent of Proprietary Innovation in Agriculture. *Annual Review of Resource Economics* 9:53-74.

Griliches, Z. 1992. The Search for R&D Spillovers. *The Scandinavian Journal of Economics*, 94, Supplement, pp.S29-S47.

Huffman, W.E., and R.E. Evenson. 2006. *Science for Agriculture: A Long-Term Perspective*. Hoboken, NJ: Blackwell Publishing.

Nagaoka, S., K. Motohashi, and A. Goto. 2010. Patent Statistics as an Innovation Indicator. In the *Handbook of the Economics of Innovation*, eds., B.H. Hall and N. Rosenberg. North Holland Publishing.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman. 2017. Technological Innovation, Resource Allocation, and Growth. *The Quarterly Journal of Economics* 132(2): 665-712.

Pardey, P., B. Koo, J. Drew, J. Horwich, and C. Nottenburg. 2013. The evolving landscape of plant varietal rights in the United States, 1930-2008. *Nature Biotechnology* 31: 25-29.

Pray, C.E., and K.O. Fuglie. 2015. Agricultural Research by the Private Sector. *Annual Review of Resource Economics* 7:399-424.

Conflict of Interests: We have no conflict of interest to report.