Combinations of Technology in US Patents, 1926-2009:

A Weakening Base for Future Innovation?

Matthew S. Clancy[1]

Abstract: In combinatorial models of innovations, new technologies are built from combinations of pre-existing technological components. Researchers learn which components work well together by observing previously successful combinations and the pool of ideas can be "fished out", i.e., exhausted, if it is not "restocked" by the discovery of novel connections. We first show US patents have made increasingly less novel connections among technological constituents since the 1950s, and that the number of technological fields to which these connections are applicable has stopped growing since the 1980s. We then estimate the parameters of an ideas production function, and find parameter estimates consistent with technology fields being fished out if not continually restocked by the discovery of novel connections between technological components. We use the ideas production function to estimate the number of new patent applications induced by each patent granted between 1926 and 2001, and show this number has trended downward since the 1940s.

Keywords: Innovation, Patents, Combinatorial growth, R&D, Novelty

JEL Codes: O30, O31, Y1

[1]US Department of Agriculture, Economic Research Service. The views expressed are those of the author and should not be attributed to the Economic Research Service or USDA.

matthew.clancy@ers.usda.gov

The rate of innovation is pulled in opposite directions by two opposing forces. As has been long observed, knowledge is cumulative: the more we know, the more areas of research are opened for exploration. On the other hand, the number of ideas may be finite and exhaustible. In this article, WE refer to the first effect as the restocking effect and the second as the fishing out effect. When the pool of ideas is restocked faster than it is fished out, the set of new ideas expands and finding ideas expected to be worth their R&D cost is easier. When the fishing out effect outpaces the restocking effect, the set of new ideas contracts and it becomes harder to find ideas worth pursuing. This paper uses a combinatorial framework to highlight stylized facts about innovation, as revealed by US patent data, and to assess which of these opposing effects has dominated over 1926-2009.

Combinatorial models of discovery start with the observation that all ideas and technologies can be broken down into constituent elements. Take the internal combustion engine as a representative example. While it is a single idea, it can also be viewed as a combination of pistons, crankshafts, flywheels, and so on. Each of these elements existed prior to the engine, and its discovery was about assembling pre-existing constituent elements into a combination not previously known. This perspective on the innovation process has a long history. Abbott Payson Usher's A History of Mechanical Inventions (1929) notes "Invention finds its distinctive feature in the constructive assimilation of preexisting elements into new syntheses, new patterns, or new configurations of behavior" (Usher, 1929, pg. 11). More recently, this perspective has been articulated in economics by Weitzman (1998), Olsson and Frey (2002), Olsson (2005), Arthur (2009), Feinstein (2011), Ghiglino (2012), and Akcigit, Kerr and Nicholas (2013).

A straightforward interpretation of "fishing out" follows from combinatorial models of innovation. If innovations are combinations drawn from a fixed set of elements, then the number of such combinations is finite and each idea draws down the stock of potential future innovations (i.e., there are only so many ways to combine pistons, crankshafts, flywheels, and so on to obtain an engine). Weitzman (1998) and Akcigit, Kerr, and Nicholas (2013) both incorporate versions of fishing out. Combinatorial models can also model the cumulative nature of knowledge in various ways.[1] Most relevant for this paper is Arthur (2009)'s concept of 'clumps,' in which some components (such as pistons and crankshafts) are understood to go together naturally because they "repeatedly form subparts of useful combinations" (Arthur 2009, pg. 70).

We take the set of technologies available for recombination as exogenously given at any time, and direct our attention to how these components are recycled and recombined. In particular, our interpretation is that researchers learn how any two pre-existing technologies can be fruitfully combined from the example of previously successful combinations. For example, inventors of the internal combustion engine knew how crankshafts and pistons interact from their previous use together (even in otherwise unrelated inventions, such as a waterwheel).

This framework for thinking about discovery can be mapped to patents by using the US Patent and Trademark office's (USPTO) US Patent Classification System (USPCS). The USPCS was developed to organize patent and other technical documents by common subject matter. Subject matter can be divided into a major component, called a class, and a minor component, called a subclass. Whereas "a class generally delineates one technology from another," subclasses "delineate processes, structural features, and functional features of the subject matter

---

[1] E.g., as arising from the discovery of new components (Weitzman 1998, Akcigit Kerr and Nicholas), or the discovery of new combinations that can be repeated to diminishing effect (Akcigit, Kerr and Nicholas 2013), or the discovery of ideas that bridge distant technological spaces (Olsson 2005 and Feinstein 2011) and reveal the quality of "nearby" ideas (Jovanovic and Robb 1990, Kauffman, Lobo and Macready 2000, Auerswald et al. 2000). Of course, it is also possible to model the cumulative nature of innovation other ways (Scotchmer 2004 provides an entry-level survey). See Galasso and Schankerman (2015) for a recent example.

encompassed within the scope of a class" (USPTO 2012). Subclasses are a natural candidate for the building blocks of combination, out of which researchers build new ideas, and classes are a natural candidate for distinct technological fields. We use the patent classification system to proxy for the technological elements being recycled and recombined across different fields, and to assess whether the restocking or fishing out effect dominates.

The paper proceeds in five steps. Each section builds on the proceeding to introduce a new stylized fact about innovation in the US over the 20th century. Section 1 introduces our patent data and our measure of patent novelty. In this section we document the decline of novel connections between technological components since the 1950s. Section 2 introduces our measure of technological fields and documents that the number of technological fields for which a patent's connections are relevant has plateaued since the 1980s. Sections 3 uses these metrics to estimate field-specific knowledge production functions. We find parameter estimates consistent with technology fields being fished out if not continually restocked by novel connections. In Section 4, we use this model to estimate the number of new patent applications induced by each patent granted between 1926 and 2001, and show this number has trended downward since the 1940s. Section 5 examines these patent impacts by field. The remainder of the paper performs robustness checks for the model estimated in section 3, discusses the paper's limitations, and relates the findings to other literature on science and technology in the 20th century.

The picture that will emerge is of a weakening base for future innovation, albeit with important caveats discussed in the conclusion. Consistent with prior work in this vein (see Clancy 2017), we find support for the general framework that innovation slows if it is not periodically replenished with novel connections. The extent to which the US private sector is generating novel connections with wide applicability across many fields, however, has been declining for a decades now. Indeed, in some of our model estimates, we find evidence that each patent now fishes out more ideas than it restocks. That said, this paper merely documents these trends, and does not attempt to determine whether current R&D strategies are socially suboptimal.

This work is most closely connected with other empirical papers that study combinatorial features of innovation. Akcigit, Kerr, and Nicholas (2013) and Clancy (2017) use the same dataset to explore combinatorial features of innovation. There is also a large literature that relies on the citations made by patents and academic papers to study combinatorial innovation[2] and spillovers.[3] This paper differs from the preceding in two ways. First, it focuses on the productivity of entire technological fields, and attempts to quantify whether the fishing out or restocking effect has been dominate in the aggregate. Second, it documents long-term changes in the conduct of R&D that are only visible from a combinatorial perspective, but which complement alternative perspectives (discussed in the conclusion).

1. Patent Data and Novelty

My data draws on the full set of US utility patents granted between 1836 and 2009 (7.6 million patents), but we focus on 1926-2009, the period for which patent application dates are available (figures going back to 1836 for data that does not depend on application dates are available in the supplemental materials). There are more than 450 classes and more than 150,000

[2] Fleming (2001), Keijl et. al (2016), Nemet (2012), Nemet and Johnson (2012), Schilling (2011), and Shoenmakers (2010) explore how combinatorial features impact citations to patents and academic journal articles.
[3] See Feldman and Kogler (2010) for a survey, and Acemoglu, Akcigit, and Kerr (2016) and Agrawal, Galasso and Oettl (2017) for more recent examples.

subclasses in the USPCS. To take two examples, class 014 corresponds to “bridges,” and class 706 corresponds to “data processing (artificial intelligence).” The subclasses are nested within each class and correspond to more fine-grained technological characteristics. The uppermost subclass is called a mainline subclass, hereafter “**mainline**.” For example, the subclasses 014/3 (bridge; truss) and 706/15 (data processing (artificial intelligence); neural network) are both mainlines. The subclass nested one level down is said to be “one indent” in from the mainline. Within these one-indent subclasses may be still further subclasses, called two indent subclasses, and so on, each delineating a more specific technological component.

Both Fleming (2001) and Akcigit, Kerr, and Nicholas (2013) use the raw subclasses assigned to patents as proxies for the technological elements combined in a patent. However, one drawback to this approach is that subclasses correspond to different levels of specificity. For example, subclass 706/29 is indented in from subclass 706/15, and both relate to neural networks (but at different levels of specificity). In contrast, neither subclass is indented in from 706/45, which is not associated with neural networks at all. Without looking at the USPC index, it is impossible to know there is a relationship between some of the subclasses, but not others. To obtain a consistent, exhaustive, non-overlapping set of technological components, we use technology mainlines as our primary elements of combination as in Aharonson and Schiling (2016) and Clancy (2017). This identifies a set comprising 13,517 technological elements that are recycled and combined in patents.

We observe the subclasses assigned to every patent and collapse each technology subclass down to the mainline to which it belongs. Not all patents have multiple mainlines assigned to them, but most do.[4] Over our sample of 1926-2009, 68% of patents were assigned more than one mainline, with a mean of 2.5 mainlines per patent. Out of 91.3 million possible pairs of mainlines, 1.75 million pairs are actually assigned to at least one patent over the period 1926-2009. Viewed through a combinatorial innovation lens, of the 91.3 million possible pairs of technological components, researchers have discovered how to usefully combine only 2%. Over the same period, the mean number of patents each pair belongs to over the entire period is 10.1, but the distribution is highly skewed: 51.2% of observed pairs are only ever assigned to one patent, but 46.1% of all pair assignments belong to 1% of pairs.

Two dates are relevant for each patent. The application date marks when the patent application was filed, and the patent grant date marks when the USPTO granted the patent. Patents are sequentially numbered as they are granted, so that the year any patent is granted can be inferred from the patent number (USPTO 2014a). However, only the application years for US patent granted between 1926-2009 are available (Kogan et al. 2017). Moreover, the number of patents for which application years are available begins to rapidly decline after 2001 due to a truncation effect (we only observe patent applications if they are granted by 2010). Thus, although we use patent data from 1836 to construct measures of researcher knowledge, we only examine patenting behavior for the period 1926-2001. There were 4.75 million patents granted during this period.

In this paper, learning underpins the restocking effect. This paper introduces an empirically tractable measure of what researchers learn from other inventions derived from the

[4] Because mainlines were not designed to track the underlying elements of technological combination, there is no guarantee that every patent will be assigned multiple mainlines. I discuss this limitation in section 7.1.

extent to which a particular pair of elements has successfully been combined in the past. To measure researcher knowledge about how to combine different mainlines, define the *affinity* of a pair of mainlines, where a pair has high affinity if researchers generally know how to integrate the two mainlines to do something useful. Every time a patent successfully combines two mainlines, researchers' assessment of the pair's affinity is increased. To compute a proxy for affinity, let n_{jt} be the number of patents granted in year t or earlier that are assigned both mainlines in the pair indexed by j. The affinity of pair j at time t is defined to be:

$$a_{jt} = \frac{n_{jt}}{50 + n_{jt}} \tag{1}$$

The functional form of our measure of affinity is derived from a logit regression of the form u_{jt} = logit($\theta + \beta$log n_{jt}), where u_{jt} is a dummy variable equal to 1 if the pair j is assigned to any patent application in year t. Using the definition of the logit function, this can be rewritten as:

$$E\left[u_{jt}\right] = \frac{n_{jt}^{\beta}}{e^{-\theta} + n_{jt}^{\beta}} \tag{2}$$

To select θ we run a logit regression on a sample of 10,000 randomly selected mainline-pairs used at least once over 1926-2009. For each pair, we compute n_{jt} for each year where $n_{jt} > 0$. Estimating equation (2) we obtain $e^{-\theta} = 50.01$ and $\beta = 0.96$, which we approximate with equation (1). Consistent with related work (Clancy 2017), equation (1) embeds the assumption that each successful instance of pair j being used in a patent is less informative than the previous one. For large n_{jt}, additional instances of the elements being integrated successfully tell researchers nothing new, because they are already certain this can be done.

I use *affinity* to measure of the extent of novel connections between mainlines made by a patent. For reasons that will become clear in section 3, define the metric of patent novelty as:

$$\Delta AFF_{it}(p) = \sum_{j \in p} \left\{ \frac{n_{jt}}{50 + n_{jt}} - \frac{n_{jt} - 1}{50 + n_{jt} - 1} \right\} \quad (3)$$

Where $j \in p$ indicates that pair j connects two mainlines assigned to patent p and the subscript i indicates the field to which patent p belongs (discussed in the next section).

This measure is largest when n_{jt} is small (i.e., the patent makes novel connections) and when a patent combines more mainlines. Note $\Delta AFF_{it}(p) = 0$ if a patent is assigned less than two mainlines. Figure 1 illustrates the average value of $\Delta AFF_{it}(p)$ for patents granted in year t, as well as the average number of mainlines assigned to patents granted in the same year.

The average $\Delta AFF_{it}(p)$ of a patent grant rises from under 0.02 in 1926 to a peak of 0.033 in 1947, an increase partially attributable to a shift towards more complex patents over this period (from 1.85 to 2.43 mainlines per patent). But thereafter the number of mainlines per patent increases from 2.43 to a peak of 2.65 between 1947 and 1988, while the average $\Delta AFF_{it}(p)$ falls from 0.033 to 0.026. The number of mainlines per patent decreased from 2.65 to 2.24 by 2009, i.e., to levels seen in 1936, while the average $\Delta AFF_{it}(p)$ accelerated its declines, reaching nearly 0.01, well below its 1936 level. When we restrict the sample to patents with the same number of mainlines, there is a steady downward trend from the start, implying researchers have been shifting away from more novel combinations of technologies throughout the period.

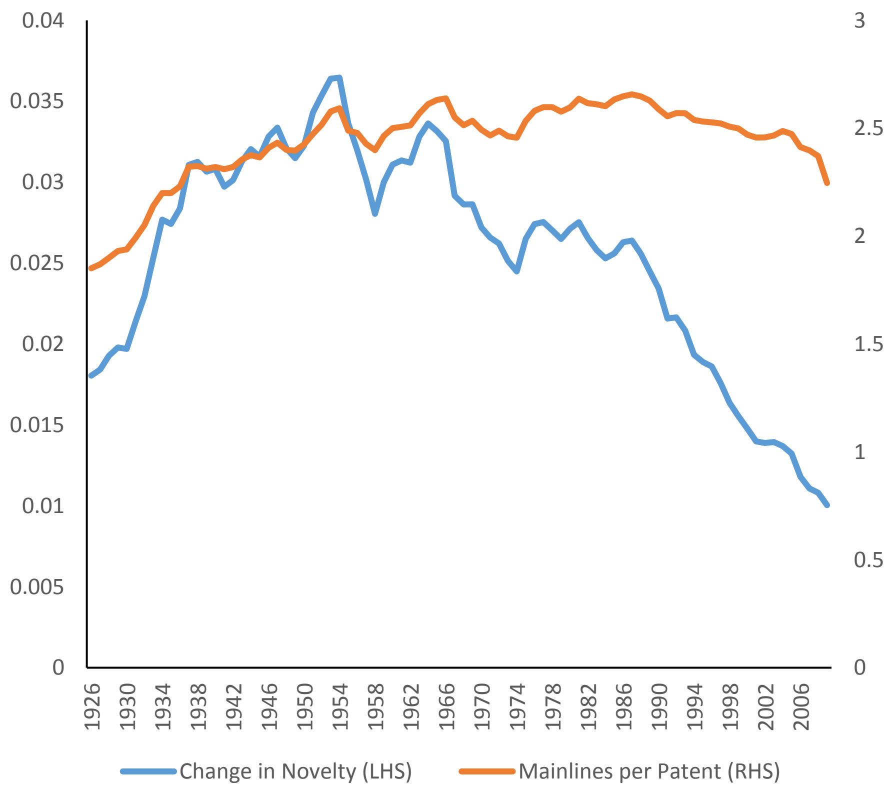

Figure 1. Average Novelty per Patent and Mainlines per Patent (1926-2009)

2. Patent Fields and Spillovers

For each patent, one and only one of the assigned subclasses is denoted the primary classification, which is based on the main inventive concept. We use primary classifications to group patents together into technological fields. This results in 428 technology fields after we exclude classes with no patents assigned to them.

Researchers in other fields also benefit from learning that technological components can be fruitfully combined. Note that a_{jt} is a function of the *total* number of patents combining the components of pair j, not merely the number of such patents in the same technological field. All else equal, learning a pair of mainlines can be combined has a bigger impact if the mainline-pair is used by many different technological fields. We define a spillover from patent *p* to be a technology field other than patent *p*'s primary classification that contains patents using at least one mainline-pair assigned to patent *p*. For example, if patent *p* is granted in 1990, belongs to field 014 and uses mainlines 014/3 and 706/15, then it has a spillover to any other field that also contains patents (granted prior to or in 1990) using 014/3 and 706/15.

The average number of spillovers per patent, as well as the number of mainlines per patent, is presented in Figure 2.

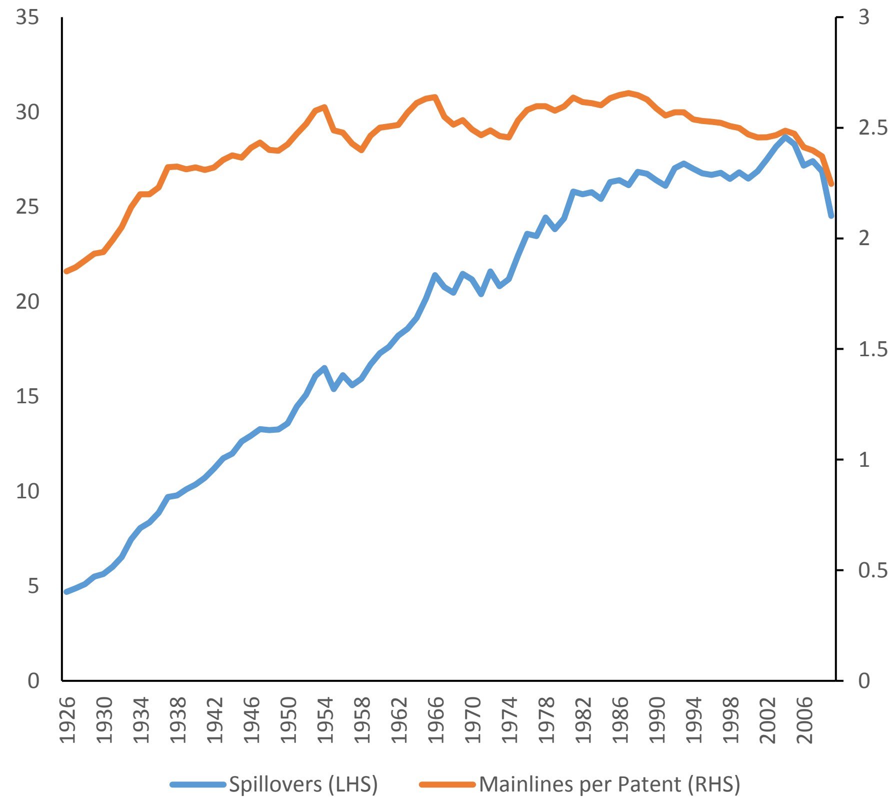

Figure 2. Spillover Fields per Patent and Mainlines per Patent (1926-2009)

The number of spillover fields steadily increases at a rate of 0.35 per year through 1986, from just under 5 in 1926 to a bit over 26 in 1986. Thereafter, the number of spillover fields plateaus, oscillating between 24 and 29 from 1984 to 2009. This may be connected to the slight decrease in the number of mainlines per patent that we observe after 1987.

3. Estimating Field Specific Knowledge Functions

To quantify the impact of the trends highlighted in Figures 1 and 2, we estimate an ideas production function. These functions links metrics of field-specific researcher "knowledge stocks" to the production of new ideas (Porter and Stern 2000, Popp 2002, Zucker et al. 2007, Madsen 2008). In this paper, we use the following functional form:

$$y_{it} = A_{it}\left(AFF_{it}\right)^{\alpha}\left(PAT_{it}\right)^{\beta} \tag{4}$$

where y_{it} signifies patent applications in field i in year t, A_{it} is a field/year specific scalar and AFF_{it} and PAT_{it} signify field-specific knowledge stocks.

3.1 *Constructing Knowledge Stocks*

Let PAT_{it} signify the patent stock, a count of all patents with field i as their primary classification and granted up through year t. AFF_{it} signifies the affinity stock, given as the sum of the affinities of all pairs belonging to patents assigned to a field:

$$AFF_{it} = \sum_{j \in i(t)} a_{jt} \tag{5}$$

where $i(t)$ indicates the set of pairs used by patents belonging to field i granted in or prior to year t. We can now also observe that the measure of novelty developed in section 1, $\Delta AFF_{it}(p)$, is equivalent to the increase in a field's affinity stock that can be attributed to a single patent.

Constructing AFF_{it} requires determining which mainline-pairs form the knowledge base to which researchers in the field have access. Clearly any mainline-pair j that is assigned to a patent in field i belongs to field i itself. Moreover, it is not important to assign never-used mainline pairs to any field because they have $a_{jt} = 0$ and so their inclusion or exclusion has no impact on AFF_{it}.

However, suppose there is a mainline-pair *j'* that has a history of use in several fields, so that $a_{j't} > 0$. In period *t* a patent in field *i'* is assigned *j'*, but no other patent in *i'* has ever been assigned *j'*. Was *j'* always part of field *i'* knowledge, latent and waiting to be used? If so, $AFF_{i't}$ should include $a_{j't}$ in periods prior to *t*. Alternatively, it may be that prior to period *t*, no researchers recognized the relevance of pair *j'* for field *i'* applications. If this is the case, then $AFF_{i't}$ should not include $a_{j't}$ in periods prior to *t*.

For estimation reasons, it is not a good idea to assume $a_{j't}$ was always part of field *i'* knowledge. Simply including all pairs that are ever used by a field will introduce bias into our estimates, because fields that are more productive in the future will tend to accumulate more mainline-pairs. Including these pairs in the calculation of AFF_{it} for early periods will skew AFF_{it} upwards for fields that will be more productive later on, introducing a misleading correlation between AFF_{it} and future patent applications. Instead, we base our definition of AFF_{it} on information available only in period *t*.

To begin, we assume any mainline-pairs assigned to patents with a primary classification in field *i* and granted through period *t* form the set of mainline-pairs necessary to construct AFF_{it}. Thus, AFF_{it} will grow to include more pairs over time as patents using new mainline-pairs are granted. However, in our estimation strategy, we use the first difference of each observation. This allows me to hold constant the set of mainline-pairs used in a given field for a field. That is, we can use as an observation:

$$\Delta AFF_{it} = \sum_{j \in i(t)} \left\{ a_{j,t} - a_{j,t-1} \right\} \tag{6}$$

Note that in equation (6) the set of pairs are those used through period *t* and we measure only the changes in their affinity (in the robustness checks, we show that we obtain qualitatively similar results if we use the set of pairs used through period *t* – 1). This measure corresponds to the interpretation that the first time a mainline-pair is used by a field, others in the field were also aware of its potential utility for applications in the field. Note that in all cases, we define the set of pairs available in period *t* by the set of patents *granted* through period *t*, rather than the set of patent *applications* through period *t*. Using patent applications to define the set of mainline-pairs in use would introduce a simultaneity problem wherein AFF_{it} would tend to jump in any year with many patent applications using novel mainline-pairs.

Histograms for all major variables are available in the supplemental materials.

3.2 *Estimation Methodology*

To estimate equation (4), we begin by taking logs:

$$\log y_{it} = \log A_{it} + \alpha \cdot \log AFF_{it} + \beta \cdot \log PAT_{it} \qquad (7)$$

Where $\log A_{it}$ is represented by a set of field and year fixed effects, and an idiosyncratic error term. Each observation is a technology field-year and we only include years that a field is 'active' (that is, every year after the first patent in it is granted). This gives me an unbalanced panel over 428 fields.

Estimating equation (4) in this way requires that the count nature of patent applications can be suitably approximated as continuous. This is true for the vast majority of observations (86.4% of observations are greater than 10). The advantage of a log-transformation over an alternative model specifically tuned to count data (such as a poisson or negative binomial model) is that we can better control for serial correlation and omitted variable bias. The dependent variable in equation (7) is trending over time, and we use a panel augmented Dickey-Fuller test, correcting for cross correlation among panels (Constantini and Lupi 2012) to test for unit roots. As indicated in Table 1, we fail to reject the null hypothesis of a unit root in levels and first-differenced data when we use the entire sample (1926-2009).

However, when we repeat the tests for 1926-2001, we reject the null hypothesis for the first difference. Data from 2002-2009 is severely impacted by truncation bias; the annual number of patent applications plummets from over 200,000 per year to nearly zero by 2009. This drop in patenting is illusory, and reflects the increasingly scarce number of patents applications that are granted by 2010, the year our patent application data was collected. We drop these observations, and so our estimates are based on first-differenced logged observations from 1926-2001 (in our robustness checks, we include estimates based on 1926-2009). When differencing AFF_{it} we use the set of mainline-pairs assigned to patents granted up through year t, so that the change in AFF_{it} is entirely due to changes in a_{jt}, not changes in which mainline-pairs belong to a field.

A potential problem in our dataset stems from survivorship bias. Each panel corresponds to a technological field, but these fields are labels invented by the USPTO to help organize their search for relevant technology. The USPTO is unlikely to invent a new label for a field to which only a very small number of patents are ever attributed – these patents are likely to be assigned instead to the closest existing technology field. For this reason, very young technology fields are likely to exhibit sustained growth, relative to the explanatory variables, because if they did not they would not have been broken out as new fields by the USPTO. To account for this possibility, we include a dummy and interaction term for "young fields." To define a young field, we note the smallest value of PAT_{it} for the year 2009 (when each field's patent stock is at its maximum) is 5, but this field is a significant outlier, and the next smallest field has PAT_{it} = 179.

We define a young field to have a patent stock of half this value (89.5). By this definition, 1,440 observations (4.1%) are classified as belonging to young fields. While the inclusion or exclusion of a young-field interaction term is important for our results (see the robustness checks for results excluding dummies for young fields), the exact definition of a young field is not. For example, we obtain qualitatively similar results when we define young fields as any with a patent stock below 1,000 (see robustness checks).

I supplement the differencing of data with two-way fixed effects to help control for field and time specific factors in the differenced data. As indicated in Table 1, a lagrange multiplier test fails to reject the null that these fixed effects are not necessary (Croissant and Millo 2008). The equation we estimate in our preferred specifications is:

$$\begin{aligned}\Delta \log y_{it} = \quad & \gamma_i + \psi_t + \alpha \cdot \Delta \log AFF_{it} + \beta \cdot \Delta \log PAT_{it} + e_{it} \\ & + \tilde{\theta} \cdot YF_{it} + \tilde{\alpha} \cdot YF_{it} \cdot \Delta \log AFF_{it} + \tilde{\beta} \cdot YF_{it} \cdot \Delta \log PAT_{it}\end{aligned} \tag{8}$$

Where γ_i and ψ_t are field and year-specific factors and YF_{it} is a dummy variable equal to 1 when the observation corresponds to a young field (PAT_{it} < 89.5). Lastly, a Breusch-Godfrey test fails to reject the null of serial correlation in the idiosyncratic error term e_{it}, so we use White robust standard errors clustered by field.

Table 1. Test Statistics

Test	Null hypothesis	Test-Statistic	p-Value
Panel augmented Dickey-Fuller test, correcting for cross correlation among panels: $\log(y_{it} + 1)$ for 1926-2009	The data series contains a unit root	7.043	1.000
Panel augmented Dickey-Fuller test, correcting for cross correlation among panels: $\Delta\log(y_{it} + 1)$ for 1926-2009	The first-difference data series contains a unit root	6.209	1.000
Panel augmented Dickey-Fuller test, correcting for cross correlation among panels: $\log(y_{it} + 1)$ for 1926-2001	The data series contains a unit root	-1.489	0.068
Panel augmented Dickey-Fuller test, correcting for cross correlation among panels: $\Delta\log(y_{it} + 1)$ for 1926-2001	The first-difference data series contains a unit root	-6.409	0.000
Lagrance Multiplier test	Two-way fixed effects are unnecessary	138.14	0.000
Breusch-Godfrey/Wooldridge test for serial correlation in panel models	No serial correlation idiosyncratic errors	3,417.5	0.000

3.3 *Results*

The parameters α and β are estimated, and once we account for issues of serial correlation, $\alpha > 0$ and $\beta < 0$. We interpret α as measuring the "restocking" effect, as it measures how increases in the affinity stock are translated into the flow of new ideas, and β as the "fishing out" effect, as it measures the rate at which a field's ideas are exhausted. Note, unlike traditional ideas production functions, we do not need to include measures of knowledge developed in rival fields, because AFF_{it} already embeds knowledge spillovers from other fields. Neither do we need to determine the knowledge depreciation rate, as the exhaustion of knowledge is instead captured directly by the patent stock.

I report three estimates in Table 2. Because our estimation methodology uses logs, in specifications 1 and 2 we omit observations with 0 patent applications, which accounts for nearly 5% of our sample. In specification 3, we restore these missing observations by replacing the dependent variable with $\Delta\log(1 + y_{it})$. Another potential issue with our data is the change in patenting trends. Over 1926-1983, patent applications per annum remained in a range between 25,000 and 75,000. Beginning in 1984, patent applications per year began to rise rapidly, peaking at 205,338 in 2001. It is unclear if this acceleration in annual applications reflects an increase in innovations (perhaps because firms shifted away from pure science and towards commercially viable R&D), or merely an extension of patenting to innovations that previously would not have been patented (see Lerner and Kortum 1998, and Gallini 2002 for discussion). To assess the significance of this change, specification 2 relies only on data from 1926-1983.

Estimates for α are positive: a 1% increase in a field's affinity stock is correlated with a 0.75-0.9% increase in the annual number of patent applications. That is, a given field applies for more patents if the field has more knowledge over how to combine pairs of technological elements, or, in the parlance of our first-differenced approach, there is a jump in patent applications in the same year that there is a jump in the affinity of pairs the field uses. At the same time, a 1% increase in the patent stock is correlated with a 0.2-0.6% decrease in the annual number of patent applications.

Table 2. Ideas production function
Time and field fixed effects are included in each regression. All variables are log-transformed and first differenced. White robust standard errors, clustered by group, are in parentheses.

*p<0.1; **p<0.05; ***p<0.01

Explanatory Variable	(1)	(2)	(3)
Affinity Stock : AFF_{it}	0.901***	0.897**	0.757***
	(0.274)	(0.357)	(0.237)
Patent Stock: PAT_{it}	-0.303***	-0.620***	-0.232***
	(0.099)	(0.152)	(0.080)
Young Field: YF_{it}	0.102***	0.070**	0.044**
	(0.027)	(0.031)	(0.018)
$AFF_{it} \times YF_{it}$	-0.850**	-0.735*	-0.705**
	(0.336)	(0.422)	(0.344)
$PAT_{it} \times YF_{it}$	0.133	0.388**	0.140
	(0.152)	(0.195)	(0.138)
Dependent Variable	y_{it}	y_{it}	$1+y_{it}$
Years	1926-2001	1926-1983	1926-2001
Observations	30,069	22,637	31,226

4. Average Patent Impact

Let Y_t be the number of patent applications across all technology fields in year t. Given estimates for α and β from the preceding section, we can calculate the marginal effect of a single patent grant on Y_t via the patent's impact on knowledge stocks. For each patent in our dataset, define $AFF'_{it}(p)$ and $PAT'_{it}(p)$ to be the value of AFF_{it} and PAT_{it} if patent p did not exist. The patent stock $PAT'_{it}(p)$ is the simpler measure, as it is equal to $PAT_{it}(p)$ - 1 if patent p belongs to field i and PAT_{it} otherwise. To compute $AFF'_{it}(p)$, for every pair of mainlines contained in patent p and field i, we swap the old value of a_{jt} for a new value, which we obtain by reducing n_{jt} by one. The affinity stock for multiple fields may be impacted in this way, because patents typically have spillovers to many fields.

Let $y'_{it}(p)$ be the expected number of patent applications in field i in year t if patent p did not exist, so that $y'_{it}(p)$ is the left-hand side variable of equation (4) if $AFF'_{it}(p)$ and $PAT'_{it}(p)$ take the place of AFF_{it} and PAT_{it}. If we use equation (4) to solve for A_{it} in terms of y_{it}, AFF_{it}, and $PAT'_{it}(p)$, then we can write:

$$y'_{it}(p) = y_{it}\left(\frac{PAT'_{it}(p)}{PAT_{it}}\right)^{\alpha}\left(\frac{AFF'_{it}(p)}{AFF_{it}}\right)^{\beta} \tag{9}$$

The net impact of a patent, across all fields is defined to be:

$$Y_t - Y'_t(p) = \sum_i y_{it}\left\{1 - \left(\frac{PAT'_{it}(p)}{PAT_{it}}\right)^{\alpha}\left(\frac{AFF'_{it}(p)}{AFF_{it}}\right)^{\beta}\right\} \tag{10}$$

Note that a patent has a large impact if its mainline-pairs are used by many fields, if it makes novel connections, and if the existing Affinity and Patent stocks are small. Figure 3 plots the average net impact of a patent using the three parameter estimates presented in Table 2.

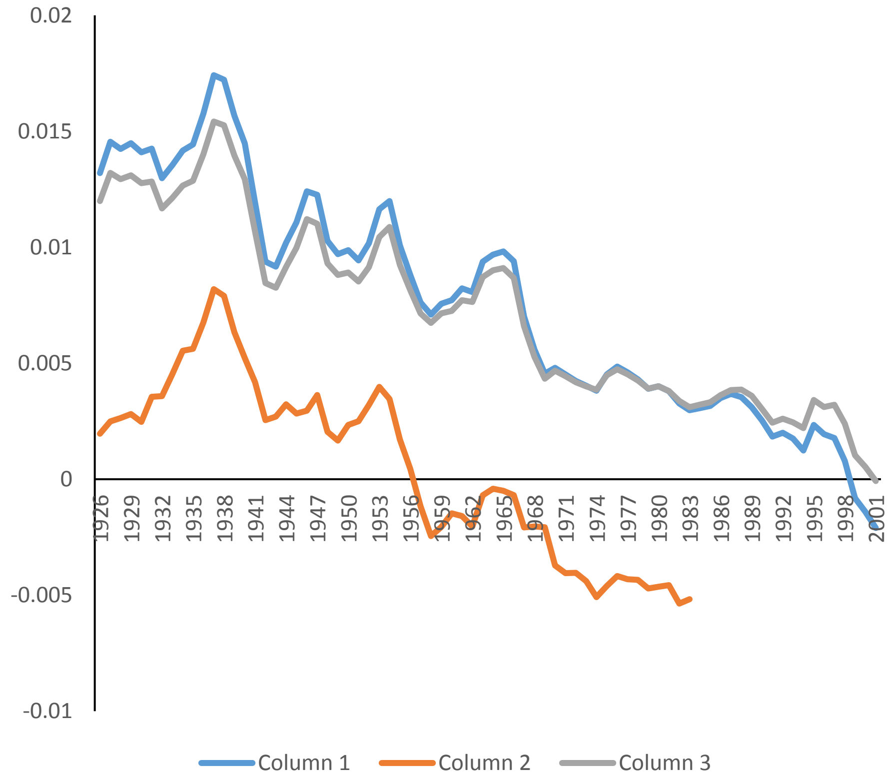

Figure 3. New Patent Applications Induced by Patent Grant (1926-2001)

The impact of a patent on future applications is initially positive, but trending downward across all three specifications. In 1926, every patent grant increases the number of new patent applications by an average of 0.002-0.013 per year (1 every 77-500 years). By the end of our sample, additional patent grants had a slightly negative impact. More substantively, in our 1926-1983 sample (Table 2, column 2), the net impact of patents is negative after 1956. For the period 1957-1983, the fishing out effect dominates the restocking effect.

5. High Impact Fields

To get a better sense of the fields whose patents have the most impact on subsequent innovation, Figure 4 charts the shares of broad NBER technology categories (as defined in Hall, Jaffe, and Trajtenberg 2001) among the top 5% of highest impact patent grants.

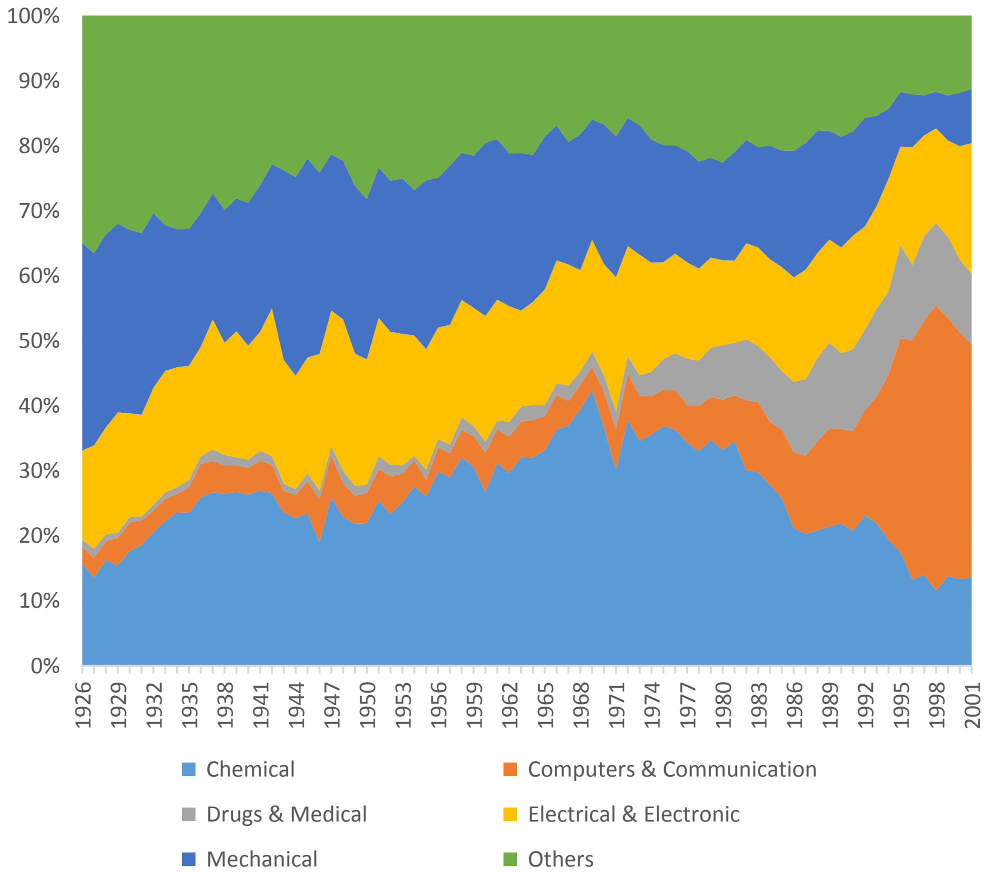

Figure 4. Top 5% Highest Impact Patents, Share by Field (1926-2001)

From 1926-1940, the Other category has the largest share. Thereafter, the Chemical category mostly dominates (occasionally handing off to the Mechanical or Other categories over 1943-1953) until 1994, at which point Computers and Communication take the lead through the end of 2001. However, if we combine the Electrical & Electronic category and Mechanical categories, this is the dominant category over 1926-1967 and 1984-1994. Broadly speaking, over our sample, the technologies where each patent most restocked the pool of ideas were electricity and machines at the beginning, chemistry in the middle, and computers at the end.

6. Robustness Checks

Table 3 presents a series of robustness checks, demonstrating the dependence of the results on the inclusion of a young field effect and differencing the data while holding the pairs used by a field constant. Columns 1, 3, and 5 omit the young field interactions. Columns 1 and 2 do not take the first difference of the data and therefore do not correct for unit roots in the data. Note the estimated parameter on the patent stock is positive when we do not difference the data. Columns 3 and 4 do not hold the set of pairs defining the affinity stock constant. This corresponds to the interpretation that the first time a mainline-pair is used by a field, it alerts the rest of the research community to the previously unknown utility of the mainline-pair. While the signs of α and β are in the correct direction, their magnitudes are attenuated and the coefficient on the affinity stock is not statistically distinguishable from zero in the absence of a young field dummy variable. Columns 5 and 6 hold the set of mainline-pairs used in a field constant, as in the main body of the paper. Note that, without the young-field effect, our estimates of α and β are not statistically distinguishable from zero. Column 6 is identical to Column 1 of Table 2.

Table 3. Robustness Checks 1

Notes: Time and field fixed effects are included in each regression. All variables are log-transformed. White robust standard errors, clustered by field, are in parentheses.
*p<0.1; **p<0.05; ***p<0.01

	Dependent Variable: Patent Applications					
	(1)	(2)	(3)	(4)	(5)	(6)
Affinity Stock: AFF_{it}	0.213*** (0.068)	0.370*** (0.080)	0.052 (0.072)	0.205** (0.104)	0.283 (0.209)	0.901*** (0.274)
Patent Stock: PAT_{it}	0.641*** (0.063)	0.597*** (0.071)	-0.130* (0.074)	-0.215** (0.098)	-0.157 (0.105)	-0.303*** (0.099)
Young Field: YF_{it}		1.219*** (0.310)		0.094*** (0.028)		0.102*** (0.027)
$AFF_{it} \times YF_{it}$		-0.243** (0.099)		-0.196 (0.134)		-0.850** (0.336)
$PAT_{it} \times YF_{it}$		-0.149 (0.122)		0.048 (0.127)		0.133 (0.152)
Differenced?	No	No	Yes	Yes	Yes	Yes
Affinity Stock Measure	AFF_{it}	AFF_{it}	ΔAFF_{it}	ΔAFF_{it}	$\sum_{j \in i(t)} \Delta a_{jt}$	$\sum_{j \in i(t)} \Delta a_{jt}$
Observations	30,756	30,756	30,069	30,069	30,069	30,069

Table 4 keeps the young-field dummy variable and interaction terms, as in section 3, but modifies measures of the knowledge stock and the data sample. In column 1, we replace the affinity stock with a count of the number of unique mainline-pairs cumulatively used by a field's patents. In the first difference regression, each observation is the number of mainline-pairs not previously used by that field. Our estimate of α using this alternative measure of researcher's knowledge about how to combine technological components in a field is not significantly different from zero. This suggests an important role for measuring the depth of knowledge about how to combine technological components, which we capture with a pair's affinity.

Table 4. Robustness Checks 2

Notes: Time and field fixed effects are included in each regression. All variables are log-transformed and first differenced. White robust standard errors, clustered by group, are in parentheses.
*p<0.1; **p<0.05; ***p<0.01

	Dependent Variable: Patent Applications				
	(1)	(2)	(3)	(4)	(5)
Affinity Stock: AFF_{it}	0.115	0.215	1.029***	1.043***	0.672**
	(0.106)	(0.160)	(0.278)	(0.300)	(0.316)
Patent Stock: PAT_{it}	-0.134	-0.139***	-0.290***	-0.273***	-0.485***
	(0.092)	(0.038)	(0.098)	(0.094)	(0.137)
Young Field: YF_{it}	0.091***	0.075***	0.128***	0.123***	0.056***
	(0.028)	(0.025)	(0.026)	(0.031)	(0.013)
$AFF_{it} \times YF_{it}$	-0.030	-0.246	-0.944***	-1.615***	-0.513
	(0.157)	(0.214)	(0.339)	(0.621)	(0.349)
$PAT_{it} \times YF_{it}$	-0.092	0.126	0.127	0.153	0.321**
	(0.137)	(0.077)	(0.158)	(0.141)	(0.148)
Observations	30,069	30,069	33,024	30,069	30,069
How model differs from Table 1, Column 1	ΔAFF_{it} replaced by $\Delta PAIR_{it}$	Depreciating Knowledge Stocks	Years expanded to 1926 - 2009	$\sum_{j \in i(t)} \Delta a_{jt}$ replaced by $\sum_{j \in i(t-1)} \Delta a_{jt}$	Young Field defined as $PAT_{it} < 1000$

It may be that our model fails to properly control for the depreciation of knowledge (the extent to which older knowledge grows less relevant to new problems). In column 2, we therefore construct analogues of PAT_{it} and AFF_{it} that include depreciation to illustrate how our results differ from this standard practice. Define the depreciated patent stock in the standard way:

$$dPAT_{it} = \hat{y}_{it} + \left(1-\delta\right) dPAT_{it-1} \tag{11}$$

where $\hat{y}_{it}$ is the number of patents granted to field i in year t. To construct a depreciated version of AFF_{it} we assume the value of any given pair to a particular field depreciates at a constant exponential rate 1 - δ. Let $t_0(i,j)$ denote the first year any granted patent in field i uses mainline-pair j. The depreciated affinity stock is:

$$dAFF_{it} = \sum_{j \in i(t)} \left(1-\delta\right)^{t-t_0(i,j)} a_{jt} \tag{12}$$

Following standard practice (see Aghion et al 2016 for one example), we set δ = 20%. With depreciation, our estimates are in the expected direction, but their magnitude is attenuated and the coefficient on the affinity stock is statistically indistinguishable from zero.

After 2001, patent applications fall rapidly in our dataset due to a truncation effect, as we only observe patent applications that are subsequently granted by 2010. Throughout the paper, we exclude data after 2001. Column 3 assesses the impact of these years by restoring them to the sample, but the estimated coefficients are qualitatively similar to those in Table 1, column 1.

In Table 2, we measure the change in the affinity stock of a field, holding constant the set of mainline-pairs the field draws on. The set of mainline-pairs used is defined by those belonging to patents granted in period t or earlier. It might be that new knowledge diffuses with a lag though. In column 4, we measure the change in the affinity stock of a field, holding constant the set of mainline-pairs used as those assigned to any patents in the field granted prior to or during period t - 1. Our results are qualitatively unchanged by this alternative definition.

Finally, in column 5 we define a young field to be any with PAT_{it} < 1,000. By this definition, 13.5% of observations are defined as belonging to young fields. Coefficients are still in the expected directions and associated *p*-values remain high, but the coefficient on the affinity stock is reduced and the coefficient on the patent stock more negative than in Table 1.

7. Conclusion

This paper presents five stylized facts. First, patents have been making increasingly making less novel connections among technology classes since 1947. Second, the extent to which these connections are useful to other technology fields grew at a steady rate through the early 1980s, but has plateaued since then. Third, the annual flow of patent applications in a given technology field decreases over time, but this can be offset by the discovery of novel connections. Fourth, using a calibrated knowledge production function, the number of subsequent patents induced per patent has been mostly declining since 1926 and may be negative. Fifth, the patents that had the biggest positive impact on subsequent patenting tended to be electrical/mechanical at the beginning of our period, chemical in the middle, and computer-related in the end.

These results are dependent on correctly controlling for serial correlation, the inclusion of explanatory variables related to "young" fields, and measuring the depth of knowledge about how to combine a pair of knowledge. They are robust to changes in the years under consideration, the precise definition of a "young field", and the way we measure the change in a field's "affinity stock." In this conclusion, we discuss the limitations of this paper, and relate it to existing literature.

7.1. *Limitations*

The above findings have several important caveats. In terms of the motivating theory, as noted at the outset, models of combinatorial innovation frequently involve R&D to discover new technological elements for combination (Weitzman 1998, Akcigit, Kerr, Nicholas 2013). This article has taken the set of technological components available for combination as given, and therefore is unable to assess whether the rate of new technological components is increasing to offset the decline in novel combinations. Relatedly, the theoretical framework set up in this model has little to say about the formation of new technological fields. Given the fishing out effect, the continuous creation of new fields (with low patent stocks) represents one way to continue the innovation process indefinitely. More broadly, combinatorial models of discovery, while useful, are surely incomplete. Other factors driving innovation may well offset the negative trends we highlighted. Finally, as noted in the introduction, this paper has not attempted to determine whether the R&D trends documented here are socially suboptimal.

Second, data limitations may also limit the robustness of our findings. Data on R&D spending by technology field is unavailable, and so this paper misses an important input of the ideas production function. However, it is notable that recent work by Bloom et al. (2016) finds it is taking more and more R&D spending to get the same quality improvements across a range of sectors, consistent with this article's cautionary take on the future of innovation. Furthermore, patents are themselves an imperfect measure of innovation: not all ideas are patented (especially those developed outside the private sector) and those that are vary tremendously in their value. Most patents are worth little, but a small number are very valuable. Clancy (2017) shows patents that forge more novel connections between their mainlines are more likely to be renewed, and therefore more valuable to the holders. This could imply that patented inventions may be becoming less valuable, in addition to less novel. This paper, however, does not adjust for the "quality" of a patent, as it is not clear whether patents with low value contribute less to the knowledge stocks developed in this paper. Still, it may be that an increase in low-quality “copycat” patents is responsible for the decline in *average* measured impact, but that we continue to produce just as many novel, high-quality patents as ever. This argument is particularly compelling after 1984, when patent applications expanded rapidly. However, our results are actually most pessimistic when we restrict our sample to the period 1926-1984.

Lastly, as patents are an imperfect proxy for innovation, mainlines are an imperfect proxy for pre-existing technological components. For example, as noted in section 1, many patents are assigned less than two mainlines. We interpret this as a failure of data to identify all the components used by the patent, not as a refutation of the idea that combination is inherent in all innovations. Locating more precise measures of technological components (i.e., words in patents, citations, or subclasses at a more disaggregated level) could clarify if the explanatory power of these results is biased by measurement error.

7.2 *Will Innovation Get Harder?*

The evidence presented may presage a slowdown in innovation in the coming decades: patents are increasingly less novel, the number of spillovers has stopped increasing, and their net impact on future patent applications is trending to zero, if not already negative. The extent to which we are pulling ideas out of the pool is growing relative to the restocking rate, and may well have surpassed it.

These findings align with Arora et al. (2015), which suggests the private sector has been withdrawing from basic science to focus more exclusively on applied R&D, since at least 1980. Basic science is believed to generate the most novel, disruptive ideas, and to have the most spillovers and follow-on applications. A reduction in this kind of research is consistent with our findings. Note one hypothesis put forward by Arora et al. (2015) for the private sector’s withdrawal from basic science is the narrowing scope of firms, as firms with a narrower range of expertise are less likely to benefit from the unpredictable applications stemming from basic research. This paper suggests there may be a complementary supply side effect. When innovation is about making new connections between previously disparate technological disciplines, firms with expertise in a wider array of disparate disciplines will be better placed to make such connections.

This work also supports the hypothesis that innovation's outlook is deteriorating because the supply of innovations from major "industrial revolutions" are being exhausted, an argument developed at length in Gordon (2016). The so-called "second industrial revolution" consisted of the innovations stemming from electricity, chemicals, plastics, and a host of other improvements in infrastructure and health, while the "third industrial revolution" is associated with information technology. Figure 4 is broadly consistent with the idea that applications of the second industrial revolution (electricity, chemicals, and plastics) drove innovation for much of the 20th century, while the applications of the third industrial revolution are ongoing. A long-run decline in patent impacts is also consistent with the broad theme that the second industrial revolution's gifts have been petering out. The small uptick in patent impacts in the late 1990s is also consistent with the argument that the third industrial revolution is real, but not as potent as the second.

References

Acemoglu, Daron, Ufuk Akcigit, and William R. Kerr. 2016. Innovation Network. Proceedings of the National Academy of Science 113(41): 11483-11488.

Aghion, Philippe, Antoine Dechezlepretre, David Hemous, Ralf Martin, John Van Reenen. 2016. Carbon Taxes, Path Dependency, and Directed Technical Change: Evidence from the Auto Industry. J. Pol. Econ. 124(1), 1-51.

Agrawal, Ajay, Alberto Galasso, and Alexander Oettl. 2017. Roads and Innovation. Review of Economics and Statistics 99(3): 417-434.

Aharonson, Barak S., and Milissa A. Schilling. 2016. Mapping the Technological Landscape: Measuring technology distance, technological footprints, and technology evolution. Research Policy 45(1): 81-96.

Akcigit, Ufuk, William R. Kerr, Tom Nicholas. 2013. The Mechanics of Endogenous Innovation and Growth: Evidence from Historical U.S. Patents. Working Paper.

Arora, Ashish, Sharon Belenzon, and Andrea Patacconi. 2015. Killing the Golden Goose? The Decline of Science in Corporate R&D. NBER Working Paper 20902.

Arthur, Brian. 2009. The Nature of Technology: What It Is and How It Evolves. New York, NY: Free Press.

Auerswald, Philip, Stuart Kauffman, Jose Lobo, and Karl Shell. 2000. The Production Recipes Approach to Modeling Technological Innovation: An Application to Learning By Doing. Journal of Economic Dynamics and Control 24: 389-450.

Bloom, Nicholas, Charles I. Jones, John Van Reenen, and Michael Webb. 2016. Are Ideas Getting Harder to Find? Preliminary and incomplete working paper.

Clancy, Matthew. 2017. Inventing by Combining Pre-Existing Technologies. Working Paper.

Costantini, Mauro, and Claudio Lupi. 2012. A Simple Panel-CADF Test for Unit Roots. Inst. for Adv. Studies, Vienna, Econ. Series 261.

Croissant, Yves, and Giovanni Millo. 2008. Panel Data Econometrics in R: The plm Package. J. Stat. Software 27(2).

Feinstein, Jonathan. 2011. Optimal Learning Patterns for Creativity Generation in a Field. American Economic Review Papers and Proceedings: 1-5.

Feldman, Maryann P., and Dieter F. Kogler. 2010. “Stylized Facts in the Geography of Innovation.” Chapter 8 in Handbook of the Economics of Innovation, Volume 1, edited by Bronwyn Hall and Nathan Rosenberg, 381-410. Oxford, UK: Elsevier.

Fleming, Lee. 2001. "Recombinant Uncertainty in Technological Search." Management Science 47(1) 117-132.

Galasso, A., and Mark Schankerman. 2015. Patents and Cumulative Innovation: Causal Evidence from the Courts. *Quarterly Journal of Economics* 130(1): 317-369.

Ghiglino, Christian. 2012. Random Walk to Innovation: Why Productivity Follows a Power Law. Journal of Economic Theory 147: 713-737.

Gordon, Robert J. 2016. The Rise and Fall of American Growth. Princeton University Press, Princeton.

Hall, Bronwyn H., Adam .B. Jaffe, Michael Trajtenberg. 2001. The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools, NBER Work. Pap. 8498.

Jovanovic, Boyan and Rafael Rob. 1990. Long Waves and Short Waves: Growth Through Intensive and Extensive Search. Econometrica: 1391-1409.

Kauffman, Stuart, Jose Lobo, and William G. Macready. 2000. Optimal Search on a Technology Landscape. Journal of Economic Behaviour and Organization 43: 141-166.

Keijl, S., V.A. Gilsing, J. Knoben, and G. Duysters. 2016. "The Two Faces of Inventions: The Relationship Between Recombination and Impact in Pharmaceutical Biotechnology." Research Policy 45(5): 1061-1074.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman. 2017. Technological Innovation, Resource Allocation and Growth. Quarterly Journal of Economics 132(2): 665-712.

Madsen, Jakob B. 2008. Semi-endogenous versus Schumpeterian growth models: testing the knowledge production function using international data. J. Econ. Growth 13, 1-26.

Nemet, Gregory F. 2012. "Inter-technology Knowledge Spillovers for Energy Technologies." Energy Economics 34 1259-1270.

Nemet, Gregory F., and Evan Johnson. 2012. "Do Important Inventions Benefit from Knowledge Originating in Other Technological Domains?" Research Policy 41 190-200.

Olsson, Ola. 2005. Technological Opportunity and Growth. Journal of Economic Growth 10(1): 35-57.

Olsson, Ola, and Bruno S. Frey. 2002. Entrepreneurship as Recombinant Growth. Small Business Economics 19(2): 69-80.

Popp, David. 2002. Induced Innovation and Energy Prices. Am. Econ. Rev. 92(1), 160-180.

Porter, Michael E., Scott Stern. 2000. Measuring the "Ideas" Production Function: Evidence From International Patent Output. NBER Work. Pap. 7891.

Schilling, Melissa A., and Elad Green. 2011. "Recombinant Search and Breakthrough Idea Generation: An Analysis of High Impact Papers in the Social Sciences." Research Policy 40 1321-1331.

Schoenmakers, Wilfred and Geert Duysters. 2010. "The Technological Origins of Radical Inventions." Research Policy 39 1051-1059.

Scotchmer, Suzanne. 2004. Standing on the Shoulders of Giants: Protecting Cumulative Innovators. Chapter 5 in Innovation and Incentives. MIT Press.

US Patent and Trademark Office. 2012. "Overview of the U.S. Patent Classification System (USPC)." The United States Patent and Trademark Office. 12. Accessed 10 15, 2014. http://www.uspto.gov/patents/resources/classification/overview.pdf.

—. 2014a. Table of Issue Years and Patent Numbers, for Selected Document Types Issued Since 1836. March 26. Accessed October 16, 2014. http://www.uspto.gov/web/offices/ac/ido/oeip/taf/issuyear.htm.

—. 2014b. "U.S. Manual of Classification File (CTAF)." USPTO Data Sets. Accessed August 2014. http://patents.reedtech.com/classdata.php.

—. 2014c. "U.S. Patent Grant Master Classification File (MCF)." USPTO Data Sets. Accessed August 2014. http://patents.reedtech.com/classdata.php.

Usher, Abbott Payson. 1929. A History of Mechanical Inventions. New York, NY: McGraw-Hill.

Weitzman, Martin L. 1998. Recombinant Growth. Quarterly Journal of Economics 2: 331-360.

Zucker, Lynne G., Michael R. Darby, Jonathan Furner, Robert C. Liu, Hongyan Ma. 2007. Minerva Unbound: Knowledge Stocks, Knowledge Flows, and New Knowledge Production. Res. Pol. 36(6), 850-863.

Supplemental Figures

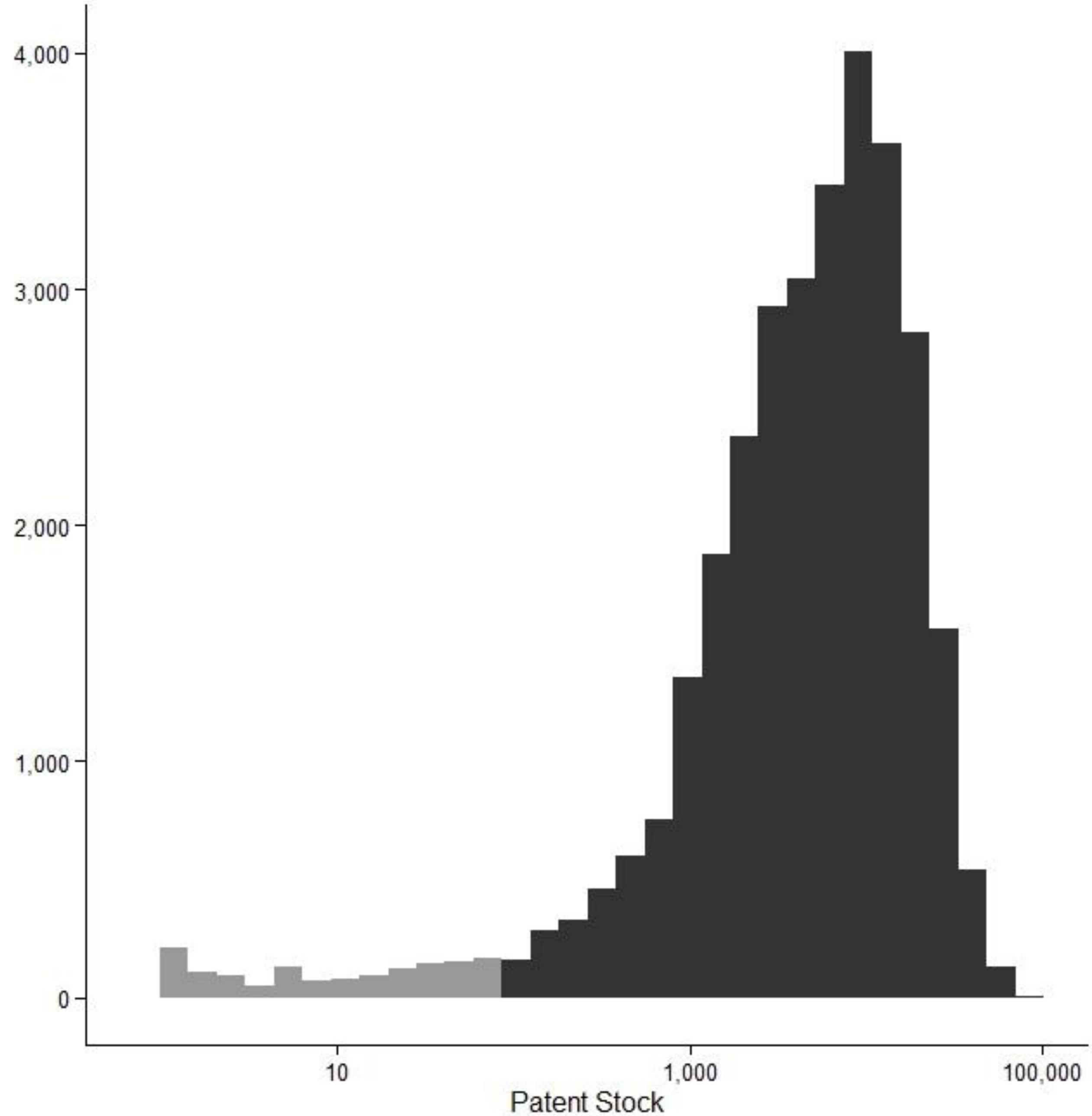

Figure S1: Histogram – Patent Stock, horizontal axis log-scale

Note: Young field observations ($PAT_{it} < 89.5$) are shaded in lighter gray.

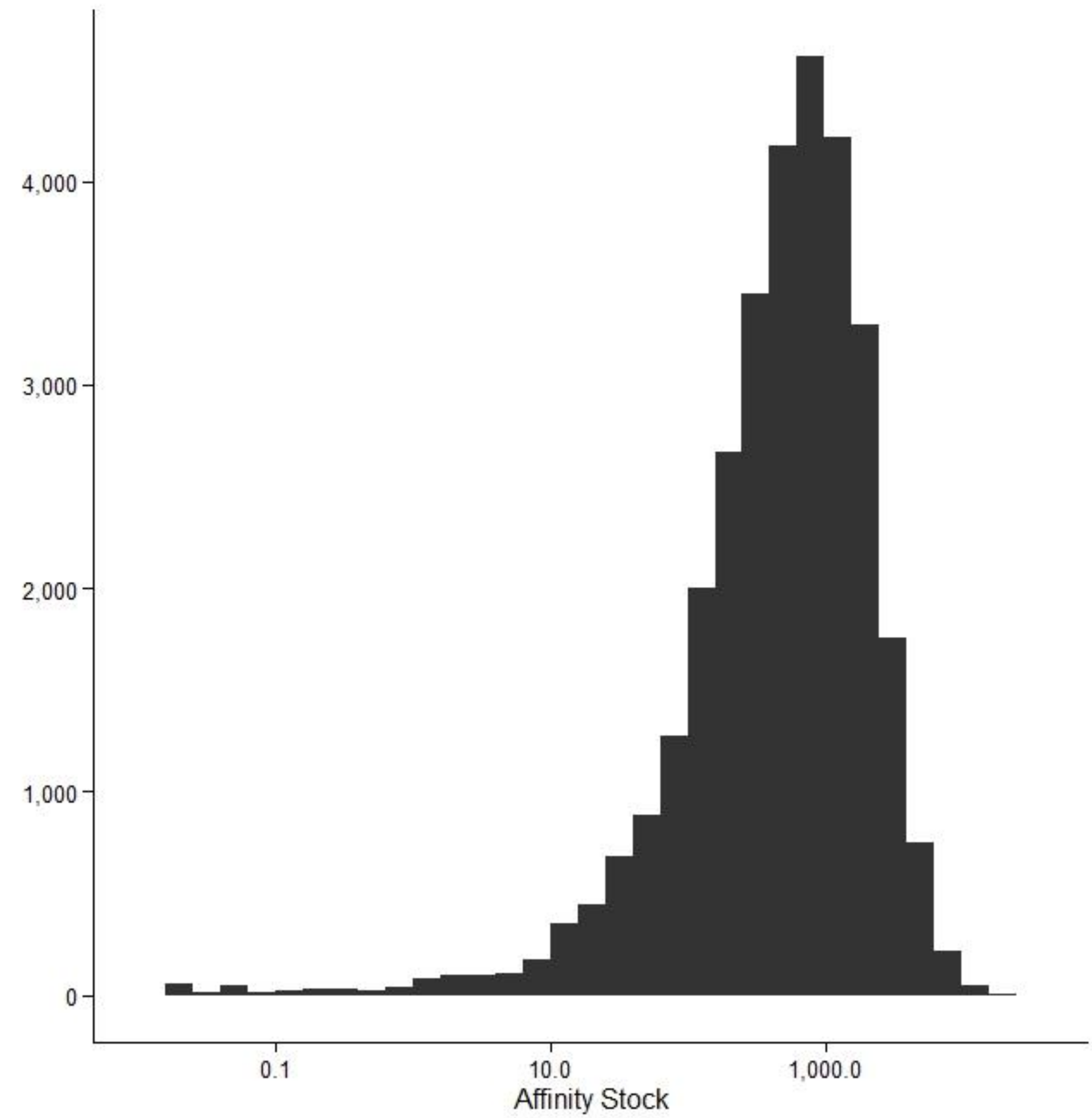

Figure S2: Histogram – Affinity Stock, horizontal axis log-scale

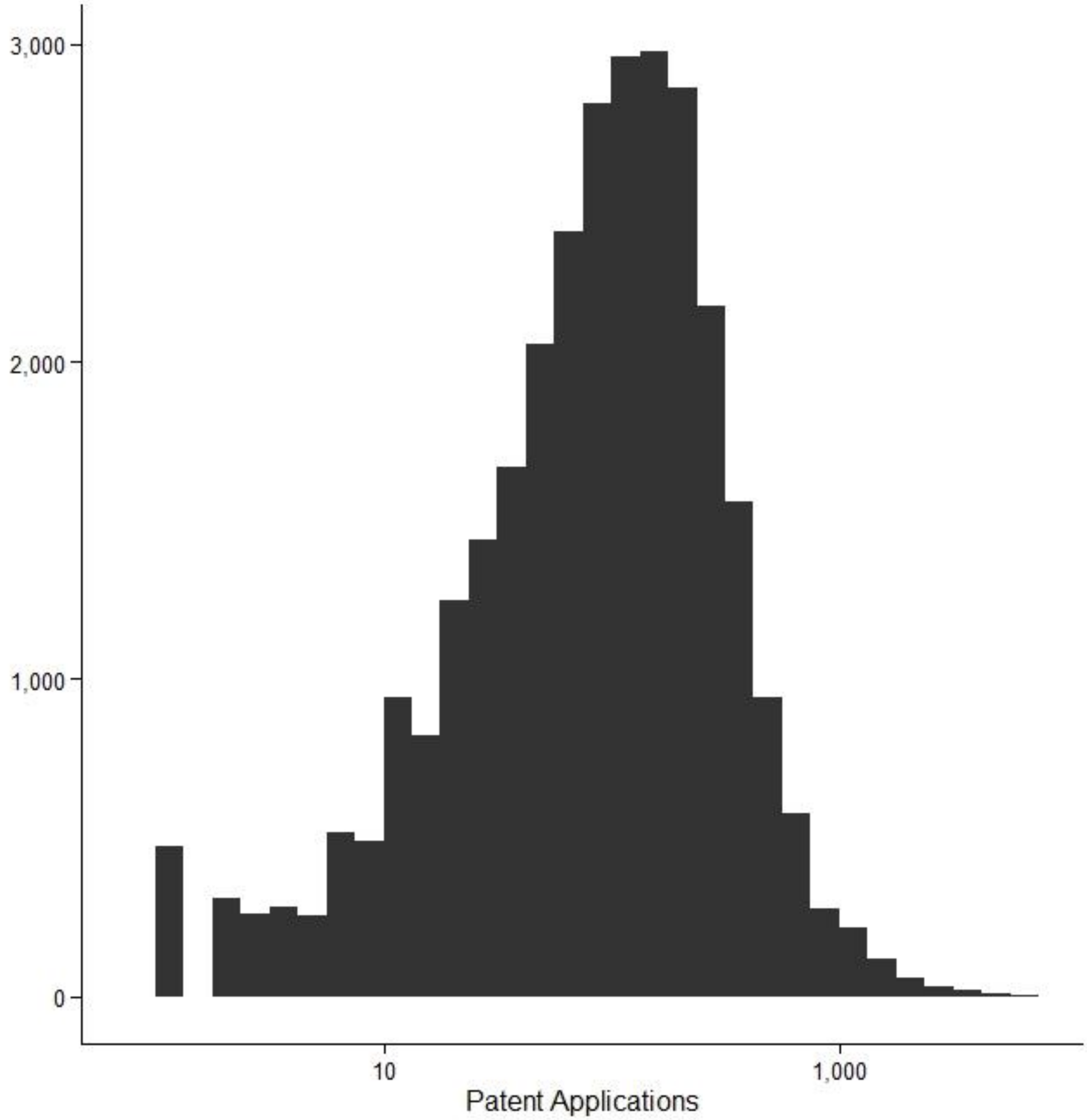

Figure S3: Histogram – Patent Applications, horizontal axis log-scale

Note: There are an additional 1,361 observations with y_{it} = 0.

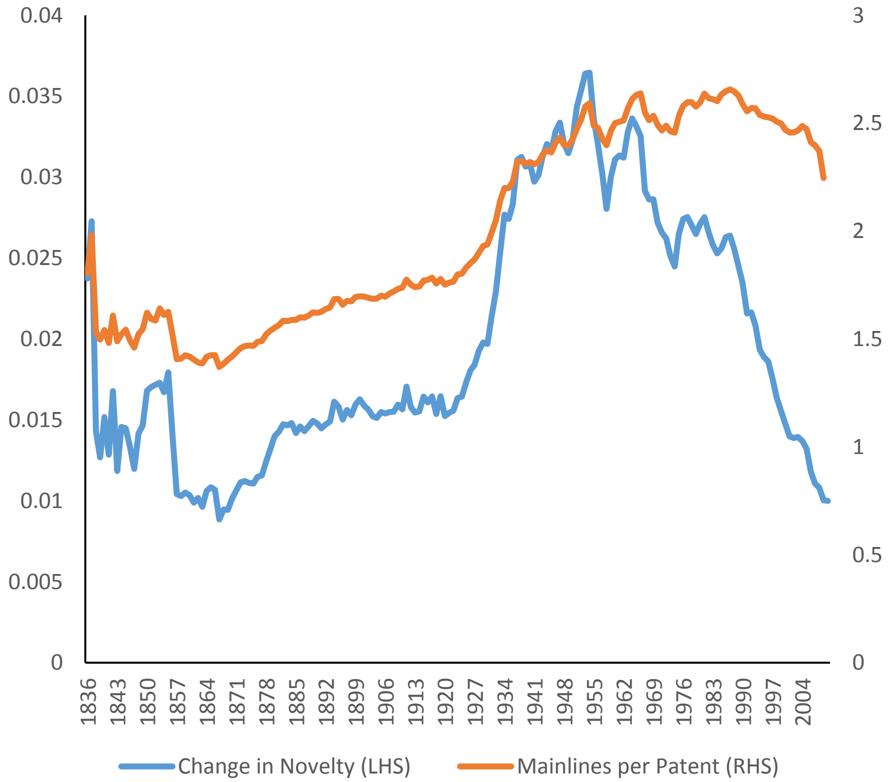

Figure S4. Average Novelty per Patent and Mainlines per Patent (1836-2010)

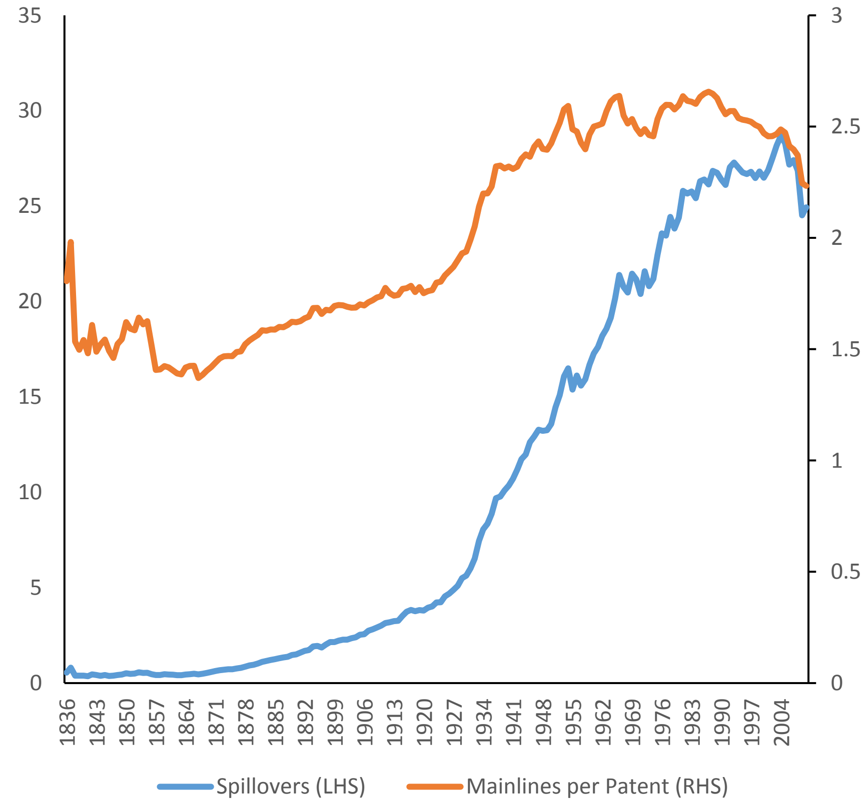

Figure S5. Spillover Fields per Patent and Mainlines per Patent (1836-2010)