Watson
Innovation analytics

D4.3 Rational Industrial Policy: Standing on the shoulders of Giant Gnomes?

This project has received funding from the European Union's Horizon 2020 research & innovation programme under grant agreement No. 777439

Watson
Innovation analytics

AUTHOR NAME	**Organization**
CHARLOTTE GUILLARD	IMPERIAL COLLEGE LONDON
RALF MARTIN	IMPERIAL COLLEGE LONDON/LSE
PIERRE MOHNEN	UNU MERIT
CATHERINE THOMAS	LSE
DENNIS VERHOEVEN	LSE

REVIEWER NAME	**Organization**
SIMONA COCUZ	INVENTYA
ZAHRA SHAH	KAPITALIZE
SUZANNE HAMILTON	STAKEHOLDER PANEL
GEORGE SUCIU JR.	STAKEHOLDER PANEL

VERSION	**Date**	**Modifications**
1	17/09/2019	First version
2	14/10/2019	Minor clarifications in text throughout

This project has received funding from the European Union's Horizon 2020 research & innovation programme under grant agreement No. 777439

Rational Industrial Policy: Standing on the shoulders of Giant Gnomes? (Watson Project Deliverable D4.3)

Charlotte Guillard, Ralf Martin, Pierre Mohnen, Catherine Thomas, Dennis Verhoeven

August 2019

Abstract

There is renewed appetite for vertical industrial policy and strategy in advanced industrialized countries. The motivation is typically a desire to promote economic growth. We suggest that a possible mechanism to guide strategy is the large and persistent differences in the knowledge spillovers generated by different technologies. By directing innovation support to technologies and sectors generating the highest spillovers, governments can have a positive effect on growth. We develop a new way to measure innovation spillovers from patent and balance sheet data called Patent Rank. This builds on Google's Page Rank but instead of a probability measure it provides a value measure of spillovers. It takes into account both direct and indirect spillovers and differentiates between global and national spillovers. We argue that the latter are relevant for a national industrial strategy. We show that there is significant and economically meaningful variation in the value of spillovers generated by different technology areas. Our framework also uncovers considerable variation between different countries, between measures of comparative advantage and measures of spillovers as well as between national and global spillovers. There is also considerable variation in the amount of internalization of spillovers at the country level; e.g. the US, Germany, and Japan are internalizing spillovers that are nearly twice as valuable as the spillovers internalized by countries such as the UK or France. Taken together, our results leave little hope for a one-size-fits-all industrial strategy, but the proposed framework allows for a tailored approach to better address market failures implied by positive externalities from knowledge spillovers.

1 Introduction

Vertical industrial policy consists of targeted government intervention in specific industries. The policy objectives can be varied and include aims such as addressing climate change, or promoting national security. Often, however, industrial policy is viewed as a mechanism for overall economic growth. With growth as the objective, policy makers turn to the challenges of policy design: which sectors should a government target and what form should interventions take?

Economic theory tells us that existing market imperfections create opportunity for welfare-enhancing interventions. A natural starting point for policy design is therefore to identify industries where markets deliver suboptimal outcomes. In many cases, this process requires substantial industry expertise, and policy makers lacking relevant information become open to economic capture by specialists.

In this document, we introduce a method that generates an industry-specific measure of scope for efficient growth-promoting subsidies. The method relies on publicly-available data and

is hence relatively immune to capture. The economic mechanism that underpins this method is based on three facts: First, the current stock of knowledge is not only a valuable input into production but is also an input to new knowledge creation that generates productivity and output growth. Second, knowledge has the public good properties of being non-rival and non-excludable, which means that the current stock of knowledge is available as an input to all potential innovators (in the absence of IP mechanisms). Third, potential innovators compare the cost of innovating with the benefits that will accrue to them. They do not consider the benefits that spill over to other innovators who could use new knowledge as an input to production or further innovation.

The implication of these three facts is that an inefficiently low level of new knowledge will be created, and current output and output growth will be below optimal levels. An industrial policy that helps align the private costs with the total social benefit of innovation will raise output and growth rates. The method described here ranks industries by the extent of current incentive misalignment, which is equivalent to ranking industries by the potential value of the proposed policy intervention.

To calculate this empirical measure, we take advantage of one aspect of the institutional infrastructure designed to mitigate these market imperfections: the patenting system. The fundamental goal of a patent is to create incentives to innovate—and invent something that is worth patenting—by granting innovators monopoly rights over that knowledge and licensing fees by allowing others to use the knowledge in production. Turning to further growth, i.e. the innovation production function, patenting systems also require formal acknowledgement of the stock of knowledge that has been used, or embodied, in new knowledge creation. When an inventor applies for a patent, she cites the patents that have contributed to the innovation process. This paper analyzes the information on 17 Million innovations that were patented between 2005 and 2014 across 90 patent authorities available in PATSTAT[1], the world's most comprehensive patent database provided by the European Patent Office.[2] Throughout, we conduct our analysis at the innovation rather than the patent level. For that we use PATSTAT's innovation family database that allows identifying when the same inventive step is patented multiple times in various jurisdictions.

Patent citations create a network where links document knowledge flows between innovations. As such, they provide information that can be used to measure the extent to which knowledge spills over between patenting organizations to increase productivity and promote growth. There are both direct and indirect connections between existing patents recording the current state of knowledge and new patents that record innovation and productivity growth. For example, one invention may have proved to be a critical input into another invention that was, itself, instrumental in creating an entire new technological area. While direct inputs are cited in new patent applications, the indirect contributions to growth are somewhat obscured.

The insight that led to the method developed here was the observation that internet search algorithms use analogous information networks to rank the webpages returned in search results. Specifically, the Google algorithm relies on the structure present in hypertext to provide high quality search results. It calculates a Web page's "PageRank", which is an objective measure of its relevance using the information that is contained in the citation (link) graph of the Web, and weighs a citation more heavily if the citing page is itself of a high PageRank. As described in Brin and Page (1998), an intuitive description is that a web page receives a high PageRank if there are many pages that point to it or if there are some pages that point to it that have a high PageRank. Web pages with high PageRanks appear first in search results.

[1] https://www.epo.org/searching-for-patents/business/patstat.html

[2] For some of our analysis we also consider changes over time. For, that we repeat some of the analysis for the 1975 to 1984, 1985 to 1994 and 1995 to 2004 time windows

This paper describes PatentRank (or P-Rank), where a patent plays the role of a web page, and the network of patent citations is analogous to the network of hypertext links in the Web. To make the analogy complete, we need a measure of patent value that corresponds to the quality of a web page. In our context, this is a measure of the total economic value of the knowledge embodied in a patent.

Here we depart from PageRank in approach by decomposing the total economic value of the knowledge embodied in a patent into two parts. We assert that the value of a patent consists of the sum of its value to the innovator responsible for the knowledge creation and the external value it creates for other innovators when used as an input to further valuable knowledge creation.[3]

We measure the private value for a large subset of patents using a new methodology proposed by Kogan et al. (2017). They measure the private value of a patent as the change in the value of the patenting firm's market value at the time when the firm is first granted the patent. The change in firm value is computed relative to the contemporaneous change in the value of a set of control firms that did not receive a patent grant, and is referred to as an abnormal stock return. The main drawback of this approach to measuring the private value of returns is that it provides information only for publicly-listed firms' patents. This makes up 2.4% of the patents in the patent network data relevant to our purposes. To estimate the private value of non-listed firms' patents, we extrapolate from the listed firms' patent sample. In this version of the paper, we extrapolate based on a highly disaggregated technology category level. We assign all patents in a narrow technology class the mean value found for listed firms' patents in that class.[4]

The network of patent citations and the estimates of patent private values provide the inputs to P-Rank allowing us to value the direct and indirect linkages between patents. As for PageRank, P-Rank can be calculated using a simple iterative algorithm, and corresponds to the principal eigenvector of the normalized link matrix of the patent citations. It sums up the values of the direct and indirect spillovers across the entire patent network, and adds this to the private value of each patent to estimate its overall value.

The decomposition of P-Rank into private value and external value underpins the policy design aspect of the paper. Some patents will have generated a large private value but contributed little to the further valuable knowledge creation as embodied in future patents, either because the original patent was not highly cited or because the patents that cite the original patent were themselves of low value. However, there are many patents whose external value is very large. Of particular interest are those patents with a large external value and small private value. Knowledge creation is inefficiently low if too few of these innovations take place. Given that innovators bear the full cost of innovation and capture only the private benefit of doing so, too few innovations with a small private value and large external value are undertaken. Directing subsidies to innovation in technology classes where the marginal external value is high relative to the private value will maximize the policy's efficiency-enhancing potential.

This idea gives rise to the second indicator described in this paper. We introduce IStra-X, an industrial strategy index, which maximizes the societal gain from a total subsidy value of S by assigning each marginal unit of subsidy to the technology category where it generates the highest marginal social return.

The usefulness of this instrument comes from the observation that P-Rank tells us which patents have generated the largest external values, and in summing across patents in a particular group we can value the external spillovers of patents from a given technology class or country. However, we can do better than simply directing all subsidies into the patent group with the highest total or highest average external value by allocating subsidies across areas so that each

[3]The web page analogy of private value would be some objective page quality measure that is separate from the measure of quality attributed from frequency and quality of citations by other pages.

[4]Other well-known limitations of event studies are mentioned later in the paper.

dollar of subsidy has the largest marginal return. Indeed, subsidy returns do not only depend on the total value generated by innovations undertaken as a consequence of the subsidy, but are also a function of the expected cost of pursuing one innovation in the area.

To address this, we need some way to identify patents where the private cost c exceeds the private benefit but where c is outweighed by the sum of private and external benefits. These are the patents that would not happen in the absence of a subsidy. IStra-X relies on two further ingredients: estimates of the parameters that govern the distribution of arrival of ideas that are essential inputs to the innovation production function and an estimate of the private cost of innovating. These estimates are likely to be technology-class, country-, and time-specific, and are unobserved. Nonetheless, the data allow us to estimate these variables by using structural assumptions and calibrating unobserved values from the observed distributions of technology-class-specific private values.

Armed with these estimates of the idea arrival process and private innovation costs, we can estimate the marginal effect of higher subsidies in a technology area, which will vary with the level of subsidy already received, if any. Subsidizing any area with positive IStra-X will be efficiency-enhancing and directing any fixed amount of subsidy into the areas with the highest IStra-X will maximize the efficiency gains from the subsidy.

Because the external value of any one patent is the sum of the value of knowledge spillovers to all other individual patents, it is possible to measure the extent to which knowledge spillovers are retained within any disaggregated technology class or within any geographical region. The methods developed in this paper allow us to construct the "balance of knowledge spillovers" for a given group, telling us whether, for example, a particular technology generates more spillovers for other industries than the spillovers it enjoys from other industries. Similar indexes are found for individual countries. Policy makers may choose to direct subsidies to those industries that generate the highest levels of social return in terms of local spillovers.

Our results suggest large and statistically significant differences in social returns between technology fields as well as between countries. Crucially, what generates the highest returns - both in terms of global or localized/national spillovers - differs greatly from country to country. Globally, chemistry and pharmaceutical innovations generate among the highest values of spillover. This is because innovations in these fields generate among the highest private values as well and spillover flows are concentrated within ones own field. However, in terms of IStra-X, our approach suggests that Robotics, Aerospace and Wireless technologies generate social returns on the order of 100% for public investment, whereas most chemical technology categories generate returns of less than 50%. This is because in those areas investment costs are higher and responsiveness of innovations to subsidies is lower. It is also instructive to look at the P-Rank of individual innovations. In line with the aggregate results, we find primarily chemical industry innovations ranked highest. However, the top 10 list also notably includes Apple's patent number US2006197753(A1) that outlines the functionality of the iPhone.[5] Comparing spillover flows measured with P-Rank between countries, we note that the US are with 45% responsible for the bulk of global knowledge spillover flows. However, most of this value arises within the US. In terms of knowledge spillovers, Japan turns out to be the most generous country with net exports of spillovers amounting to over $4000 per capita.

Patent Citations have been used for a long time to measure either spillovers or the quality of innovations. Compared to simple citation counts, our methodology takes into account, firstly, private value differences between innovations and, secondly, indirect spillovers. To explore the value added of this, we examine the notion of Giant Gnomes[6]: Innovations who are suggested to

[5] "Multifunctional hand-held device" https://patentswarm.com/patents/US20060197753A1

[6] In honour of Newton's famous quote in 1675: "If I have seen further it is by standing on the shoulders of Giants."

be more eminent using P-Rank than with simple citation counts. Innovations where the reverse is true we dub Illusory Giants instead. We discover that the largest share of external value derives from Giant Gnomes (74%). This is because Giant Gnomes are concentrated among high P-Rank Innovations. An implications of this is that a policy targeting spillovers measured via citation counts would rank projects incorrectly in particular when it matters most.

Further, we look at the power of P-Rank as a measure of spillovers to predict future innovations by country and technology. We distinguish between spillover received and provided by a given area. We consider both the number and value of future innovations. This exercise indicates that P-Rank better explains future innovations than citation counts or simple innovation counts. Moreover, by decomposing P-Rank (received), we show that knowledge imported from other areas tend to have a stronger effect on future innovations than local knowledge.

Overall, the benefits of this methodology include the objectivity coming from data combined with the theoretical properties of knowledge as an input to further innovation. However, there are associated caveats to its use. Regarding the way the measure is implemented, it is critical to note that much of the knowledge stock is not embodied in patents. Moreover, we capture only the value of spillovers that are ultimately reflected in private values.

Another caveat is that these estimates of private and external values are subject to the effects of differences in local institutions. Patent systems are themselves distorting, and some national systems may permit more or less private capture of benefits. Also, factors such as local industry structure and local regulation of prices affect patents' private values. Finally, the IStra-x policy instrument implicitly assumes that there are no subsidies, made in the absence of any institutional change, that affects existing incentives for innovation.[7]

Furthermore, structural estimates of σ, ϕ, α, δ, c, are all open to debate. Likewise, past values of distribution of ideas and cost need to remain constant. We provide some evidence that computed P-Rank values for countries and technological areas are very stable over time. But future research will require further validation of this pattern.

The paper proceeds as follows: Section 2 describes how P-Rank is computed. Section 3 presents some output from P-rank, decomposing patent values into private value, direct external value and indirect external value, across technology classes and countries. It also assesses P-Rank's ability to predict future patenting activity, in contrast to traditional citations methods. Section 4 describes how IStra-x is calculated and presents some calibrated technology area case studies. Section 5 discusses and concludes.

2 Patent Rank

2.1 Motivation

An innovation makes up part of the stock of knowledge and is described in a patent. Each patent i cites N_i other patents given by set B_i. Patent i may also be one of the patents that make up B_j, i.e. be a citation for patent j, where j is any other patent in the data. The set of all patents that cite i is F_i.[8]

P-Rank is the value of patent i, V_i, and is made up of two parts. First is the private value of patent i, PV_i. Second is the value that patent i has created when used as an input to the production of patent j for all $j \in F_i$, denoted $EV_i = \sum_{j \in F_i} f_{ij}(V_j)$. Therefore:

[7] That said, we could build in pre-existing subsidies if such data is available.

[8] B refers to "backwards citations" and F refers to "forwards citations".

$$V_i = PV_i + EV_i = PV_i + \sum_{j \in F_i} f_{ij}(V_j) \tag{1}$$

An important aspect of Patent Rank is the nature of the function $f_{ij}(V_j)$. In this version of the paper, we assume that this function contains two parameters that are properties of the innovation production function. Suppose that the value V_j derives from a production function where the inputs include the R&D investment from the firm (RD_j) and also the stock of knowledge embodied in prior patents B_j. For example, if we assume that there exists an innovation value production function and it has a Cobb-Douglas form with efficiency shifter A_j, then:

$$V_j = A_j B_j^{\sigma} RD_j^{1-\sigma} \tag{2}$$

In this formulation, the parameter σ measures the relative contribution of prior knowledge to future innovation, and does not vary with j. By differentiating equation(2) with respect to each cited innovation $i \in B_j$, and then substituting in for V_j, we derive an expression for the marginal contribution of citation i to V_j.

$$\frac{\partial V_j}{\partial i} = V_j \sigma \frac{\partial B_j}{\partial i} \frac{1}{\partial B_j} \tag{3}$$

We denote as ϕ_{ij} the term $\frac{\partial B_j}{\partial i} \frac{1}{\partial B_j}$. This term measures patent i's contribution to the stock of knowledge used in the production of patent j. For now, we assume that each of the N_j patents in the set B_j contributes equally, therefore this term simplifies to $\phi_{ij} = \frac{1}{N_j}$ for each $i \in B_j$ and zero for all $i \neq B_j$.[9] We are now able to write the P-Rank of patent i as:

$$V_i = PV_i + \sigma \sum_{j \in F_i} \phi_{ij} V_j = PV_i + \sigma \sum_{j \in F_i} \frac{1}{N_j} V_j \tag{4}$$

The parameter σ determines what share of patent i's value is attributable to the value of the spillovers that it creates. Because of the recursive nature of the P-Rank measure, it weighs the value of indirect forward citations by σ to the power of the level of indirectness. As such, it can be viewed as a distance decay parameter (e.g. the value of a patent that contains a backward citation of a backward citation of patent i will be weighted by σ^2 in V_i). Below we find that while the overall level of V_i is highly sensitive to σ, it has little impact on the ranking of innovations in terms of value.[10] This is re-assuring as our primary objective is to rank innovation relative to each other rather than come up with an estimate of their value overall. Nevertheless, it would be preferable to come up with an estimate for σ. In the appendix, we show some experimental results on how this could be done.

2.2 Estimating PV_i

We refer to D4.1 for a detailed description of the methodology to estimate private returns to innovations. Generally, our approach is to leverage private return estimates developed in Kogan et al. (2017) based on the stock market reaction to a patent grant to obtain an approximation of PV_i. They measure the private value of a patent as the change in the value of the patenting firm's market value at the time when the firm is first granted the patent. The change in firm value is computed relative to the contemporaneous change in the value of a set of control firms that did not receive a patent grant, and is referred to as an abnormal stock return. The main drawback

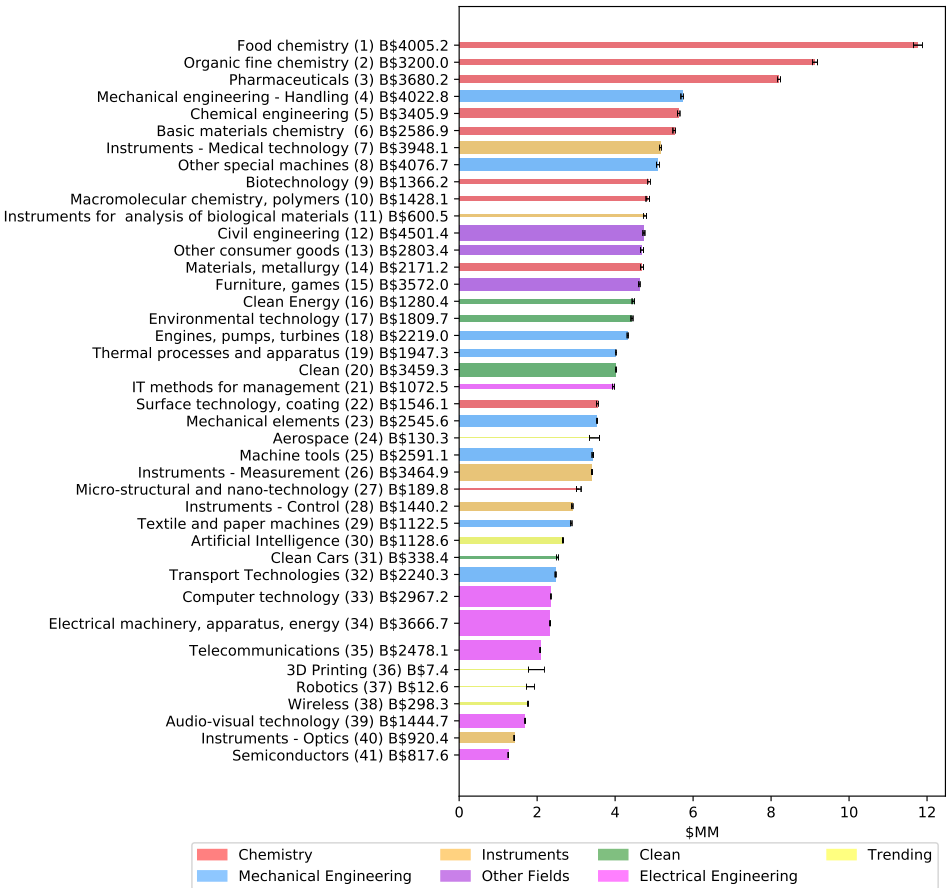
[9] A more general specification would be to assume that the contribution of innovation i to the knowledge used in j is a function of the characteristics of i and j, for example, whether both patents are in a similar technology class. We could denote this $\phi_{ij} = \phi(x_i, x_j)$, where the arguments of the function are patent characteristics.

[10] We present results corresponding to different values of σ, 0.25, 0.50, and 0.75.

of this approach to measuring the private value of returns is that it provides information only for a relatively small set of US publicly-listed firms' patents. These patents amount to only 2.4% of the relevant population for our purposes. To estimate the private value of non-listed firms' patents, we extrapolate from the listed firms' patent sample computed and made available by Kogan et al. (2017) in order to derive estimates for (nearly) the entire population. To do this, we regress the estimated private value on a set of patent characteristics that are plausible predictors of private value (technology classes, time of application cohorts, country and patent family size). We then fit various models — depending on which data is available for a patent family — to the relevant population based on these observed characteristics.

Figure 1 examines average private values across technology categories.[11] Chemical industry and pharmaceutical innovations lead the field with average private values of mor than $ 8 Million. At the bottom end we find technologies such as seminconducters or optics private values averaging at around $ 2 Million. Below we will argue that some of this variation reflects variations in the R&D costs between these fields. D4.1 provides further results strengthening our belief that our estimates of PV_i reflect relevant heterogeneity of private returns to innovation across different segments of innovative activity.

Figure 1: Private Returns by Technology - All 2005-2014 Innovations



Notes: Diagram of the average private returns in millions of CPI-adjusted 1982 US dollars (x-axis) by technology field (y-axis). Width of each bar represents the number innovations in the field. Area of each bar (in billions $) represents total private returns in the technology field and is printed next to y-axis labels.

[11] Details on our technology categories are provided in the Appendix.

2.3 Computing P-Rank

We collect the private values in the vector $\boldsymbol{PV}$, where the number of elements is equal to the total number of patents in the data, N. We also construct the $[N \times N]$ matrix $\boldsymbol{\Phi}$, where the element (i, j) is equal to B_j if patent j cites patent i.

We can write the vector of P-Rank values as V, which is equal to:

$$\boldsymbol{V} = \boldsymbol{PV} + \sigma \boldsymbol{\Phi} \boldsymbol{V}, \tag{5}$$

which can be rearranged as:

$$\boldsymbol{V^*} = (\boldsymbol{I} - \sigma \boldsymbol{\Phi})^{-1} \boldsymbol{PV}. \tag{6}$$

This equation can be estimated using the following recursive procedure: Starting with an arbitrary set of initial values $V_i^{(0)}$[12]. We compute a set of new values $V_i^{(n)}$ as:

$$V_i^{(n)} = PV_i + \sigma \sum_{j \in F_i} \phi_j V_j^{(n-1)} \tag{7}$$

In the appendix we prove that equation 7 has V^* as a fixed point given our assumptions about $\boldsymbol{\Phi}$.[13]

Armed with the P-Rank estimates, $\boldsymbol{V}$, we can find the external value of every patent as the vector:

$$\boldsymbol{EV} = \boldsymbol{V} - \boldsymbol{PV} = \left[(\boldsymbol{I} - \sigma \boldsymbol{\Phi})^{-1} - \boldsymbol{I} \right] \boldsymbol{PV}. \tag{8}$$

The external value of a patent represents the knowledge spillovers that it generates for the benefit of the rest of the innovation network. Suppose we segment all patents into areas, where an area, A, can be a technology class, a, a geographical area c or the intersection of the two groups, $a - c$. The sum of the external values by patents belonging to a given area A constitute the spillovers *generated* by this area.

$$ST_A^{out} = \sum_{i \in A} EV_i \tag{9}$$

Furthermore, part of the external value generated by a patent is transmitted to a specific area. To measure the total spillovers an area *receives* from other patents, we need to slightly alter the calculation of value V such that:

$$\tilde{V}_{i,A}^{(n)} = \tilde{PV}_{i,A} + \tilde{EV}_{i,A} = \tilde{PV}_{i,A} + \sigma \sum_{j \in F_i} \frac{V_{i,A}^{(n)}}{N_j} \tag{10}$$

With $\tilde{V}_{i,A}^{(0)} = \tilde{PV}_{i,A}$ and $\tilde{PV}_{i,A} = \begin{cases} PV_i, & \text{if } i \in A, \\ 0, & \text{otherwise.} \end{cases}$

The total spillovers area A receives from other patents can then be easily calculated as follows:

$$ST_A^{in} = \sum_i \tilde{EV}_{i,A}^{(n)} \tag{11}$$

[12] A natural choice is $V_i^{(0)} = PV_i$

[13] In our results we compare V^* to a simpler measure of spillovers based on direct linkages only. This is the first iteration of 7; i.e. the estimate of the direct value of innovation i, DV_i, is $DV_i = V_i^{(1)} = V_i + \sum_j \phi_{ij} V_j^{(0)} = V_i + \sigma \sum_j \phi_{ij} V_j$

This value can be further decomposed into those spillovers from patents in the same area and those from patents outside the area. We call these quantities L_A and I_A, respectively the values of local and of imported spillovers.

$$L_A^{in} = \sum_{i \in A} \tilde{EV}_{i,A}^{(n)} \tag{12}$$

$$I_A^{in} = \sum_{i \notin A} \tilde{EV}_{i,A}^{(n)} \tag{13}$$

3 P-Rank Results

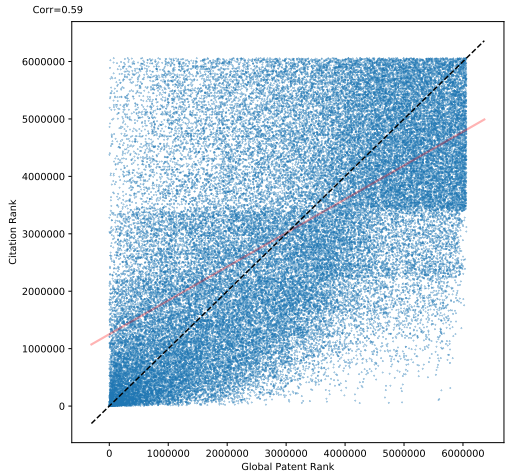
3.1 Comparisons across Patents

We first compare a patent's P-Rank with its citation count, which is the number of forward citations that the patent receives—i.e. the size of the set F_i. Forward citations have been used in many settings as a measure of the quality of the patent. In comparison to the P-Rank measure, it counts the number of direct spillovers to new innovations, whereas P-Rank includes both direct and indirect spillovers and attaches a value to each spillover.

Figure 2 is a scatter plot with a patent's global P-Rank (including all spillovers to all innovation areas) on the x-axis and its citation count on the y-axis. There is a positive correlation of 0.59 between the two measures, reflecting the fact that patents that are cited by many future patents also tend to be valuable. However, there are clearly many deviations—observations that lie far from the 45 degree line in Figure 2. Patents that lie above the 45 line are ranked higher using P-Rank than their citation count alone would suggest. This means that their actual value would be underestimated by the citation count measure. We call these patents "Giant Gnomes" as the extent of their influence makes them giants in knowledge creation but their true value is hidden and they appear as gnomes. Patents that lie below the 45 degree line have the opposite property in that their citation count rank overstates their relative P-Rank. We call these patents "Illusory Giants"—they appear to be giants in knowledge creation but actually have relatively little spillover value.[14]

[14] To learn more about Illusory giants see Michael Ende, "Jim Knopf und Lukas der Lokomotivführer".

Figure 2: Patent Rank vs Simple citations



Notes: Scatter plot of ranking all innovations with non zero citations in 2005 to 2014 by either Patent Rank (with $\sigma = 50\%$) or by simple citation count. A rank implies a higher citation count or external value. To ease computation of the scatter plot we draw a random 1% sample of the data. The correlation figure is based on the full sample.

In a regression of P-Rank on citation count, the slope of the estimated relationship between the two is significantly less than 1 (see Figure 2). This tells us that Giant Gnomes are more prevalent among patents with a particularly high P-Rank and Illusory Giants are more prevalent among patents with a low P-Rank. This proves insightful from the perspective of policy design. For example, among those patents with high P-Rank, a government agency subsidizing innovations that were expected to have a large number of forward citations would end up subsidizing a higher share of Illusory Giants than Giant Gnomes.

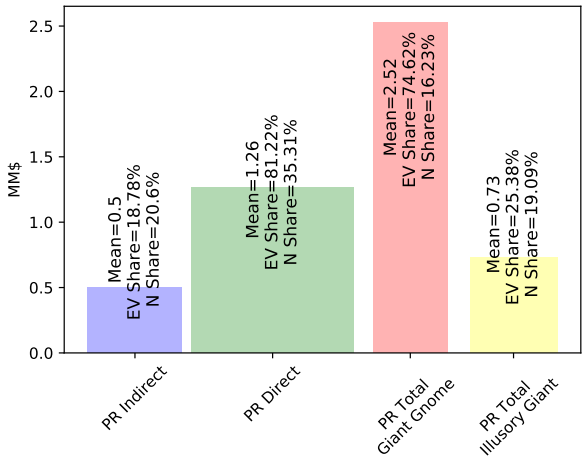
Recall that the value of a patent can be decomposed into private and external value (i.e. spillovers). The latter can be further broken down into direct and indirect spillovers. Figure 3 describes some aspects of the external value component of P-Rank. In particular, it explores the relationship between direct and indirect external value. The share of innovations with indirect Patent Rank (20%) is naturally smaller than the share of innovations with direct patent rank.[15] Also, the average value of indirect spillovers is much smaller (less than half) than the average direct value. Consequently, a smaller fraction of total external value (19%) is indirect.

The third and fourth bars in Figure 3 compare the average Patent Rank for Giant Gnomes compared to Illusory Giants. Panel (a) presents these mean values for our preferred specification of $\sigma = 0.5$. We see that the mean P-Rank of Giant Gnomes is \$2.52 Million, which is more than three times the mean value for Illusory Giants (at \$0.73 Million). Therefore, even though Giant Gnomes are somewhat less prevalent[16] as they account for a much larger fraction of the total external value generated: 65% as opposed to 35% for Illusory Giants.

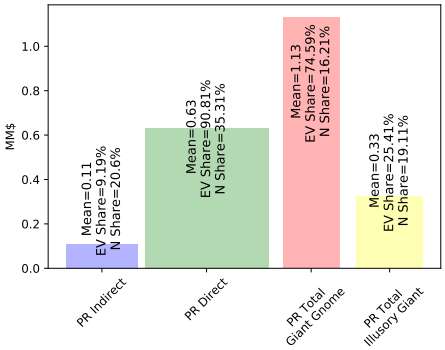
[15] An innovation cannot have indirect spillovers without direct ones.

[16] They account for 16% of innovations, whereas Illusory Giants account for 10%. Note that most innovations - i.e. 65% - do not generate any external value (at least as measured by citations) at all.

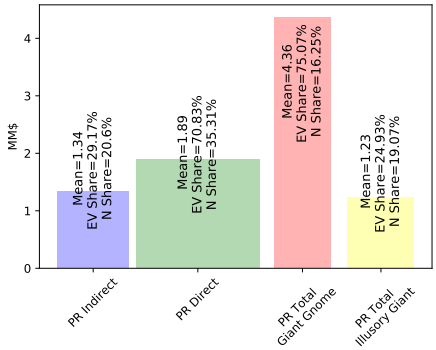
Figure 3: Indirect and Direct Spillovers, Giant Gnomes and Illusory Giants



(a) $\sigma = 50\%$



(b) $\sigma = 25\%$

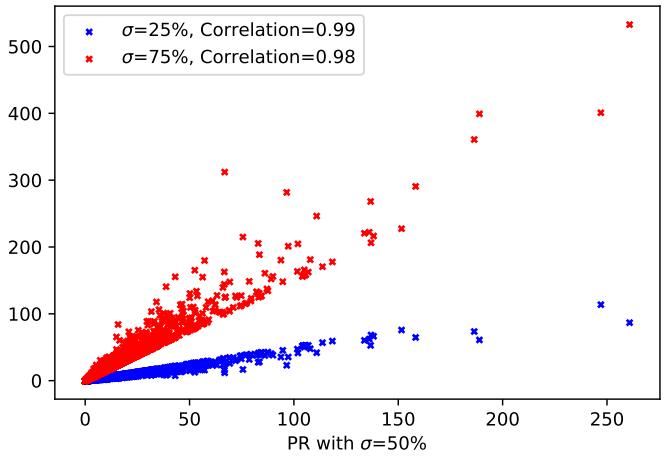


(c) $\sigma = 75\%$

Panels (b) and (c) of Figure 3 explore what happens if we vary the decay parameter σ. Because this parameter is the weighting of the external value to private value in P-Rank, and also determines the importance of indirect spillover values, it has a strong impact on the magnitude of the calculated external values and on P-rank value. For instance, with $\sigma = 25\%$, external values are about half as large as when we set $\sigma = 50\%$. With $\sigma = 75\%$, external values are 50% larger. We note that σ has a disproportionately large impact on indirect spillover value and, hence, on the ratio of indirect to direct spillover value.

Because σ enters the external value part of P-Rank multiplicatively but the external value enters P-rank additively, the value of σ has very little impact on the P-Rank patent ordering. Figure 4 illustrates that there is a near-perfect correlation between Patent Ranks computed under different values for σ. Consequently, the choice of σ value has little impact on which innovation is classified as a Giant Gnome or as an Illusory Giant or on the share of the value that derives from each of the two groups.

Figure 4: Patent rank with different σ's

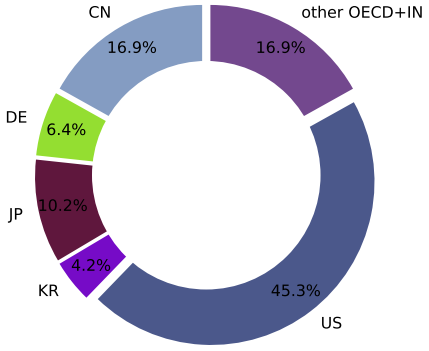


Notes: Correlations of PRANKs with different σ's.

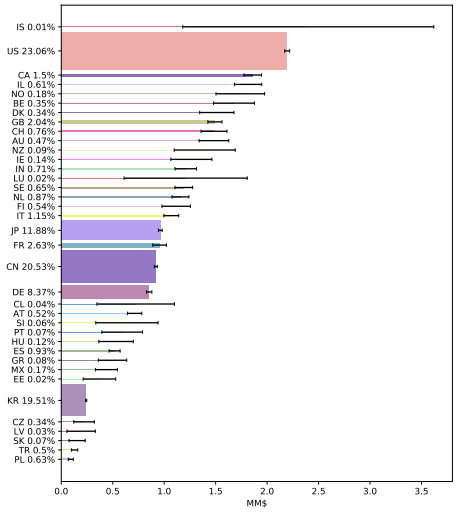
The fact that the choice of σ affects the level of P-Rank but not the ordering has implications for policy design. Specifically, it creates a caveat to statements that quantify the benefits of these policies. Nonetheless, we can be confident that recommendations based on the relative merits of different technologies, sectors, countries or other groupings of innovations are robust to the choice of σ.

We now turn to illustrate how P-Rank output allows comparisons among subsets of innovations. Panel (a) in Figure 5 presents a breakdown of the total external value created by spillover origin country (ST_c^{out}). That is, it shows the share of the total external value that is generated by the patents granted to innovators based in the country. US patents create the largest share of spillover value, at just over 45% of the total. China is the second largest spillover generator, at nearly 17% of the total. Japan makes the third largest contribution, at 10%, Germany creates just over 6% of the total, and Korea creates just over 4%. All other countries combined contribute to just under 17% of the total, with any individual country making up 3% or less of the total.

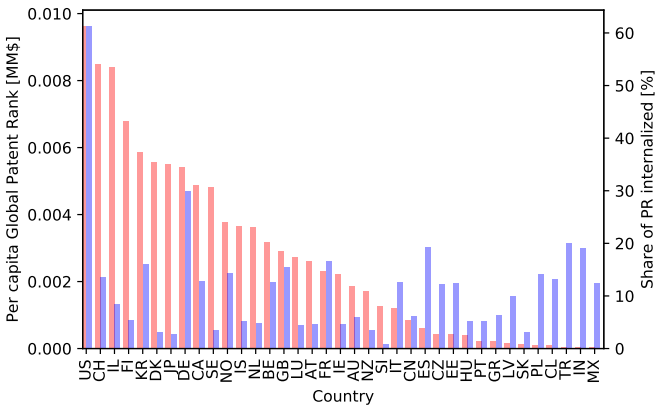
Figure 5: Global versus National Spillovers ($\sigma = 50\%$)



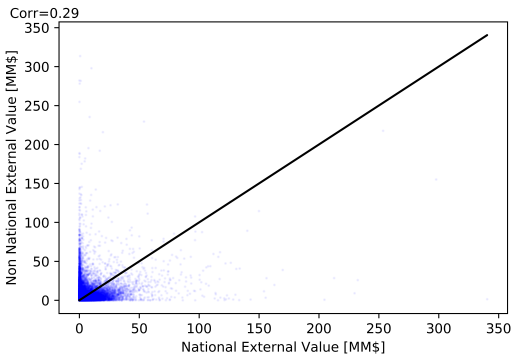
(a) Share of global EV by country



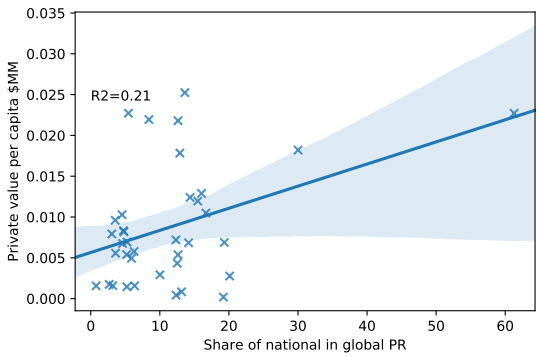
(b) EV across countries and share of global innovations



(c) EV across countries and Share Internalized



(d) National vs International EV



(e) Per capita private value of innovation vs Share Internalized

Panel (b) of Figure 5 shows the mean external value per patent ranked by country of origin. The width of each bar represents the total number of patents from that country. The average patent from Israel creates the highest external value, at more than $2 million per patent. Patents from the US and Canada also create an average of above $2 million per patent.

Panel (c) shows two country rankings. First, on the left-hand side y-axis (and plotted in

red), is the total external value created per capita. The US also tops this ranking, despite being a relatively large country, generating spillovers valued at nearly $10,000 per person. The US is followed by Switzerland and Israel.

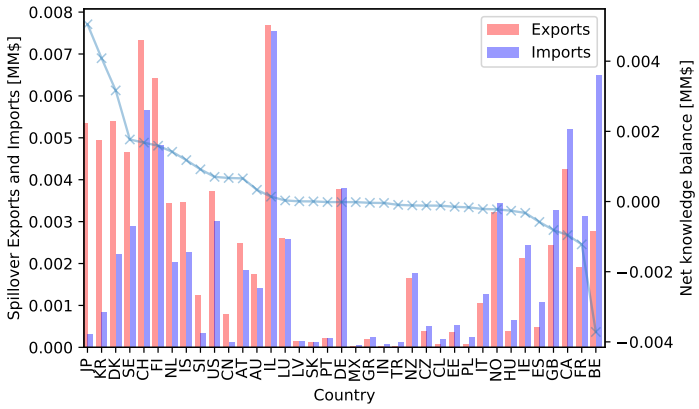
Second, the right-hand side y-axis of Panel (c) introduces the novel measure of the share of external value that is internalized at the country level. To compute this statistic, we take the total external value created by patents in each country and decompose those spillovers into those that reflect value gained—directly or indirectly—by other patents in the same country and those that benefit new patents outside the original country. The figure plots the former divided by the total. The first fact to note is that most of the innovations' external value is created outside the country of the original patent. Only the US internalizes more than 50% of the external values it creates. The next highest-ranked country is Germany where about 30% of spillovers are internalized. All other countries have much lower internalization rates.

Panel (d) presents a scatter plot at the patent level of the external value that flows abroad versus the external value that is retained in the country of innovation.

Panel (e) compares the share of external value that is internalized within the origin country shown in panel (c) with the mean per capita private value of patents in the country. There is a significant and positive correlation between a country's retention share (x-axis) and the per capita private value of innovation (y-axis). While this is not a causal estimate, it is suggestive of the idea that countries might be able to increase their innovative output - and thereby ultimately economic growth - by making sure their innovation environment facilitates a high level of within-country internalisation of innovation spillovers. Measures such as P-RANK can be useful to design policies with that objective.

Figure 6 provides a further perspective on knowledge spillover flows between countries by computing the value of spillover exports and imports. This permits a country-level measure of the "external value balance", which is analogous to the current account balance of trade that is the sum of the value of country exports and the negative of the value of a country imports. This balance is plotted on the right-hand side y-axis.

Figure 6: The balance of knowledge, in per capita terms ($\sigma = 50\%$)



Notes: Imports and exports spillovers consist in the knowledge spillovers a country generates and transfers to innovations within the country and outside of it, respectively. The net knowledge balance is calculated by subtracting imports to exports. All figures are in per capita terms

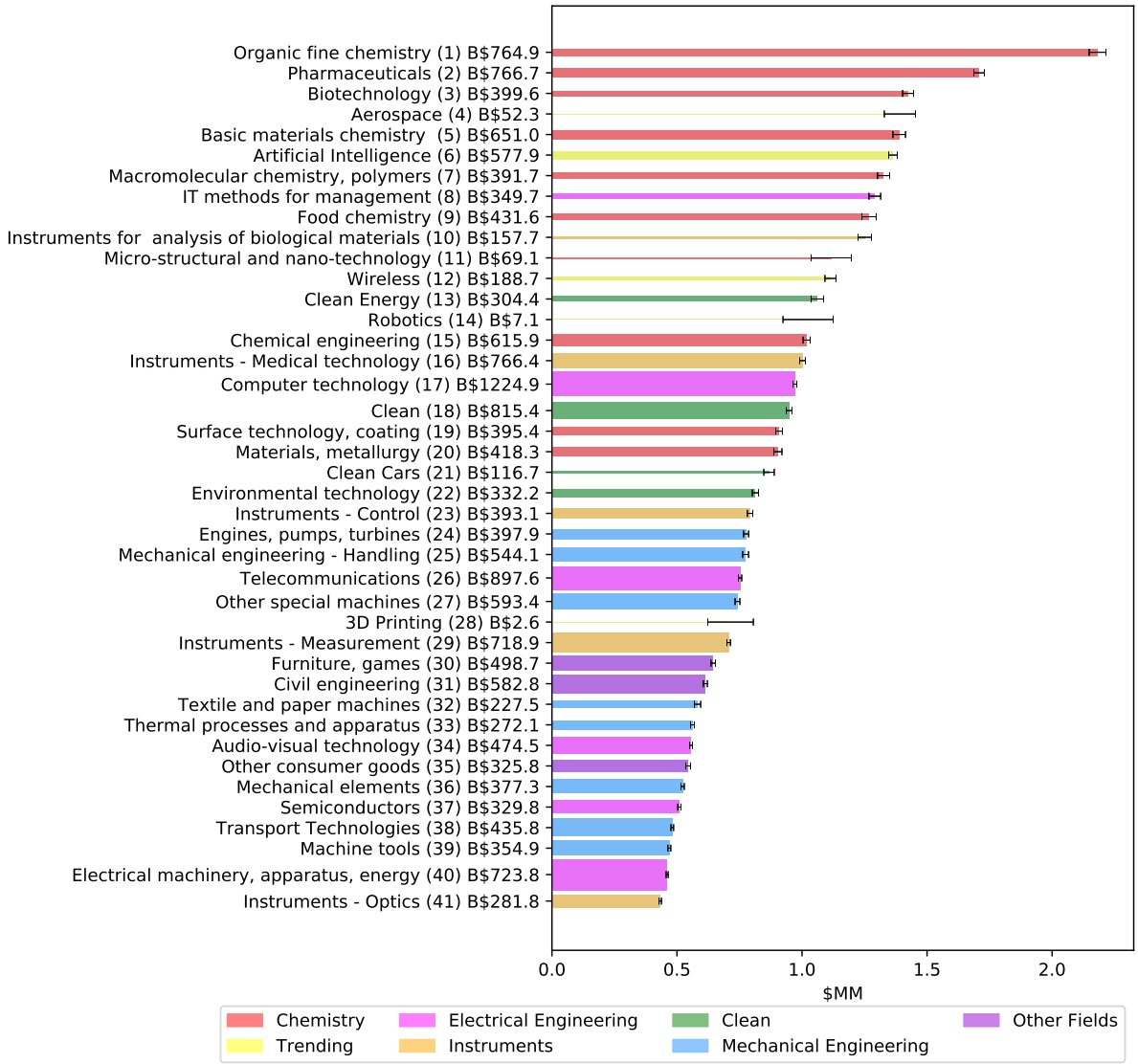
This leads to somewhat different country ranking than we have seen in the other figures in this section. We observe that Japan, Korea and Denmark are the biggest net providers of spillovers. Note that all figures are in per capita terms. Hence, every Japanese citizen provides about $5,000 worth of spillovers to other countries over the ten-year period (2005-2014) used

in this calculation. At the other end of the spectrum, we have the countries such as Belgium, France and Canada who are —per capita— the biggest net importers. Germany emerges as the country in the middle with a net balance that is virtually zero.[17]

Similarly, we can compare external values derived from the P-Rank methodology by technology fields. Figure 7 suggest that there are large and statistically significant differences. We find in particular that Chemistry patents generate large amounts of external value (on average more than $1.5 Million), and the lowest number is that of optics with less than $0.5 Million. Note that the ranking in external values is somewhat in line with the ranking of private values found in Figure 1. This is not entirely surprising: most spillover flows happen within technology categories so that categories with higher private values will also tend to have more valuable spillovers. That said, there are also notable exceptions from this rule.For instance, Aerospace innovations appear more in the bottom end of the ranking of private values whereas they are at the top end of external values. Semiconductors also rank higher in terms of external values.

[17]If one was to conjure up an international system of compensation for spillovers, this measure would provide guidance on the size and direction of country-level transfers.

Figure 7: Global External Value – Categories – All 2005-2014 innovations



Notes: Diagram of the average global external value in millions of CPI-adjusted 1982 US dollars (x-axis) by technology field (y-axis). Calculations are based on the citation network of all innovations for which a patent application was filed in the period 2005-2014. As such, values reflect external values from and to all innovations globally in that period. Width of each bar represents the number innovations in the field. Area of each bar (in billions $) represents total external value in the technology field and is printed next to y-axis labels.

It is also instructive to look at the external value measured via P-Rank for individual innovations. In Table 1, we report the 10 innovations with the highest external value (as measured by P-Rank) in the 2005-2014 period. In line with the average results, we find primarily chemical industry innovations ranked highest, suggesting this industry does not only dominate in terms of average external value generated, but also with respect to the far upper tail of the distribution. However, the top 10 list also notably includes Apple's patent number US2006197753(A1) that outlines the functionality of the iPhone.[18].

[18] "Multifunctional hand-held device" https://patentswarm.com/patents/US20060197753A1

Table 1: Top 10 innovations in terms of EV – All 2005-2014 innovations

Rank	EV	Patent	Applicant (1st)	Title
1	2228.21	JP2009057568(A)	NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL SCIENCE AND TECHNOLOGY	METHOD FOR PRODUCING ZINC OXIDE NANOPARTICLE FOR ULTRAVIOLET LIGHT EMITTING BODY
2	1901.26	US2010076236(A1)	SHELL INTERNATIONALE RESEARCH MAATSCHAPPIJ	PROCESS FOR PRODUCING PARAFFINIC HYDROCARBONS
3	1248.34	US2008305065(A1)	GOLDSCHMIDT	POLYSILOXANES HAVING QUATERNARY AMMONIUM GROUPS, A PROCESS FOR THE PREPARATION THEREOF AND THE USE THEREOF IN CLEANING AND CARE FORMULATIONS
4	1169.84	CN101875855(A)	CHINA PETROLEUM & CHEMICAL CORPORATION SINOPEC BALING COMPANY	METHOD FOR HYDROGENATION AND CATALYTIC CRACKING OF RESIDUAL OIL
5	885.97	US2006269496(A1)	GILLETTE COMPANY	REDUCTION OF HAIR GROWTH
6	842.40	US2006186020(A1)	PETROLEO BRASILEIRO S.A. - PETROBRAS	VEGETABLE OIL HYDROCONVERSION PROCESS
7	790.25	US2006173985(A1)	NEWSILIKE MEDIA GROUP	ENHANCED SYNDICATION
8	783.38	US2006286259(A1)	CADBURY ADAMS USA	TASTE POTENTIATOR COMPOSITIONS AND BEVERAGES CONTAINING SAME
9	738.22	US2006197753(A1)	APPLE	MULTI-FUNCTIONAL HAND-HELD DEVICE
10	658.88	CN101608132(A)	PETROCHINA COMPANY	DELAYED COKING PRODUCTION METHOD UTILIZING ETHYLENE CRACKING TAR AS RAW MATERIAL

3.2 P-Rank's ability to predict future innovation

The motivation for the P-Rank measure is that a patent contributes to the stock of knowledge that can be used as an input in future innovation. This introduces the concept of an innovation production function, which can be used to evaluate whether or not P-Rank does indeed quantify the value of this knowledge input.

If P-Rank captures the input that is valuable for future innovation then we expect there to be a positive association between the P-Rank measure and future patenting activity. Because our data span a long time period, we can test for the presence of this association.

We design a test as follows: we suppose that future patents in technology class a and country c during time period τ are associated with spillovers *received* from the $a - c$ area, and those *generated* by that area, measured by P-rank. We can also allow for future patents in $a - c - \tau$ to be related to the number of previous patents in the $a - c$ area, and the private values of those patents. Further, we can also include a more traditional measure of innovation spillovers—citation counts. In doing so we can evaluate the potential additional explanatory power of the novel elements of P-Rank, that it accounts for, and places values on, both direct and indirect spillovers.[19] Finally, both measures of spillovers received are also broken down into between and within spillovers.

We use the periods 1985-1994, 1995-2004 and 2005-2014 to construct the explanatory variables required for this test. These are the patent count as well as patent value in the area $a - c$, spillovers into the area and spillovers provided by area $a - c$. Two measures of spillovers are included in the estimations: (i) P-Rank, which accounts for direct and direct spillovers and their value, and, (ii) citation counts, which only considers direct spillovers. The dependent variable

[19]This set up is similar to Acemoglu et al. (2016), who evaluate the role of inward and outward citation counts.

is calculated for each year t, with $t \in \tau$. Innovations emerging at year t in $a - c$ are explained by the sum of innovations or spillovers in $a - c$ in the previous period $\tau - 1$.

Specifically, we estimate the following model using data from 1681 $a - c$ cells based on 41 technology classes (as detailed in the appendix), a, and 41 EU and/or OECD countries, c:

$$\begin{aligned}\log\left(\frac{N_{c,a,t\in\tau}}{N_{c,a,\tau-1}}\right) =& \beta_0 + \beta_N \log\left(N_{c,a,\tau-1}\right) + \beta_P \log\left(\frac{PV_{c,a,\tau-1}}{N_{c,a,\tau-1}}\right) \\ &+ \beta_{C^{in}} \log\left(\frac{C^{in}_{c,a,\tau-1}}{N_{c,a,\tau-1}}\right) + \beta_{C^{out}} \log\left(\frac{C^{out}_{c,a,\tau-1}}{N_{c,a,\tau-1}}\right) \\ &+ \beta_{ST^{in}} \log\left(\frac{ST^{in}_{c,a,\tau-1}}{N_{c,a,\tau-1}}\right) + \beta_{ST^{out}} \log\left(\frac{ST^{out}_{c,a,\tau-1}}{N_{c,a,\tau-1}}\right) \\ &+ \alpha_c + \beta_i + \epsilon_{c,a,t}\end{aligned} \tag{14}$$

where $N_{c,a,t\in\tau}$ is the number of innovations in time period τ, for a country c and technology class a. As explanatory variables we include the number of innovations in the previous time window ($N_{c,a,\tau-1}$), the private value of innovations in the previous period ($P_{c,a,\tau-1}$), the number of backward citations—i.e. innovations in an area citing another innovation—($C^{in}_{c,a,\tau-1}$), the number of forward citations ($C^{out}_{c,a,\tau-1}$)—i.e. innovations in an area being cited. We also include the equivalent of that in terms of patent rank; i.e. we include the spillover flows that an area receives($ST^{in}_{c,a,\tau-1}$) as well as the spillover flows an area generates ($ST^{out}_{c,a,\tau-1}$). We also include country and technology area fixed effects.

The dependent and explanatory variables are normalised by the number of innovations in time window $\tau - 1$. We also consider a version of equation 14 where we use the private value of an area $P_{c,a,t}$ for the construction of the dependent variable. Finally, we explore specifications where we distinguish between spillover value received from within the same innovation area ($SW^{in}_{c,a,\tau-1}$) and those received between areas ($SB^{in}_{c,a,\tau-1}$)

The results of these regressions are give in Tables 2 and 3, where the dependent variable is the number of innovations and then the private value of innovations, respectively. The estimated coefficients merit detailed discussion.

Table 2: Estimating the effects of spillovers on the number of future innovations

	Number of innovations			
	(1)	(2)	(3)	(4)
Innovations	−0.320***	−0.310***	−0.306***	−0.269***
	(0.006)	(0.006)	(0.006)	(0.006)
Innovation value	−0.247***	−0.227***	−0.337***	−0.284***
	(0.008)	(0.008)	(0.009)	(0.009)
Citation counts in		−0.029***	−0.088***	−0.060***
		(0.007)	(0.008)	(0.008)
Citation counts out		−0.042***	−0.093***	−0.066***
		(0.006)	(0.009)	(0.009)
Total Prank in (sigma = 0.5)			0.184***	
			(0.009)	
Between Prank in (sigma = 0.5)				0.150***
				(0.009)
Within Prank in (sigma = 0.5)				0.094***
				(0.003)
Prank out (sigma = 0.5)			0.060***	0.040***
			(0.008)	(0.008)
Constant	0.689***	0.632***	0.809***	0.976***
	(0.038)	(0.039)	(0.038)	(0.038)
Observations	50,430	50,430	50,430	50,430
R^2	0.561	0.562	0.570	0.579
Adjusted R^2	0.560	0.561	0.570	0.578

Notes: *p<0.1; **p<0.05; ***p<0.01.
All variables are in log. Prank was calculated based on network without self citation.
The dependent variable per year is normalized by the number of innovation
during the previous period.

Table 3: Estimating the effects of spillovers on the value of future innovations

	Average private value of innovations			
	(1)	(2)	(3)	(4)
Innovations	−0.324***	−0.319***	−0.315***	−0.326***
	(0.010)	(0.010)	(0.010)	(0.010)
Innovation value	−0.010	−0.007	−0.108***	−0.126***
	(0.014)	(0.014)	(0.017)	(0.017)
Citation counts in		0.016	−0.025	−0.035**
		(0.013)	(0.015)	(0.016)
Citation counts out		−0.041***	−0.132***	−0.141***
		(0.012)	(0.018)	(0.018)
Total Prank in (sigma = 0.5)			0.161***	
			(0.016)	
Between Prank in (sigma = 0.5)				0.173***
				(0.017)
Within Prank in (sigma = 0.5)				−0.029***
				(0.007)
Prank out (sigma = 0.5)			0.107***	0.113***
			(0.016)	(0.016)
Constant	1.686***	1.658***	1.832***	1.779***
	(0.072)	(0.072)	(0.073)	(0.074)
Observations	50,430	50,430	50,430	50,430
R^2	0.270	0.270	0.274	0.274
Adjusted R^2	0.269	0.269	0.273	0.273

Notes: *p<0.1; **p<0.05; ***p<0.01.
All variables are in log. Prank was calculated based on network without self citation.
The dependent variable per year is normalized by the number of innovation
during the previous period.

Column 1 includes the number and value of innovations in the previous period. Both variables are negative, which is consistent with mean reversion. That is, those areas that had above average innovation in the previous period still have higher rates of innovation in the subsequent period but not to the same extent. In column 2, we include inward and outward citation counts. The citation count coefficients are significant and negative, which is somewhat

surprising. We would expect counts of past citations to be positively associated with future innovation, as found in Acemoglu et al. (2016). We note that the R^2 value for the regression is very similar in columns 1 and 2, suggesting that including past citations has little explanatory effect in these regressions.

In column 3, we include both P-Rank received and generated in an $a - c$ area in the prior period. Both these variables are positive and significant, evidence that P-Rank does indeed measure the value of knowledge inputs. The coefficient magnitudes suggest that a 10% increase in spillovers received is associated with a 1.84% increase in the number of patents and a 1.61% increase in the value of patents. Moreover, the R^2 measure in column 3 is greater than in columns 1 and 2, suggesting that these variables increase the explanatory power of the regressions.

Finally, in column 4, we differentiate between spillovers received from local (within $a - c$ area) and from imported (between $a - c$ area) knowledge. The coefficient estimates suggest that the knowledge imported from other $a - c$ areas has a larger positive effect on the number and value of future innovations in area $a - c$. In fact, the within effect (local spillovers) is negative in the case of the private value of innovations.[20]

To conclude, P-Rank significantly predicts future innovation, which is consistent with the idea that knowledge spillovers, as measured using the P-Rank methodology, are a valuable input to future innovation production. This measure performs better than direct citation counts.

4 IStra-X

This section describes how P-Rank can inform policy design. The goal is an instrument that directs a given total amount of subsidy into different areas in a way that maximizes the efficiency-enhancing impact of the subsidy. This instrument is called IStra-X (Industrial Strategy Index), and takes the form of a distribution across technology classes.

The question, then, is how should any subsidy amount s be distributed? To answer this question we need to quantify the marginal impact of a small amount of subsidy in every sector, which is likely a function of the level of subsidy it is already receiving.

4.1 A model of innovation

To estimate this impact we need to understand the innovation process within any one technology class. We impose some structural assumptions on how inventors make decision in to order to estimate levels of innovation, and the value of innovation, at each level of subsidy (including the status quo levels of zero new subsidy).

Assume that a new innovation first requires an idea. Project ideas in a given technology class are heterogeneous in quality, distributed as a Pareto distribution with the following probability density function (pdf):

$$f(\delta) = \begin{cases} \frac{\alpha\mu^\alpha}{\delta^{\alpha+1}} & \text{if } \delta > \mu \\ 0 & \text{if otherwise} \end{cases} \tag{15}$$

and the support of this quality distribution is $[\mu, \infty)$. An inventor that has an idea will try to innovate using the idea if it generates private financial gain for her. Her payoff at the time of deciding whether to pursue the idea includes a fixed cost c and also takes into account that the outcome is uncertain. In this version of the paper, we assume that the probability of

[20] As argued in Arthur (2009), the combination of knowledge from different fields tends to lead to more radical innovations. This could explain that spillovers from the same field are less valuable.

innovation success is independent of the idea quality and is a draw from a uniform distribution on the interval $[0, \kappa)$ where $\kappa < 1$. The expected private benefit from innovating conditional on having an idea of quality δ is $PV = \epsilon \times \delta$. Which, because the expected value of ϵ is $\frac{\kappa}{2}$, gives $E\{PV|\delta\} = \frac{\kappa}{2}\delta$.

An inventor chooses to innovate if:

$$E\{PV|\delta\} \geq c$$

Consequently, they will only pursue ideas where $\frac{\kappa}{2}\delta \geq c$. We define λ as the lowest quality idea that will be developed where

$$\lambda = \frac{2c}{\kappa} \tag{16}$$

We are interested in the distribution of idea quality conditional on idea development, which can be written as:

$$f(\delta_i|\delta > \lambda) = \frac{f(\delta_i)}{P(\delta > \lambda)} = \begin{cases} \frac{\alpha\lambda^\alpha}{\delta_i^{\alpha+1}} & \text{if } \delta > \lambda \\ 0 & \text{if otherwise} \end{cases}$$

where $P(\delta > \lambda)$ is the likelihood that any new idea is above the minimum quality required to be developed:

$$P(\delta > \lambda) = \int_\lambda^\infty \frac{\alpha\mu^\alpha}{\delta_i^{\alpha+1}} d\delta_i = \frac{\mu^\alpha}{\lambda^\alpha} \tag{17}$$

This structure forms the benchmark, or status quo, in the industry. We can think of α, κ, and c as area-specific parameters that can be estimated from the data.

We have now sufficient information to estimate the distribution of the private values of the ideas that will be developed. This is given by:

$$P(PV_i = v|\delta > \lambda) = \int \phi(PV_i = v|\delta) f(\delta|\delta > \lambda) d\delta$$

where

$$\phi(PV_i = v|\delta) = \begin{cases} \frac{1}{\delta\kappa} & \text{if } v < \kappa\delta \\ 0 & \text{if } v > \kappa\delta \end{cases} \tag{18}$$

is the density of PV conditional on δ. Together, these expressions yield:[21]

$$P(PV_i = v|\delta > \lambda) = \int_{\max\{\lambda, \frac{v}{\kappa}\}}^\infty \frac{f(\delta)}{\delta\kappa} d\delta = \int_{\max\{\lambda, \frac{v}{\kappa}\}}^\infty \frac{\alpha\lambda^\alpha}{\kappa\delta^{\alpha+2}} d\delta$$

Consequently

$$P(PV_i = v|\delta > \lambda) = \left[-\frac{\alpha\lambda^\alpha}{(\alpha+1)\kappa\delta^{\alpha+1}}\right]_{\max\{\lambda, \frac{v}{\kappa}\}}^\infty = \begin{cases} \frac{\alpha}{(\alpha+1)\kappa\lambda} & \text{if } \lambda > \frac{v}{\kappa} \\ \frac{\alpha\lambda^\alpha\kappa^\alpha}{(\alpha+1)v^{\alpha+1}} & \text{if } \lambda < \frac{v}{\kappa} \end{cases} = \begin{cases} \frac{\alpha}{(\alpha+1)2c} & \text{if } 2c > v \\ \frac{\alpha 2^\alpha c^\alpha}{(\alpha+1)v^{\alpha+1}} & \text{if } 2c < v \end{cases}$$

where the last equality follows from equation 16.

Notice that the density of PV given $\delta > \lambda$ depends only on c and α. This is because c is a sufficient statistic for the combined effect of κ and μ on the density. Hence, we are able to estimate parameters α and c using maximum likelihood using the observed PV_i values derived in Section 2.2.[22] The appendix illustrates how this density function varies with the values of α and c.

[21] Note that $f(\delta_i|\delta > \lambda)$=0 if $\delta < \lambda$ from 15. However, we also have that $\phi(PV_i = v|\delta) = 0$ if $\delta < \frac{v}{\kappa}$. This means that $\phi(PV_i = v|\delta)f(\delta|\delta > \lambda)$ will be zero if either of those conditions is binding. This is the reason for the $\max\{\lambda, \frac{v}{\kappa}\}$ expression that is the lower bound of integration.

[22] Because the likelihood function won't be differentiable, we rely on a genetic algorithm to find the optimum.

We can also work out the expected value of the distribution of conditional private values. This is:

$$E\{PV_i|\delta > \lambda\} = \left[\frac{\alpha}{\alpha+1}\frac{v^2}{4c}\right]_0^{2c} + \left[-\frac{\alpha 2^\alpha c^\alpha}{(\alpha+1)(\alpha-1)v^{\alpha-1}}\right]_{2c}^{\infty} = \frac{\alpha c}{\alpha+1} + \frac{\alpha 2c}{(\alpha-1)(\alpha+1)} = \frac{\alpha c}{\alpha-1} \tag{19}$$

The cumulative density is given by:

$$P(PV_i \leq v|\delta > \lambda) = \begin{cases} \frac{\alpha v}{(\alpha+1)2c} & \text{if } 2c > v \\ \frac{\alpha}{(\alpha+1)} + \int_{2c}^{v} \frac{2^\alpha \alpha c^\alpha}{(\alpha+1)w^{\alpha+1}} dw & \text{if } 2c < v \end{cases}$$

We note that

$$\int_{2c}^{v} \frac{2^\alpha \alpha c^\alpha}{(\alpha+1)w^{\alpha+1}} dw = \left[-\frac{2^\alpha c^\alpha}{(\alpha+1)w^\alpha}\right]_{2c}^{v} = \frac{1}{(\alpha+1)} - \frac{2^\alpha c^\alpha}{(\alpha+1)v^\alpha},$$

which means that the cumulative density is

$$P(PV_i \leq v|\delta > \lambda) = \Phi^{PV}(v) = \begin{cases} \frac{\alpha v}{(\alpha+1)2c} & \text{if } 2c > v \\ 1 - \frac{2^\alpha c^\alpha}{(\alpha+1)v^\alpha} & \text{if } 2c < v \end{cases}$$

We can invert this to find quantiles of the distribution. Note that $\Phi^{PV}(2c) = \frac{\alpha}{(\alpha+1)}$. Hence, the p quantile is given by:

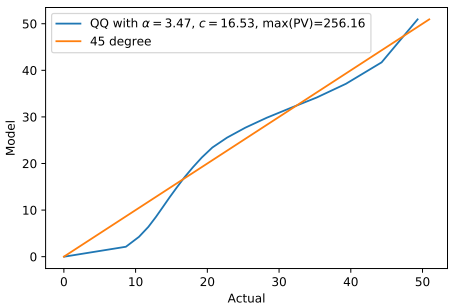
$$Q^{PV}(p) = \begin{cases} p\frac{(\alpha+1)2c}{\alpha} & \text{if } \frac{\alpha}{(\alpha+1)} > p \\ \frac{2c}{(\alpha+1)^{\frac{1}{\alpha}}(1-p)^{\frac{1}{\alpha}}} & \text{if } \frac{\alpha}{(\alpha+1)} < p \end{cases}$$

The data give us a p-quantile value for every technology class. Hence, we can estimate parameter values for α and c by matching the model quantiles with the data.

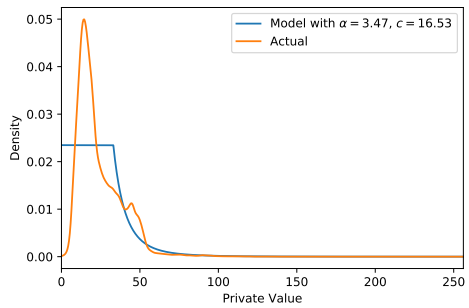
To illustrate this process, the figures here focus on German Chemistry innovations. The estimated parameter values for this area produce the blue lines in the graphs. The orange lines present the actual data.

Figure 8: Actual vs modeled PV distributions (Chemistry,DE)

(a) QQ plot



(b) Density



4.2 Quantifying the marginal impact of a subsidy

We now turn to policy instruments. We can think of a unit of subsidy s as changing the private cost of idea development. Hence, s directed to a given set of inventors becomes a variable that

will affect the quality distribution of ideas that are developed. We let $c' = c - s < c$, be the subsidized cost of idea development. This gives us a new minimum quality threshold, $\lambda' < \lambda$. More ideas will be developed. Our goal is to quantify the total value of these news ideas. To do this we develop a measure of the marginal social value, where rather than evaluate values at c and c', we measure the marginal impact of any change in c.

A policy maker cares about the change in the sum of total private value and total external value, less the idea development costs of the marginal innovations. We can write this as follows:

$$E\{V\} = E\{PV - c + EV|\delta > \lambda\} P(\delta > \lambda) \tag{20}$$

To examine the effect of (further) subsidies we need to quantify

$$\begin{aligned} \frac{\partial E\{V\}}{\partial s} &= \frac{\partial E\{V\}}{\partial c}\frac{\partial c}{\partial s} \\ &= \left[\left(\frac{\partial E\{PV|\delta > \lambda\}}{\partial c} + \frac{\partial E\{EV|\delta > \lambda\}}{\partial c} - 1\right) P(\delta > \lambda)\right. \\ &\quad \left. + E\{PV - c + EV|\delta > \lambda\} \frac{\partial P(\delta > \lambda)}{\partial c}\right] \frac{\partial c}{\partial s} \end{aligned} \tag{21}$$

We can show that the marginal net total social value with respect to the subsidy level is:

Proposition 4.1. *Derivative of Expected Social Value*

$$\frac{\partial E\{V\}}{\partial s} = E\{c + EV(\alpha - \alpha \times \mathbb{I}\{v > 2c\} + \mathbb{I}\{v < 2c\} - \alpha)|\delta > \lambda\} \frac{P(\delta > \lambda)}{c} \tag{22}$$

Appendix C contains the proof of this proposition.

One final consideration is that equation 22 only reports the marginal benefit of increasing the policy threshold. We note that the same change of support level s will require different amounts of actual support depending on the likelihood of worthwhile ideas emerging. Hence, ex ante expected government costs S of a hypothetical ex post (after idea generation) support level of s will amount to

$$E\{S\} = P(\delta > \lambda)s$$

with

$$\frac{\partial E\{S\}}{\partial s} = P(\delta > \lambda) + \frac{\partial c}{\partial s}\frac{\alpha}{c}P(\delta > \lambda)s = P(\delta > \lambda) - \frac{\alpha}{c}P(\delta > \lambda)s$$

where we are using the result in equation 32. Note that this depends on the level of support already granted. If existing support is non existing it simplifies to $\frac{\partial E\{S\}}{\partial s} = P(\delta > \lambda)$

This allows us to work out a harmonised measure of the expected net benefit of a fixed amount of government spending across different technology areas, and this is the Industrial Strategy Index:

$$Section 4.1 illustrated how to derive estimates of$$

α and c. We can calibrate the probability of having a worthwhile idea for a technology area a as the share of innovations in the area relative to the other areas; i.e.

$$P_a(\hat{\delta} > \lambda) = \rho_a = \frac{\#a}{\sum_{a \in A} \#a}$$

where $\#a$ is the size of technology area a (number of innovations in a) and A is the set of all technology areas.

Hence we can easily estimate $\frac{\partial E\{PV|\delta>\lambda\}}{\partial c}$ by the average private values divided by our estimate of c

$$\left.\frac{\partial E\left\{P\hat{V}|\delta>\lambda\right\}}{\partial c}\right|_a = \frac{1}{c}\frac{1}{\#a}\sum_{i\in A} PV_i$$

Similarly we can compute

$$E\left\{EV \times \mathbb{I}\{P\hat{V} > 2c\}|\delta > \lambda\right\}\Big|_a = \frac{1}{\#a}\sum_{i\in A} EV_i \times \mathbb{I}\{PV_i > 2c\}$$

and

$$E\left\{EV \times \mathbb{I}\{P\hat{V} < 2c\}|\delta > \lambda\right\}\Big|_a = \frac{1}{\#a}\sum_{i\in A} EV_i \times \mathbb{I}\{PV_i < 2c\}$$

Combining, we obtain:

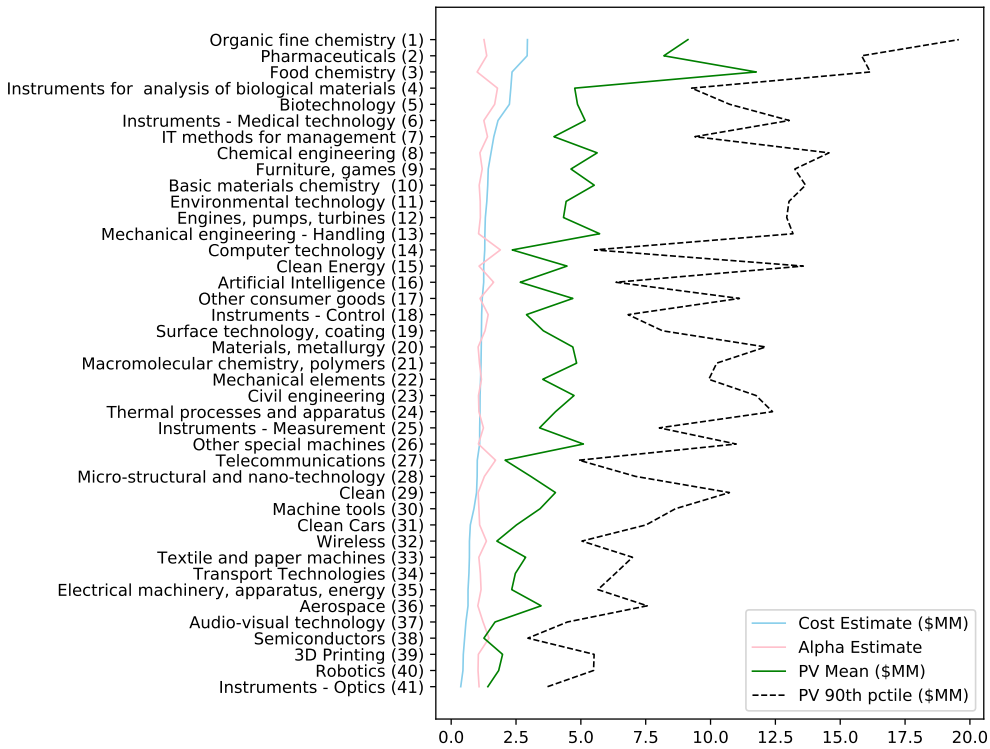
$$\left.\frac{\partial E\hat{\{V\}}}{\partial s}\right|_a = \frac{\rho_a}{\hat{c}_a}\frac{1}{\#a}\sum_{i\in A}\hat{\alpha}_a\left(PV_i + EV_i\right) - \quad PV_i - EV_i\left(\hat{\alpha}_a \times \mathbb{I}\{PV_i > 2\hat{c}_a\} - \mathbb{I}\{PV_i < 2\hat{c}_a\}\right)$$

This brings to a close our description of the IStra-X policy instrument.

4.3 IStra-X Results

IStra-X combines the external value information illustrated in Section 3 with estimates of the responsiveness of innovation to governments subsidies, which depends on the R&D costs associated with an innovative step across technology fields. Figure 9 summarizes our estimates of these costs. As explained above, we derive these from a simple structural model of innovation that we fit to the distribution of observed private values within a sector-country cell. Hence, Figure 9 also reports various statistics on private values. The figure orders technologies by the estimated R&D cost of a research project. This shows that there is a positive relationship between cost and average private value. In particular, the chemical and pharmaceutical sectors rank highly. However, we also see that the relationship is far from monotone.

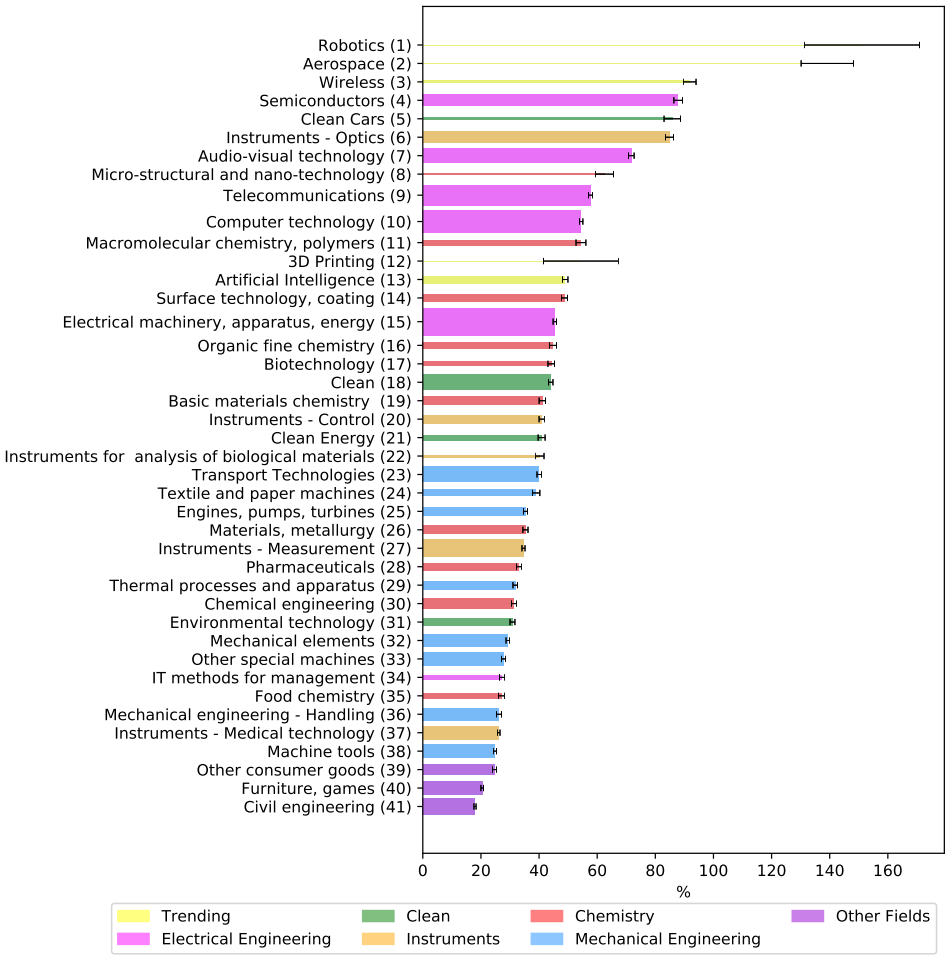
Figure 9: Cost Diagram – Categories – All 2005-2014 innovations



Notes: Diagram of various estimates relevant to the calculation of the IStra-X indicator. The blue line shows the estimated cost (in millions of CPI-adjusted 1982 US dollars) of pursuing an innovation idea by technology field (y-axis). The green and dashed lines show the mean and 90th percentile of the estimated private returns distribution in each field (in in millions of CPI-adjusted 1982 US dollars). The pink line shows the estimate of parameter alpha for each technology field. Calculations are based on innovations for which a patent application was filed in the period 2005-2014.

Figure 10 reports calculations of IStra-X across all innovations by technology field. This suggests a rather different ranking from either cost estimates or external value figures. Categories such as Robotics, Aerospace and Wireless are now at the top with returns of public money of more than 100%.

Figure 10: Global IStra-X – Categories – All 2005-2014 innovations



Notes: Diagram of the global IStra-X indicator representing the return on public spending (x-axis) by technology field (y-axis). Calculations are based on the global population of innovations for which a patent application was filed in the period 2005-2014. The x-axis gives the expected total welfare effect – as a return on investment – of a decrease in the cost of pursuing an innovation idea, broken down by technology field. IStra-X values reflect the difference between the expected increase in total value (private returns as well as external values from knowledge spillovers) generated by innovations in a field and the expected cost of the subsidy, scaled by the expected cost of the subsidy. Width of each bar represents the number innovations in the field.

While the global perspective of Figure 10 is instructive, it is not the level at which industrial policy is undertaken. Much of actual policy making occurs at the national level, although there are also efforts at the supranational level – e.g. within the EU – or at the sub-national/regional level. Hence, to relate the results to actual policy initiatives, we report in the following results at the country level as well. In this report, we only illustrate country-specific results for a selected set of economies. [23] For each country, we report two versions of IStra-X: In the first case, taking into account global external values, and, in the second, we only take into account national external values (although they might be indirect as discussed above). For comparison, we also report the following results for each country separately:

- R&D cost estimates
- Estimates of Relative Technological Advantage (RTA). This is the share of a country's

[23] Because reporting results for all countries or even all OECD and EU countries would dramatically inflate this report, we will make complete country-by-country results available online.

innovation in a technology relative to the global share of that category. This measure provides information on whether a category is more or less important in a country than globally and can be interpreted as a measure of specialization or comparative advantage.

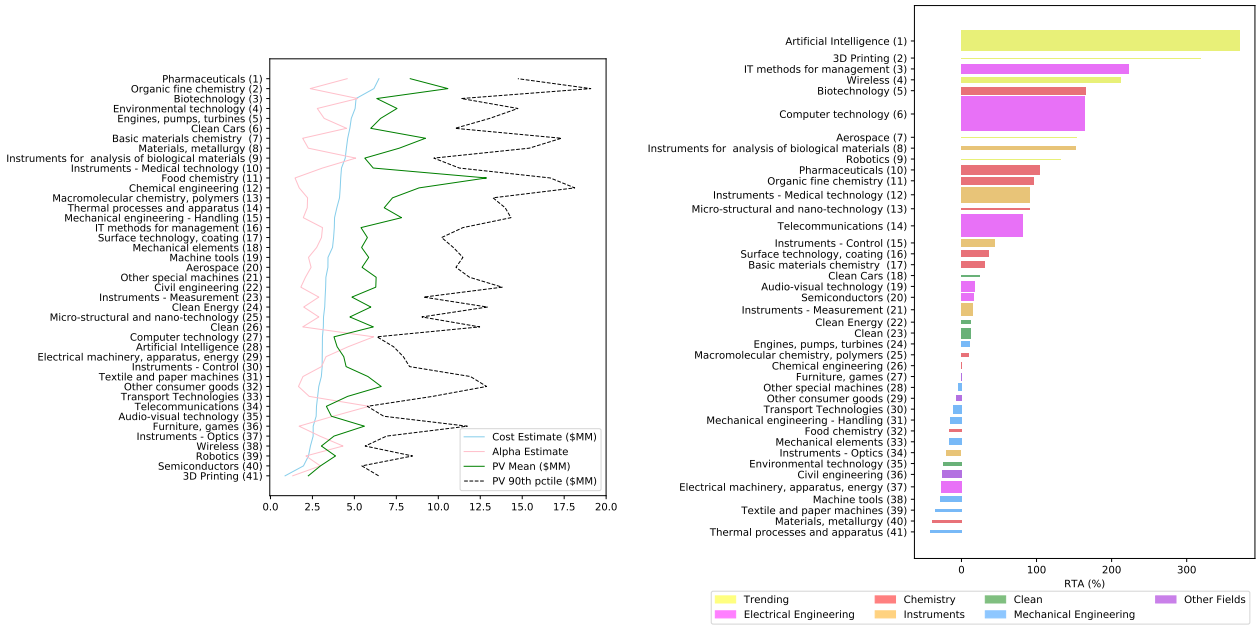
- Global as well as national external values
- Global as well as national IStra-X

Overall, this shows that there is considerable heterogeneity and variation both between countries but also between different measures. Technologies where countries have a comparative advantage are rarely the technologies that generate high levels of spillovers. Similarly, the marginal return to government subsidies (IStra-X) is not necessarily highest in the categories that generate the highest levels of spillover. Finally, what makes sense from a global point view is rarely equivalent to national priorities.

4.3.1 Technological Fields - United States

Here we examine P-RANK and IStra-X results for the US along with related indicators. Figure 11 (Panel b) shows that the US is particularly specialized in sectors such as AI, 3D printing as well as various other ICT technologies where US shares of patenting are several times higher than the global average (e.g. more than 300% higher for AI).

Figure 11: Cost Diagram and RTA – Categories – United States



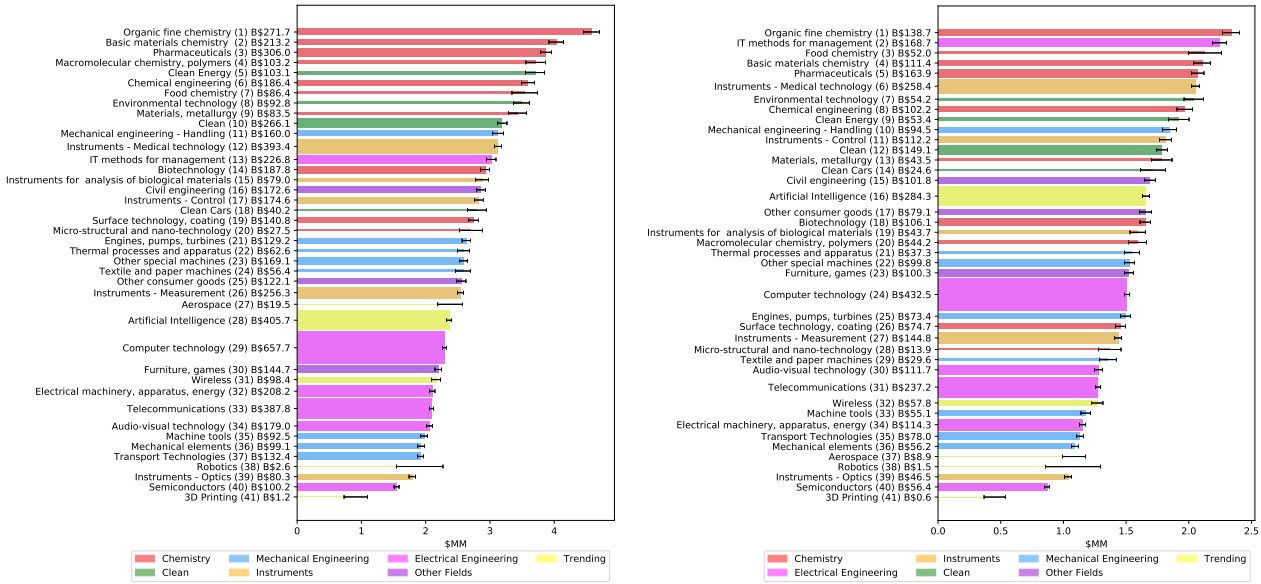
(a) Costs, Alpha & PV (mean and 90th pct.)

(b) Relative Technological Advantage (RTA)

Notes: Calculations are based on innovations originating in the United States for which a patent application was filed in the period 2005-2014. The left-hand panel shows the diagram of various estimates relevant to the calculation of the IStra-X indicator. The blue line shows the estimated cost (in millions of CPI-adjusted 1982 US dollars) of pursuing an innovation idea by technology field (y-axis). The green and dashed lines show the mean and 90th percentile of the estimated private returns distribution in each field (in in millions of CPI-adjusted 1982 US dollars). The pink line shows the estimate of parameter alpha for each technology field. The right-hand panel shows the Relative Technological Advantage of the United States for each technology field. It is calculated as the ratio between the share of the technology field in all innovations from the United states and the share of the technology field in all innovations globally. The resulting ratio is subtracted by one and multiplied by 100 to obtain a percentage value.

Figure 12 by contrast reports the average external value both globally and within the US (i.e. nationally). This leads to a very different ranking. Globally, various chemical industry categories provide the highest average value. Clean energy is also notably high on position 5. Nationally, things look quite a bit different, although Organic Fine Chemistry comes first in both cases. However, we now find IT for management in second position with an average of about $2 Million. This is somewhat in line with RTA figures, although the other ICT technologies are much further behind — either in terms of global or national EV.

Figure 12: External Value – Categories – United States



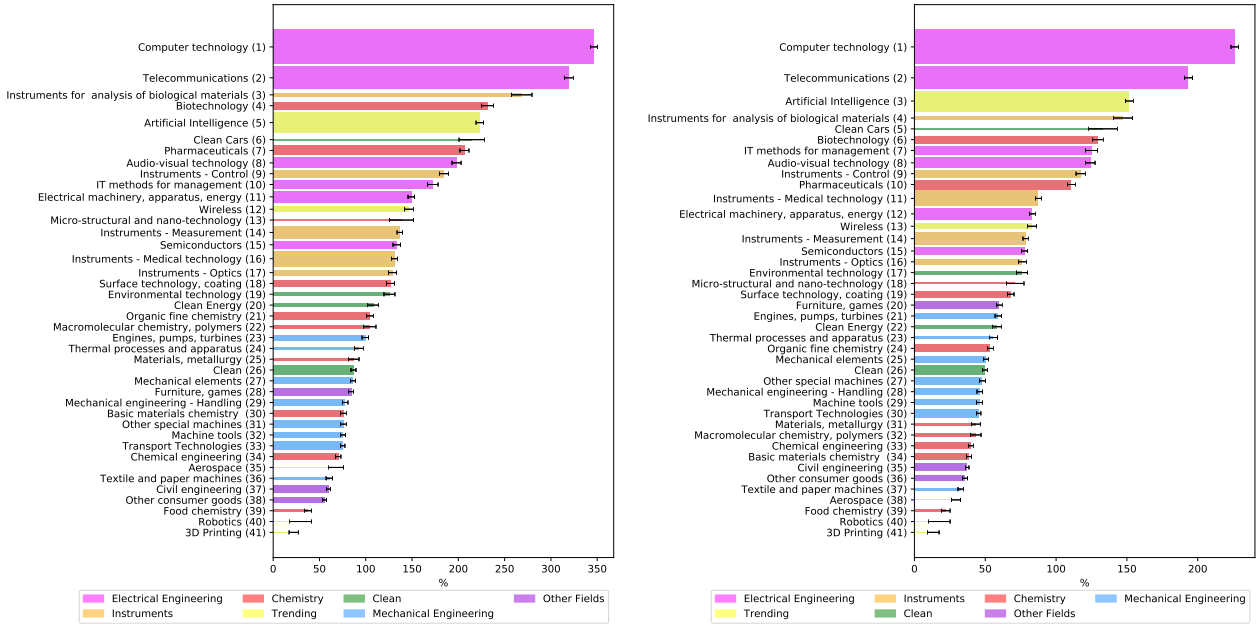
(a) Global

(b) National

Notes: Diagrams of the average external value in millions of CPI-adjusted 1982 US dollars (x-axis) by technology field (y-axis) in the United States. Calculations in the left-hand panel are based on the external value to the global population of innovations generated by innovations originating in the United States (for the period 2005-2014). External values in the right-hand panel are restricted to external value generated *within* the United States. Width of each bar represents the number innovations in the field. Area of each bar (in billions $) represents total external value in the technology field and is printed next to y-axis labels.

Moving on to IStra-X in Figure 13, we find the highest returns among ICT-related sectors again, although the precise ranking is markedly different from the RTA figures. From a national industrial strategy point of view, the numbers suggest a particular focus on general Computer Technology, Telecommunications as well as ICT with returns in excess of 150% in terms of social value generated.

Figure 13: IStra-X – Categories – United States



(a) Global

(b) National

Notes: Diagrams of the global IStra-X indicator representing the return on public spending (x-axis) by technology field (y-axis) in the United States. Calculations are based on the global population of innovations for which a patent application was filed in the period 2005-2014. For the left-hand panel, calculations are based on external value of spillovers to all innovations worldwide during this period. The right-hand panel shows the results when only taking into account external value generated for innovations *within* the United States. The x-axis gives the expected total welfare effect – as a return on investment – of a decrease in the cost of pursuing an innovation idea, broken down by technology field. IStra-X values reflect the difference between the expected increase in total value (private returns as well as external values from knowledge spillovers) generated by innovations in a field and the expected cost of the subsidy, scaled by the expected cost of the subsidy. Width of each bar represents the number innovations in the field.

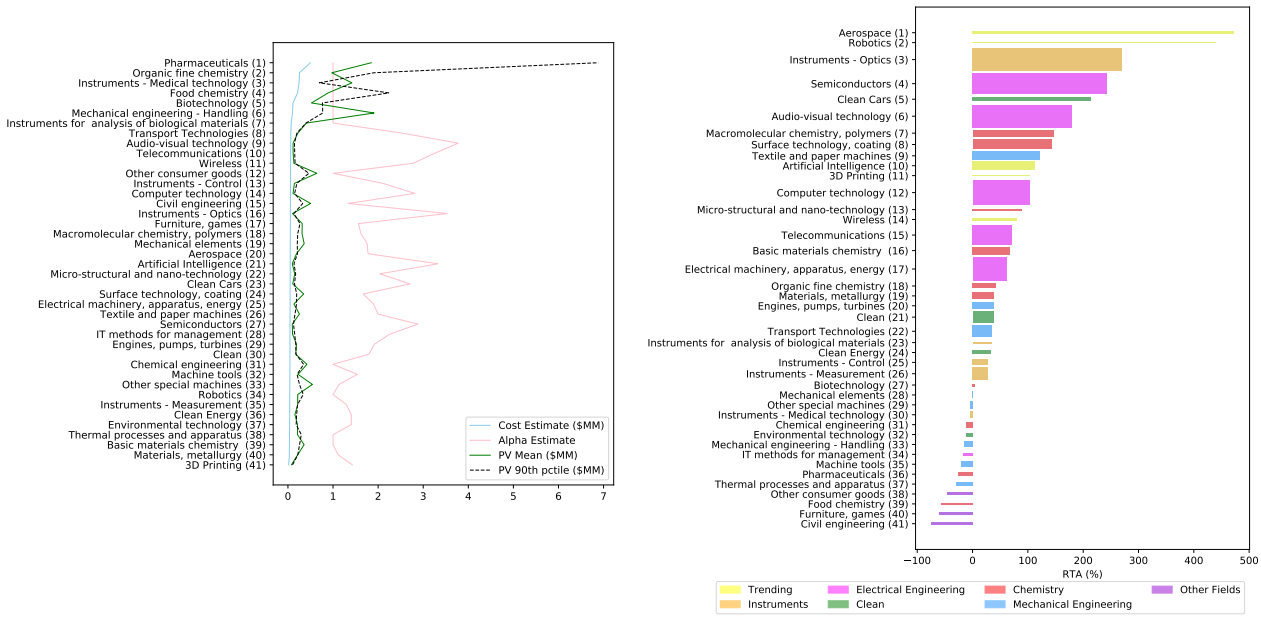
4.3.2 Technological Fields - Japan

Not entirely surprisingly, Japan has a comparative advantage in sectors such as Robotics, Optics and Semiconductors. Somewhat surprisingly, Aerospace is also ranked highly. Looking at IStra-X (Figure 16), it is also dominated by various ICT technologies, with some notable exceptions; e.g. clean cars generates very high global returns, whereas optics is better from a national perspective. Japan is unique in that it generates dramatic returns when considering the global return of R&D subsidies (Panel a of Figure 16), where we find returns in excess of 1000%. Returns are more in line with what we find in other countries when only considering national spillovers (Panel b). Hence, there would be a strong case for innovation subsidies for inventors in Japan from other countries. Presumably, this is a consequence of the very high net spillover balance found in Figure 6.

Table 4: Top 10 innovations in terms of EV – United States

Rank	EV	Patent	Applicant (1st)	Title
1	885.97	US2006269496(A1)	GILLETTE COMPANY	REDUCTION OF HAIR GROWTH
2	790.25	US2006173985(A1)	NEWSILIKE MEDIA GROUP	ENHANCED SYNDICATION
3	738.22	US2006197753(A1)	APPLE	MULTI-FUNCTIONAL HAND-HELD DEVICE
5	575.93	US2007004978(A1)	COMPUMEDICS	SENSOR ASSEMBLY WITH CONDUCTIVE BRIDGE
6	564.96	US2007197695(A1)	SIGMA-ALDRICH COMPANY	STABILIZED DEUTEROBORANE-TETRAHYDROFURAN COMPLEX
7	488.12	US2007171921(A1)	CITRIX SYSTEMS	METHODS AND SYSTEMS FOR INTERACTING, VIA A HYPERMEDIUM PAGE, WITH A VIRTUAL MACHINE EXECUTING IN A TERMINAL SERVICES SESSION
8	480.06	US2007082929(A1)	AUSPEX PHARMACEUTICALS	INHIBITORS OF THE GASTRIC H+, K+-ATPASE WITH ENHANCED THERAPEUTIC PROPERTIES
9	471.07	US2009163545(A1)	UNIVERSITY OF ROCHESTER	METHOD FOR ALTERING THE LIFESPAN OF EUKARYOTIC ORGANISMS
10	463.45	US2006188457(A1)	AQUEA SCIENTIFIC CORPORATION	BODYWASHES CONTAINING ADDITIVES

Figure 14: Cost Diagram and RTA – Categories – Japan

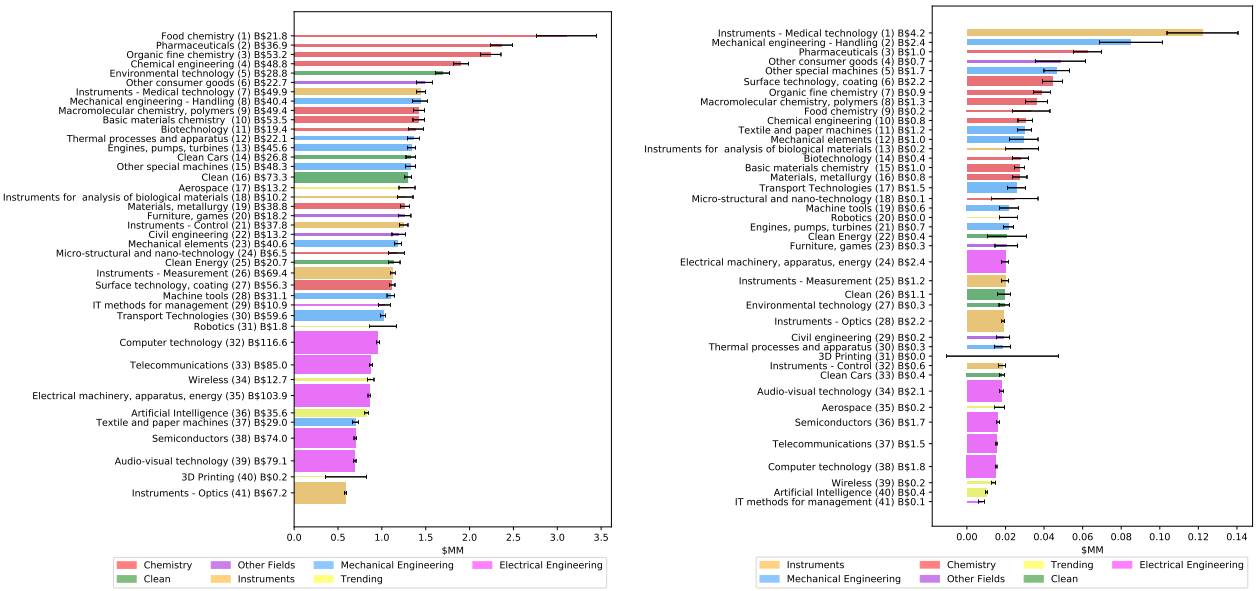


(a) Costs, Alpha & PV (mean and 90th pct.)

(b) Relative Technological Advantage (RTA)

Notes: Calculations are based on innovations originating in Japan for which a patent application was filed in the period 2005-2014. The left-hand panel shows the diagram of various estimates relevant to the calculation of the IStra-X indicator. The blue line shows the estimated cost (in millions of CPI-adjusted 1982 US dollars) of pursuing an innovation idea by technology field (y-axis). The green and dashed lines show the mean and 90th percentile of the estimated private returns distribution in each field (in in millions of CPI-adjusted 1982 US dollars). The pink line shows the estimate of parameter alpha for each technology field. The right-hand panel shows the Relative Technological Advantage of Japan for each technology field. It is calculated as the ratio between the share of the technology field in all innovations from Japan and the share of the technology field in all innovations globally. The resulting ratio is subtracted by one and multiplied by 100 to obtain a percentage value.

Figure 15: External Value – Categories – Japan

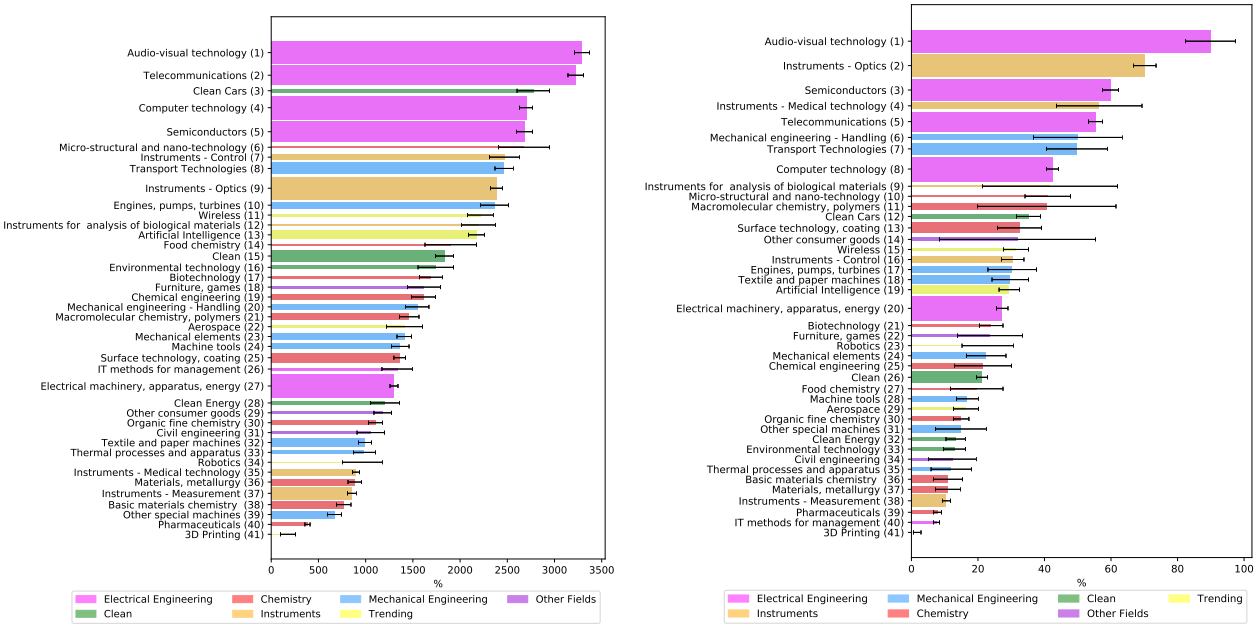


(a) Global

(b) National

Notes: Diagrams of the average external value in millions of CPI-adjusted 1982 US dollars (x-axis) by technology field (y-axis) in Japan. Calculations in the left-hand panel are based on the external value to the global population of innovations generated by innovations originating in Japan (for the period 2005-2014). External values in the right-hand panel are restricted to external value generated *within* Japan. Width of each bar represents the number innovations in the field. Area of each bar (in billions $) represents total external value in the technology field and is printed next to y-axis labels.

Figure 16: IStra-X – Categories – Japan



(a) Global (b) National

Notes: Diagram of the global IStra-X indicator representing the return on public spending (x-axis) by technology field (y-axis) in Japan. Calculations are based on the global population of innovations for which a patent application was filed in the period 2005-2014. For the left-hand panel, calculations are based on external value of spillovers to all innovations worldwide during this period. The right-hand panel shows the results when only taking into account external value generated for innovations *within* Japan. The x-axis gives the expected total welfare effect – as a return on investment – of a decrease in the cost of pursuing an innovation idea, broken down by technology field. IStra-X values reflect the difference between the expected increase in total value (private returns as well as external values from knowledge spillovers) generated by innovations in a field and the expected cost of the subsidy, scaled by the expected cost of the subsidy. Width of each bar represents the number innovations in the field.

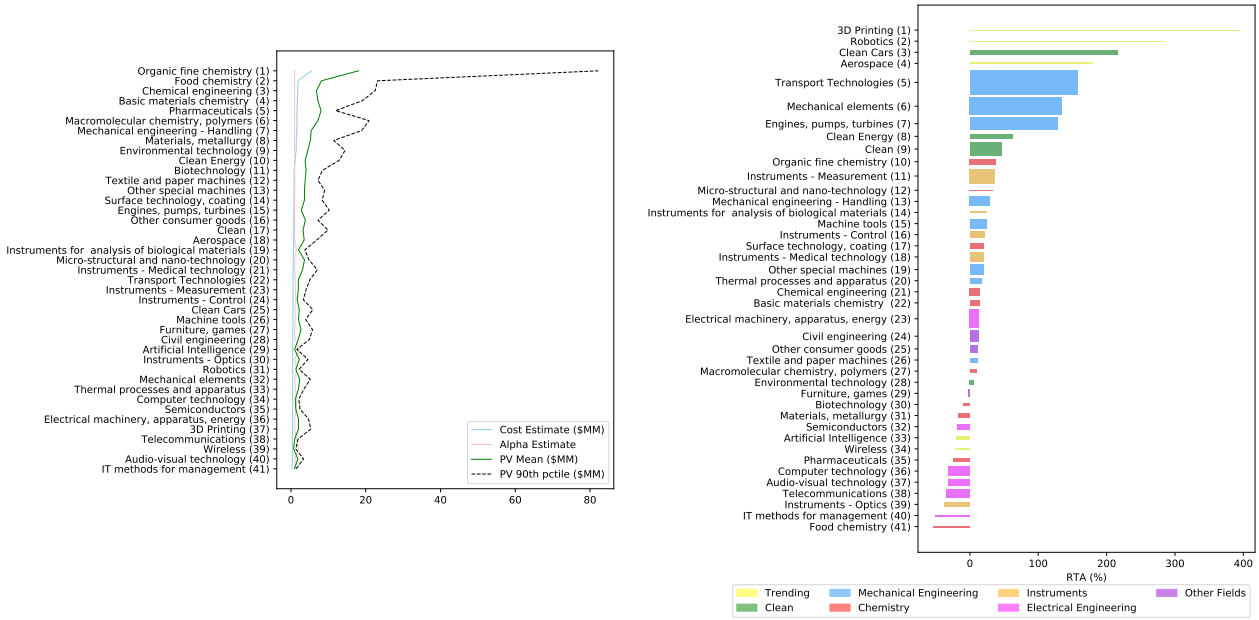
Table 5: Top 10 innovations in terms of EV – Japan

Rank	EV	Patent	Applicant (1st)	Title
1	558.29	US2010087692(A1)	SHOWA DENKO	HYDROGENATION METHOD AND PETROCHEMICAL PROCESS
2	516.42	US2012021110(A1)	KIRIN BEER	FLAVOR OF UNFERMENTED BEER-FLAVORED MALT BEVERAGE
3	431.28	US2007047851(A1)	OSHIO INDUSTRY COMPANY	SELF-STANDING BAG AND MANUFACTURING METHOD THEREOF
4	325.44	US2007003679(A1)	SATO PHARMACEUTICAL COMPANY	SWEETENER COMPRISING A STEVIA-DERIVED SWEET SUBSTANCE
5	296.04	CN101382920(A)	KYORAKU INDUSTRIAL COMPANY	ACCESS CONTROL DEVICE, ACCESS CONTROL METHOD AND ACCESS CONTROL PROGRAM
6	292.28	US2011047512(A1)	TASHIRO, KOICHI	DISPLAY DEVICE AND DISPLAY METHOD
7	282.77	JP2006219233(A)	MURATA MACHINERY	CARRYING DEVICE
8	280.60	WO2007063879(A1)	NIPPON OIL CORPORATION	HYDROREFINING PROCESS AND HYDROREFINED OIL
9	278.41	US2009182091(A1)	KANEKA CORPORATION	CURABLE COMPOSITION
10	274.38	WO2007058315(A1)	ASYST SHINKO	STOCKER

4.3.3 Technological Fields - Germany

Figure 17 shows that comparative advantage for Germany tends to be in mechanical engineering. External value rankings are dominated by chemical technologies (Figure 18). This carries over to IStra-X only when a national perspective is adopted (Figure 19). By contrast, when looking at Global IStra-X, Germany generates the highest returns in various ICT technologies, such as Artificial Intelligence in particular and Computer Technology in general.

Figure 17: Cost Diagram and RTA – Categories – Germany

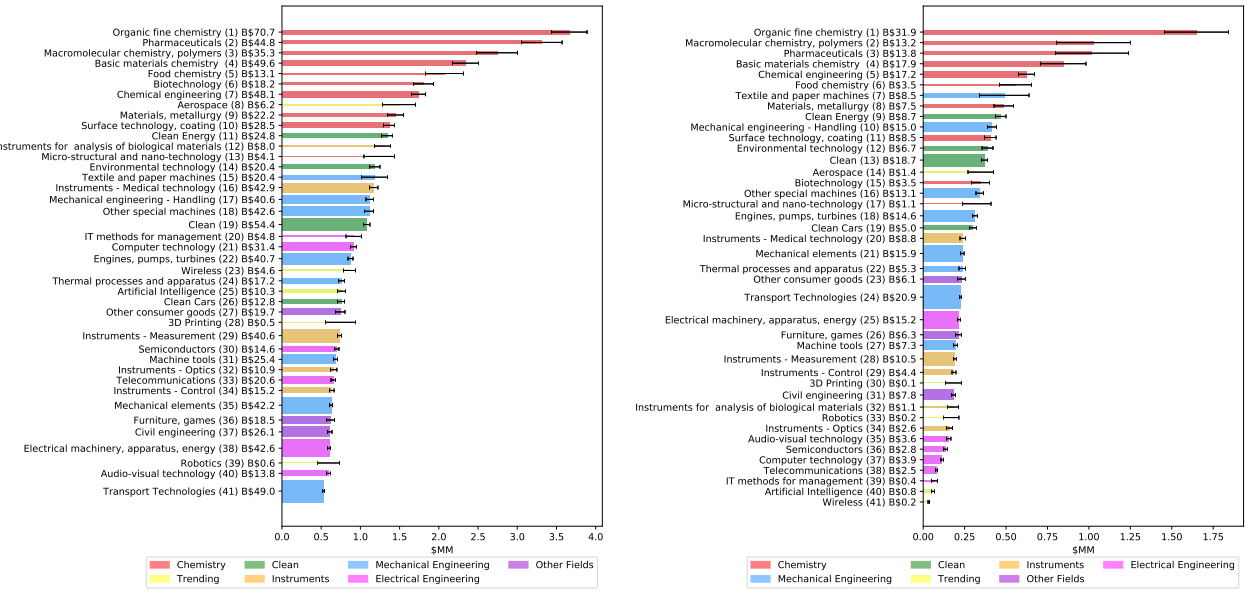


(a) Costs, Alpha & PV (mean and 90th pct.)

(b) Relative Technological Advantage (RTA)

Notes: Calculations are based on innovations originating in Germany for which a patent application was filed in the period 2005-2014. The left-hand panel shows the diagram of various estimates relevant to the calculation of the IStra-X indicator. The blue line shows the estimated cost (in millions of CPI-adjusted 1982 US dollars) of pursuing an innovation idea by technology field (y-axis). The green and dashed lines show the mean and 90th percentile of the estimated private returns distribution in each field (in in millions of CPI-adjusted 1982 US dollars). The pink line shows the estimate of parameter alpha for each technology field. The right-hand panel shows the Relative Technological Advantage of Germany for each technology field. It is calculated as the ratio between the share of the technology field in all innovations from Germany and the share of the technology field in all innovations globally. The resulting ratio is subtracted by one and multiplied by 100 to obtain a percentage value.

Figure 18: External Value – Categories – Germany

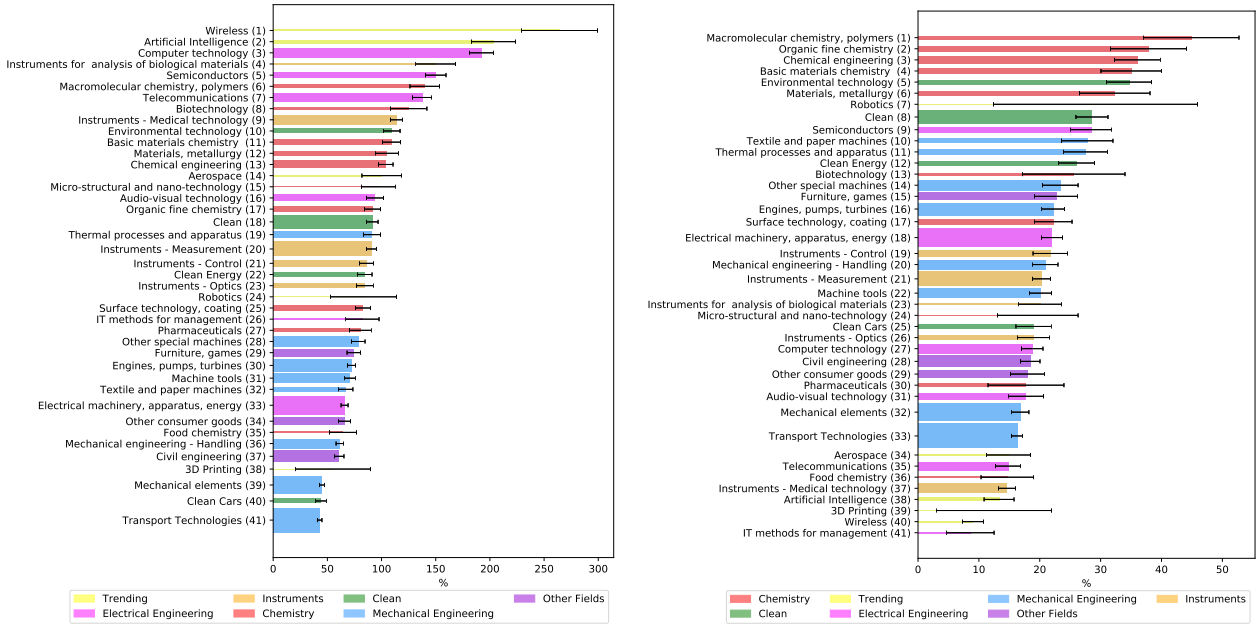


(a) Global

(b) National

Notes: Diagrams of the average external value in millions of CPI-adjusted 1982 US dollars (x-axis) by technology field (y-axis) in Germany. Calculations in the left-hand panel are based on the external value to the global population of innovations generated by innovations originating in Germany (for the period 2005-2014). External values in the right-hand panel are restricted to external value generated *within* Germany. Width of each bar represents the number innovations in the field. Area of each bar (in billions $) represents total external value in the technology field and is printed next to y-axis labels.

Figure 19: IStra-X – Categories – Germany



(a) Global

(b) National

Notes: Diagram of the global IStra-X indicator representing the return on public spending (x-axis) by technology field (y-axis) in Germany. Calculations are based on the global population of innovations for which a patent application was filed in the period 2005-2014. For the left-hand panel, calculations are based on external value of spillovers to all innovations worldwide during this period. The right-hand panel shows the results when only taking into account external value generated for innovations *within* Germany. The x-axis gives the expected total welfare effect – as a return on investment – of a decrease in the cost of pursuing an innovation idea, broken down by technology field. IStra-X values reflect the difference between the expected increase in total value (private returns as well as external values from knowledge spillovers) generated by innovations in a field and the expected cost of the subsidy, scaled by the expected cost of the subsidy. Width of each bar represents the number innovations in the field.

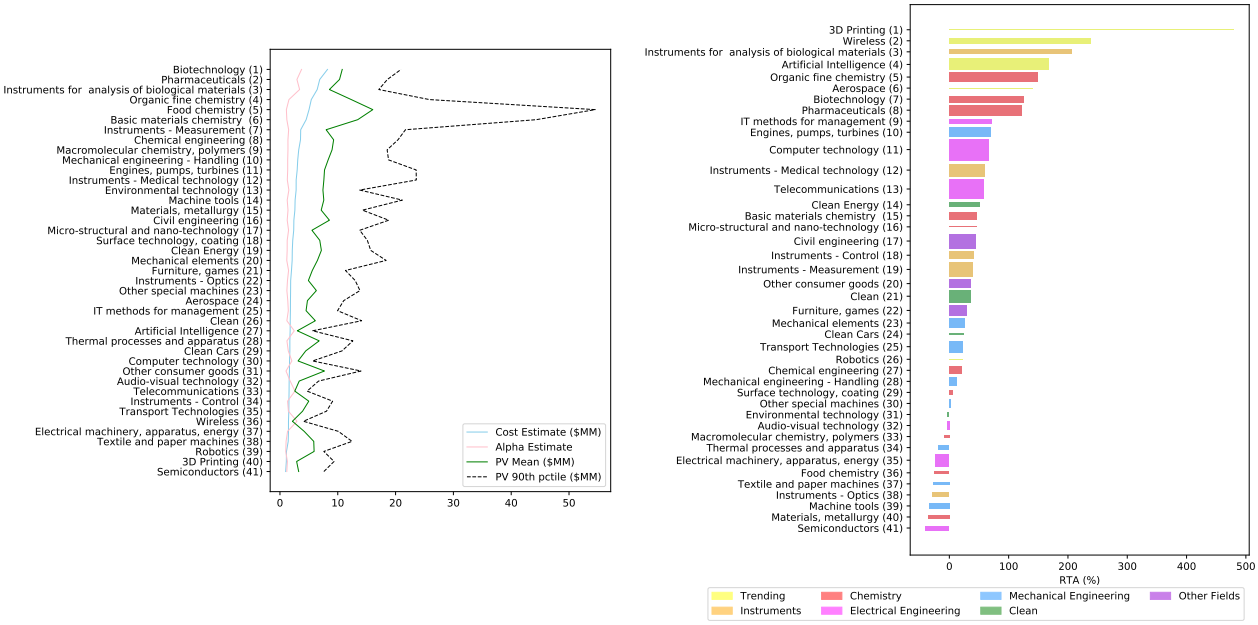
Table 6: Top 10 innovations in terms of EV – Germany

Rank	EV	Patent	Applicant (1st)	Title
1	1248.34	US2008305065(A1)	GOLDSCHMIDT	POLYSILOXANES HAVING QUATERNARY AMMONIUM GROUPS, A PROCESS FOR THE PREPARATION THEREOF AND THE USE THEREOF IN CLEANING AND CARE FORMULATIONS
2	488.38	EP1810589(A2)	LEKISPORT	STICK KNOB
3	433.12	EP1942864(A1)	BEIERSDORF	COSMETIC PREPARATION IN THE FORM OF AN OIL-IN-WATER EMULSION CONTAINING 1,2-ALKANEDIOL(S)
4	362.75	WO2006074633(A1)	VORDERMAIR, ENGELBERT	STICK HAVING AN ADJUSTABLE GRIP
6	342.23	US2011003010(A1)	CLARIANT	PHOSPHORIC ACID ESTERS CONTAINING PHOSPHORUS ATOMS BRIDGED BY DIOL UNITS
7	287.00	US2009068136(A1)	DSM IP ASSETS (DUTCH STATE MINES IP ASSETS)	HAIR CARE COMPOSITIONS
8	281.57	US2011160478(A1)	BAYER MATERIAL TECHNOLOGY TRADING (SHANGHAI) COMPANY	CATALYST FOR THE SYNTHESIS OF ALKYL CARBAMATES, THE METHOD FOR PREPARING THE SAME AND THE USE THEREOF
9	280.80	US2008146852(A1)	ARKEMA	PROCESS FOR DEHYDRATING GLYCEROL TO ACROLEIN
10	277.77	US2009306210(A1)	AQUANOVA	SOLUBILIZATES OF PRESERVATIVES AND METHOD FOR PRODUCING THE SAME

4.3.4 Technological Fields - United Kingdom

As documented in Figure 20, UK's comparative advantage lies in various ICT technologies, followed by chemical technologies. As is the case for Germany, external values are dominated by chemical technologies, while global IStra-X returns are highest for ICT technologies (Figures 21 and 22). National returns are highest in Biotech and Pharma, while — notably — Clean Energy technologies rank fourth.

Figure 20: Cost Diagram and RTA – Categories – United Kingdom

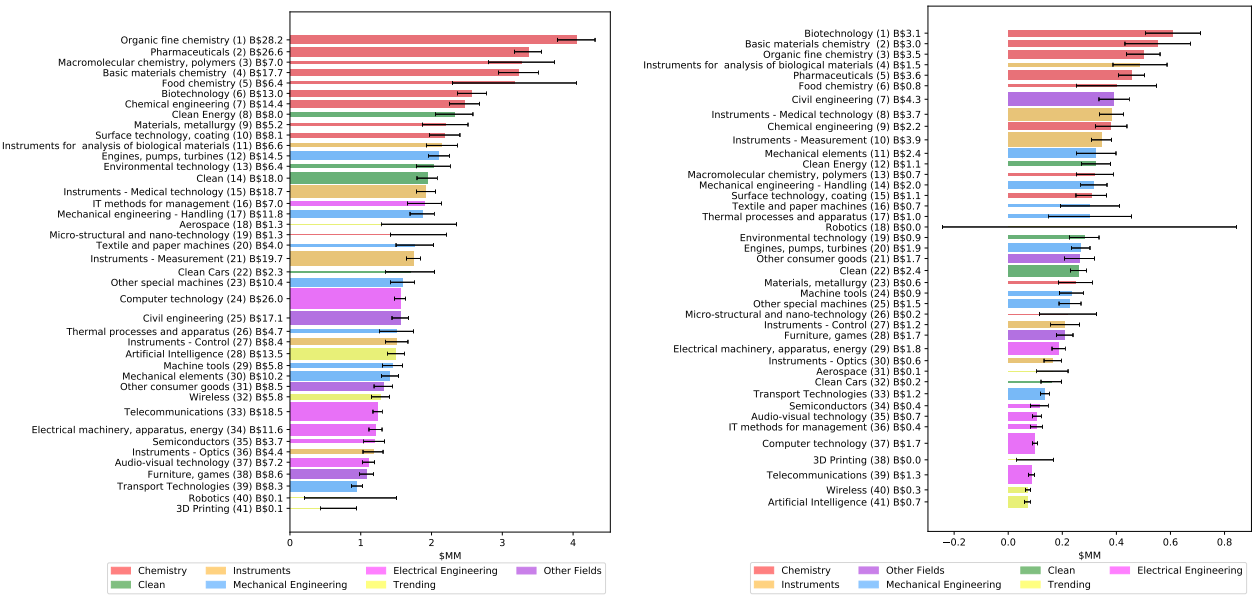


(a) Costs, Alpha & PV (mean and 90th pct.)

(b) Relative Technological Advantage (RTA)

Notes: Calculations are based on innovations originating in the United Kingdom for which a patent application was filed in the period 2005-2014. The left-hand panel shows the diagram of various estimates relevant to the calculation of the IStra-X indicator. The blue line shows the estimated cost (in millions of CPI-adjusted 1982 US dollars) of pursuing an innovation idea by technology field (y-axis). The green and dashed lines show the mean and 90th percentile of the estimated private returns distribution in each field (in in millions of CPI-adjusted 1982 US dollars). The pink line shows the estimate of parameter alpha for each technology field. The right-hand panel shows the Relative Technological Advantage of the United Kingdom for each technology field. It is calculated as the ratio between the share of the technology field in all innovations from the the United Kingdom and the share of the technology field in all innovations globally. The resulting ratio is subtracted by one and multiplied by 100 to obtain a percentage value.

Figure 21: External Value – Categories – United Kingdom

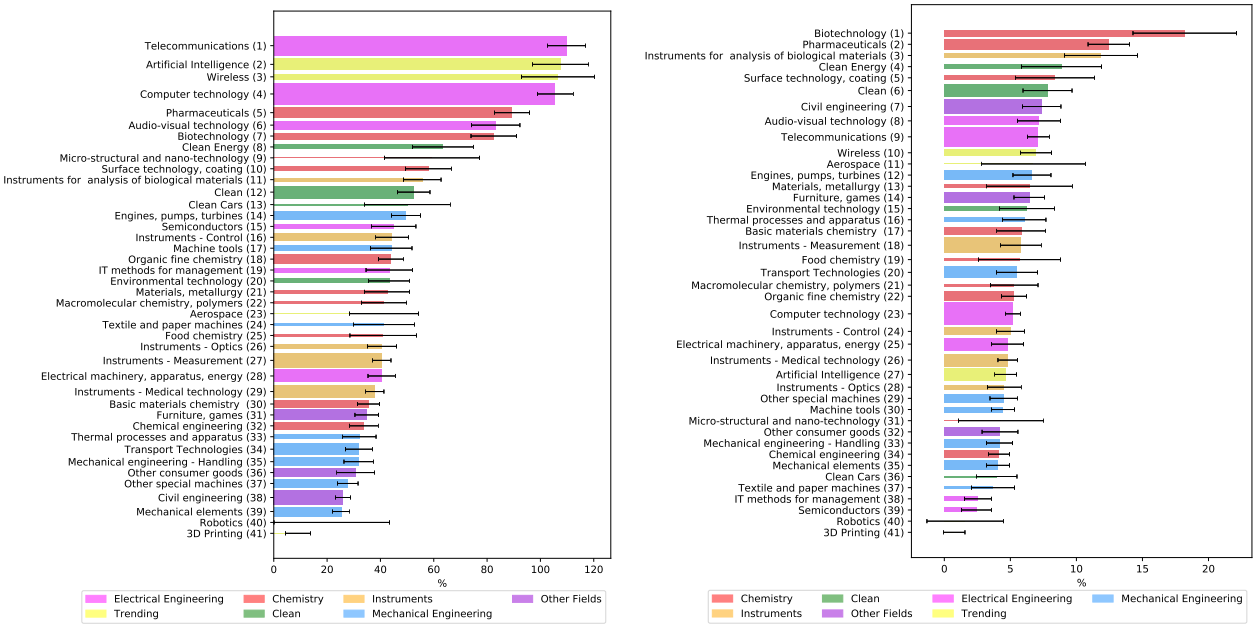


(a) Global

(b) National

Notes: Diagrams of the average external value in millions of CPI-adjusted 1982 US dollars (x-axis) by technology field (y-axis) in the United Kingdom. Calculations in the left-hand panel are based on the external value to the global population of innovations generated by innovations originating in the United Kingdom (for the period 2005-2014). External values in the right-hand panel are restricted to external value generated *within* the United Kingdom. Width of each bar represents the number innovations in the field. Area of each bar (in billions $) represents total external value in the technology field and is printed next to y-axis labels.

Figure 22: IStra-X – Categories – United Kingdom



(a) Global

(b) National

Notes: Diagram of the global IStra-X indicator representing the return on public spending (x-axis) by technology field (y-axis) in the United Kingdom. Calculations are based on the global population of innovations for which a patent application was filed in the period 2005-2014. For the left-hand panel, calculations are based on external value of spillovers to all innovations worldwide during this period. The right-hand panel shows the results when only taking into account external value generated for innovations *within* the United Kingdom. The x-axis gives the expected total welfare effect – as a return on investment – of a decrease in the cost of pursuing an innovation idea, broken down by technology field. IStra-X values reflect the difference between the expected increase in total value (private returns as well as external values from knowledge spillovers) generated by innovations in a field and the expected cost of the subsidy, scaled by the expected cost of the subsidy. Width of each bar represents the number innovations in the field.

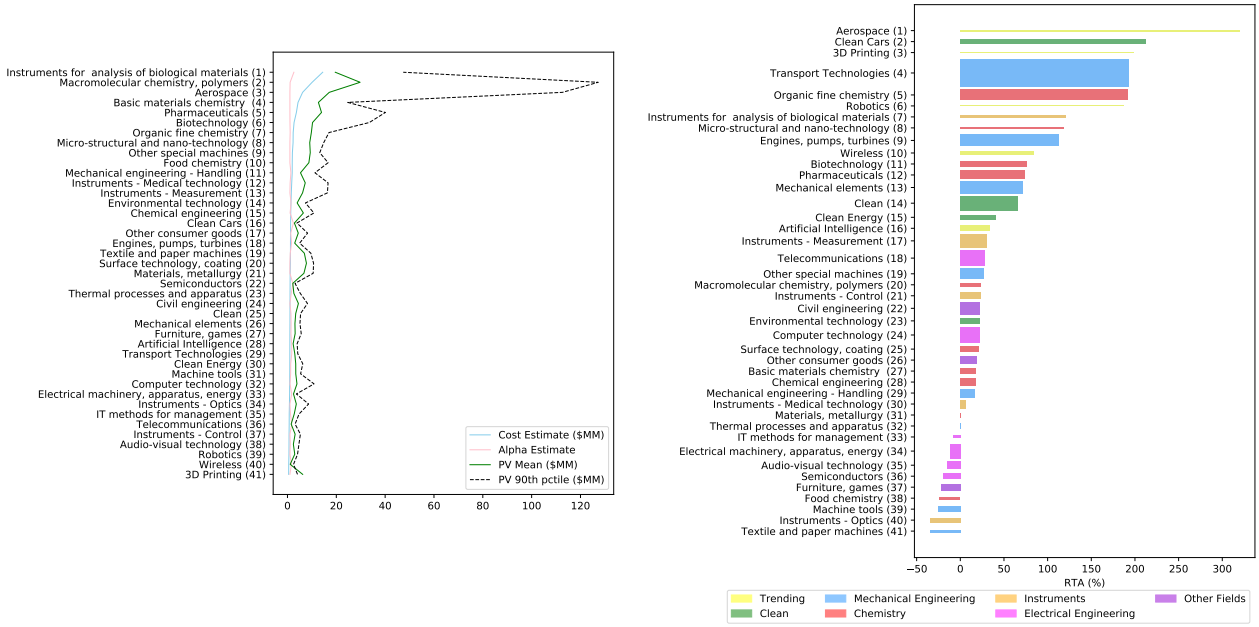
Table 7: Top 10 innovations in terms of EV – United Kingdom

Rank	EV	Patent	Applicant (1st)	Title
1	783.38	US2006286259(A1)	CADBURY ADAMS USA	TASTE POTENTIATOR COMPOSITIONS AND BEVERAGES CONTAINING SAME
3	247.20	US2008204253(A1)	RIWA	PEST MONITORING SYSTEM
4	240.06	GB0513847(D0)	REHAU & COMPANY	UNDERFLOOR HEATING
6	216.89	US2006224491(A1)	DE NOVO MARKETS	TRADING AND SETTLING ENHANCEMENTS TO THE STANDARD ELECTRONIC FUTURES EXCHANGE MARKET MODEL LEADING TO NOVEL DERIVATIVES INCLUDING ON EXCHANGE ISDA TYPE CREDIT DERIVATIVES AND ENTIRELY NEW RECOVERY PRODUCTS INCLUDING NOVEL OPTIONS ON THESE
7	207.66	US2010069287(A1)	QUEST INTERNATIONAL SERVICES	FRAGRANCE COMPOUNDS
8	204.65	US2009197791(A1)	RHODIA RECHERCHES ET TECHNOLOGIES	COPOLYMER CONTAINING ZWITTERIONIC UNITS AND OTHER UNITS, COMPOSITION COMPRISING THE COPOLYMER, AND USE
9	189.67	US2007202063(A1)	APPLETON PAPERS	BENEFIT AGENT CONTAINING DELIVERY PARTICLE
10	185.85	US2008208096(A1)	ARMITAGE, KENNETH	MOBILITY AIDS

4.3.5 Technological Fields - France

France follows a pattern even closer to Germany than the UK (with some exceptions). As we can see from Figure 23, France's comparative advantage lies in mechanical engineering, while external value is dominated by chemical technologies (Figure 24. National IStra-X returns are highest in chemicals as well, whereas Global IStra-X is, by and large, highest for ICT-related technologies (Figure 25).

Figure 23: Cost Diagram and RTA – Categories – France

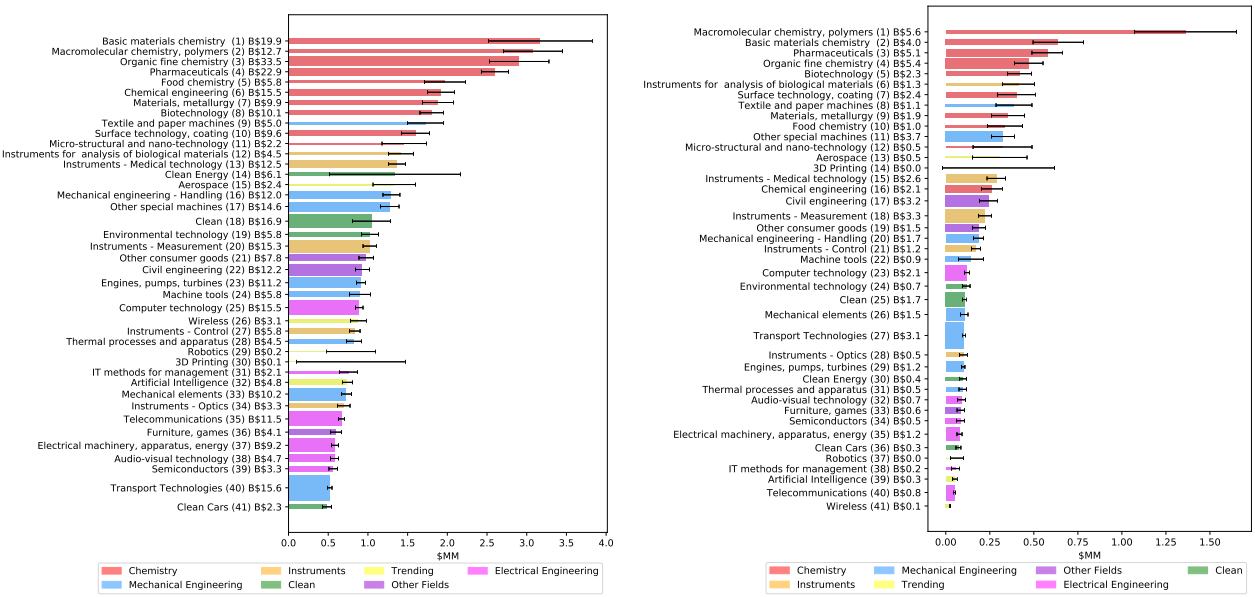


(a) Costs, Alpha & PV (mean and 90th pct.)

(b) Relative Technological Advantage (RTA)

Notes: Calculations are based on innovations originating in France for which a patent application was filed in the period 2005-2014. The left-hand panel shows the diagram of various estimates relevant to the calculation of the IStra-X indicator. The blue line shows the estimated cost (in millions of CPI-adjusted 1982 US dollars) of pursuing an innovation idea by technology field (y-axis). The green and dashed lines show the mean and 90th percentile of the estimated private returns distribution in each field (in in millions of CPI-adjusted 1982 US dollars). The pink line shows the estimate of parameter alpha for each technology field. The right-hand panel shows the Relative Technological Advantage of France for each technology field. It is calculated as the ratio between the share of the technology field in all innovations from France and the share of the technology field in all innovations globally. The resulting ratio is subtracted by one and multiplied by 100 to obtain a percentage value.

Figure 24: External Value – Categories – France

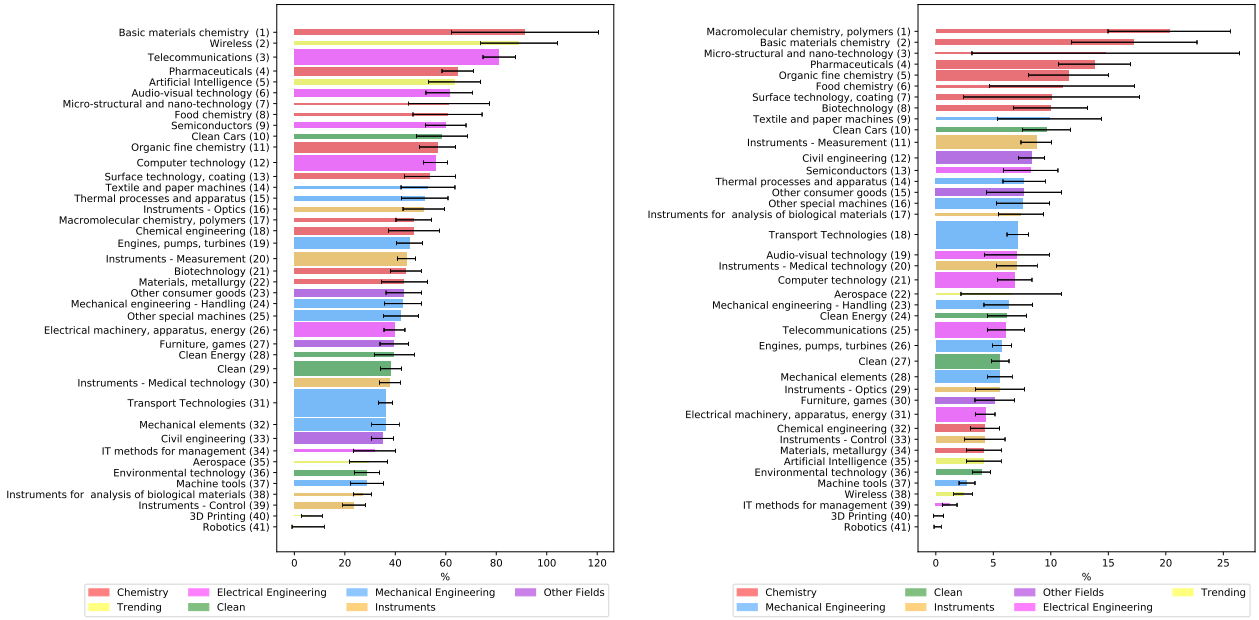


(a) Global

(b) National

Notes: Diagrams of the average external value in millions of CPI-adjusted 1982 US dollars (x-axis) by technology field (y-axis) in France. Calculations in the left-hand panel are based on the external value to the global population of innovations generated by innovations originating in France (for the period 2005-2014). External values in the right-hand panel are restricted to external value generated *within* France. Width of each bar represents the number innovations in the field. Area of each bar (in billions $) represents total external value in the technology field and is printed next to y-axis labels.

Figure 25: IStra-X – Categories – France



(a) Global

(b) National

Notes: Diagram of the global IStra-X indicator representing the return on public spending (x-axis) by technology field (y-axis) in France. Calculations are based on the global population of innovations for which a patent application was filed in the period 2005-2014. For the left-hand panel, calculations are based on external value of spillovers to all innovations worldwide during this period. The right-hand panel shows the results when only taking into account external value generated for innovations *within* France. The x-axis gives the expected total welfare effect – as a return on investment – of a decrease in the cost of pursuing an innovation idea, broken down by technology field. IStra-X values reflect the difference between the expected increase in total value (private returns as well as external values from knowledge spillovers) generated by innovations in a field and the expected cost of the subsidy, scaled by the expected cost of the subsidy. Width of each bar represents the number innovations in the field.

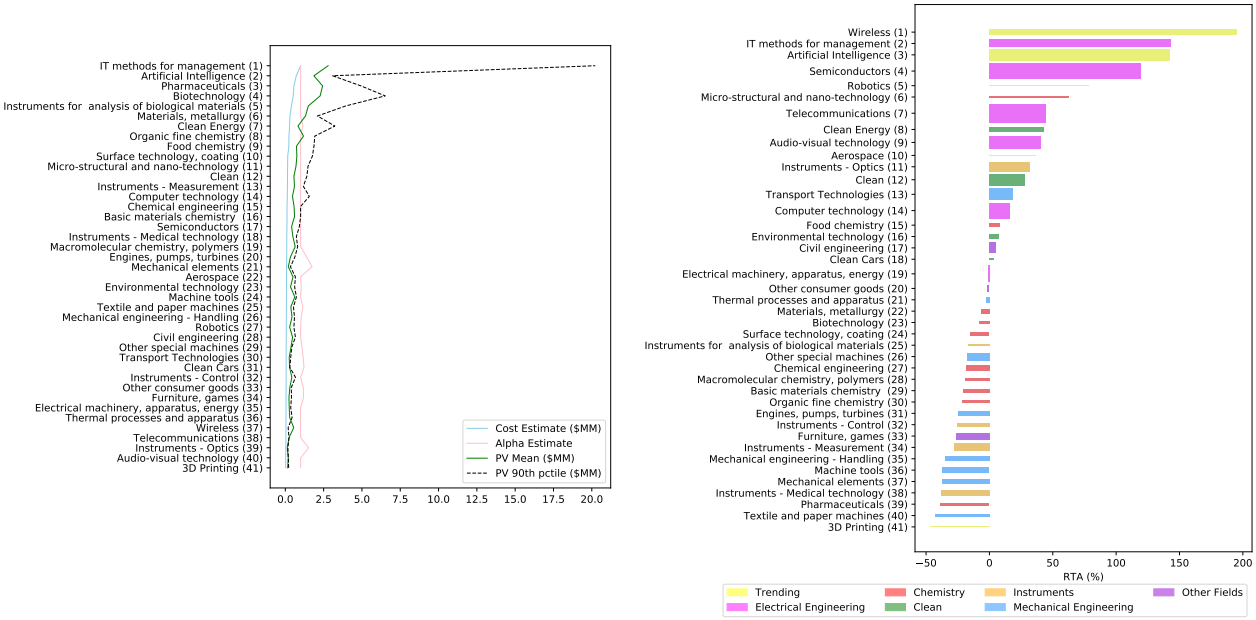
Table 8: Top 10 innovations in terms of EV – France

Rank	EV	Patent	Applicant (1st)	Title
1	1901.26	US2010076236(A1)	SHELL INTERNATIONALE RESEARCH MAATSCHAPPIJ	PROCESS FOR PRODUCING PARAFFINIC HYDROCARBONS
2	296.52	US2009214453(A1)	CIBA SPECIALTY CHEMICALS HOLDING	USE OF TRANSMISSION DYES FOR PROTECTING HUMAN SKIN FROM BROWNING AND AGEING
3	265.25	US2009116343(A1)	LEVINGSTON, GIDEON	BALANCE SPRING, REGULATED BALANCE WHEEL ASSEMBLY AND METHODS OF MANUFACTURE THEREOF
4	234.94	US2010081751(A1)	RHODIA	USE OF AN ORGANOPOLYSILOXANE COMPOSITION VULCANIZABLE FROM ROOM TEMPERATURE TO FORM A SELF-ADHESIVE ELASTOMER
5	227.58	US2006236614(A1)	EIFFAGE TRAVAUX PUBLICS	METHOD OF MANUFACTURING A BITUMINOUS COATED MATERIAL
6	226.79	US2008159081(A1)	MONTRES BREGUET	MULTIFUNCTION COAXIAL CORRECTOR DEVICE
7	225.37	US2007077221(A1)	RHODIA	COSMETIC COMPOSITION COMPRISING AN AMPHOLYTIC COPOLYMER AND ANOTHER AGENT
8	202.97	US2008291781(A1)	TOTAL	PROCESS AND PROGRAM FOR CHARACTERISING EVOLUTION OF AN OIL RESERVOIR OVER TIME
9	194.81	US2010145089(A1)	RHODIA OPERATIONS	PREPARATION OF (POLY)SULFIDE ALKOXYSILANES AND NOVEL INTERMEDIATES THEREFOR
10	181.14	FR2922446(A1)	L'OREAL	COMPOSITION, USEFUL FOR OXIDATION COLORING, LIGHTENING DIRECT DYEING, AND BLEACHING OF THE KERATIN FIBERS, COMPRISES POLYLYSINES, AND ADDITIONAL ALKALINE AGENTS COMPRISING E.G. AMMONIUM SALTS AND ALKALI METAL OR ALKALINE EARTH CARBONATES

4.3.6 Technological Fields - Korea

Korea, by contrast, follows a pattern markedly different than the one of the European economies. Comparative advantage clearly lies in ICT technologies (Figure 26). Only global external value is dominated by chemicals and pharmaceuticals, whereas national external value is topped by ICT technologies (Figure 27). Figure 28 shows that both national and global IStra-X returns are dominated by ICT-related technologies, with some important exceptions; e.g. nano-technology is leading on national IStra-X returns.

Figure 26: Cost Diagram and RTA – Categories – Korea

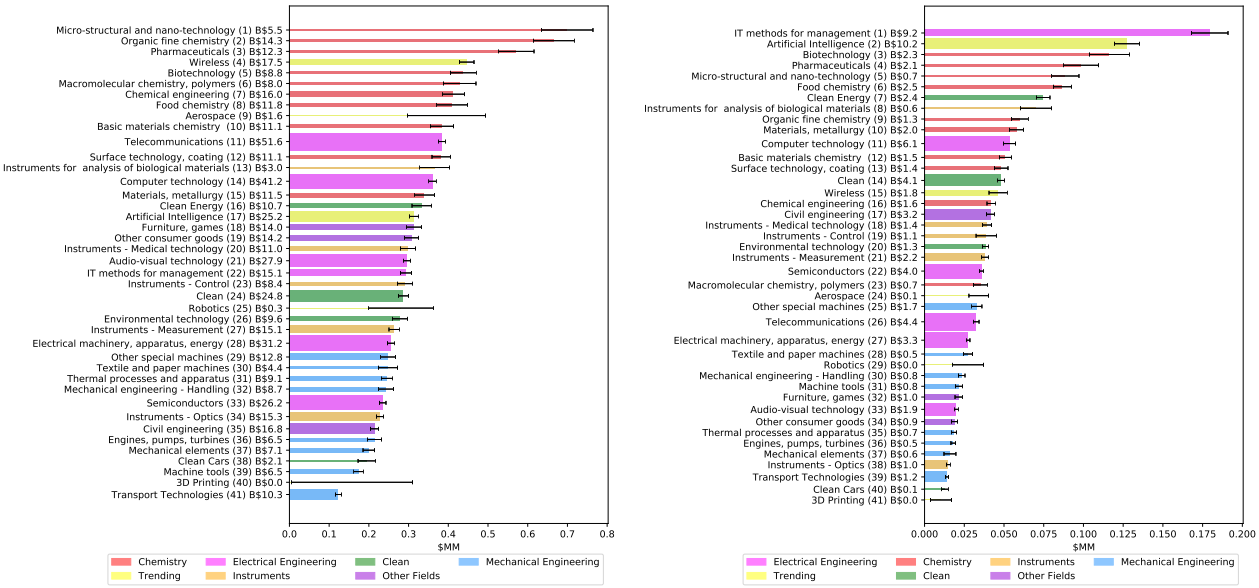


(a) Costs, Alpha & PV (mean and 90th pct.)

(b) Relative Technological Advantage (RTA)

Notes: Calculations are based on innovations originating in Korea for which a patent application was filed in the period 2005-2014. The left-hand panel shows the diagram of various estimates relevant to the calculation of the IStra-X indicator. The blue line shows the estimated cost (in millions of CPI-adjusted 1982 US dollars) of pursuing an innovation idea by technology field (y-axis). The green and dashed lines show the mean and 90th percentile of the estimated private returns distribution in each field (in in millions of CPI-adjusted 1982 US dollars). The pink line shows the estimate of parameter alpha for each technology field. The right-hand panel shows the Relative Technological Advantage of Korea for each technology field. It is calculated as the ratio between the share of the technology field in all innovations from Korea and the share of the technology field in all innovations globally. The resulting ratio is subtracted by one and multiplied by 100 to obtain a percentage value.

Figure 27: External Value – Categories – Korea

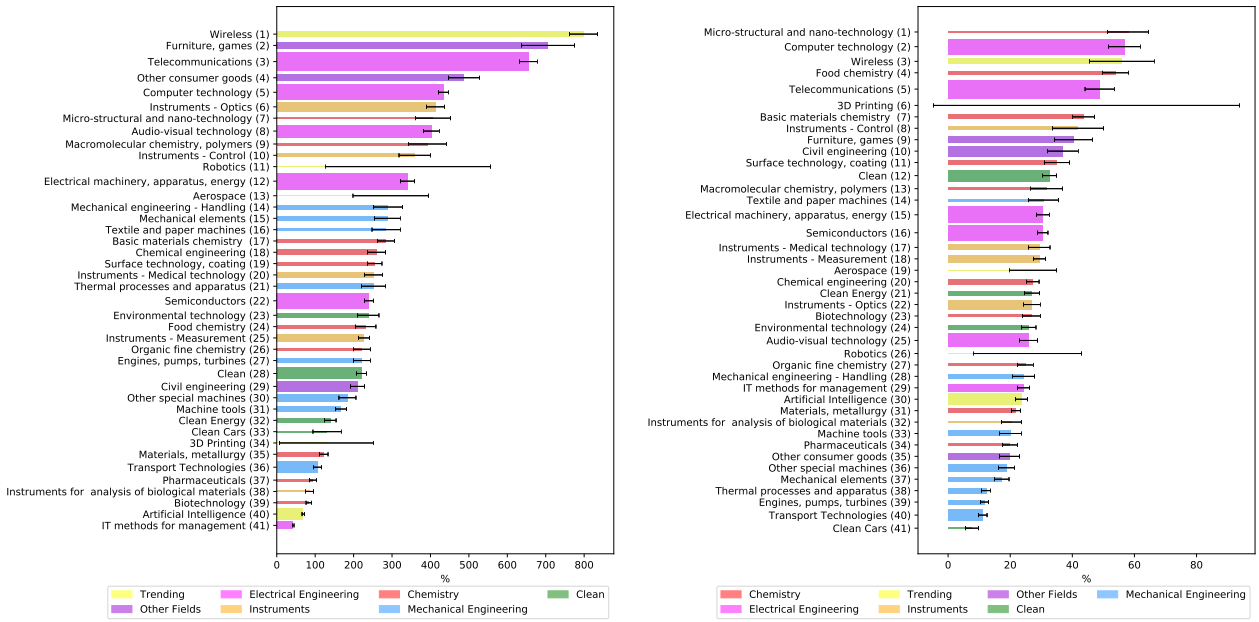


(a) Global

(b) National

Notes: Diagrams of the average external value in millions of CPI-adjusted 1982 US dollars (x-axis) by technology field (y-axis) in Korea. Calculations in the left-hand panel are based on the external value to the global population of innovations generated by innovations originating in Korea (for the period 2005-2014). External values in the right-hand panel are restricted to external value generated *within* Korea. Width of each bar represents the number innovations in the field. Area of each bar (in billions $) represents total external value in the technology field and is printed next to y-axis labels.

Figure 28: IStra-X – Categories – Korea



(a) Global

(b) National

Notes: Diagram of the global IStra-X indicator representing the return on public spending (x-axis) by technology field (y-axis) in Korea. Calculations are based on the global population of innovations for which a patent application was filed in the period 2005-2014. For the left-hand panel, calculations are based on external value of spillovers to all innovations worldwide during this period. The right-hand panel shows the results when only taking into account external value generated for innovations *within* Korea. The x-axis gives the expected total welfare effect – as a return on investment – of a decrease in the cost of pursuing an innovation idea, broken down by technology field. IStra-X values reflect the difference between the expected increase in total value (private returns as well as external values from knowledge spillovers) generated by innovations in a field and the expected cost of the subsidy, scaled by the expected cost of the subsidy. Width of each bar represents the number innovations in the field.

Table 9: Top 10 innovations in terms of EV – Korea

Rank	EV	Patent	Applicant (1st)	Title
1	204.53	US2006185714(A1)	SAMSUNG ELECTRONICS COMPANY	FLEXIBLE SOLAR CELL AND METHOD OF PRODUCING THE SAME
2	195.92	US2006229387(A1)	SAMSUNG TOTAL PETROCHEMICALS COMPANY	POLYPROPYLENE RESIN COMPOSITION WITH ANTI-SCRATCH CHARACTERISTICS
3	188.89	WO2006088314(A1)	SK CORPORATION	PROCESS FOR PRODUCING ULTRA LOW SULFUR AND LOW AROMATIC DIESEL FUEL
4	159.60	KR20100022203(A)	PARK, JI WOONG	STREET LIGHT HAVING MEANS FOR ENTICING HARMFUL INSECTS
5	156.84	KR20080007942(A)	JANG, MAN SIK	A STRUCTURE FOR CLEANING A WINDOW GLASS
6	156.60	KR101040206(B1)	SHIN, BONG SUK	AQUEUS POWER OF FLAX SEED OIL COMPRISING OMEGA 3-FATIC ACID
7	146.15	US2006177341(A1)	KOREA ATOMIC ENERGY RESEARCH INSTITUTE	ZIRCONIUM BASED ALLOYS HAVING EXCELLENT CREEP RESISTANCE
8	136.86	WO2011099665(A1)	KCI	ANTIMICROBIAL COMPOSITION CONTAINING EXTRACTS FROM NATURAL INGREDIENTS, NATURAL COMPOSITE ANTISEPTICS, AND METHOD FOR MANUFACTURING SAME
9	131.38	US2008194828(A1)	HANMI PHARM. COMPANY	METHOD OF PREPARING ESOMEPRAZOLE AND SALTS THEREOF
10	130.53	KR100666966(B1)	YANG, DAE KEUN	AUTO PILING APPARATUS FOR CONTAINER

5 Conclusion

This report develops a new framework to measure both direct and indirect knowledge spillovers from patent data. It also proposes a new way to quantify the monetary value of spillovers. Different versions of this can be used to differentiate between global and national knowledge flows or to examine other realms of internalisation of knowledge flows.[24] Moreover, we develop a new methodology to assess the marginal social impact of government subsidies on different sectors or technology areas of an economy. We apply this to the question of vertically differentiated industrial policy; i.e. targeted support by government for specific sectors or technologies. Varying degrees of knowledge spillovers between different sectors can in principle justify such policies. However, when such differences exist, identifying the set of sectors that deserve special government attention is an empirical question. This paper shows that there is a substantial and statistically significant variation in the social returns to government support. We also show that the set of technologies or sectors that should be supported varies greatly from country to country. It also depends greatly on the desired level of internalization of externalities. Hence, our results provide an interesting starting point for a discussion about specific national, supranational or indeed sub-national designs of industrial policy. The framework also suggests a set of indicators that could continuously be computed to monitor ongoing efforts by governments.

[24] An interesting future application could include the degree the merging of companies could increase internalization of spillovers.

References

Acemoglu, D., Akcigit, U., and Kerr, W. R. (2016). Innovation network. *Proceedings of the National Academy of Sciences of the United States of America*, 113(41):11483–11488.

Arthur, W. B. (2009). *The nature of technology: What it is and how it evolves.* Simon and Schuster.

Brin, S. and Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems*, 30(1-7):107–117.

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

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288.

A Computing Patent Rank Recursively

To compute patent rank in practice we need to rely on a recursive procedure rather than inverting outright as in equation 6. Here we show that our recursive procedure converges to the actual solution. We can write the vector of social values at a given iteration as

$$V^{(n)} = V^* + \Delta^{(n)} \tag{24}$$

where $\Delta^{(n)}$ captures the difference of the social value at iteration n relative to the actual solution. Using 7 we can write

$$\boldsymbol{V}^* + \Delta^{(n)} = (\boldsymbol{V}^* + \boldsymbol{\Delta}^{(n-1)})\sigma\boldsymbol{\Phi} + PV = V^*\sigma\boldsymbol{\Phi} + \boldsymbol{PV} + \boldsymbol{\Delta}^{(n-1)}\sigma\boldsymbol{\Phi} \tag{25}$$

Because $\boldsymbol{V}^*$ is the actual solution of the equation system we have $\boldsymbol{V}^* = \boldsymbol{V}^*\sigma\boldsymbol{\Phi} + \boldsymbol{PV}$, hence we can re-write equation 25 as

$$\boldsymbol{\Delta}^{(n)} = \boldsymbol{\Delta}^{(n-1)}\sigma\boldsymbol{\Phi} \tag{26}$$

We now need to show that 26 is a contraction mapping. For that consider the characteristic element

$$\Delta_i^{(n)} = \sigma \sum_j \phi_{ij}\Delta_j^{(n-1)} \tag{27}$$

Note that all elements ϕ_{ij} are positive but some elements $\Delta_j^{(n-1)}$ could be negative. Hence

$$|\Delta_i^{(n)}| \leq \sigma \sum_j \phi_{ij}|\Delta_j^{(n-1)}| \tag{28}$$

Summing over all innovations yields

$$\sum_i |\Delta_i^{(n)}| \leq \sigma \sum_i \sum_j \phi_{ij}|\Delta_j^{(n-1)}| = \sigma \sum_j |\Delta_j^{(n-1)}| \sum_i \phi_{ij} \tag{29}$$

Note that $\sum_i \phi_{ij} = 1$ so that we can write

$$\sum_i |\Delta_i^{(n)}| \leq \sigma \sum_j |\Delta_j^{(n-1)}| \tag{30}$$

We can think of the LHS of this as an index of the total error we are making in using the n-th iteration of the the algorithm rather than the actual solution V^*. The equation consequently suggests that the error at iteration n is smaller than at iteration $n-1$ by virtue of $\sigma < 1$. Hence, the error will tend exponentially to zero.

B Further Estimation results

B.1 Investigating the importance of indirect spillovers (σ)

To compute Patent Rank, we need to make an assumption about the value of σ. As seen above, σ can be interpreted as the extent to which indirect spillovers contribute to the overall value of an innovation. The lower it is, the lower is the contribution of indirect spillovers, and even more so for innovations are that distant from the innovations receiving the spillovers. What value of σ should we use? We can test this by measuring the spillover flows between technology classes and countries using Patent rank methodology and include the spillover flows as factors in an innovation production function. Including all variables at the same time leads to only insignificant coefficients. However, a joint hypothesis tests suggests that jointly they are highly significant. Hence, the individual insignificance is likely due to multi-collinearity between the measures with different values of σ. Hence, in column (2) and (4), we use a LASSO approach to narrow the field (Tibshirani, 1996). More specifically, we use this method to test the value of sigma in Patent rank measures, i.e. spillovers received and provided by a given field, that leads to the best prediction of future innovations. For this reason, we only allow for the exclusion of these variables of interest (patentRank). The number and private value of innovations, traditional spillover measures and fixed effects are always included in the estimations. Results from the LASSO suggests that a distance decay parameter of 25% for between spillovers received and spillovers provided, and of 75% (or 50% with private value as dependent variable) for within spillovers received is more appropriate. Thereby, we decide to use the in-between option, $\sigma = 0.5$, in the OLS regressions and in the paper. The point estimates in the LASSO are also very close to the ones when selecting $\sigma = 0.5$.

Table 10: Estimating the effects of spillovers on future innovations in EU and OECD countries

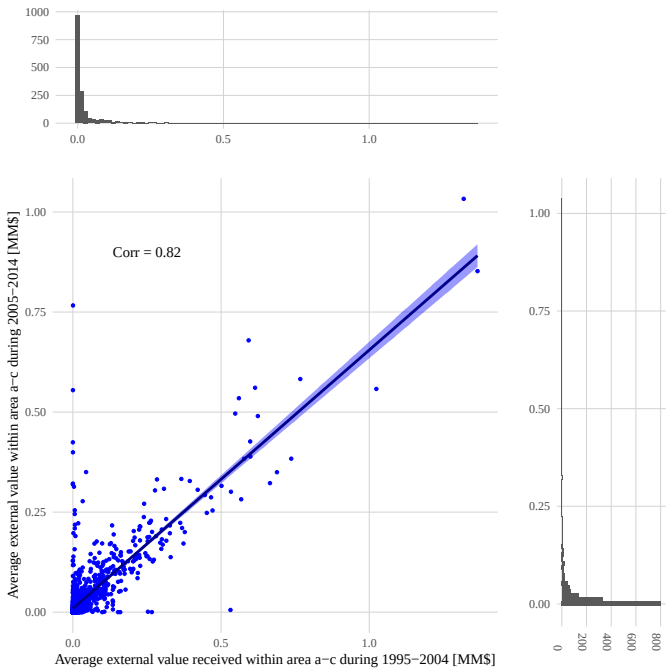
	Number of innovations		Innovation private value	
	OLS	LASSO	OLS	LASSO
	(1)	(2)	(3)	(4)
Innovations	−0.258***	−0.274***	−0.334***	−0.323***
	(0.006)	(0.006)	(0.011)	(0.010)
Innovation value	−0.128***	−0.281***	−0.056***	−0.100***
	(0.009)	(0.008)	(0.019)	(0.016)
Citation counts in	0.021**	−0.051***	0.025	−0.022
	(0.008)	(0.008)	(0.017)	(0.015)
Citation counts out	0.040***	−0.078***	−0.090***	−0.116***
	(0.009)	(0.008)	(0.019)	(0.017)
Between Prank in (sigma = 0.25)	1.024***	0.211***	0.156	0.205***
	(0.130)	(0.009)	(0.264)	(0.017)
Between Prank in (sigma = 0.5)	−1.253***		0.626	
	(0.284)		(0.579)	
Between Prank in (sigma = 0.75)	0.400**		−0.581*	
	(0.160)		(0.327)	
Within Prank in (sigma = 0.25)	0.005		−0.052	
	(0.035)		(0.079)	
Within Prank in (sigma = 0.5)	−0.010		−0.074	−0.050***
	(0.069)		(0.157)	(0.007)
Within Prank in (sigma = 0.75)	0.026	0.067***	0.076	
	(0.037)	(0.003)	(0.086)	
Prank out (sigma = 0.25)	0.358***	0.073***	−0.880***	0.112***
	(0.096)	(0.008)	(0.195)	(0.017)
Prank out (sigma = 0.5)	−0.211		2.098***	
	(0.201)		(0.406)	
Prank out (sigma = 0.25)	−0.121		−1.169***	
	(0.111)		(0.223)	
Constant	1.281***	1.093***	1.907***	1.879***
	(0.037)	(0.037)	(0.076)	(0.074)
Observations	50,430	50,430	50,430	50,430
R^2	0.598	0.586	0.277	0.276
Adjusted R^2	0.597	0.585	0.276	0.274

Notes: *p<0.1; **p<0.05; ***p<0.01.
All variables are in log. Prank was calculated based on network without self citation. The dependent variables per year are normalized by the number of innovation during the previous period.

B.2 Present and future external values

Figure indicates that there is a high correlation in the average spillovers received by a given area $a - c$ between the periods 1995-2004 and 2005-2014.

Figure 29: Present and future external value received within each area $a - c$ ($\sigma = 50\%$)



Note: Each group corresponds to a country and a technological category ($a - c$ area). Corr corresponds to the correlation of external values between the periods 1995-2004 and 2005-2014 across areas $a - c$.

C Proof of Proposition 4.1

Proposition 4.1 provided a simple expression for the marginal effect of changes in government innovation support on S on innovation value V. Here we derive this expression.

Proof. We look at thee various elements of equation 21 in turn:

Marginal effect on probability of having worthwile idea $\frac{\partial P(\delta>\lambda)}{\partial c}$*:*

Note that

$$P(\delta > \lambda) = \frac{\mu^\alpha}{\lambda^\alpha} = \frac{(\kappa\mu)^\alpha}{(2c)^\alpha} \tag{31}$$

Hence

$$\frac{\partial P(\delta > \lambda)}{\partial c} = -\alpha \frac{(\kappa\mu)^\alpha}{2^\alpha c^{\alpha+1}} = -\frac{\alpha}{c} P(\delta > \lambda) \tag{32}$$

Hence the derivative is equal to the probability of having a good idea times the ratio between α and c.

Marginal effect on expected private value profits $\frac{\partial E\{PV|\delta>\lambda\}}{\partial c}$*:*

$$\frac{\partial E\{PV|\delta > \lambda\}}{\partial c} = \frac{\alpha}{\alpha - 1} = \frac{1}{c} E\{PV|\delta > \lambda\}$$

Marginal effect on expected external value $\frac{\partial E\{EV|\delta>\lambda\}}{\partial c}$*:*

Firstly note that

$$E\left\{EV|\delta>\lambda\right\}=E\left\{E\left\{EV|PV\right\}|\delta>\lambda\right\}=\int E\left\{EV|v\right\}P(v|\delta>\lambda)dv$$

i.e. we can compute the expected value of EV via iterated expectation. Moreover

$$E\left\{EV|PV,c\right\}=E\left\{EV|PV\right\}$$

i.e. external values depend on the cost threshold c via private values. We can consequently write

$$E\left\{EV|\delta>\lambda\right\}=\int_0^{2c}E\left\{EV|v\right\}P(v|\delta>\lambda)dv+\int_{2c}^{\infty}E\left\{EV|v\right\}P(v|\delta>\lambda)dv \tag{33}$$

Let's look at the derivative of each integral in 33 in turn. For the first integral we get

$$\frac{\partial}{\partial c}\left[\int_0^{2c}E\left\{EV|v\right\}P(v|\delta>\lambda)dv\right]=E\left\{EV|2c\right\}P(2c|\delta>\lambda)\times 2+\int_0^{2c}E\left\{EV|v\right\}\frac{\partial P(v|\delta>\lambda)}{\partial c}dv$$

$$\begin{aligned}\int_0^{2c}E\left\{EV|v\right\}\frac{\partial P(v|\delta>\lambda)}{\partial c}dv &= -\frac{1}{c}\int_0^{2c}E\left\{EV|v\right\}P(v|\delta>\lambda)dv\\ &= -\frac{1}{c}E\left\{EV|\delta>\lambda,v<2c\right\}P(v<2c|\delta>\lambda)\\ &= =-\frac{1}{c}E\left\{EV\times\mathbb{I}\{v<2c\}|\delta>\lambda\right\}\end{aligned}$$

where $\mathbb{I}\{\cdot\}$ is the indicator function. For the second integral we get

$$\frac{\partial}{\partial c}\left[\int_{2c}^{\infty}E\left\{EV|v\right\}P(v|\delta>\lambda)dv\right]=-E\left\{EV|2c\right\}P(2c|\delta>\lambda)\times 2+\frac{\alpha}{c}\int_{2c}^{\infty}E\left\{EV|v\right\}P(v|\delta>\lambda)dv$$

$$\int_{2c}^{\infty}E\left\{EV|v\right\}P(v|\delta>\lambda)dv=E\left\{EV\times\mathbb{I}\{v>2c\}|\delta>\lambda\right\}$$

Combining yields

$$\frac{\partial E\left\{EV|\delta>0\right\}}{\partial c}=\frac{\alpha}{c}E\left\{EV\times\mathbb{I}\{PV>2c\}|\delta>\lambda\right\}-\frac{1}{c}E\left\{EV\times\mathbb{I}\{PV<2c\}|\delta>\lambda\right\}$$

Consequently we can write the marginal effect on the total innovation value as

$$\begin{aligned}\frac{\partial E\left\{V\right\}}{\partial c} &= \left[E\left\{PV|\delta>\lambda\right\}-c+\alpha E\left\{EV\times\mathbb{I}\{v>2c\}|\delta>\lambda\right\}-E\left\{EV\times\mathbb{I}\{v<2c\}|\delta>\lambda\right\}\right.\\ &\quad\left.-\alpha E\left\{PV+EV-c|\delta>\lambda\right\}\right]\times\frac{P(\delta>\lambda)}{c}\\ &= E\left\{PV-c+EV\left(\alpha\times\mathbb{I}\{v>2c\}-\mathbb{I}\{v<2c\}\right)-\alpha\left(PV+EV-c\right)|\delta>\lambda\right\}\frac{P(\delta>\lambda)}{c}\\ &= E\left\{(1-\alpha)\left(PV-c\right)+EV\left(\alpha\times\mathbb{I}\{v>2c\}-\mathbb{I}\{v<2c\}-\alpha\right)|\delta>\lambda\right\}\frac{P(\delta>\lambda)}{c}\\ &= E\left\{-c+EV\left(\alpha\times\mathbb{I}\{v>2c\}-\mathbb{I}\{v<2c\}-\alpha\right)|\delta>\lambda\right\}\frac{P(\delta>\lambda)}{c}\end{aligned}$$

where the last equation follows from $E\{PV|\delta>\lambda\}=\frac{\alpha c}{\alpha-1}$ (see equation 19).Finally becase $\frac{\partial c}{\partial s}=-1$ we get the expression in 22:

$$\frac{\partial E\left\{V\right\}}{\partial s}=E\left\{c+EV\left(\alpha-\alpha\times\mathbb{I}\{v>2c\}+\mathbb{I}\{v<2c\}-\alpha\right)|\delta>\lambda\right\}\frac{P(\delta>\lambda)}{c}$$

□

D Defining technological fields

Table 11: Concordance between technological fields and IPC/CPC classes

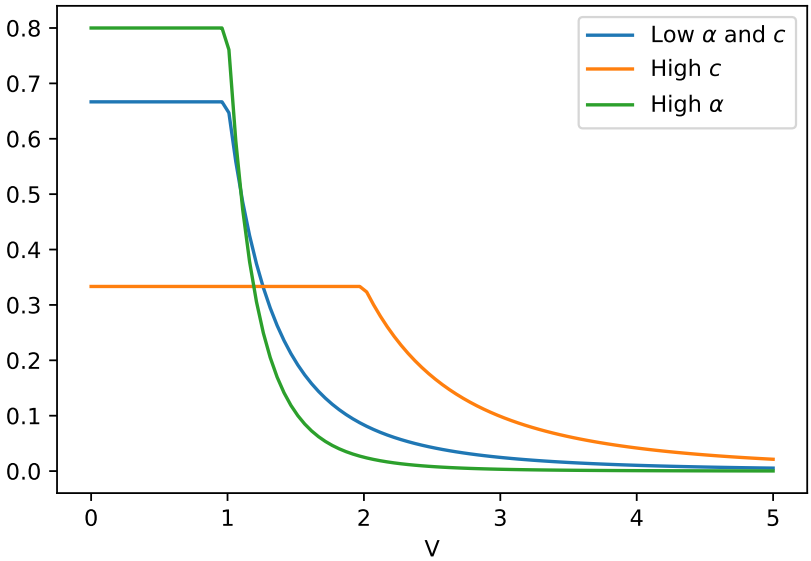
Label	Field	Classes	Scheme
Electrical Engineering	Electrical machinery, apparatus, energy	F21H, F21K, F21L, F21S, F21V, F21W, F21Y, H01B, H01C, H01F, H01G, H01H, H01J, H01K, H01M, H01R, H01T, H02B, H02G, H02H, H02J, H02K, H02M, H02N, H02P, H02S, H05B, H05C, H05F, H99Z	IPC
Electrical Engineering	Audio-visual technology	G09F, G09G, G11B, H04N 3, H04N 5, H04N 7, H04N 9, H04N 11, H04N 13, H04N 15, H04N 17, H04N 19, H04N 101, H04R, H04S, H05K	IPC
Electrical Engineering	Telecommunications	G08C, H01P, H01Q, H04B, H04H, H04J, H04K, H04M, H04N 1, H04Q, H04L, H04N 21, H04W, H03B, H03C, H03D, H03F, H03G, H03H, H03J, H03K, H03L, H03M	IPC
Electrical Engineering	Computer technology	G06C, G06D, G06E, G06F, G06G, G06J, G06K, G06M, G06N, G06T, G10L, G11C	IPC
Electrical Engineering	IT methods for management	G06Q	IPC
Electrical Engineering	Semiconductors	H01L	IPC
Instruments	Instruments - Optics	G02B, G02C, G02F, G03B, G03C, G03D, G03F, G03G, G03H, H01S	IPC
Instruments	Instruments - Measurement	G01B, G01C, G01D, G01F, G01G, G01H, G01J, G01K, G01L, G01M, G01N 1, G01N 3, G01N 5, G01N 7, G01N 9, G01N 11, G01N 13, G01N 15, G01N 17, G01N 19, G01N 21, G01N 22, G01N 23, G01N 24, G01N 25, G01N 27, G01N 29, G01N 30, G01N 31, G01N 35, G01N 37, G01P, G01Q, G01R, G01S, G01V, G01W, G04B, G04C, G04D, G04F, G04G, G04R, G12B, G99Z	IPC
Instruments	Instruments for analysis of biological materials	G01N 33	IPC
Instruments	Instruments - Control	G05B, G05D, G05F, G07B, G07C, G07D, G07F, G07G, G08B, G08G, G09B, G09C, G09D	IPC
Instruments	Instruments - Medical technology	A61B, A61C, A61D, A61F, A61G, A61H, A61J, A61L, A61M, A61N, G16H, H05G	IPC
Chemistry	Organic fine chemistry	A61K 8, A61Q, C07B, C07C, C07D, C07F, C07H, C07J, C40B	IPC
Chemistry	Biotechnology	C07G, C07K, C12M, C12N, C12P, C12Q, C12R, C12S	IPC
Chemistry	Pharmaceuticals	A61K 6, A61K 9, A61K 31, A61K 33, A61K 35, A61K 36, A61K 38, A61K 39, A61K 41, A61K 45, A61K 47, A61K 48, A61K 49, A61K 50, A61K 51, A61K 101, A61K 103, A61K 125, A61K 127, A61K 129, A61K 131, A61K 133, A61K 135, A61P	IPC
Chemistry	Macromolecular chemistry, polymers	C08B, C08C, C08F, C08G, C08H, C08K, C08L	IPC
Chemistry	Food chemistry	A01H, A21D, A23B, A23C, A23D, A23F, A23G, A23J, A23K, A23L, C12C, C12F, C12G, C12H, C12J, C13B 10, C13B 20, C13B 30, C13B 35, C13B 40, C13B 50, C13B 99, C13D, C13F, C13J, C13K	IPC
Chemistry	Basic materials chemistry	A01N, A01P, C05B, C05C, C05D, C05F, C05G, C06B, C06C, C06D, C06F, C09B, C09C, C09D, C09F, C09G, C09H, C09J, C09K, C10B, C10C, C10F, C10G, C10H, C10J, C10K, C10L, C10M, C10N, C11B, C11C, C11D, C99Z	IPC
Chemistry	Materials, metallurgy	B22C, B22D, B22F, C01B, C01C, C01D, C01F, C01G, C03C, C04B, C21B, C21C, C21D, C22B, C22C, C22F	IPC
Chemistry	Surface technology, coating	B05C, B05D, B32B, C23C, C23D, C23F, C23G, C25B, C25C, C25D, C25F, C30B	IPC
Chemistry	Micro-structural and nano-technology	B81B, B81C, B82B, B82Y	IPC
Chemistry	Chemical engineering	B01B, B01D 1, B01D 3, B01D 5, B01D 7, B01D 8, B01D 9, B01D 11, B01D 12, B01D 15, B01D 17, B01D 19, B01D 21, B01D 24, B01D 25, B01D 27, B01D 29, B01D 33, B01D 35, B01D 36, B01D 37, B01D 39, B01D 41, B01D 43, B01D 57, B01D 59, B01D 61, B01D 63, B01D 65, B01D 67, B01D 69, B01D 71, B01F, B01J, B01L, B02C, B03B, B03C, B03D, B04B, B04C, B05B, B06B, B07B, B07C, B08B, C14C, D06B, D06C, D06L, F25J, F26B	IPC
Clean	Environmental technology	A62C, B01D 45, B01D 46, B01D 47, B01D 49, B01D 50, B01D 51, B01D 52, B01D 53, B09B, B09C, B65F, C02F, E01F 8, F01N, F23G, F23J, G01T	IPC

Mechanical Engineering	Mechanical engineering - Handling	B25J, B65B, B65C, B65D, B65G, B65H, B66B, B66C, B66D, B66F, B67B, B67C, B67D	IPC
Mechanical Engineering	Machine tools	A62D, B21B, B21C, B21D, B21F, B21G, B21H, B21J, B21K, B21L, B23B, B23C, B23D, B23F, B23G, B23H, B23K, B23P, B23Q, B24B, B24C, B24D, B25B, B25C, B25D, B25F, B25G, B25H, B26B, B26D, B26F, B27B, B27C, B27D, B27F, B27G, B27H, B27J, B27K, B27L, B27M, B27N, B30B	IPC
Mechanical Engineering	Engines, pumps, turbines	F01B, F01C, F01D, F01K, F01L, F01M, F01P, F02B, F02C, F02D, F02F, F02G, F02K, F02M, F02N, F02P, F03B, F03C, F03D, F03G, F03H, F04B, F04C, F04D, F04F, F23R, F99Z, G21B, G21C, G21D, G21F, G21G, G21H, G21J, G21K	IPC
Mechanical Engineering	Textile and paper machines	A41H, A43D, A46D, B31B, B31C, B31D, B31F, B41B, B41C, B41D, B41F, B41G, B41J, B41K, B41L, B41M, B41N, C14B, D01B, D01C, D01D, D01F, D01G, D01H, D02G, D02H, D02J, D03C, D03D, D03J, D04B, D04C, D04G, D04H, D05B, D05C, D06G, D06H, D06J, D06M, D06P, D06Q, D21B, D21C, D21D, D21F, D21G, D21H, D21J, D99Z	IPC
Mechanical Engineering	Other special machines	A01B, A01C, A01D, A01F, A01G, A01J, A01K, A01L, A01M, A21B, A21C, A22B, A22C, A23N, A23P, B02B, B28B, B28C, B28D, B29B, B29C, B29D, B29K, B29L, B33Y, B99Z, C03B, C08J, C12L, C13B 5, C13B 15, C13B 25, C13B 45, C13C, C13G, C13H, F41A, F41B, F41C, F41F, F41G, F41H, F41J, F42B, F42C, F42D	IPC
Mechanical Engineering	Thermal processes and apparatus	F22B, F22D, F22G, F23B, F23C, F23D, F23H, F23K, F23L, F23M, F23N, F23Q, F24B, F24C, F24D, F24F, F24H, F24J, F24S, F24T, F24V, F25B, F25C, F27B, F27D, F28B, F28C, F28D, F28F, F28G	IPC
Mechanical Engineering	Mechanical elements	F15B, F15C, F15D, F16B, F16C, F16D, F16F, F16G, F16H, F16J, F16K, F16L, F16M, F16N, F16P, F16S, F16T, F17B, F17C, F17D, G05G	IPC
Mechanical Engineering	Transport Technologies	B60B, B60C, B60D, B60F, B60G, B60H, B60J, B60K, B60L, B60M, B60N, B60P, B60Q, B60R, B60S, B60T, B60V, B60W, B61B, B61C, B61D, B61F, B61G, B61H, B61J, B61K, B61L, B62B, B62C, B62D, B62H, B62J, B62K, B62L, B62M, B63B, B63C, B63G, B63H, B63J, B64B, B64C, B64D, B64F, B64G	IPC
Other Fields	Furniture, games	A47B, A47C, A47D, A47F, A47G, A47H, A47J, A47K, A47L, A63B, A63C, A63D, A63F, A63G, A63H, A63J, A63K	IPC
Other Fields	Other consumer goods	A24B, A24C, A24D, A24F, A41B, A41C, A41D, A41F, A41G, A42B, A42C, A43B, A43C, A44B, A44C, A45B, A45C, A45D, A45F, A46B, A62B, A99Z, B42B, B42C, B42D, B42F, B43K, B43L, B43M, B44B, B44C, B44D, B44F, B68B, B68C, B68F, B68G, D04D, D06F, D06N, D07B, F25D, G10B, G10C, G10D, G10F, G10G, G10H, G10K	IPC
Other Fields	Civil engineering	E01B, E01C, E01D, E01F 1, E01F 3, E01F 5, E01F 7, E01F 9, E01F 11, E01F 13, E01F 15, E01H, E02B, E02C, E02D, E02F, E03B, E03C, E03D, E03F, E04B, E04C, E04D, E04F, E04G, E04H, E05B, E05C, E05D, E05F, E05G, E06B, E06C, E21B, E21C, E21D, E21F, E99Z	IPC
Trending	Robotics	B25J 9	CPC
Trending	Wireless	H04W	CPC
Trending	3D Printing	B29C64	CPC
Trending	Artificial Intelligence	G06F17, F06N5, G06N3, G10L15, G06F3, G06Q10, G06Q30, G06F9, G06Q50	CPC
Trending	Aerospace	C22F1, C08K3, B64G1, C08G59, C22C21, B64C1, C22C1, C08G73	CPC
Clean	Clean Energy	Y02E	CPC
Clean	Clean Cars	Y02T10	CPC
Clean	Clean	Y02	CPC

E Conditional Private Value Distribution

Notice increasing c will push the kink to of the distribution to the left, whereas increasing α will make the distribution more skewed.

Figure 30: Modelling the Private Value distribution



Notes: This figure represents the distribution of private values for different values of the cost (c), and the shape parameter of the Pareto distribution of innovation quality, α.