Local and Global Spillovers in Knowledge Networks

Greg Morrison, Eleftherios Giovanis,
Massimo Riccaboni, Fabio Pammolli

IMT Institute for Advanced Studies, Lucca, Italy

First Annual KnowEscape Conference, 19.11.2013

Introduction

- Knowledge spillovers (KS) are an important source of economic growth.
- One channel through which KS work are traded goods (rent spillovers, such as in Grossman & Helpman, 1991).
- Pure KS, on the other hand, are transmitted by channels such as scientific literature, labor mobility, patent information or pure imitation.

Introduction

- Knowledge spillovers (KS) are an important source of economic growth.
- One channel through which KS work are traded goods (rent spillovers, such as in Grossman & Helpman, 1991).
- Pure KS, on the other hand, are transmitted by channels such as scientific literature, labor mobility, patent information or pure imitation.

Introduction

- Knowledge spillovers (KS) are an important source of economic growth.
- One channel through which KS work are traded goods (rent spillovers, such as in Grossman & Helpman, 1991).
- Pure KS, on the other hand, are transmitted by channels such as scientific literature, labor mobility, patent information or pure imitation.

Introduction

- Knowledge spillovers (KS) are an important source of economic growth.
- One channel through which KS work are traded goods (rent spillovers, such as in Grossman & Helpman, 1991).
- Pure KS, on the other hand, are transmitted by channels such as scientific literature, labor mobility, patent information or pure imitation.

Technology flow matrices

- One possible way of analyzing KS is the use of so-called technology flow matrices.
- Technology flow matrices, as developed e.g. by Scherer (1982) and Putnam and Evenson (1994) and Verspagen (1997), describe how technological knowledge developed in one sector of the economy spills over to other sectors.
- Technology being the main source of long-run economic growth, the economic performance of nations is related to the ability to generate new knowledge domestically and the ability to apply this knowledge, as well as knowledge generated abroad, in the economy.
- Technology policy, especially in the somewhat larger countries, as well as at the international level (e.g., the EU technology programmes) has traditionally focused on the domestic generation of knowledge

Technology flow matrices

- One possible way of analyzing KS is the use of so-called technology flow matrices.
- Technology flow matrices, as developed e.g. by Scherer (1982) and Putnam and Evenson (1994) and Verspagen (1997), describe how technological knowledge developed in one sector of the economy spills over to other sectors.
- Technology being the main source of long-run economic growth, the economic performance of nations is related to the ability to generate new knowledge domestically and the ability to apply this knowledge, as well as knowledge generated abroad, in the economy.
- Technology policy, especially in the somewhat larger countries, as well as at the international level (e.g., the EU technology programmes) has traditionally focused on the domestic generation of knowledge

Technology flow matrices

- One possible way of analyzing KS is the use of so-called technology flow matrices.
- Technology flow matrices, as developed e.g. by Scherer (1982) and Putnam and Evenson (1994) and Verspagen (1997), describe how technological knowledge developed in one sector of the economy spills over to other sectors.
- Technology being the main source of long-run economic growth, the economic performance of nations is related to the ability to generate new knowledge domestically and the ability to apply this knowledge, as well as knowledge generated abroad, in the economy.
- Technology policy, especially in the somewhat larger countries, as well as at the international level (e.g., the EU technology programmes) has traditionally focused on the domestic generation of knowledge

Technology flow matrices

- One possible way of analyzing KS is the use of so-called technology flow matrices.
- Technology flow matrices, as developed e.g. by Scherer (1982) and Putnam and Evenson (1994) and Verspagen (1997), describe how technological knowledge developed in one sector of the economy spills over to other sectors.
- Technology being the main source of long-run economic growth, the economic performance of nations is related to the ability to generate new knowledge domestically and the ability to apply this knowledge, as well as knowledge generated abroad, in the economy.
- Technology policy, especially in the somewhat larger countries, as well as at the international level (e.g., the EU technology programmes) has traditionally focused on the domestic generation of knowledge

Technology flow matrices

- One possible way of analyzing KS is the use of so-called technology flow matrices.
- Technology flow matrices, as developed e.g. by Scherer (1982) and Putnam and Evenson (1994) and Verspagen (1997), describe how technological knowledge developed in one sector of the economy spills over to other sectors.
- Technology being the main source of long-run economic growth, the economic performance of nations is related to the ability to generate new knowledge domestically and the ability to apply this knowledge, as well as knowledge generated abroad, in the economy.
- Technology policy, especially in the somewhat larger countries, as well as at the international level (e.g., the EU technology programmes) has traditionally focused on the domestic generation of knowledge

A network view of knowledge spillovers

Case study: Solar and wind innovation reflected in booming patents (Bettencourt et al., 2013). The Obama administration distributed $9 billion in economic stimulus funds to solar and wind projects in 2009-11. Knowledge spillovers are found to stimulate innovative activity, especially foreign spillovers (Braun et al., 2011)

- What is the importance of foreign vs domestic sources of knowledge?
- How does knowledge diffuse within and across borders?

A network view of knowledge spillovers

Case study: Solar and wind innovation reflected in booming patents (Bettencourt et al., 2013). The Obama administration distributed $9 billion in economic stimulus funds to solar and wind projects in 2009-11. Knowledge spillovers are found to stimulate innovative activity, especially foreign spillovers (Braun et al., 2011)

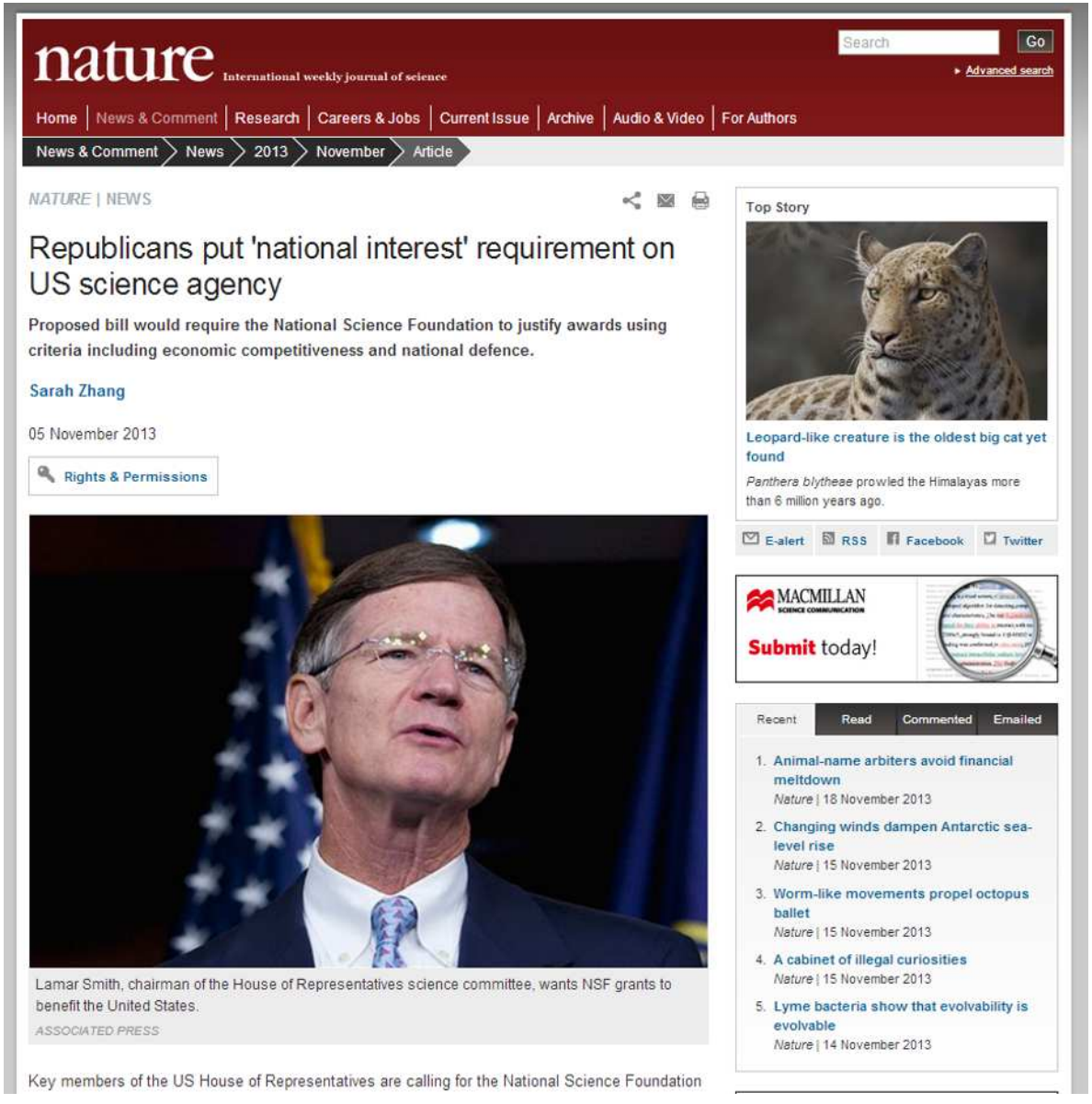
- What is the importance of foreign vs domestic sources of knowledge?
- How does knowledge diffuse within and across borders?

A network view of knowledge spillovers

Case study: Solar and wind innovation reflected in booming patents (Bettencourt et al., 2013). The Obama administration distributed $9 billion in economic stimulus funds to solar and wind projects in 2009-11. Knowledge spillovers are found to stimulate innovative activity, especially foreign spillovers (Braun et al., 2011)

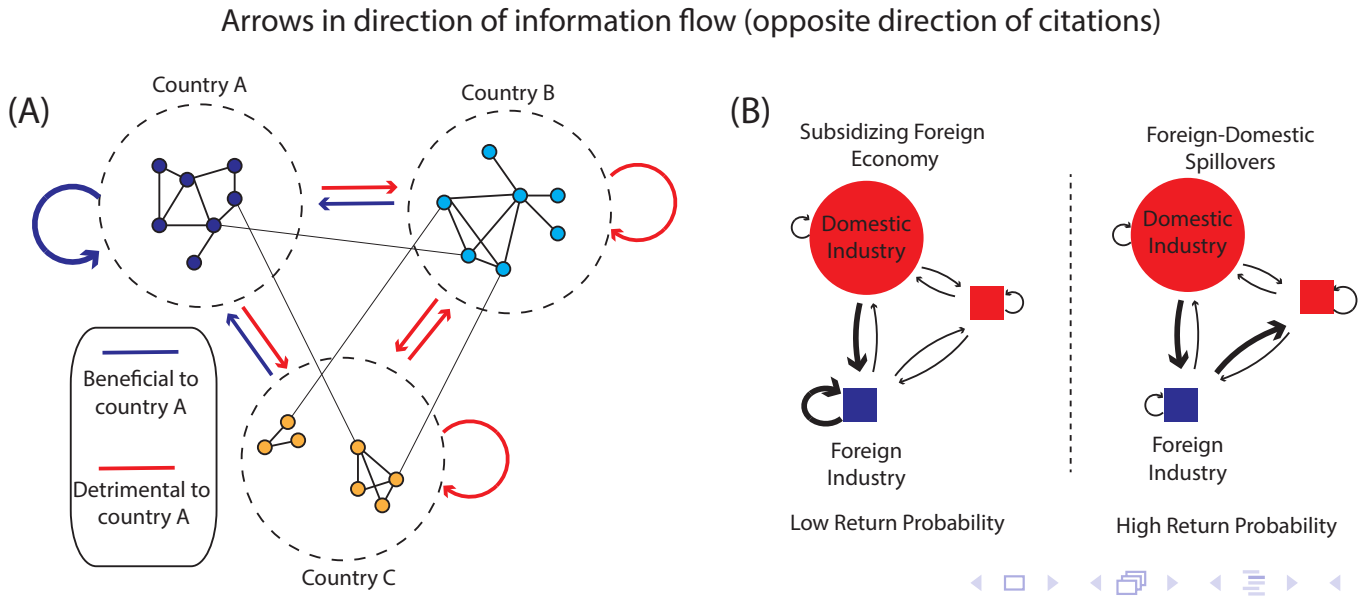
- What is the importance of foreign vs domestic sources of knowledge?
- How does knowledge diffuse within and across borders?

National interest and knowledge diffusion



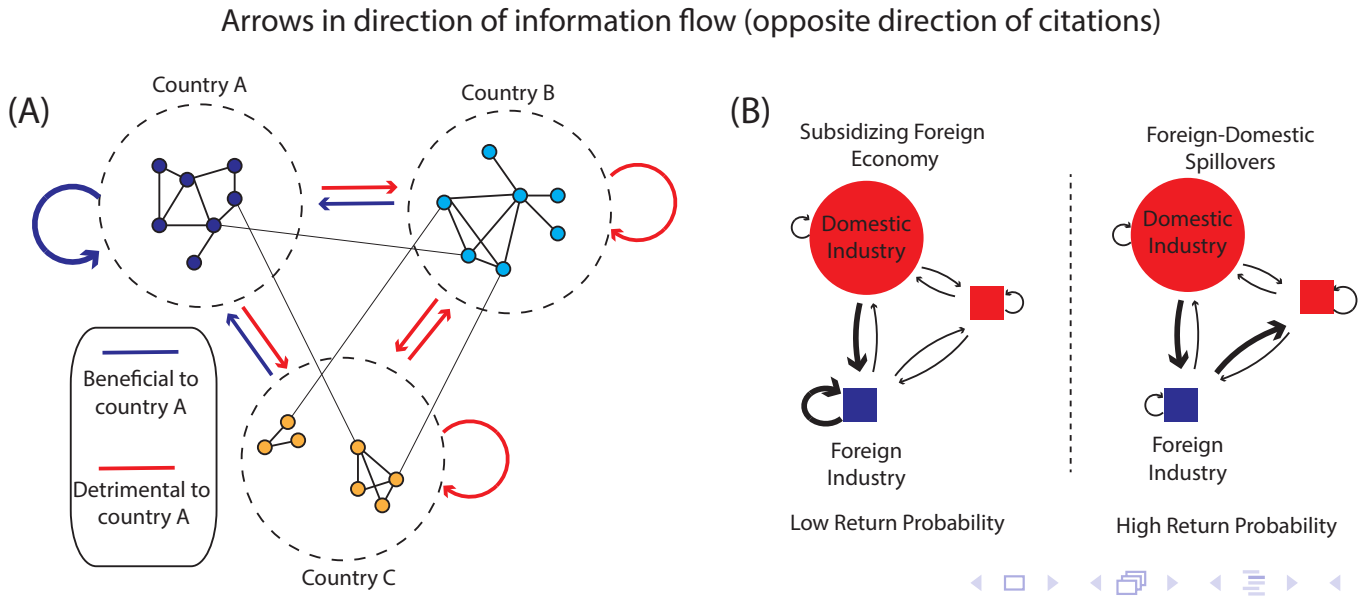
Competition and cooperation in knowledge networks

- The sectors of each national economy are nodes and citations between the sectors are edges.
- Information flow that benefits foreign classes (red arrows) is more likely to be detrimental to the domestic economy than information flow within the domestic economy.
- If a domestic industry is funded and is heavily cited by a foreign sector, the information will likely move across the border (the thick arrow in B) to the detriment of the domestic economy.
- The likelihood of the information returning to the domestic economy depends on the network topology.



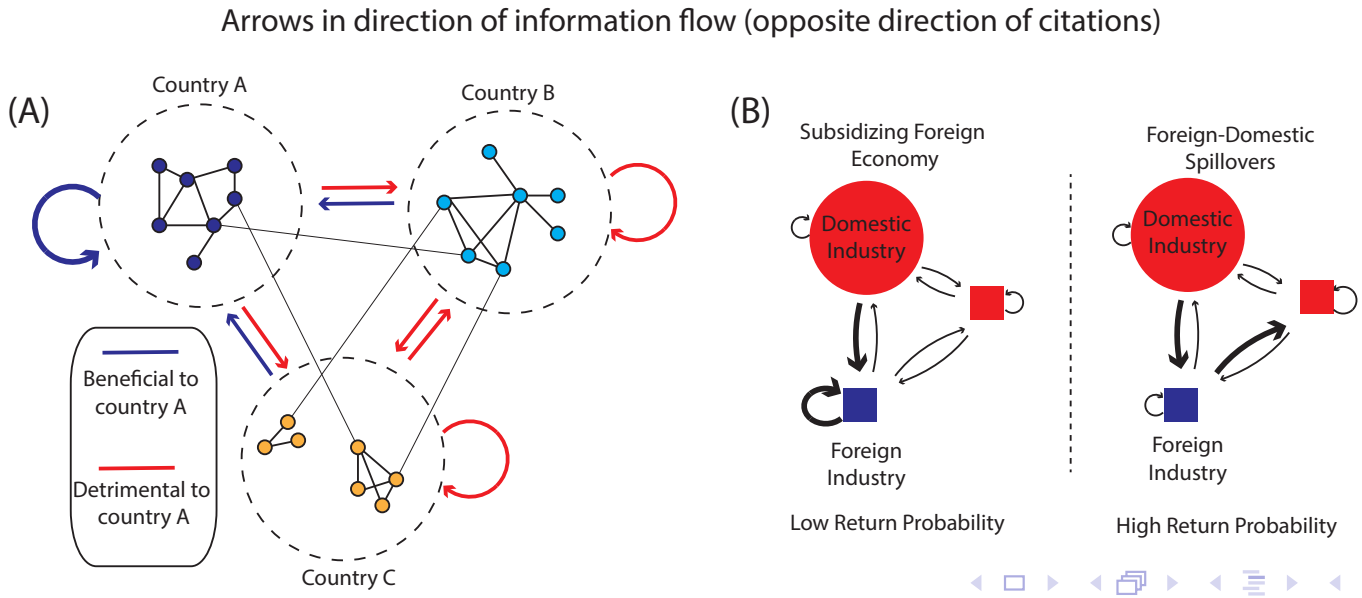
Competition and cooperation in knowledge networks

- The sectors of each national economy are nodes and citations between the sectors are edges.
- Information flow that benefits foreign classes (red arrows) is more likely to be detrimental to the domestic economy than information flow within the domestic economy.
- If a domestic industry is funded and is heavily cited by a foreign sector, the information will likely move across the border (the thick arrow in B) to the detriment of the domestic economy.
- The likelihood of the information returning to the domestic economy depends on the network topology.



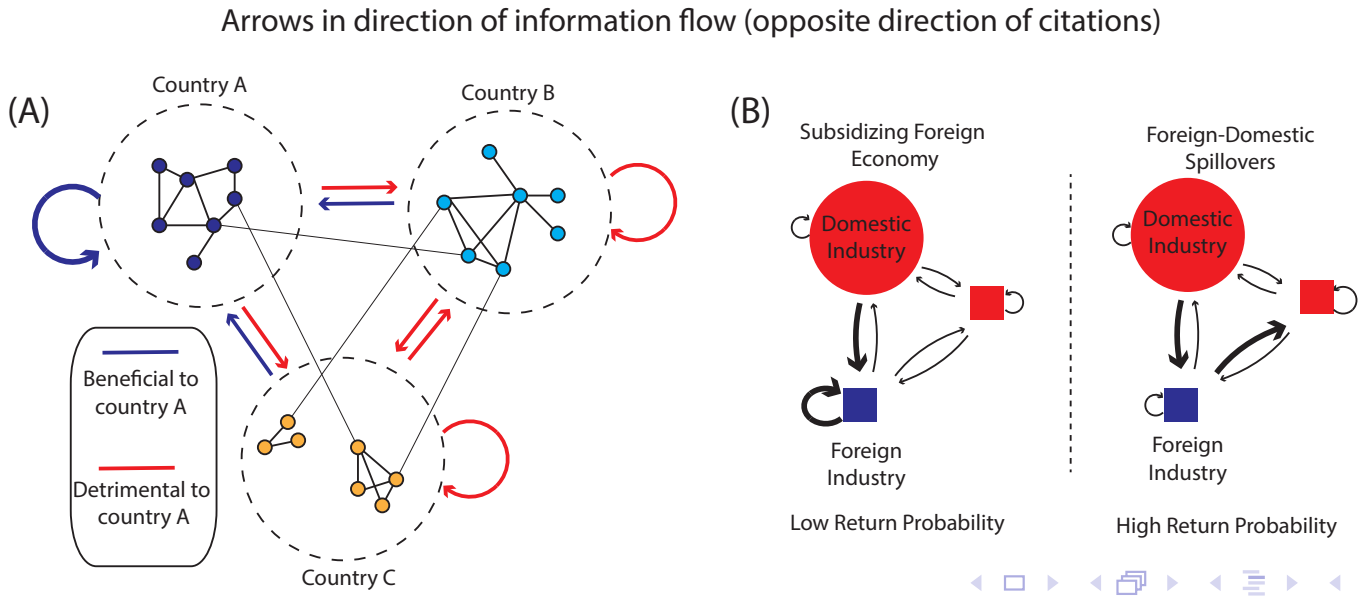
Competition and cooperation in knowledge networks

- The sectors of each national economy are nodes and citations between the sectors are edges.
- Information flow that benefits foreign classes (red arrows) is more likely to be detrimental to the domestic economy than information flow within the domestic economy.
- If a domestic industry is funded and is heavily cited by a foreign sector, the information will likely move across the border (the thick arrow in B) to the detriment of the domestic economy.
- The likelihood of the information returning to the domestic economy depends on the network topology.



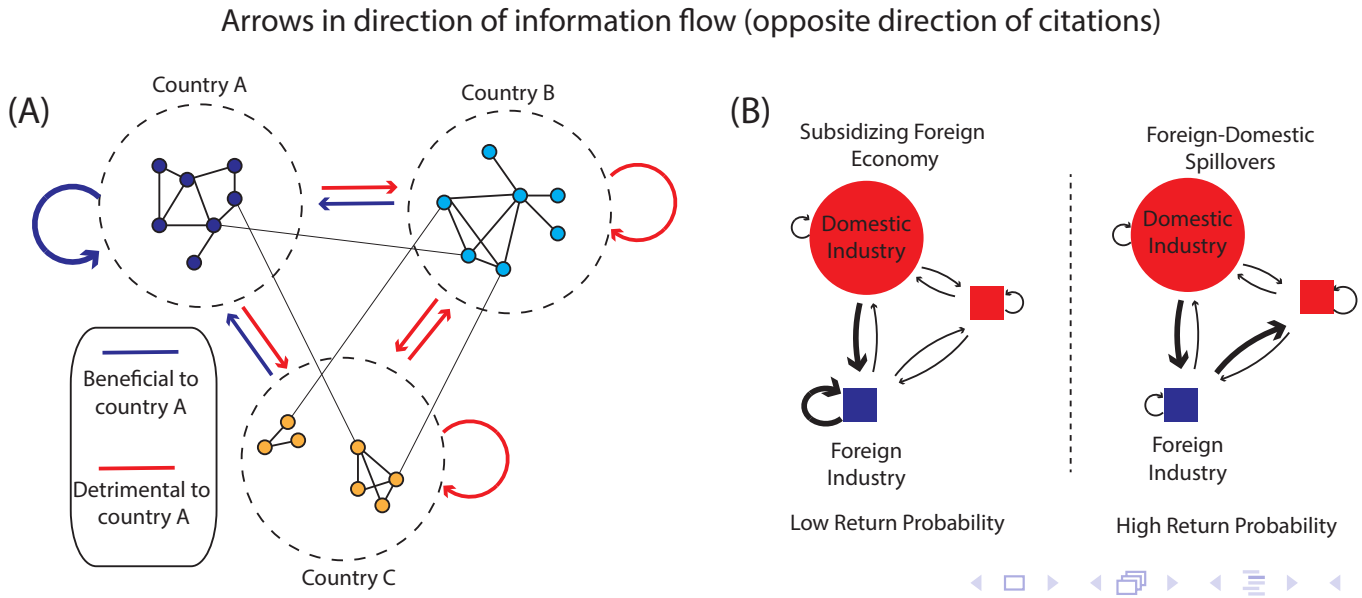
Competition and cooperation in knowledge networks

- The sectors of each national economy are nodes and citations between the sectors are edges.
- Information flow that benefits foreign classes (red arrows) is more likely to be detrimental to the domestic economy than information flow within the domestic economy.
- If a domestic industry is funded and is heavily cited by a foreign sector, the information will likely move across the border (the thick arrow in B) to the detriment of the domestic economy.
- The likelihood of the information returning to the domestic economy depends on the network topology.



Competition and cooperation in knowledge networks

- The sectors of each national economy are nodes and citations between the sectors are edges.
- Information flow that benefits foreign classes (red arrows) is more likely to be detrimental to the domestic economy than information flow within the domestic economy.
- If a domestic industry is funded and is heavily cited by a foreign sector, the information will likely move across the border (the thick arrow in B) to the detriment of the domestic economy.
- The likelihood of the information returning to the domestic economy depends on the network topology.



Information flow and Random Walks

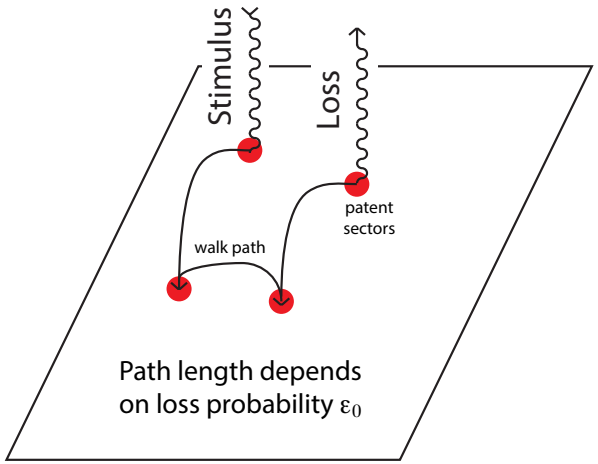
The probability that a patent filed in sector i in year t provides information to class j in the future is assumed to have the form (Zhirov et. al., 2010)

$$Pr(j \text{ benefits} \mid \text{info at } i) \propto c_{i \leftarrow j}(t)$$

Information flow through the network can be viewed as a random walk (Newman, 2005) with transition probability

$$p^{(0)}_{i \rightarrow j} = \frac{c_{i \leftarrow j}}{\sum_k c_{i \leftarrow k}}$$

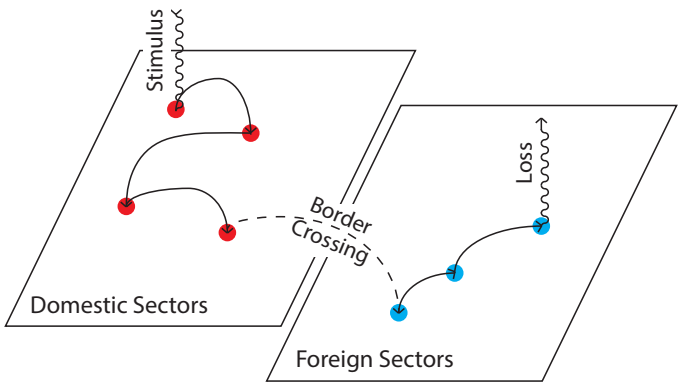
Random walks and loss



The information at class i spills over to j with probability

$$p_{i\to j}^{(loss)} = (1-\epsilon_0)\frac{c_{i\leftarrow j}}{\sum_k c_{i\leftarrow k}}$$

with ϵ_0 the probability of the walk ending without a spillover.

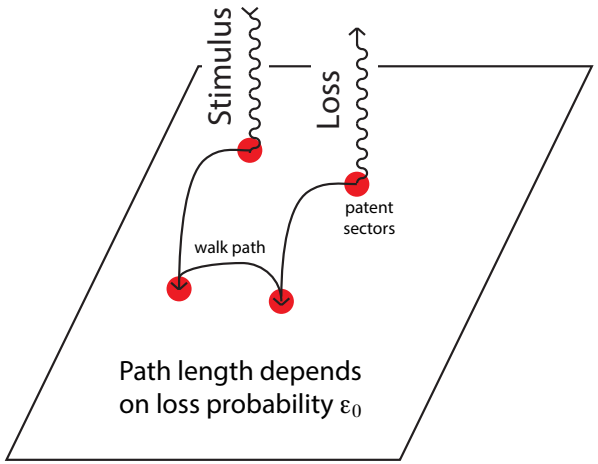


If borders matter,

$$p_{i\to j}^{(loss)} = (1-\epsilon_{i\to j})\frac{c_{i\leftarrow j}}{\sum_k c_{i\leftarrow k}}$$

with $\epsilon_{i\to j}$ a loss probability that depends on the start and end location.

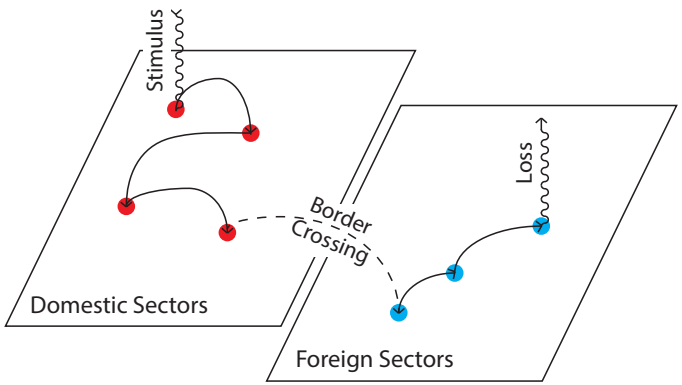
Random walks and loss



The information at class i spills over to j with probability

$$p_{i\to j}^{(loss)} = (1-\epsilon_0)\frac{c_{i\leftarrow j}}{\sum_k c_{i\leftarrow k}}$$

with ϵ_0 the probability of the walk ending without a spillover.



If borders matter,

$$p_{i\to j}^{(loss)} = (1-\epsilon_{i\to j})\frac{c_{i\leftarrow j}}{\sum_k c_{i\leftarrow k}}$$

with $\epsilon_{i\to j}$ a loss probability that depends on the start and end location.

Asymmetric discount at borders

From the perspective of a country, information that flows *out* may represent value lost, but information that flows *in* is beneficial. From the domestic perspective,

- Domestic citations are good: $\epsilon_{d \to d} = \epsilon_0$
- Domestic-to-foreign citations are good: $\epsilon_{f \to d} = \epsilon_0$
- Foreign-to-domestic citations are less beneficial: $\epsilon_{d \to f} > \epsilon_0$
- Foreign-to-foreign citations are less beneficial: $\epsilon_{f \to f} > \epsilon_0$

We define ϵ as the elevated probability of loss for information flow in a foreign economy:

$$\epsilon_{d \to f} = \epsilon_{f \to f} = 1 - (1 - \epsilon_0)(1 - \epsilon) = \bar{\epsilon}$$

PageRank and Domestic Reinsertion

In PageRank (Franceschet, 2011), a random walker that becomes lost is re-inserted into the network:

$$p_{i\to j}^{(PR)} = (1-\epsilon_0)p_{i\to j}^{(0)} + \frac{\epsilon_0}{N}$$

Nations may be interested in the importance of *domestic* sectors, suggesting reinsertion within the domestic economy only (Haveliwala, 2003):

$$\begin{aligned} p_{i\to j}^{x\to d} &= (1-\epsilon_0)p_{i\to j}^{(0)} + \frac{1}{n}\left[\epsilon_0 + (1-\epsilon_0)\epsilon f_i^{(d)}\right] \\ p_{i\to j}^{x\to f} &= (1-\bar{\epsilon})p_{i\to j}^{(0)} \end{aligned}$$

with $f_i^{(d)}$ is the fraction of domestic citations i receives.

Steady State Probability and Ranking

Information will flow readily through central classes, so we can rank domestic sectors through the steady state probability of occupancy:

$$R_i^{(d)}(\epsilon) = \sum_{j \in \mathbf{D}} p_{j \to i}^{d \to d}(\epsilon) R_j^{(d)}(\epsilon) + \sum_{l \in \mathbf{F}} p_{l \to i}^{f \to d}(\epsilon) R_l^{(f)}(\epsilon)$$

$$R_k^{(f)}(\epsilon) = \sum_{j \in \mathbf{D}} p_{j \to k}^{d \to f}(\epsilon) R_j^{(d)}(\epsilon) + \sum_{l \in \mathbf{F}} p_{l \to i}^{f \to f}(\epsilon) R_k^{(f)}(\epsilon).$$

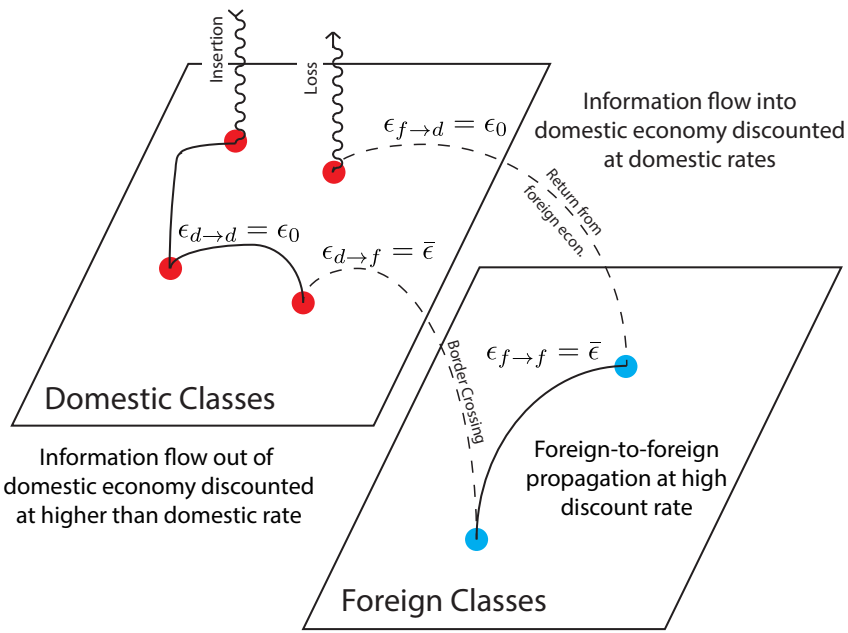
Rankings from by direct spillovers within the domestic economy and higher-order spillovers through foreign economies:

$$\mathbf{R}^{(d)}(\epsilon) = \left(\mathbf{P}_\epsilon^{d \to d} + \mathbf{P}_\epsilon^{f \to d} [\mathbf{1} - \mathbf{P}_\epsilon^{f \to f}]^{-1} \mathbf{P}_\epsilon^{d \to f} \right) \mathbf{R}^{(d)}(\epsilon),$$

with $(\mathbf{P}_\epsilon^{x \to y})_{ji} = p_{i \to j}^{x \to y}$.

Border-sensitive, asymmetric random walk

In our model, the walker is always inserted into the domestic economy, since the national government will only fund domestic classes directly. While propagating within the national economy a low discount rate is assumed, but once the information flows into the foreign economy the discount rate is elevated. The discounting of future value is reflected in the teleportation probability, with the random walk suddenly ending and the walker being re-inserted as a new investment in the domestic economy.



Data

- To calculate ϵ two main databases have been used.
- The first is the OECD Triadic Patent Families (TPF) database which covers patent applications filed to EPO), JPO and USPTO. The data was compiled using patent linkages provided in Patent Statistical (PATSTAT) Database of April 2013.
- The second data base is the OECD Citations database which covers citations of patent and non-patent literature (NPL) in patents taken at the EPO or filed through the Patent Cooperation Treaty (PCT). The data is derived from EPOs PATSTAT database, April 2013.
- The public and private R&D expenditures as well as the value added have been derived by OECD in 2005 US dollars constant prices.
- The map of the data has as follows: R&D expenditures, value added and are assigned to sectors i at time-year t.
- The data frequency is annual. The period examined in this study is 1987-2005.

Data

- To calculate ϵ two main databases have been used.
- The first is the OECD Triadic Patent Families (TPF) database which covers patent applications filed to EPO), JPO and USPTO. The data was compiled using patent linkages provided in Patent Statistical (PATSTAT) Database of April 2013.
- The second data base is the OECD Citations database which covers citations of patent and non-patent literature (NPL) in patents taken at the EPO or filed through the Patent Cooperation Treaty (PCT). The data is derived from EPOs PATSTAT database, April 2013.
- The public and private R&D expenditures as well as the value added have been derived by OECD in 2005 US dollars constant prices.
- The map of the data has as follows: R&D expenditures, value added and are assigned to sectors i at time-year t.
- The data frequency is annual. The period examined in this study is 1987-2005.

Data

- To calculate ϵ two main databases have been used.
- The first is the OECD Triadic Patent Families (TPF) database which covers patent applications filed to EPO), JPO and USPTO. The data was compiled using patent linkages provided in Patent Statistical (PATSTAT) Database of April 2013.
- The second data base is the OECD Citations database which covers citations of patent and non-patent literature (NPL) in patents taken at the EPO or filed through the Patent Cooperation Treaty (PCT). The data is derived from EPOs PATSTAT database, April 2013.
- The public and private R&D expenditures as well as the value added have been derived by OECD in 2005 US dollars constant prices.
- The map of the data has as follows: R&D expenditures, value added and are assigned to sectors i at time-year t.
- The data frequency is annual. The period examined in this study is 1987-2005.

Data

- To calculate ϵ two main databases have been used.
- The first is the OECD Triadic Patent Families (TPF) database which covers patent applications filed to EPO), JPO and USPTO. The data was compiled using patent linkages provided in Patent Statistical (PATSTAT) Database of April 2013.
- The second data base is the OECD Citations database which covers citations of patent and non-patent literature (NPL) in patents taken at the EPO or filed through the Patent Cooperation Treaty (PCT). The data is derived from EPOs PATSTAT database, April 2013.
- The public and private R&D expenditures as well as the value added have been derived by OECD in 2005 US dollars constant prices.
- The map of the data has as follows: R&D expenditures, value added and are assigned to sectors i at time-year t.
- The data frequency is annual. The period examined in this study is 1987-2005.

Data

- To calculate ϵ two main databases have been used.
- The first is the OECD Triadic Patent Families (TPF) database which covers patent applications filed to EPO), JPO and USPTO. The data was compiled using patent linkages provided in Patent Statistical (PATSTAT) Database of April 2013.
- The second data base is the OECD Citations database which covers citations of patent and non-patent literature (NPL) in patents taken at the EPO or filed through the Patent Cooperation Treaty (PCT). The data is derived from EPOs PATSTAT database, April 2013.
- The public and private R&D expenditures as well as the value added have been derived by OECD in 2005 US dollars constant prices.
- The map of the data has as follows: R&D expenditures, value added and are assigned to sectors i at time-year t.
- The data frequency is annual. The period examined in this study is 1987-2005.

Data

- To calculate ϵ two main databases have been used.
- The first is the OECD Triadic Patent Families (TPF) database which covers patent applications filed to EPO), JPO and USPTO. The data was compiled using patent linkages provided in Patent Statistical (PATSTAT) Database of April 2013.
- The second data base is the OECD Citations database which covers citations of patent and non-patent literature (NPL) in patents taken at the EPO or filed through the Patent Cooperation Treaty (PCT). The data is derived from EPOs PATSTAT database, April 2013.
- The public and private R&D expenditures as well as the value added have been derived by OECD in 2005 US dollars constant prices.
- The map of the data has as follows: R&D expenditures, value added and are assigned to sectors i at time-year t.
- The data frequency is annual. The period examined in this study is 1987-2005.

Econometrics

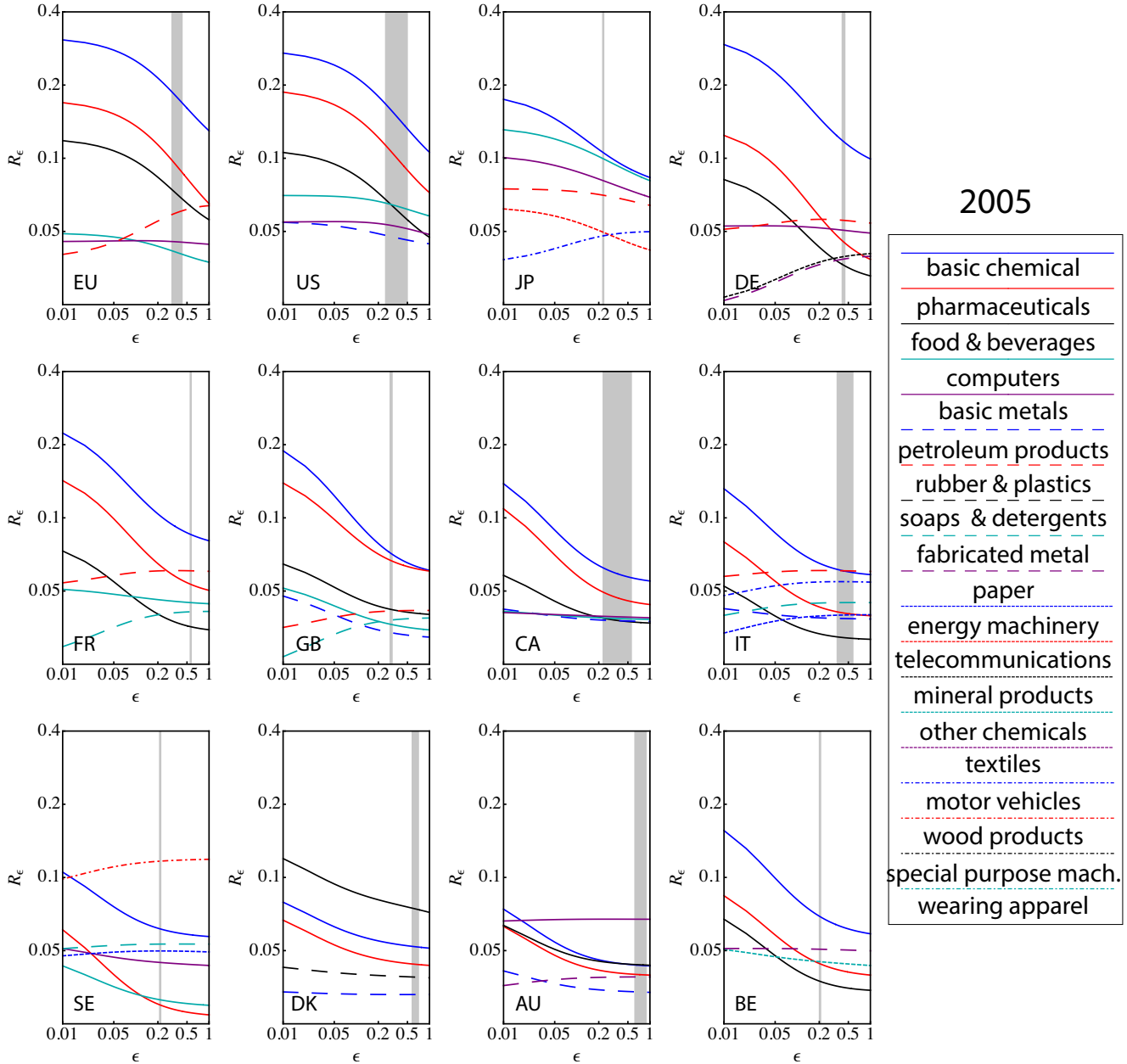
To link the centrality to the economic output of sectors of innovation in a nation's economy, we perform regressions to determine its relationship to public R&D and value added:

$$r_i(\epsilon; t) \sim \beta_{RD}(\epsilon) E_i(t) + \alpha_{RD}(\epsilon) + \boldsymbol{\gamma}_{RD}(\epsilon) \cdot \mathbf{x} + u_i(t) \quad (1)$$

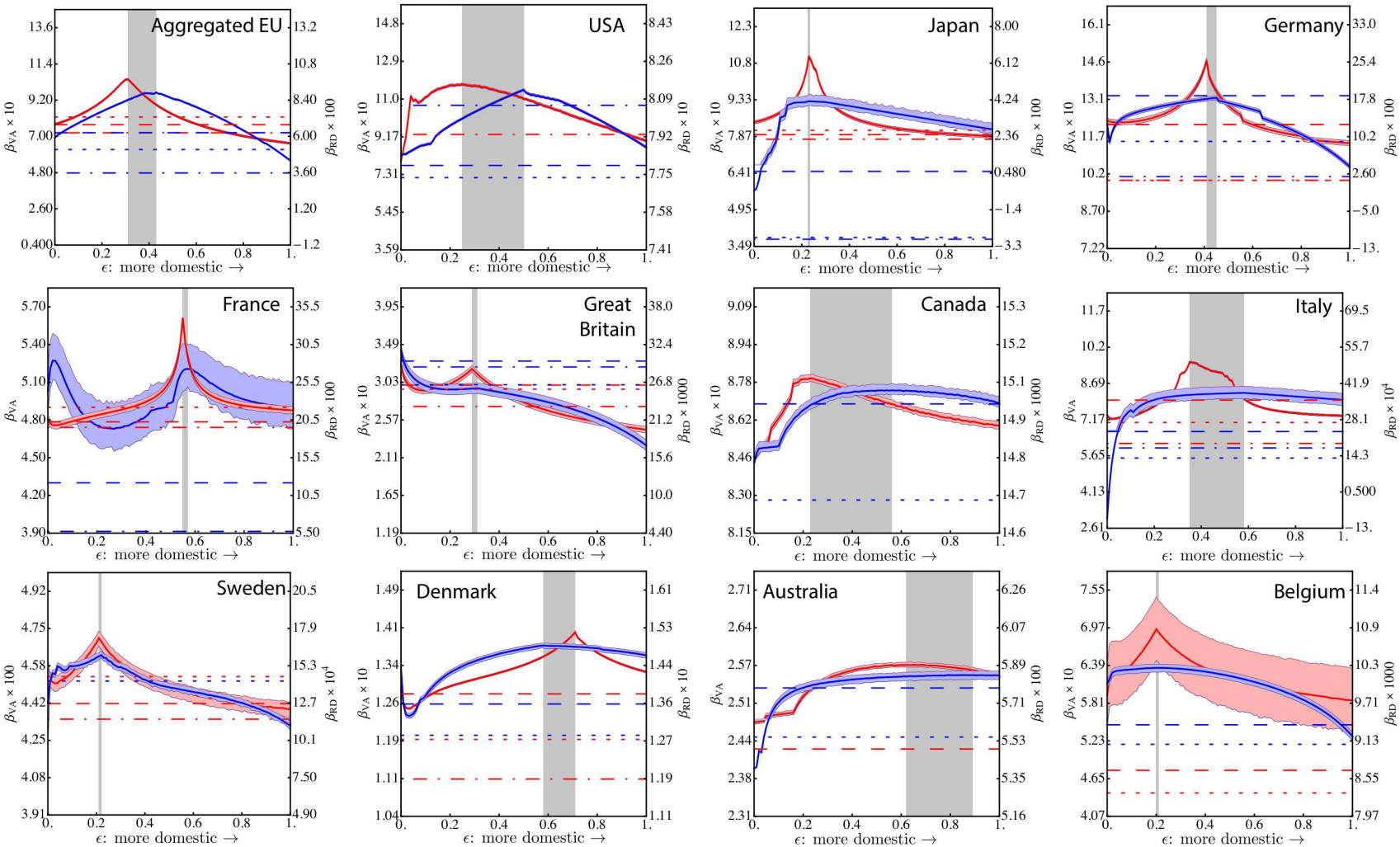
$$V_i(t) \sim \beta_{VA}(\epsilon) r_i(\epsilon; t) + \alpha_{VA}(\epsilon) + \gamma_{VA}(\epsilon) \cdot \mathbf{y} + e_i(t) \quad (2)$$

where $r_i(\epsilon; t)$ is the logarithm of the centrality of a sector i for a single country in year t ($r_i(\epsilon; t) = \log[R_i(\epsilon; t)]$ for a chosen value of our free parameter ϵ, $E_t(t)$ is the logarithm of the public R&D expenditure of a country in sector i in year t, $u_i(t)$ and $e_i(t)$ are noise terms, $\mathbf{x}$ and $\mathbf{y}$ are control parameters (both including private R&D and the number of patents filed and $\mathbf{y}$ including public R&D as well) and the α, β, and γ terms are all regression coefficients that depend on ϵ, but not i or t. Since independent evaluation of eq. (1) and (2) is problematic (the dependent variables and the noise terms are correlated), we use Seemingly Unrelated Regressions (SURE).

Results, I

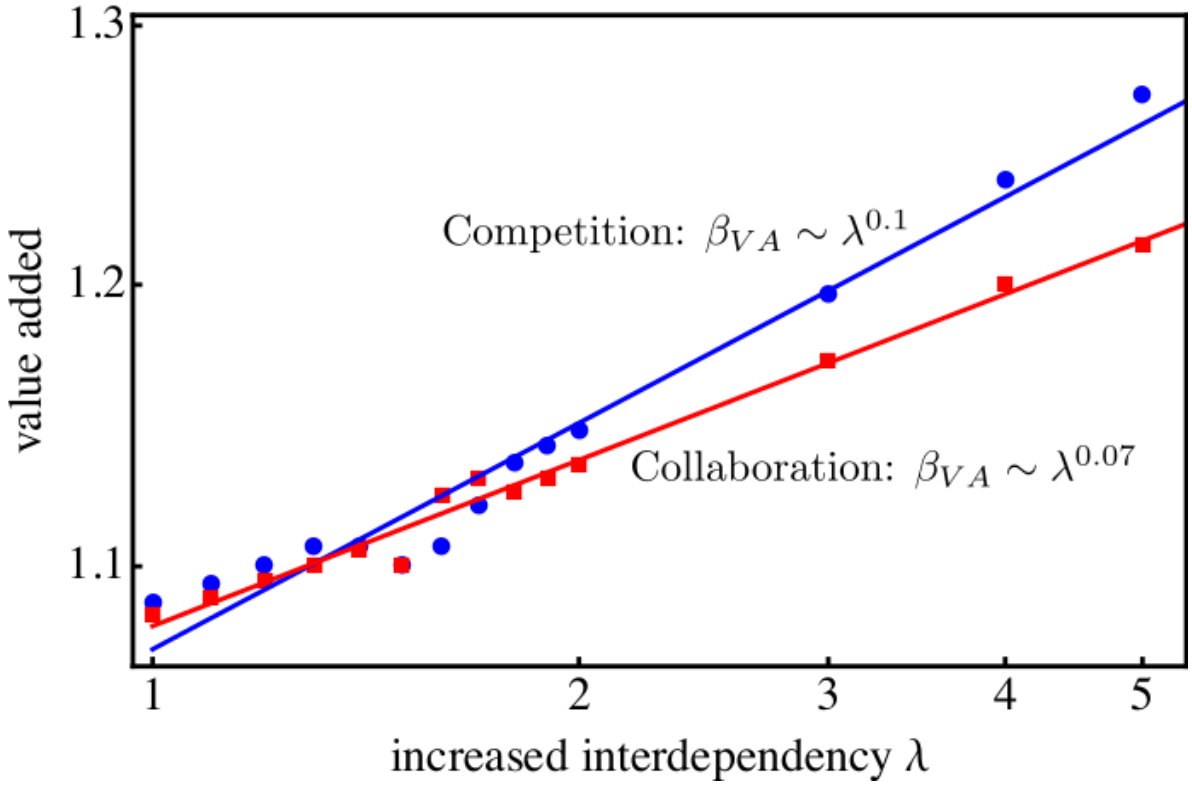

EU
US
JP
DE
FR
GB
CA
IT
SE
DK
AU
BE
R_ϵ
ϵ
0.4
0.2
0.1
0.05
0.01 0.05 0.2 0.5 1
2005
basic chemical
pharmaceuticals
food & beverages
computers
basic metals
petroleum products
rubber & plastics
soaps & detergents
fabricated metal
paper
energy machinery
telecommunications
mineral products
other chemicals
textiles
motor vehicles
wood products
special purpose mach.
wearing apparel

Results, II



The regression coefficients with β_{VA} (blue) and β_{RD} (red). Lines show the regression coefficient if centrality is measured by: N_i (dashed), the global PageRank P_i (dotted),the autarchy PageRank P_i' (dashed-dotted), with red and blue corresponding to R&D and value added respectively.

European Research Area



The value added as the simulated interdependency within the EU is increased (with the number of citations between EU countries increased by a factor λ with other citations unchanged) at $\epsilon = 0.2$. A 10% increase in the intra-EU citation rate will increase value added by $\sim 0.5\%$, and a 100% increase in the citation rate increases value added by $\sim 5\%$.

Conclusion

- we present a new method for determining the central sectors for a nation in the global innovation network which distinguishes between classes of global vs. domestic importance using a single parameter
- by implementing the centrality as a knowledge multiplier we show that this measure is typically a better predictor of value added than other common measures of centrality
- we show that the national value added is greatest for strategies that balance both international cooperation with competition
- we show that raising intra-EU collaboration may become beneficial for individual nations only for significant increases in within-EU citations.