Survival of the Fittest? Using Network Methods to Assess the Diffusion of Project Design Concepts

Kenneth Chomitz, Pierre-Yves Koenig, Guy Melançon, Benjamin Renoust

Survival of the fittest? Using network methods to assess the diffusion of project design concepts

Kenneth Chomitz, Pierre-Yves Koenig, Guy Melançon, and Benjamin Renoust

JEL classification O190

Keywords development effectiveness; network analysis; project evaluation

Contents

Tables

Figures

Abbreviations and Acronyms

ADB	Asian Development Bank
AfDB	African Development Bank
EBRD	European Bank for Reconstruction and Development
IEG	Independent Evaluation Group
IFAD	International Fund for Agricultural Development
IFC	International Finance Corporation
OCR	optical character recognition

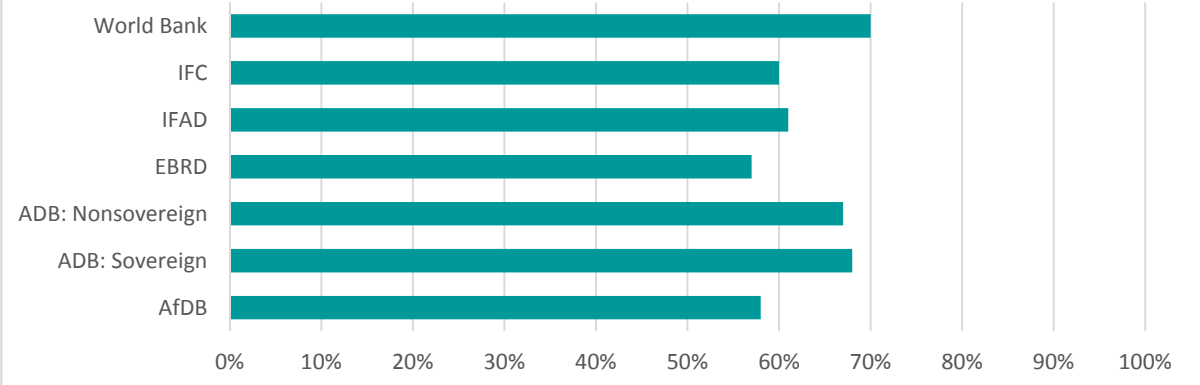
Acknowledgments

This work was undertaken at the Independent Evaluation Group under Chomitz's direction. We are grateful to Alex McKenzie for advice and assistance on data access and interpretation. Helpful comments were provided by Marie Gaarder and by participants in a seminar at William and Mary College for helpful comments. Parsing-out of the project component descriptions was undertaken by Shashi Godbole in a prior project.

1. Motivation and overview

1.1 About a third of development projects fail to achieve satisfactory outcomes, according to agencies' independent evaluation units. (Figure 1) To a large extent, these outcomes appear to be baked into projects at their inception. This is evident in Figure 2, which shows a strong correlation between project outcome[1] and quality at entry. Quality at entry, in turn, is closely associated with quality and relevance of project design, the focus of this paper [2].

Figure 1 Proportion of international donor agency projects with successful outcomes



Note: ADB = Asian Development Bank; AfDB = African Development Bank; EBRD = European Bank for Reconstruction and Development; IFC = International Finance Corporation; IFAD = International Fund for Agricultural Development. "Outcome" is an index combining efficacy, relevance, and efficiency. "Success" is a rating of four or better on a six-point scale (IFAD, IFC, World Bank); three or better on a four-point scale (ADB, EBRD); or 2.5 or better on a four-point scale (AfDB). Based on most recent available reviews of independent evaluations or independent validations of self-assessments.

Source: World Development Report 2016

1.2 This paper addresses the question: are better-designed and/or better-performing projects more likely to be emulated? One would hope so, but the answer is by no means obvious. The development business is subject to fads: hype and hope are piled onto fashionable interventions, which are widely replicated before evidence of efficacy is available. Evaluations by the World Bank's Independent Evaluation Group (IEG) have found pervasive, structural failures that inhibit learning from experience (IEG 2014, 2015). But the impact on diffusion of project ideas has not been quantified.

1.3 To address this question, this paper introduces a novel approach: a network model of the diffusion of project ideas. The motivation is as follows. Investment projects consist of

[1] "Outcome" is a rating that incorporates achievement of objectives, relevance of objectives, and efficiency.

[2] According to evaluation guidelines, quality at entry captures the degree to which project identification, preparation and appraisal supported the achievement of planned outcomes in a way consistent with the Bank's fiduciary role. It includes but goes beyond the relevance of project design.

components. For instance, a roads project may have a component devoted to intercity roads, another devoted to rural feeder roads, and a third to build capacity at the ministry of public works. Think of the written plans for these project components as the 'genes' of development interventions. (See Table 1.) They can be passed down and recombined in subsequent projects – sometimes unchanged, sometimes with mutations. A project's success depends in part on the quality and relevance of its 'genes', and in part on 'epigenetics' including the country context, the implementing agency, and the quality of Bank supervision. In a learning organization, we might expect that better-designed or better-performing components have differentially greater fitness. That is, we might expect that these components are more likely to be incorporated into subsequent projects. However, depending on organizational incentives, it is possible that characteristics aside from design quality might affect the likelihood that a component is replicated. For instance, there could be a preference for projects that can be prepared quickly or which can disburse quickly.

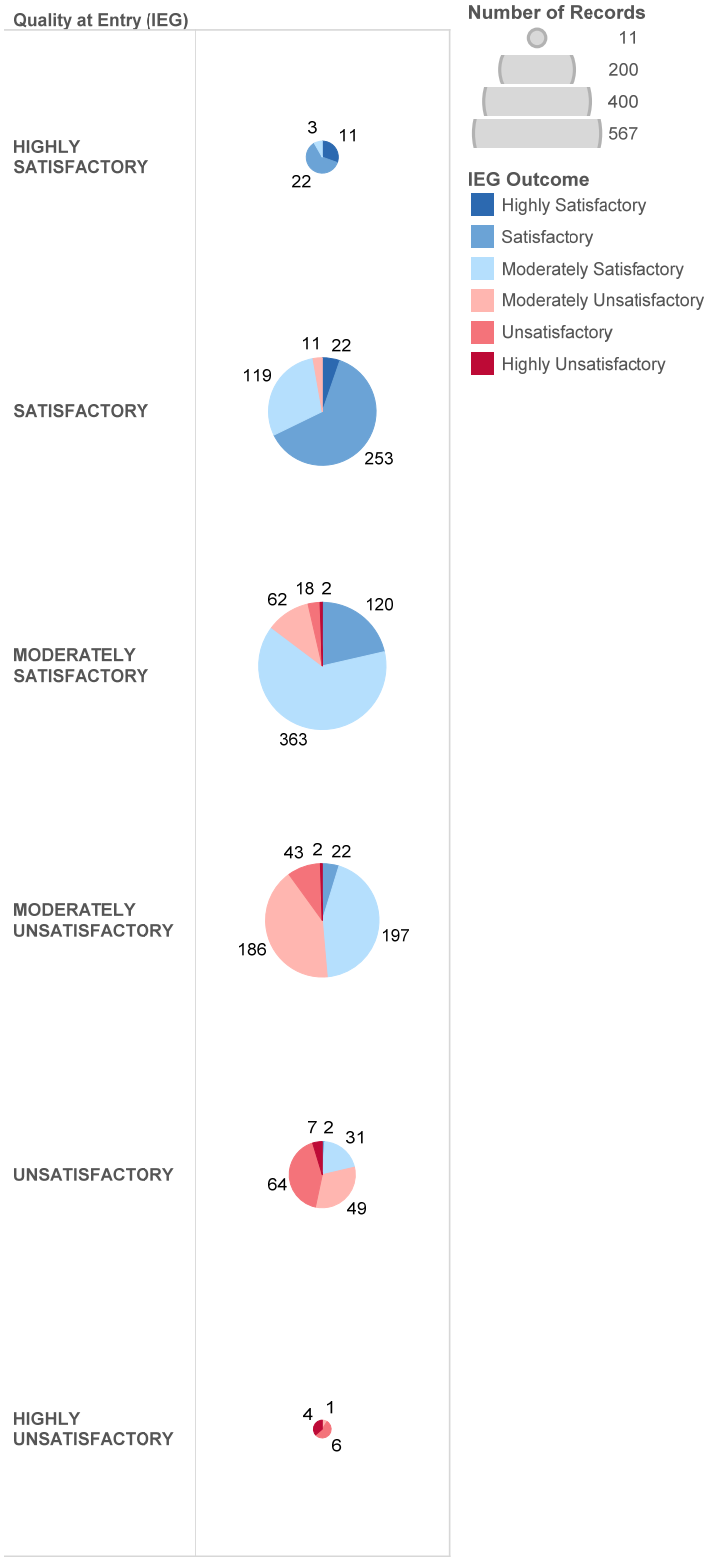
Table 1 Biological metaphors for project concept diffusion

Biology	**Economic Development**
genes	project components
chromosomes	projects
recombinations of genes via sex	remixing of components in new projects
mutations	changes in component design
epigenetic influences on gene expression	Contextual factors affecting implementation
evolutionary tree	network graph of components

1.4 To get a handle on the transmission of ideas, we construct a network graph of components of all World Bank investment projects approved over the period 1997-2012. The network graph is based on a measure of semantic similarity between each pair of components. Where that similarity exceeds a threshold, the two components are considered linked. The graph is *directed* – that is, the connecting arrow points from earlier to later components. For convenience, we will say that an earlier component has 'influenced' or 'parented' a later one if they are connected. (There is a suspicion but not a presumption of a causal connection.) We can then assess whether 'good' projects (by various measures) are more prolific in influence.

1.5 In addition to exploring the transmission of ideas, this paper serves to introduce the network graph as a tool for evaluation and research. An annex describes how to explore the graph using Tulip, an open-source software package.

Figure 2 Project outcomes as a function of Quality at Entry



Source: IEG outcomes database, downloadable at ieg.worldbankgroup.org. Investment projects evaluated FY2007-2015 based on ICR Review. **Note:** figure best viewed in color.

2. Methodology

2.1 This section describes the construction of the database containing the descriptions of project components, the derivation from it of the network graph, and the econometric analysis.

Construction of the component database

2.2 World Bank projects are described in an appraisal document, which is presented to the Bank's Board for approval. Phasing in around 1997, the Project Appraisal Document (PAD) has followed a standard format, with evolutionary change over time. These documents are publicly available in pdf or text versions. The format includes an annex entitled "Detailed Project Description" (or a variant).

2.3 To extract the components descriptions, the IEG applied text-analytic methods to all PADs for the period 1997-2012. The project-description annex was extracted from each PAD. The contents of the annex were scanned for headers describing the components. Where these could be identified, the components were parsed out and numbered sequentially. For some of the cases, components could not be distinguished, and the project was considered to have a single component. The extracted components were stored in an Access database, which included descriptors of the the associated project such as beginning and end date, sector, and country.

Construction of the network

2.4 The text data was then preprocessed in four steps. First, the text data was corrected for misspellings and other errors in the OCR process. (Table 2.) Second, words were stemmed –that is, different verb and noun forms ('visit', 'visits', 'visited') were mapped to a common root. Algorithms differ in their aggressiveness in distinguishing roots, and the choice is consequential for determining similarity between components. We chose the Porter (1980) algorithm. (Table 3.) Third, country, city and currency names were automatically identified and excluded, so that network proximity would be based on components' conceptual content rather than location. Also excluded were conjunctions, cardinal numbers, determiners, prepositions and pronouns. Finally, the entire corpus of words was manually inspected to filter out uninformative words that would not be helpful in assessing whether two components were similar in concept. Excluded words included names of currencies; most adverbs and rhetorical words ('therefore', 'however'), numbers, and generic terminology. The filtering process was subjective and carried the risk of both false negatives and false positives. (See example in Table 4.) The classification file is available on request.

2.5 Finally, components with the titles "project management", "project monitoring", "project implementation", and "project administration" were excluded from consideration. Most projects have such a component, devoted to general administration, and so their inclusion would create a confusing and meaningless cluster.

Table 2 An egregious example of poor OCR

```
Minimum Portion orTotal Cost to be Rorne by the Users for

     Devefonment and Rehabifitanon of Irriation Systems'
```

Table 3 Four stemming algorithms

WORDNET: "visa": ["visa", "visas"], "vision": ["vision", "visions"], "visit": ["visit", "visits"], "visitor": ["visitors", "visitor"], "visualization": ["visualizations", "visualization"],
SNOWBALL: "visa": ["visa", "visas"], "vise": ["vise", "vised", "vising"], "visi": ["visi", "visie"], "visibl": ["visibility", "visible", "visibly"], "vision": ["vision", "visions", "visioning"], "visit": ["visit", "visits", "visited", "visiting", "visitation", "visites", "visite", "visiter"], "visitor": ["visitors", "visitor"], "visual": ["visual", "visually", "visualizations", "visualization", "visualized", "visualizing", "visualize", "visuals"
PORTER: "visa": ["visa", "visas"], "vise": ["vise", "vised", "vising"], "visi": ["visi", "visie"], "visibl": ["visibility", "visible", "visibly"], "vision": ["vision", "visions", "visioning"], "visit": ["visit", "visits", "visited", "visiting", "visitation", "visites", "visite", "visiter"], "visitor": ["visitors", "visitor"], "visual": ["visual", "visually", "visualizations", "visualization", "visualized", "visualizing", "visualize", "visuals"],
LANCASTER: "vis": ["vision", "visibility", "visible", "vis", "visual", "visually", "visa", "visions", "visi", "visibly", "visualizations", "visualization", "visualized", "vise", "visioning", "visionary", "vised", "visualizing", "visualize", "visuals", "vising", "visat"], "visit": ["visit", "visits", "visited", "visitors", "visiting", "visitor", "visitation", "visites", "visite", "visiter"],

Table 4 Examples of manual classification of meaningful terms

Deemed meaningful	**Excluded**
Culminate Subordinated Deliberate Assessors Statewide Snow Respiratory Curb Fig *(food or abbreviation?)* Corrosion Forge	Aforementioned Meantime Southwest Feel Timescale Formerly Somewhat Kinds Lesser Distinguish Worldwide Graphic

Multidisciplinary Fold Blended Ahead Sheep Karst geotechnical	Via Merely Lac *(region or bug excretion?)*

2.6 The result is a set of components, and a set W of the keywords (i=1 to N) that occur in the entire corpus of components. Each component can be viewed as a vector c with elements corresponding to each of the N keywords. The i-th element of c takes the value:

$$c_i = (\text{frequency of i in c})*\ln(1/\text{proportion of components that include i})$$

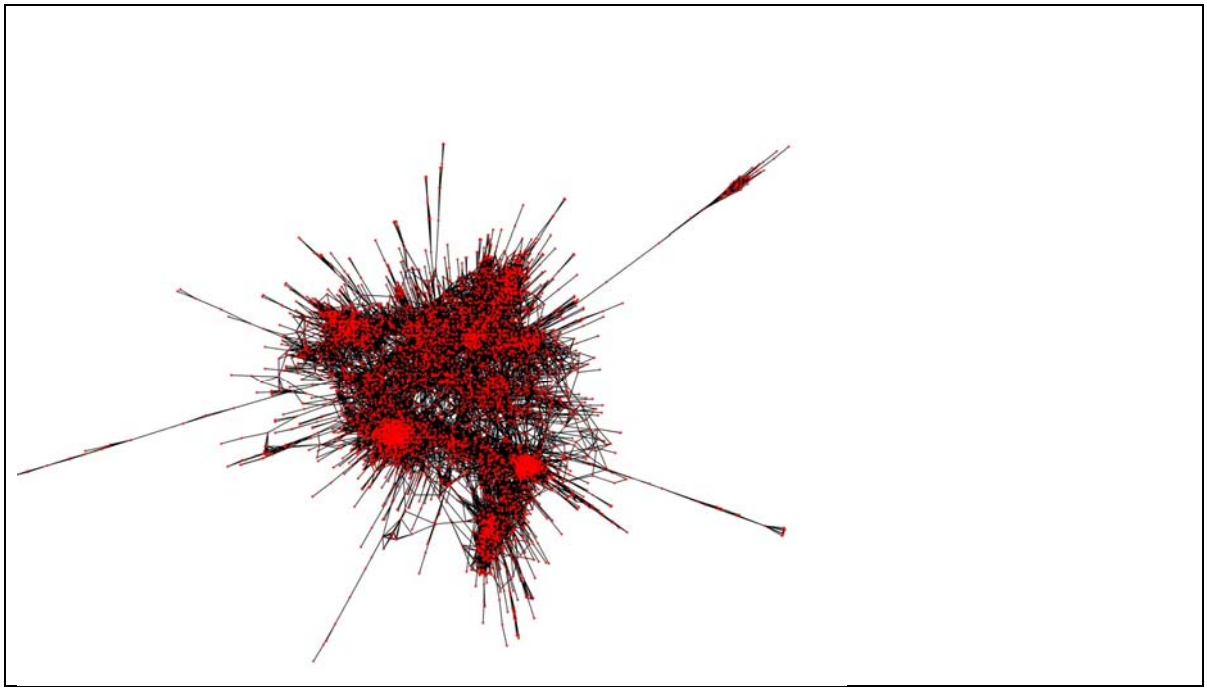
Thus c_i is 0 if word i does not occur in component c. If the word does occur in c, it receives a higher weight the more frequently it occurs within c, and the less prevalent it is in other components[3]. Thus the weight captures specificity of the word to the component.

2.7 Similarity between components is computed based on their vectors – the 'bag of words' approach. We follow Connor and Moss (2012), who apply a similarity measure based on Shannon entropy, widely used in information theory. The intuition is that the smaller the information gain from combining two vectors, the more similar they are. To ease the burden of computing the similarity matrix (which is 8913 x 8913 in size), for each component we retain only the 20 elements with the highest weight, setting the others to 0. This has the added advantage of correcting for the tendency of PADs to become more detailed and verbose over time.

2.8 The similarity matrix yielded a symmetric graph with 16,327,594 edges (links) between nodes (components), where the similarity quantified the strength of the link from 0 to 1. To aid in the visualization and manipulation of the graph, links with similarity below 0.4 were discarded, reducing the links to 74140. Even so, the resultant network is too complex for visual analysis (Figure 3.)

[3] The weighting scheme follows Salton and Buckley (1988).

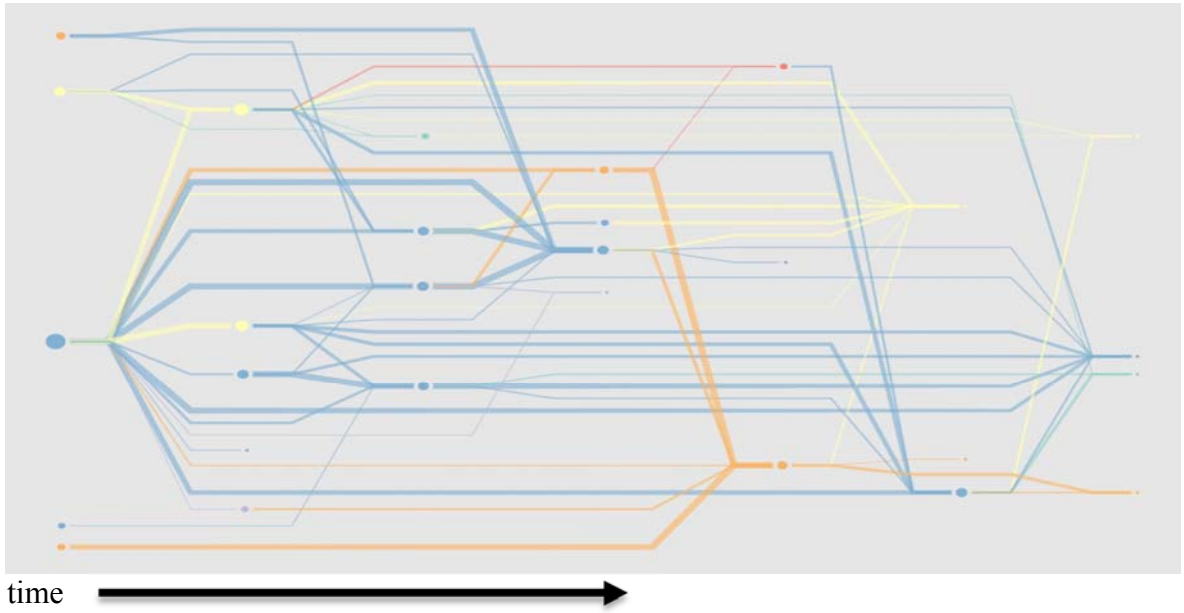
Figure 3 A visualized portion of the network



2.9 We used a layout procedure to facilitate visualization of this network. This algorithm (Noack 2004) aims at separating a network into denser sub-networks that can be individually examined. These subnetworks were expected to represent clusters of functionally similar components. Roughly speaking, the algorithm seeks to position nodes at a Euclidean distance proportional to the number of hops need to go from one node to the other in the network. We have used the Euclidean length of edges in the layout as a criterion to filter out additional edges, keeping about 95% of them. Those longer edges connect groups laid out further away from each other. We have then applied a clustering algorithm (Blondel and others 2008) using the link structure to identify groups. The Tulip database that accompanies this paper contains a pull down menu that allows screen by screen visualization of each of the derived clusters.

2.10 The graph can be arranged temporally. The Tulip visualization arranges links so that prior-approved projects are on the left. Figure 4 shows an example. The nodes (projects) can be color coded according to region, sector, or outcome rating. The width of the edges (links) corresponds to the strength of similarity. The links can be coded according to the sector, region, or rating of the source ('parent') or target ('child') component. This allows visual inspection of patterns of 'diffusion'.

Figure 4 Temporal layout of the network



Construction of 'influence' measures

2.11 For ease of terminology, we will say that component A 'influences' component B if they are similar in content and if A's approval year was prior to B's. (We recognize that causal impact, or direct diffusion of ideas, cannot be verified without additional information.) Many network-theoretic measures are available to quantify the strength of 'influence'. One could, for instance, look at connections that go from 'parents' to 'children' to 'grandchildren'. Because World Bank projects last about eight years, the number of 'generations' in our database is limited. Therefore we start with a simple measure of 'influence': the weighted sum of outlinks from a component. Project 'influence' can be computed simply as the sum of the 'influence' weights of the project's components.

Relating project quality to 'influence'

2.12 As an initial, exploratory illustration of the use of these metrics, we propose the following analysis of 'influence'. We propose three, non-mutually exclusive sets of factors that affect the degree to which a project is emulated.

Structural factors. In the Bank's matrix management system, staff and funds are organized both by region (e.g. South Asia, Sub-Saharan Africa) and sector (e.g. Agriculture, Energy). These vary substantially in staffing and lending volume. Projects in more prominent sectors and regions might be expected to have more emulators. In addition, earlier projects will have had more time to spawn emulators than later ones.

Factors related to quality. We may hypothesize that Bank staff and clients perceive the quality of projects and are more likely to build on those of higher quality or which had greater success. These perceptions may be correlated with subsequent ratings by evaluators. Perceptions may be formed at the time of appraisal, or subsequently as project results emerge. Available measures are Quality at Entry and Outcome, as independently validated by IEG at project closure[4].

Strong assumptions are needed to align these measures with a causal hypothesis linking project quality to the likelihood of emulation. For Quality at Entry, we have to assume that the *ex post* rating by IEG is strongly correlated with staff's perception of quality at time of entry itself. For Outcome, we need the even stronger assumption that early perceptions of quality are correlated with eventual rating of outcome. Of course, it is possible to abandon a causal hypothesis of emulation and still be interested to know whether 'good' projects tend to be more frequently emulated.

Factors related to ease of approval and disbursement. Multilateral development banks and their clients have strong incentives for rapid disbursal of funds. Projects which can be rapidly set up and which disburse quickly might therefore tend to attract emulators. Relevant metrics include elapsed time from concept note to first disbursement, and time from approval to disbursement of half of allocated funds.

Factors related to multiple criteria. Project size (in dollars) might plausibly affect 'influence' through all the above channels.

These considerations lead to regression models with the above factors as explanatory variables and 'influence' as the dependent variable.

Assessment of cross-region 'influence'

2.13 A question of interest is the degree to which project concepts diffuse across the Bank's organizational boundaries that correspond to global regions (such as South Asia and Sub-Saharan Africa). We have thus computed a series of indices to reveal the interplay between cross region dynamics and influence. We say that a link is *cross region* whenever it links two components implemented in distinct regions. Given a component c, we may thus compute:

- the number $r(c)$ of outgoing cross region links of component c
- similarly, we define the ratio of outgoing cross region links as $r\ (c) = r(c)/d^{+}(c)$ where $d^{+}(c)$ is the total number of outgoing links of component c.

These statistics can be computed in weighted or unweighted terms.

[4] Project closure generally coincides with the final disbursement of loan or grant funds, typically five to eight years after initiation for investment projects. Project ratings are assigned six months after closure and subsequently independently validated.

2.14 A further measure is of diversification – the degree to which links spread across many regions. For this we use Blau's diversification coefficient:

$$b(c) = 1 - \sum_{\rho \epsilon \mathcal{R}} \left(\frac{d_\rho^+(c)}{d^+(c)} \right)^2$$

where c is the component, d^+ the total number of outgoing links, and ρ indexes the region.

3. Results

Using the network for visual analysis

3.1 One goal of this work is to provide a method of exploratory data analysis. The technique allows visualization of clusters of conceptually similar projects and their evolution over time, in a way that is not constrained by predetermined sector labels or keywords. Whether or not the components are *causally* related, casual inspection of the network graph shows them to be *conceptually* related. For instance, the densely linked road management subgraph (see Figure 5, top) includes links between projects that take a 'black spot' approach to targeting accident-prone road segments for safety measures.

3.2 The Tulip representation allows tracing the precursors and successors of individual components. Figure 5 (bottom) shows how it is possible to isolate these specific strands of the spaghetti shown in the top figure -- in this case, focusing on a component of the Bosnia and Herzegovina Road Management and Safety Project that will be shown to be strongly linked using the 'influence' metrics described below.

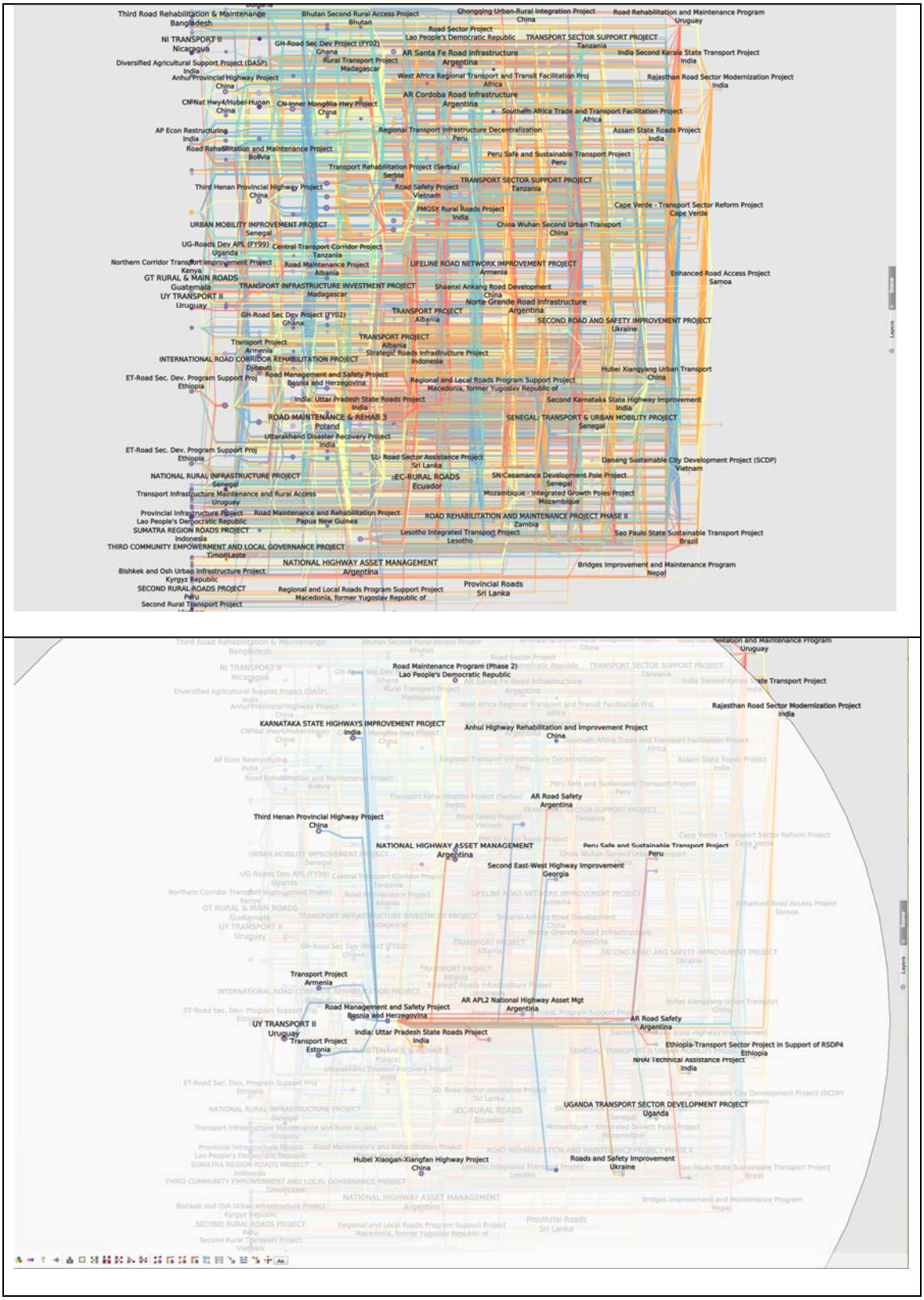
3.3 A script written for the database allows further exploration. In Figure 5 (bottom) the interior of each component node is color coded according to the Regional division of the World Bank that prepared the project. The outward link is coded with the color of the target component's region. (See the Annex.) This allows quick visual inspection of the extent to which project concepts diffuse across geographic and organizational boundaries. The exterior ring of each node is color-coded with the project's IEG outcome rating. (This is more visible when the view is zoomed.) These color assignments are customizable, so that links can be colored according to target or source, and by region or rating.

3.4 Validation of the network presents a challenge. Preliminary inspection of the derived network graph showed examples of causally linked components. For instance, the appraisal document for the Sri Lanka Renewable Energy for Rural Economic Development, component p077761_0 has the following text in the introduction (but not in the component description text contained in the database)

"The Project design aims to build a sustainable market for renewable energy through IDA and GEF support. Of the six distinct elements in the Project, the four that support grid- connected and off- grid generation of energy from renewable sources all have antecedents in the ESD project launched in 1997. Accounting for nearly 90 percent of total Project costs, these four components build on the success of the ESD project laid in small hydro, wind and solar power and in promoting energy efficiency."

The graph shows the ‘parent’ of this component to be p010498_1, which is in fact the antecedent ESD project.

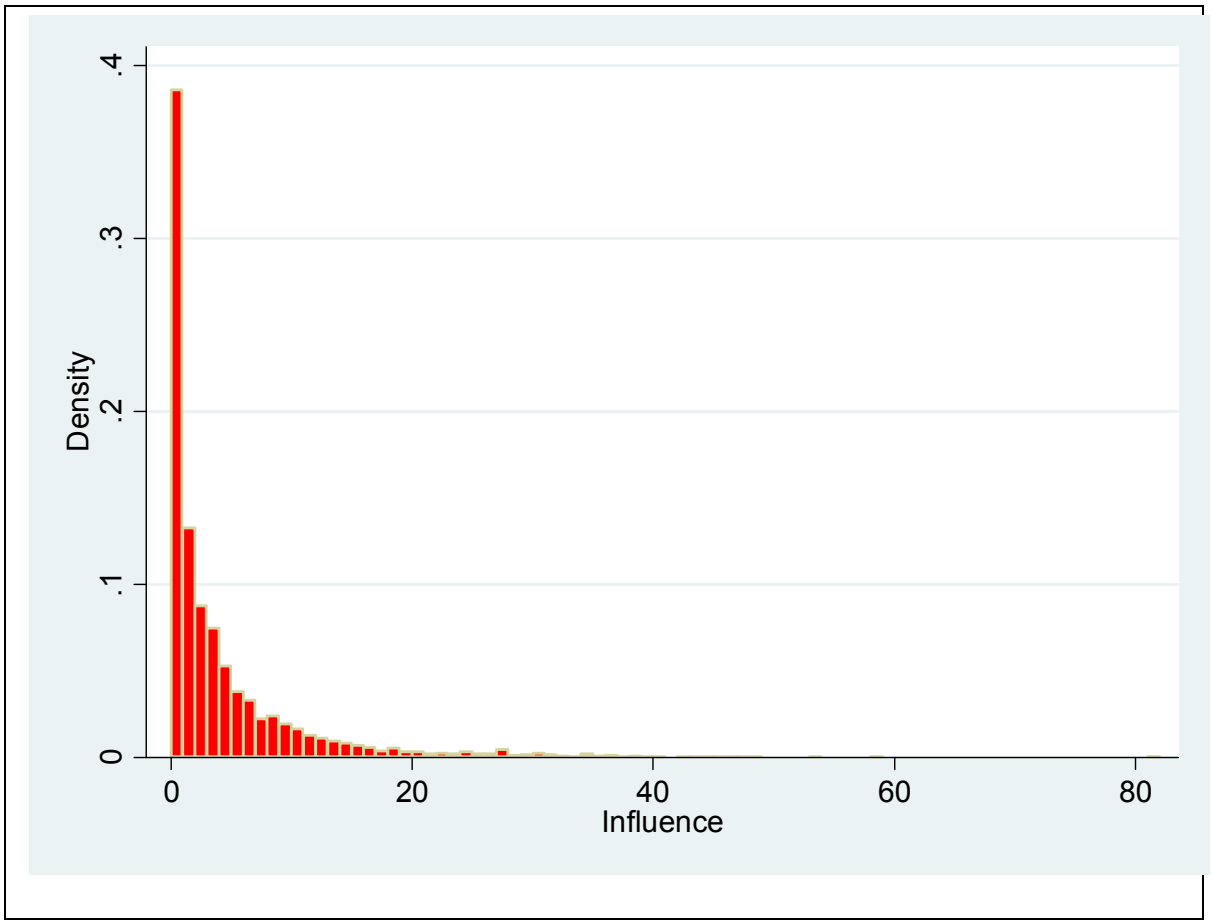
Figure 5 The road management subnetwork



‘Influence’ and its determinants

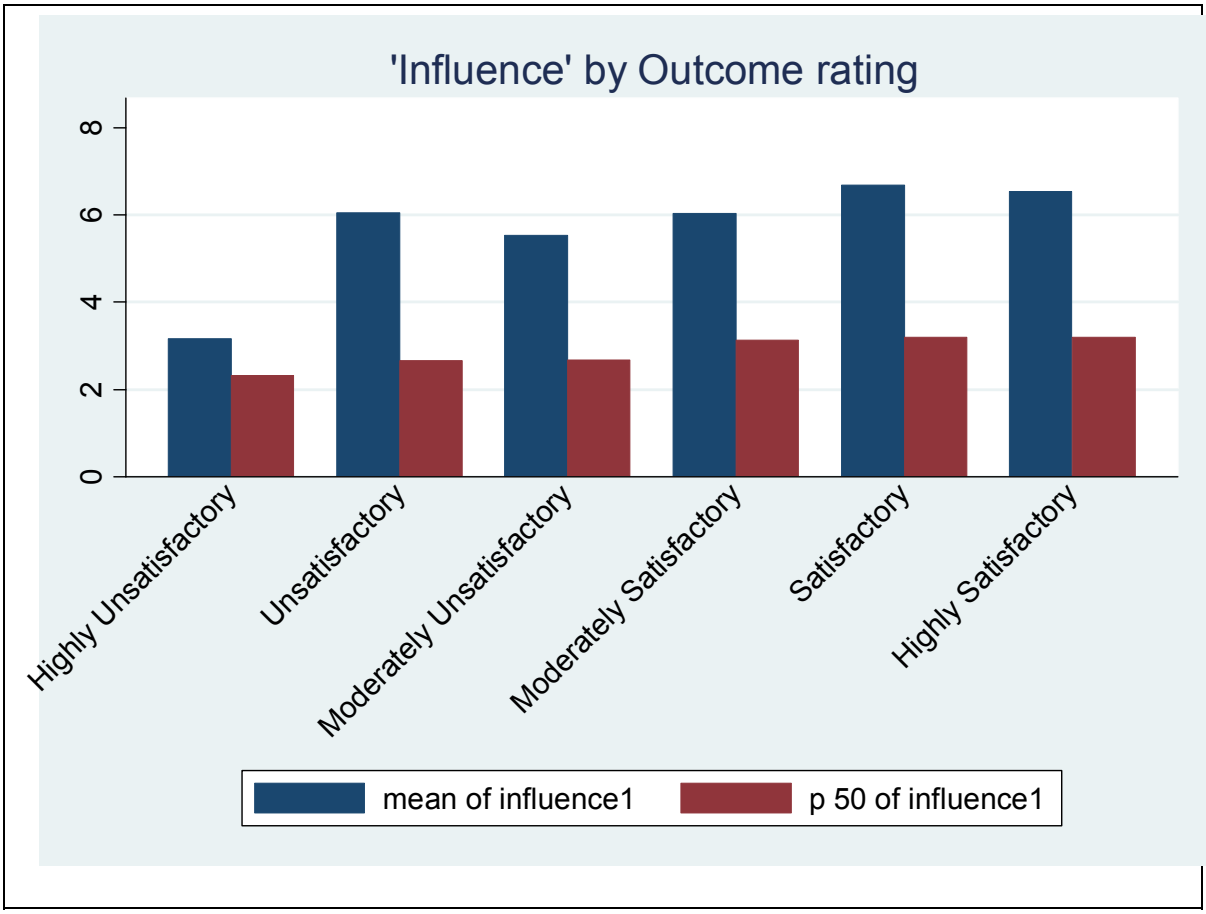
3.5 For initial analysis, we aggregated the ‘influence’ measures to the project level, allowing correlation with ratings and other project characteristics. Figure 6 shows the histogram of ‘influence’. Most projects have low ‘influence’ – the median is 2.98 – but a minority have very high levels, with a maximum of 81.97. The Bosnia/Herzegovina project mentioned earlier is an example.

Figure 6 Histogram of 'influence' at the project level



3.6 A simple cross-tabulation shows that ‘influence’ is not strongly related to IEG-rated project outcome. (Figure 7) Highly unsatisfactory projects do have markedly lower ‘influence’ but there are few such projects.

Figure 7 'Influence' is not strongly related to project outcome



3.7 There is however a relationship between 'influence' and region. (Figure 8) Projects originating from the Bank's Middle East and Northern Africa (MENA) region have markedly lower mean and median values of 'influence'. MENA has the smallest portfolio of the Bank's regions.

3.8 Figure 9 shows the bivariate relationship between project exit (closure[5]) year and 'influence.' As expected, more recent projects have lower mean 'influence' because there has been less time for them to exert influence. Less expected is the rise in 'influence' from 2000 to 2006. These effects are clearer in the multivariate analysis discussed below.

[5] Projects 'exit' or close when disbursement is complete.

Figure 8 MENA has lower 'influence'

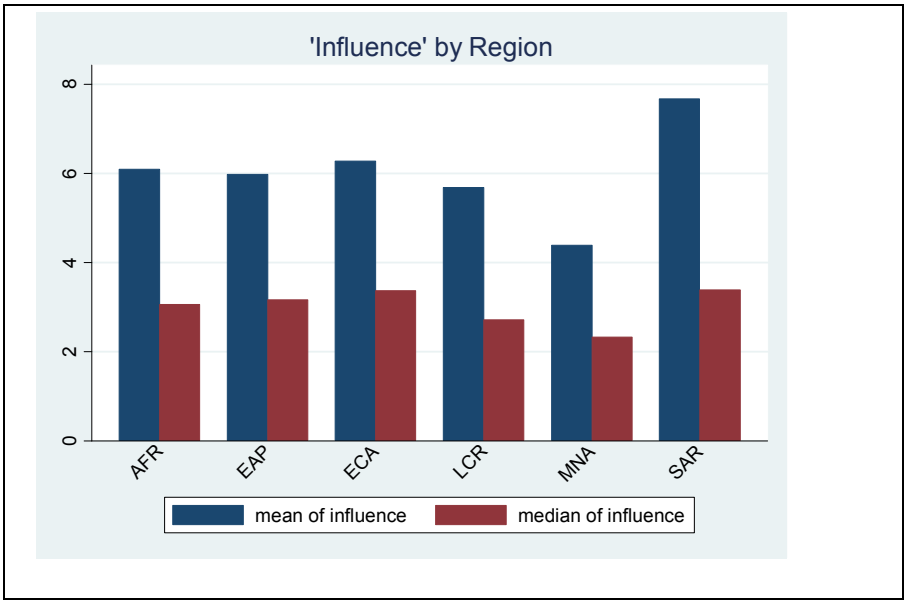


Figure 9 'Influence' by exit year

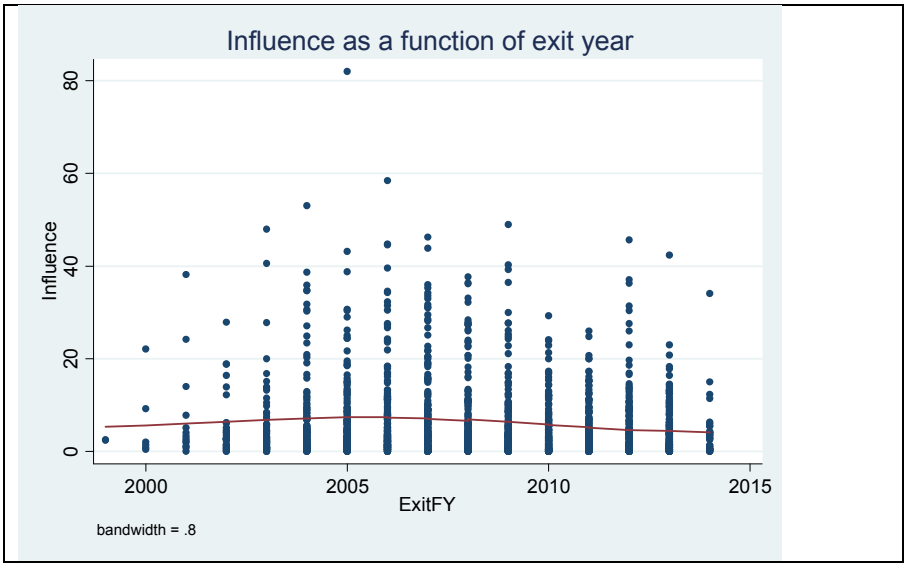


Table 5 Regression models of 'influence'

	Specification 1:Quality at Entry			Specification 2: Outcome		
	Coefficient	t	P>|t|	Coefficient	t	P>|t|
Unsatisfactory	-0.7774	-0.26	0.792	1.3688	0.68	0.495
Moderately Unsat	0.0813	0.03	0.978	1.1687	0.6	0.548
Moderately Sat	0.0218	0.01	0.994	1.5988	0.84	0.404
Satisfactory	-0.1495	-0.05	0.959	1.3921	0.72	0.472
Highly Sat	-0.6340	-0.2	0.84	0.8430	0.37	0.713
East Asia/Pacific	-0.3390	-0.51	0.609	-0.2883	-0.43	0.666
Europe/Central Asia	0.0647	0.11	0.914	0.0682	0.11	0.91
Latin America/Carib.	-0.5034	-0.84	0.402	-0.4412	-0.72	0.469
Middle East/N. Africa	-2.0695	-2.45	0.014	-2.0517	-2.4	0.017
South Asia	0.8294	1.03	0.304	1.0119	1.25	0.212
log10(project value)	20.7478	2.51	0.012	21.1543	2.52	0.012
(log project value)^2	-1.2812	-2.31	0.021	-1.3151	-2.34	0.019
exit Fiscal Year	200.1788	2.81	0.005	204.6989	2.84	0.005
(exit Fiscal Year)^2	-0.0500	-2.81	0.005	-0.0511	-2.85	0.004
elapsed time, concept note to effectiveness	0.0009	0.78	0.438	0.0010	0.8	0.423
(elapsed time)^2	0.0000	-0.28	0.777	0.0000	-0.33	0.738
constant	-200605.4	-2.8	0.005	-205173.9	-2.83	0.005
omitted: Highly Unsatisfactory, Sub-Saharan Africa						
	AdjRsq=.0496			AdjRsq=.0480		
	F(16,1653)=6.44			F(16,1636)=6.20		
	n=1670			n=1653		

Table 6 Descriptive statistics for regression

Variable	Obs	Mean	Std. Dev.	Min	Max
influence1	1670	6.068356	8.346737	0	81.97582
QAE_n	1670	3.878443	1.111065	1	6
logvalue	1670	7.453179	.5492733	5.69897	9.177178
logvaluesq	1670	55.85139	8.223799	32.47826	84.2206
exitFY	1670	2008.905	3.385184	1999	2015
exitFYsq	1670	4035710	13598.76	3996001	4060225
preptime	1670	740.0108	445.7467	65	3728
preptimesq	1670	746187.1	1131696	4225	1.39e+07

Region	Freq.	Percent
AFR	412	24.67
EAP	265	15.87
ECA	367	21.98
LCR	351	21.02
MNA	122	7.31
SAR	153	9.16
Total	1,670	100.00

3.9 Table 5 shows multivariate analysis at the project level, with descriptive statistics in Table 6. Two specifications were estimated, one using outcome rating as the measure of project quality, the other based on quality at entry. Both specifications included also dummies for region, a quadratic in total commitment value (grant plus loan) of the project, and a quadratic in the elapsed time (in days) from concept note to project effectiveness.

3.10 The estimates do not support the hypothesis that project design quality (as imperfectly proxied either by quality at entry or outcome rating) is associated with subsequent replication of project concepts – i.e., 'influence'. Nor is ease of project preparation (as proxied by time from concept note to effectiveness) associated with 'influence'. However, there is a very strong and statistically significant association between project size (in current dollars) and 'influence' (Figure 10). The increase in 'influence' from a tiny ($1 million) to a small ($10 million) project is 4.05; from a $10 million to a $100 million project the increment is 1.42, after which 'influence' plateaus. As expected, recently exiting projects have much lower 'influence' than those circa 2005, simply as a result of their youth. (Figure 11.) However, the statistically significant and substantial rise in 'influence' from 1999 to 2005 is unexplained and may be an artifact of increasingly detailed component descriptions over this period.

Figure 10 Impact of project value on 'influence'

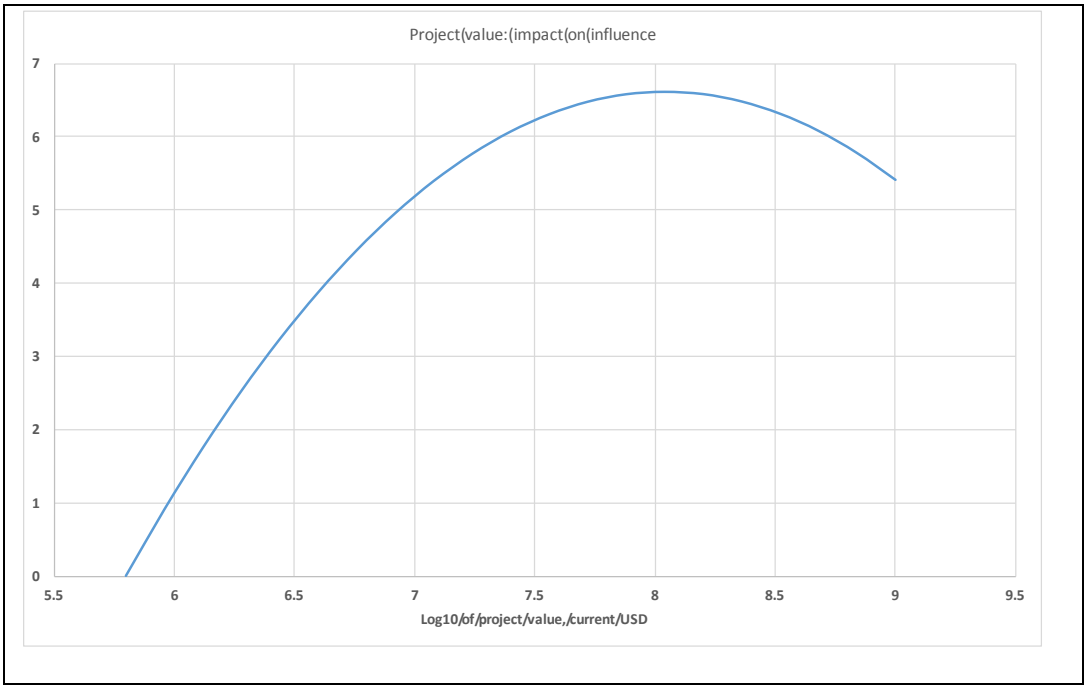
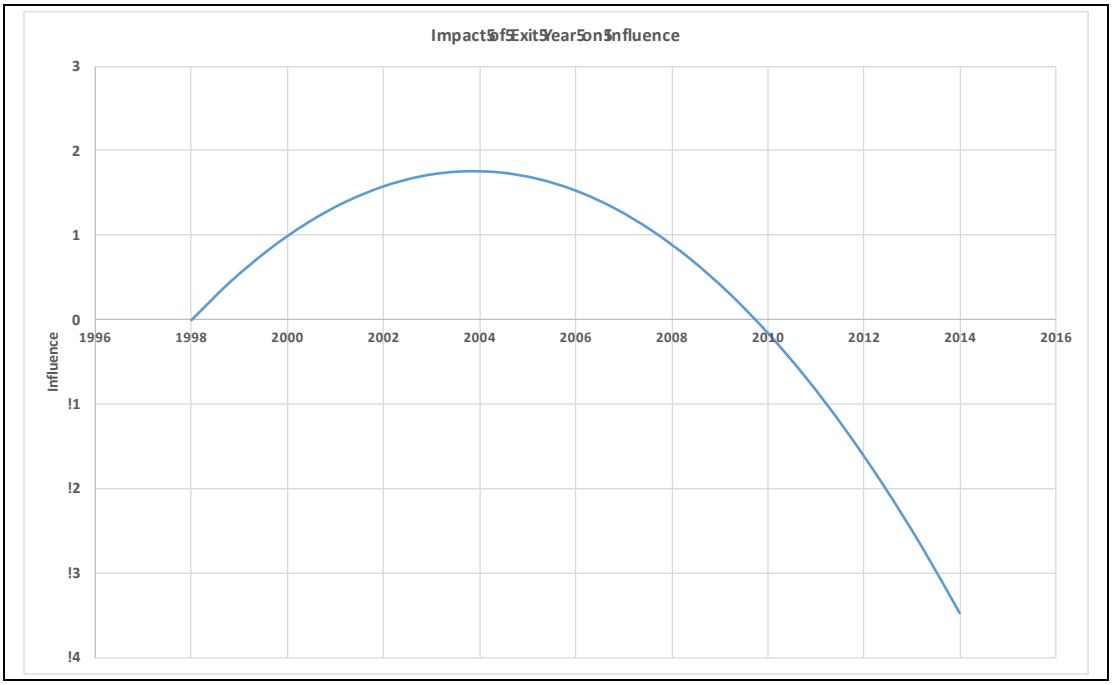


Figure 11 Impact of Exit Year on 'influence'



4. Discussion and conclusions

4.1 This paper has presented a conceptual and operational framework for analyzing the diffusion of project design concepts. The framework can be used for exploratory data analysis and for formal hypothesis testing. As an illustration of the framework's application, we test some simplistic (and non-exclusive) hypotheses about the determinants of diffusion. We find no support for the hypothesis that good quality projects are more likely to be replicated. We also find no support for the hypothesis that easy-to-prepare projects are more likely to be replicated. However, we find that small projects (<$10 million) on average are much less likely to be emulated than larger ones. This may have implications for the approach of using small projects as pilots or proving grounds for ideas worthy of scale-up. In a venture-capital type model of development, small pilot projects would be the proving ground for testing new concepts. While some might fail, one might expect the portfolio of small projects to have high influence on average. Such a model is not borne out by our results, but further investigation is needed.

4.2 We suggest that the results are of interest even in the absence of a causal story. For instance, it would be desirable if factors that favor good quality projects happened also to be conducive to emulation. Unfortunately, that is not our finding.

4.3 Both the results and the methodology presented here should be seen as exploratory and highly preliminary. Here are some issues for further investigation and experimentation.

- Validating causal interpretations of 'influence'. Is it possible to test whether the network-derived linkages represent conscious, deliberate emulation? This might be done through documentary evidence or interviews with project proponents. It would include in-depth studies of components and projects that are determined to have extraordinary 'influence.' It might also be possible to construct a network map of staff, where proximity is related to co-membership on a project team. This social network could be tested for its relation to the network of project component descriptions. Congruence of the two networks would support the hypothesis of diffusion of concepts via transmission of tacit knowledge between staff. This would accord with the findings of IEG's learning evaluations on the importance of tacit knowledge.

- Is a causal interpretation essential? For the purpose of determining the likelihood that a particular concept is emulated, do we need to posit that the concept is deliberately transmitted? Or can we consider that these are memes in common circulation? If so, do we need to aggregate across all similar memes at time t? Can this be done with network methods?

- Testing robustness of the network to text analytic methods. How sensitive is the network to, e.g., different classifications of 'meaningful' terms. Can more semantics be brought into the analysis?

- Is it possible that the finding that small projects are less informative is an artifact resulting from shorter texts or fewer components?

- Development and testing of more sophisticated hypotheses about concept transmission. Can we for instance allow for learning from, and adaptation of, failed but informative projects? Further measures of bureaucratic attractiveness could be tried, including measures of speed of disbursement.

References

Auber, D. (2004). Tulip—A huge graph visualization framework. In *Graph Drawing Software* (pp. 105-126). Springer Berlin Heidelberg.

Blau, P. M. (1970). A formal theory of differentiation in organizations. *American sociological review*, 201-218.

Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(10), P10008.

Connor, R., & Moss, R. (2012). A multivariate correlation distance for vector spaces. In *Similarity Search and Applications* (pp. 209-225). Springer Berlin Heidelberg.

IEG (2014) Learning and Results in World Bank Operations: How the Bank Learns. Evaluation I**.** Washington DC: World Bank.

IEG (2015) Learning and Results in World Bank Operations: Toward a New Learning Strategy. Washington DC: World Bank.

Noack, A. (2004, January). An energy model for visual graph clustering. In *Graph Drawing* (pp. 425-436). Springer Berlin Heidelberg.

Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information processing & management*, *24*(5), 513-523.

Annex. How to use the network database

Tulip is documented and downloadable at http://tulip.labri.fr/TulipDrupal/

Color codes for regions and ratings are shown below.

Figure 12 Legend for region and rating codes