Production efforts in large Indian cities: a network-based approach

Sebastian Bervoets[†] and Frédéric Deroïan[†]

Abstract

In this paper we construct a network of roads connecting large Indian cities and we evaluate this network's overall performance. We consider a model where the production efforts of connected cities are strategic complements, and we relate the equilibrium effort profile to a well known centrality measure, the Katz-Bonacich centrality. We then make use of this result to compute the level of efforts of different cities in the current network and identify which city contributes most to overall efforts, which existing road is the most influential and which new road should be constructed in priority. Our results shed light on the importance of relatively small cities on aggregate efforts. Our exercise illustrates how network details might generate unexpected effects.

Keywords: Network Policy, Indian National Roads Network, Strategic Complementarity, Aggregate Efforts Maximization

JEL: C72, D85

[†] *Aix-Marseille University (Aix-Marseille School of Economics), CNRS and EHESS*
E-mail addresses: sebastian.bervoets@univ-amu.fr, frederic.deroian@univ-amu.fr

1 Introduction

Indian development is partly slowed down by the low quality of the transportation network, due to lack of investments. To remedy this, the National Highway Development Project was launched in India in 2000, with the aim of upgrading, rehabilitating and widening major highways to a higher standard. This huge ongoing project is managed by the National Highway Authority, who states that "advantages of having a well developed network of world class highways are many for a nation like India". The list of such advantages is of course long, in particular for the economic development of Indian regions. Among others, better roads offer a better access to goods, services, high quality education, employment opportunities etc. Furthermore, the benefits to trade are better extracted when transportation costs are low. Having high standard roads is a direct way to reduce transportation costs, by making journeys faster and safer.

In this paper, we argue that beyond the quality of roadways linking cities together, the organization of the roads' network itself matters a great deal. The choice of the location of the roads that are to be constructed or maintained is of crucial importance and can lead to surprising conclusions. We wish to illustrate this point with a simple comparative exercise, by making use of the latest developments in economic network theory.

We consider that roads link pairs of cities together and that a link generates some synergies between the two cities. These synergies induce a positive impact on the cities' productivity, such that more production effort by a given city makes production effort more attractive to the other cities that are linked to it. This complementarity has many sources, an example of which could be the following: assume a new road is constructed between two cities, one of which has a good university, the other has industries requiring high skilled workers. The fact that the road is built creates an incentive for students of the first city to perform well at university because they anticipate they might get a job in the other city, while having high skill students available makes jobs creation more attractive for the firms of the second city. The new road thus generates spillovers for both cities.

Of course, once we assume that links between cities increase incentives to produce some effort (in education, in investments etc.), then the entire network geometry comes into consideration. This is because a city (call it A) linked to another (B) will be directly impacted by the activity of the later, but the activity of the later directly depends on its other linked cities (say C and D). Therefore, cities C and D have an indirect effect on city A and that effect, although indirect, has to be accounted for. Our model encompasses such effects and it accounts for every possible indirect effect that a city could exert on another city, however far away they are one from the other, and however small these effects might be.

Because the entire geometry of the network of roads comes into play in determining the influences of cities on one another, a social planner needs a model that predicts which road has what influence on which cities. Recent theories in the economics of networks analyse precisely how the organization of links affects efforts, in a world of complementarities between nodes (nodes being taken as cities herein). We would like to apply this theory to the Indian National Roadway network.

We assume cities are nodes in a roadway transportation network. They have a production effort to choose, the choice being the result of a trade off between the benefits of increasing production effort and the cost it entails. If cities were isolated, this could be captured by a linear quadratic utility function to be maximised. Because cities are linked together and spillovers take place, an additional term enters the model that captures the positive externalities exerted by cities that are directly linked through the network. This model is adapted from the standard Ballester, Calvó-Armengol and Zenou (2006) model. In that paper, the authors show that under small levels of interactions between nodes, there exists an unique equilibrium effort. This equilibrium effort has very nice properties because it is directly related to a specific centrality measure of the node in the network, called the Katz-Bonacich centrality. We detail this theory in this paper.

The equilibrium level of efforts being determined, it is then possible, although still complex, to conduct some comparative statics exercise on the links of the network. This allow us to answer the following questions: Which *city* has the largest impact on aggregate production efforts? Which existing *road* has the largest impact on aggregate production efforts? If a central planner wants to build a new road, which cities should he link in order to maximise production efforts?

For the purpose of illustrating the importance of network considerations, we apply these results to the case of the Indian National Roadway network, by simplifying the analysis. Indeed, we restrict our attention to the sample of major cities in India, counting more than 1.5 million people. As for the 2011 census, there are 20 cities in India satisfying this criterion. We next assume that a link exists when a road links two cities directly, without going through another city in the sample. This allows us to build the current network of roads between major cities. We can thus use that network as the basis of our analysis.

At this point, we would like to stress out that our contribution is not meant to derive realistic policy decisions. The only purpose of the exercise is to illustrate how networks matter, how tricky their effects can be, how unexpected their analysis can lead to, but at the same time, how crucial their importance is in the design of optimal public policies.

Our results are the following: First, we confirm that the organization of roads matters. Depending on its location, the effects of a new road on aggregate production levels are strongly heterogeneous. Second, small variations in the level of synergies between connected cities may also induce large differences. This implies that any correct public policy decision should be based on an accurate evaluation of the levels of synergies between cities. Third, we find that the city that has the largest impact on national production effort is Kanpur, which is only ranked 12th in the ranking of cities by population. One expects Mumbai, Delhi or Bangalore, the three largest cities in India, to be the most important cities in terms of synergies for the rest of the cities, but Kanpur happens to be the most central city in this network of roads. Indeed, when the levels of synergies is relatively high, Mumbai, Delhi and Bangalore arrive respectively in 15th, 4th and 16th position in the ranking of most contributing cities. Removing Kanpur from the map would have the largest negative aggregate effect on production efforts. Next, we find that the most important road in the current network is either the road between Delhi and Jaipur, the road between Kanpur and Jaipur or the road between Kanpur and Agra, depending on the level of interactions. This again is surprising as the cities involved, except for Delhi, are of relative small size. Last, we find that the best road a central planner should build in the future, among all possible non existing roads, would be the road between Delhi and Kanpur.

The fact that all of these results are surprising is illustrative of the fact that intuitions no longer hold once networks are seriously accounted for. There are very complex indirect effects that ought to be underestimated in models without networks, while these indirect effects may be, in aggregate, larger that the direct effects. This is the reason for which network theories should be incorporated into public policy decision models. As mentioned earlier, our objective is not to convince the reader that our precise conclusions are solid. Our analysis relies on too simplified data to be seriously considered. Rather, our objective is to illustrate how intuition can be misleading when dealing with network effects.

The rest of the paper is organised as follows: in the next section we briefly present the underlying theory that we use in order to derive our results. In section 3 we present our setting and the data we rely on to conduct our exercise, while results are presented in section 4. A conclusion in section 5 discusses extensions that the model should account for, in order to prescribe more realistic public policies.

2 A brief overview of the theory

In this section, we describe the theory that we will use to predict the overall performance of the Indian National Road Network. The theory is kept as simple as possible, more details can be found in the original paper by Ballester, Calvó-Armengol and Zenou (2006). To make reading easy, the technical details are removed from the text and appear in the appendix.

2.1 Isolated cities

Before introducing the network of relations between cities and the spillovers they exert one over another, we briefly present the model with isolated cities. This will serve as a benchmark theoretical case to better illustrate the impact of synergies.

Assume a city is represented by an agent who has productivity a [1] per unit of effort x. The benefits of exerting a level x of effort is measured by ax. Exerting effort, however, is costly and the cost function is assumed herein to be convex of the form $\frac{c}{2}x^2$. Therefore, the utility derived by the city (or by its representative agent) when exerting effort x is

$$U(x) = ax - \frac{c}{2}x^2$$

The representative individual who wishes to maximise this utility is facing a simple optimisation problem. The solution to this problem is found by setting the first order condition to zero, which leads to

$$x^* = \frac{a}{c} \tag{1}$$

When a city is isolated, the optimal effort level that it should choose is increasing in its productivity a and decreasing in the cost of effort c.

We are now ready to introduce relations between cities through a network of roads and analyse the synergies this generates.

2.2 A model of interrelated production efforts

Assume we have a set $N = \{1, 2, \cdots, n\}$ of cities, each represented by a representative agent indexed by i, having marginal productivity a_i, and achieving some level of production through a costly effort x_i. The cost of effort is c_i. In what follows, we allow cities i and j (or their respective representative agents) to differ in productivity (i.e. $a_i \neq a_j$).

[1] We will discuss further how this productivity can be estimated.

The transportation network between cities. There are bilateral connections between some of these towns. A bilateral connection between town i and town j exists when there is a direct road between those towns. The overall set of roads is usefully represented by a matrix G, where entry g_{ij} is equal to 1 if and only if there is a direct road between both towns i and j, while g_{ij} is equal to 0 if there is no direct road between both towns i and j. Notice that when a road exists (resp. does not exist) between towns i and j then it also exists (resp. does not exist) between j and i, entailing $g_{ij} = g_{ji}$. Thus the matrix G, referred to as the adjacency matrix in network theory language, is a symmetric matrix.

A path between cities in the network is a sequence of consecutive roads, including possible loops, that join both cities.

Synergies and utilities. As discussed in the introduction, roads between cities generate complementarities between the production and development of the linked cities. We model this by adding to the utility of the representative agent of an isolated city an extra term that relates the benefits of her production effort to the level of effort exerted by representative agents of towns accessible through a direct road. Assume $X = (x_1, ..., x_n)$ denotes the effort profile of all cities of the network. Then the utility of the representative agent of city i, when exerting production effort x_i, given the transportation network G, is written as follows:

$$U_i(x_i, x_{-i}; G, \delta) = a_i x_i - \frac{c_i}{2} x_i^2 + \delta \sum_{j=1}^{n} g_{ij} x_i x_j \tag{2}$$

where δ represent the intensity of interactions between linked cities. Notice that because $g_{ij} = 0$ whenever cities i and j are not linked, this model only assumes that a city i has synergies with the cities j she is directly linked to.

Example 1. *Three cities.*
Assume the country only contains three large cities, one in the North (N), one in the East (E) and one in the West (W). A road goes through the country, linking E to N and N to W.

In this case, the eastern city benefits from synergies with the northern city, but not directly with the western city because they do not have a direct

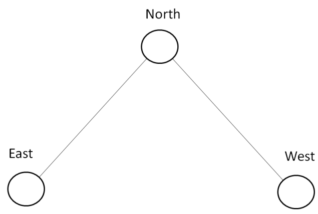

Figure 1: A three-city transportation network

road linking them. Utilities are given by

$$U_E = a_E x_E - \frac{c_E}{2} x_E^2 + \delta x_E x_N$$
$$U_N = a_N x_N - \frac{c_N}{2} x_N^2 + \delta x_N x_E + \delta x_N x_W$$
$$U_W = a_W x_W - \frac{c_W}{2} x_W^2 + \delta x_W x_N$$

Clearly the northern city will directly benefit from both cities, whereas the other cities will only directly benefit from the northern city.

In this model, synergies arise between connected towns. These synergies can have many different sources that this simple model accounts for. For instance, the synergies could come from the fact that a road decreases the cost of production of cities. If a road increases the access to some production inputs, if it helps increasing the information transmission about new production technologies etc., then synergies can be interpreted as a cost reduction and the city i's utility can be written as

$$U_i(x_i, x_{-i}; G, \delta) = a_i x_i - \underbrace{x_i \left(\frac{c_i}{2} x_i - \delta \sum_{i=1}^{n} g_{ij} x_j \right)}_{\text{cost reduction of effort } x_i}$$

Another possibility is to assume that synergies come from increased productivity, due for instance to more efficient matches between firms and workers, to innovation spillovers etc. In that case, synergies can be interpreted as an increase in a_i and utilities can be written as:

$$U_i(x_i, x_{-i}; G, \delta) = \underbrace{\left(a_i + \delta \sum_{i=1}^{n} g_{ij} x_j \right)}_{\text{productivity of agent } i} x_i - \frac{c_i}{2} x_i^2$$

In what follows, we will assume that the marginal cost of effort, c_i, is constant across cities, i.e. $c_i = c$ for all i, and we normalize this to $c = 1$. This simplifying assumption is made essentially to keep the exposition clear. Introducing too many sources of heterogeneity introduces complexities that are detrimental to the understanding of the exercise we are presenting.

2.3 Equilibria and centralities

The objective we have in mind is to determine how the structure of the network of roads affects the levels of production effort. This will allow us to answer questions such as "what is the most important city?", "what is the most important road?", "what is the next road that should be built?" etc.

One nice feature of this model is that, under some mild condition that we discuss later, it generates a unique equilibrium effort profile. We identify precisely the equilibrium effort profile $X^* = (x_1^*, ..., x_n^*)$ here. Setting the first order condition to zero [2], the equilibrium production effort x_i^* of city i is given by :

$$x_i^* = a_i + \delta \sum_{i=1}^{n} g_{ij} x_j^* \tag{3}$$

Obviously, by comparing Equations (1) and (3), the optimal production effort for each city is higher in the presence of other cities than if the city was isolated. However, this optimal production level depends not only on city i's characteristics (a_i), it also depends on city j's effort level. Of course, city j's effort will in turn depend on city j's neighbours, including city i. Because i's effort depends on j's effort which depends on i's effort and so on, the solution to this problem of finding the equilibrium effort of every city amounts to finding a fixed point to the system.

Assume that the value of δ is fixed. When city i increases its effort by one unit, city j increases its effort by δ units. But the increase in j's effort induces in turn an increase in i's effort. This increase in i's effort again induces an increase in j' effort and so on. Consequently, if δ is large, this process might not converge and lead to infinite efforts. Because infinite efforts do not make sense, the model should be restricted to low enough values of δ, such that the aggregate effects do not go to infinity.

The condition that δ should satisfy depends on the geometry of the network. This geometry is captured by what is called the *index* of the network, denoted by $\mu(G)$. It is the largest eigenvalue of the adjacency matrix G.[3] As

[2] Utilities being linear quadratic, the second order conditions are satisfied.

[3] This eigenvalue is known to be a real number when G is symmetric.

shown in Debreu and Herstein (1953), the condition $\delta < \frac{1}{\mu(G)}$ guarantees that the feedback effects are not large enough to go to infinity. It guarantees that they are reasonable enough to ensure the existence of a finite equilibrium level of efforts. In some sense, the index of a network measures the extent to which variations at one node spread across the network. We will assume throughout this paper that this condition is satisfied over all the networks we examine.

Small enough δ guarantees the existence of an equilibrium effort profile, and this same condition also guarantees that the equilibrium effort profile is unique. Uniqueness of the solution is an important feature of the model, because it makes comparative statics exercises relatively simple. Indeed, if one wishes to modify the specific values of a_i for instance, or the level of interactions δ, or else the network itself by reallocating, deleting or adding some links, then the uniqueness of the equilibrium allows us to compare the situations before and after the change.

Equilibrium effort levels and Katz-Bonacich centralities. The equilibrium of this game has been recently characterized in Ballester, Calvó-Armengol and Zenou (2006). They show that it has a very nice feature: it corresponds to the so-called Katz-Bonacich centrality measure[4].

Centrality measures are designed to measure how important individuals are in a network. Sociologists have introduced various centrality measures, such as the degree centrality, the closeness centrality, the betweenness centrality, etc (see Wasserman and Faust [1994] for more details). The Katz-Bonacich centrality measure (see Bonacich [1987]) is one specific centrality measure that happens to exactly coincide with the equilibrium of the game we are analysing. More central agents, in the sense of Katz-Bonacich centrality, will exert more effort at equilibrium than less central agents. Before describing the Katz-Bonacich centrality measure in more details, it is important to stress out that the linkage between equilibrium and centralities is crucial for the exercise we are undergoing. Indeed, when comparing individuals or cities in different networks, it is enough to compute their centrality in the respective networks and compare the number that are obtained.

Denote by $P_k(i, G)$ the number of paths of length k that arrive at node i in the network G. This number is in general difficult to determine in an arbitrary network because the number of paths increases exponentially with the length of paths considered. However, network theory tells us that this number is given by the row of the matrix G elevated at the power k. More

[4]The reader interested with the technical details will find them in the appendix.

precisely, the entry G^k_{ij} corresponds precisely to the number of paths of length k arriving at node j and starting from node i. By symmetry of G this also corresponds to the number of paths of length k starting at node j and arriving at node i. Therefore

$$P_k(i,G) = \sum_{j=1}^{n} G^k_{ij}$$

In order to measure how central an individual is in a network, the idea behind the Katz-Bonacich measure is that central individuals have many path arriving to them. However, long paths should count less than short paths, therefore paths of length k are decayed by a factor δ^k, where $0 < \delta < 1$. Furthermore, paths arriving from a highly productive city should count more than a similar path starting at a lower productive city. Therefore, the Katz-Bonacich centrality of node i depends on δ and on the vector $A = (a_1, ..., a_n)$ of productivities, and is given by

$$b(\delta, A, G) = \sum_{k=1}^{\infty} \delta^k G^k.A$$

where b is the Katz-Bonacich centrality vector.

As said earlier, the equilibrium effort x^*_i of city i in the roads' network is given by this centrality measure, i.e.

$$x^*_i = b_i(\delta, A, G) \tag{4}$$

Example 2. *Three cities - continued.*
In the three-city example, with costs normalized to unity, we get

$$\begin{cases} x^*_E = \frac{(1-\delta^2)a_E + \delta a_N + \delta^2 a_W}{1-\delta^2} \\ x^*_W = \frac{(1-\delta^2)a_W + \delta a_N + \delta^2 a_E}{1-\delta^2} \\ x^*_N = \frac{a_N + \delta(a_E + a_W)}{1-\delta^2} \end{cases}$$

This equilibrium effort profile illustrates two things: First, every city's effort is greater than the effort of the corresponding isolated city (with cost normalised to 1*, the effort of city i is $x^*_i = a_i$). The more central the city, the larger the difference between the two equilibrium efforts. Second, more productive cities exert higher efforts, and exert a larger influence on other cities efforts.*

Now that equilibrium levels of effort are known, we turn to the analysis of the Indian National Highway network.

3 Data

In this section, we describe the data that we use, as well as all the choices we have made in order to undertake our theoretical exercise.

Cities: Considering every city in India as well as every road would generate an intractable amount of information. We have therefore decided to restrict our attention to cities whose population exceeds 1.5 million people, in the 2011 Census. There are 23 such cities, among which 3 are merged with another because they are at really short distances one from the other. These cities are Thane that has been merged with Mumbai, Ghaziabad that has been merged with Delhi and Pimpri-Chinchwad that has been merged with Pune. We thus end up with a set of 20 large cities. Figure 2 presents the selected towns, with their population in census 1991, census 2001 and census 2011, and the respective states they belong to.

The roads' network: We only wish to consider roads that support a significant level of traffic. This is because we are analysing synergies that have an undisputed impact on local economies and we believe that the synergies only arise when major flows of traffic take place. For that purpose, we restrict our attention to roads that are multi lane highways, such as described in the National Highways Development Project, supervised by the National Highways Authority of India. This roads' network can be found on the Indian roadway map in Figure 8 in the appendix.

Some of the cities selected are close one to the other, while others are very distant one from the other. Also, some cities are directly linked by a road, without having to go through another large city while others are linked by a path going through other cities. We wish to take these two points into consideration. For instance, we want to treat differently the relations between Delhi and Agra from the relations between Jaipur and Kolkata.

First, the distance between Jaipur and Kolkata is about 1.400 kilometres, against less than 200 between Agra and Delhi. Irrespective of the existence of a direct road between two cities, we assume here that distance matters because direct synergies can only take place when cities are not too far away. We have arbitrarily chosen to select roads between cities separated from less than 600 kilometres. This implies that a direct road between Agra and Delhi is a candidate, while a potential direct road between Jaipur and Kolkata is excluded from the network[5].

[5]We have chosen the number 600 because it is the first distance such that the network is connected. For instance, if we had chosen 500 kilometres as the threshold, then the network would have been disconnected into two components and cities like Mumbai and

Label	Town	Census 1991	Census 2001	Census 2011	State
1	Mumbai+Thane	10 729 280	13 241 001	14 297 319	Maharashtra
2	Delhi+Ghaziabad	7 749 696	10 847 968	12 643 903	Delhi
3	Bangalore	2 908 018	4 301 326	8 425 970	Karnataka
4	Hyderabad	3 059 262	3 449 878	6 809 970	Andhra Pradesh
5	Ahmedabad	2 966 312	3 637 483	5 570 585	Gujarat
6	Pune+Pimpri-Chinchwad	2 219 459	3 550 945	4 844 790	Maharashtra
7	Chennai	3 857 529	4 343 645	4 681 087	Tamil Nadu
8	Kolkata	4 399 819	4 572 876	4 486 679	Bengale-Occidental
9	Surat	1 498 817	2 433 835	4 462 002	Gujarat
10	Jaipur	1 518 235	2 322 575	3 073 350	Rajasthan
11	Lucknow	1 619 115	2 185 927	2 815 601	Uttar Pradesh
12	Kanpur	1 879 420	2 551 337	2 767 031	Uttar Pradesh
13	Nagpur	1 624 752	2 052 066	2 405 421	Maharashtra
14	Indore	1 091 674	1 474 968	1 960 631	Madhya Pradesh
15	Bhopal	1 062 771	1 437 354	1 795 648	Madhya Pradesh
16	Visakhapatnam	752 037	982 904	1 730 320	Andhra Pradesh
17	Patna	917 243	1 366 444	1 683 200	Bihar
18	Vadodara	1 046 009	1 306 227	1 666 703	Gujarat
19	Ludhiana	1 042 740	1 398 467	1 613 878	Pendjab
20	Agra	891 790	1 275 134	1 574 542	Uttar Pradesh

Figure 2: The 20 largest Indian cities

Second, even if we had not considered a distance threshold, we must account for the fact that when travellers go from Jaipur to Kolkata by road, they have to pass through, or close to another large city of the sample (Agra and Kanpur in this case). This is not the case for a traveller who would like to go from Delhi to Agra. We therefore consider that there is a potential direct link between Delhi and Agra but that there is none between Jaipur

Delhi, belonging to separate components, would have been considered as if in two different countries. We have also run the calculations for other distance thresholds and have found no significant difference. We therefore stick to the case of 600 kilometres for the clarity of exposition.

and Kolkata (even if the distance was less than 600 kilometres).

The adjacency matrix we have built using these constraints can be found in the appendix (see Figure 9). Recall that a cell G_{ij} is set to 1 whenever there is a direct road between the cities i and j, whereas a 0 indicates there is no direct road. Some cases are straightforward. For instance, there is a direct road between Hyderabad and Bangalore, but there is none between Kolkata and Surat. However, other cases are less easy to decide and again, we have made arbitrary choices.

This is the case for instance for the Jaipur-Ludhiana connection, that we have set to 0. One could argue that the road from Jaipur to Delhi then to Ludhiana is in some sense a direct road between Jaipur to Ludhiana, but it seemed to us that the difference in kilometres between a potential direct route and the existing route is large enough to justify that the cell in the matrix is set to 0. In contrast, we considered that the direct road between Kanpur and Patna exists, although it goes through the city of Lucknow. We opted for this choice because the three cities are on a straight line and there is hardly any other option than to go through Lucknow.

Once all these choices have been made, we end up with a roadway network that contains 26 roads out of 95 possible direct roads. On average, every city has 2.5 neighbours, but there is great variability across cities. Indeed, four towns (Chennai, Surat, Visakhapatnam and Ludhiana) only have a single road, while the city of Kanpur has five connections.

Productivities: Having constructed the adjacency matrix with the cities and the roads, we need a last characteristic in order to compute the equilibrium levels of productive effort x_i^*: the productivity levels a_i. To determine these values, we rely on a well established literature in economic geography (see for instance the Handbook of Regional and Urban Economics (2004) for a recent survey of this literature) that argues that in a given economic environment, the productivity of a representative agent of a city is well approximated by some measure of the size of the city. In particular, what matters is the population of cities in determining productivities and a consensual estimation is given by

$$a_i = n_i^\alpha$$

where n_i is the population of city i and α is a parameter estimated to be close to 0.05. In the remainder we stick to these estimations and set $\alpha = 0.05$. The table of cities' productivity is given in Figure 3.

Of course, more populated cities have higher productivity than less populated cities, the maximal difference being a rough 10% between Mumbai and

Label	Town	Productivity
1	Mumbai+Thane	2,28
2	Delhi+Ghaziabad	2,27
3	Bangalore	2,22
4	Hyderabad	2,20
5	Ahmedabad	2,17
6	Pune+Pimpri-Chinchwad	2,16
7	Chennai	2,16
8	Kolkata	2,15
9	Surat	2,15
10	Jaipur	2,11
11	Lucknow	2,10
12	Kanpur	2,10
13	Nagpur	2,08
14	Indore	2,06
15	Bhopal	2,05
16	Visakhapatnam	2,05
17	Patna	2,05
18	Vadodara	2,05
19	Ludhiana	2,04
20	Agra	2,04

Figure 3: Towns' estimated productivities

Agra.

Equilibrium efforts: With G and A determined, one can compute the equilibrium levels of efforts by using Equation (4) for several values of interaction δ. We analysed three cases, with respectively low interactions ($\delta = 0.01$), median interactions ($\delta = 0.1$) and high interactions ($\delta = 0.25$).[6] The results are in Figure 4. The table reads as follows: the numbers in the cells indicate the ranking in terms of influence of a given city, for a given level of interactions δ. For instance, Bangalore is city exerting the 4th highest effort when $\delta = 0.01$, it is then ranked 11th and 15th for $\delta = 0.1$ and 0.25 respectively.

[6]The maximal admissible value of δ associated with the collected transportation network between the 20 cities is around 0.28.

Label	Town	Ranking $\delta = 0.01$	Ranking $\delta = 0.1$	Ranking $\delta = 0.25$
1	Mumbai+Thane	2	7	13
2	Delhi+Ghaziabad	1	2	4
3	Bangalore	4	11	15
4	Hyderabad	3	5	9
5	Ahmedabad	5	8	8
6	Pune+Pimpri-Chinchwad	7	12	16
7	Chennai	10	18	17
8	Kolkata	9	13	18
9	Surat	11	17	20
10	Jaipur	8	3	3
11	Lucknow	12	6	5
12	Kanpur	6	1	1
13	Nagpur	14	14	10
14	Indore	17	15	12
15	Bhopal	15	9	6
16	Visakhapatnam	19	20	19
17	Patna	16	10	7
18	Vadodara	18	16	14
19	Ludhiana	20	19	11
20	Agra	13	4	2

Figure 4: Towns' efforts

What we observe here is interesting. When interactions are small, the ranking in terms of equilibrium efforts almost coincides with the ranking in terms of productivity. This is because cities get very little feedback from the network when the interaction level is low. When $\delta = 0$ we are back to the case of isolated cities and as we observe from equation (1), a higher productivity implies a higher effort, because returns to effort are higher. Thus the ranking of effort levels of isolated cities coincides with the ranking of productivities, which itself coincides with the ranking of cities in terms of population.

However, as soon as δ is positive, even though very small, the ranking only "almost" coincides with productivities. In particular, some low ranked cities climb very quickly in the ranking. It is the case of Jaipur, Kanpur and Agra. This shows that these cities will play an important role in our analysis because they are highly central in this network.

As δ increases, the feedback effects of the network get stronger, and equilibrium efforts are shaped more and more by neighbours rather than own productivity. As we see, for the highest value of δ, the three cities highlighted just before appear to be the most central and therefore to benefit the most from their position on the network. As they benefit from efforts made by others, their own effort becomes more attractive and they choose to exert higher efforts than other more productive cities. This totally changes the picture observed for small values of δ (or even for $\delta = 0$) as the overall ranking does not at all coincide with the population ranking.

4 Network-based comparative statics

In this section, we make use of the theory exposed and the data just constructed to perform our exercise and answer three questions: 1- Given the existing network, what city has the highest contribution to overall production efforts? 2- Which road, among those existing, contributes the most to overall production efforts? 3- Which road, among the roads not yet built, would generate the highest increase in production efforts?

Once more, we insist on the fact that we will not advocate specific choices according to the specific results we obtain here, essentially because too many simplifying assumptions have been made throughout the data collection. However, we hope to convince our readers that any public policy decision involving networks should rely on the type of analysis we are presenting. Answering the three questions above is not an easy task but it is an essential task, and we provide in this section a methodology to do it.

The first question is referred to as the key city, the second as the key road and the last as the optimal road addition.

4.1 The key city

In order to know which town contributes the most to the aggregate level of efforts on the network, we remove one town at a time from the network and compute the consequences it has in terms of aggregate efforts. The town that has the largest impact when removed will be the *key city*. Although our intuition suggests that the key city is the most central city, this is not necessarily true. The most central city might have less impact on aggregate efforts than a less central city.

For each of the three values of δ (0.01, 0.1, 0.25), we computed the equilibrium efforts resulting from the removal of one city and compared the aggregate efforts of this new network with the aggregate efforts of the initial network. Results are presented in table 5.

Label	Town	Ranking $\delta = 0.01$	Ranking $\delta = 0.1$	Ranking $\delta = 0.25$
1	Mumbai+Thane	6	10	15
2	Delhi+Ghaziabad	2	4	4
3	Bangalore	11	12	16
4	Hyderabad	4	5	10
5	Ahmedabad	8	9	8
6	Pune+Pimpri-Chinchwad	12	13	17
7	Chennai	20	19	14
8	Kolkata	13	15	18
9	Surat	17	20	20
10	Jaipur	3	2	3
11	Lucknow	7	6	5
12	Kanpur	1	1	1
13	Nagpur	15	11	9
14	Indore	14	14	12
15	Bhopal	10	7	6
16	Visakhapatnam	19	18	19
17	Patna	9	8	7
18	Vadodara	16	16	13
19	Ludhiana	18	17	11
20	Agra	5	3	2

Figure 5: Key City Rankings

The table contains, for the three values of δ, the ranking of every city of the sample. Table 5 shows some interesting features. First, we find that the key city is Kanpur, whatever the value of δ. Kanpur, however, is only ranked 12 in terms of population and has therefore a medium productivity level. This finding contrasts with the intuition that large cities have more influence on aggregate. We have seen that, for high levels of interactions, Kanpur is the most central city. However, for small values of δ, Kanpur is only 6th in terms of centrality but still it is the key city. This shows that being central and contributing most to aggregate efforts does not always coincide. Kanpur is not so big a city but it represents an important node through which externalities go through and this is what captures our analysis.

Second, we find that size does not matter so much. Large cities, except for Delhi, are all ranked relatively low in terms of their contribution to aggregate efforts. For instance, when $\delta = 0.25$, Mumbai, the most populated city, is only ranked 15th while Bangalore, the third largest city, is ranked 16th. This is again surprising but reflects the position in the network of theses cities. Although they have a big weight, they are not central enough to exert large aggregate influence.

Last, we observe that the ranking of some cities is rather stable across variations of δ whereas others are very fluctuating. We identify one main group of cities that remain very influential whatever the value of δ: Kanpur, Delhi, Agra and Jaipur. They are the most influential cities and, except for some instances, their ranking is stable. Another group is formed of Surat and Visakhapatnam, that remain poorly ranked across δ. In the third group we find all the other cities, in majority the largest Indian cities, whose ranking is very dependent on the value of δ. For instance, Mumbai is ranked 6th for small interactions and only 15th for higher interactions. This illustrates the importance of correctly appraising local interactions when making policy decisions.

4.2 The key road

Here we conduct the same analysis, but instead of focusing on cities, we remove existing roads one after the other. Every time we remove a road we compute the impact this has on the aggregate efforts and again, we try to identify the key road. Table 6 presents the results for the same values of δ.

Our results are particularly interesting. First, we observe that the key road depends on the value of δ. This is also true for the rest of the ranking, as no groups of roads can be distinguished. Almost every existing road can be highly or poorly ranked according to the values of δ. This again emphasizes the role of this parameter in making the right decision. For instance, a policy maker willing to maintain some existing roads in priority, could be totally mistaken if he were considering the wrong values of interactions.

Second, and this is a consequence of the first point, the key roads do not coincide, in general, with large cities. The road between Kanpur and Agra joins cities only ranked 12 and 20 in terms of population, but happens to be the road that contributes most to aggregate efforts when $\delta = 0.25$, while the road joining Jaipur, ranked 10th, to Kanpur is the best for $\delta = 0.1$.

Third, there is some partial consistency between the ranking of the key roads and the ranking of the key cities. Indeed, the group of most contributing roads includes Kanpur, Delhi, Jaipur and Agra, which are the four most important cities according to table 5. This is somehow good news, because it

Road	Ranking δ = 0.01	Ranking δ = 0.1	Ranking δ = 0.25
2 10	1	4	5
3 4	2	13	19
2 11	3	9	10
2 20	4	5	4
1 8	5	19	23
4 6	6	14	20
10 12	7	1	2
1 9	8	25	26
5 10	9	11	12
1 14	10	21	21
3 6	11	23	25
11 12	12	6	6
4 13	13	15	16
12 20	14	2	1
2 19	15	17	14
10 20	16	3	3
12 15	17	7	7
12 17	18	8	9
4 16	19	22	24
5 14	20	18	17
5 18	21	20	18
11 17	22	12	11
15 20	23	10	8
7 17	24	24	15
8 18	25	26	22
13 15	26	16	13

Figure 6: Key road

confirms the validity of the first analysis. It makes sense that most important roads start or end at most important cities.

4.3 The optimal road addition

The third exercise consists of identifying, among all non existing roads, the one that would most contribute to aggregate efforts if a policy maker decided to build it. This is an important exercise in contexts of budget constraints, because money has to be spent efficiently. The way one chooses the next road to construct should account for all the network effects we have described throughout this paper.

To make the exercise somewhat consistent, we stick to the same distance constraint that we imposed on the adjacency matrix. Therefore only cities distant from no more than 600 kilometres can be considered to be linked. We do not consider the effects of a direct road between Mumbai and Kolkata for instance, because we believe the distance is not consistent with production's synergies going through roads. Once we impose this 600 kms constraint, we are left with 27 possible roads to be built.

We add each of these 27 roads in turn to the existing network and compute, for the same three values of δ, the effect this road has on aggregate efforts. Results are presented in Table 7.

Road	Ranking $\delta = 0.01$	Ranking $\delta = 0.1$	Ranking $\delta = 0.25$
2 12	1	1	1
1 5	2	11	16
2 15	3	5	5
4 9	4	19	24
1 18	5	18	21
5 8	6	16	17
10 11	7	2	3
5 9	8	20	19
5 15	9	8	9
10 15	10	4	4
11 20	11	3	2
10 14	12	7	11
8 9	13	25	27
8 15	14	14	15
10 18	15	10	12
11 15	16	6	6
8 14	17	22	23
10 19	18	12	8
14 20	19	9	10
9 14	20	26	25
14 15	21	15	13
9 18	22	27	26
13 14	23	21	18
15 18	24	17	14
19 20	25	13	7
14 18	26	23	20
13 16	27	24	22

Figure 7: Optimal road addition

This table shows that the results of the two first exercises are consistent, as the best road one could add to the network should be the one joining Delhi to Kanpur. This is true for every value of δ. Both Kanpur and Delhi are highly ranked in the key city table and involved in highly ranked roads.

It thus seems natural that this road should yield high synergies through the network and have a significant impact on aggregate efforts.

However, it is no longer true that optimal non existing roads involve cities that are high in the ranking of key cities. For instance, for low values of δ, some important potential roads join Mumbai to Ahmedabad (ranked 6th and 8th) or Hyderabad to Surat (ranked 4th and 17th). When δ increases, out of 27 possible roads, these same roads end up ranked 11th and 19th when $\delta = 0.1$ and 16th and 24th when $\delta = 0.25$.

4.4 A general appraisal

All the results we have obtained in the previous exercises emphasize a consistent feature: whether δ is close to 0 or far from 0 drives to radically different results. As soon as δ is far from 0, a group of six cities appears: Delhi, Jaipur, Lucknow, Kanpur, Bhopal and Agra. These cities always appear very highly ranked, according to current efforts, to key city rankings, to key road rankings and to optimal road addition rankings, both for $\delta = 0.1$ or 0.25. The interesting and surprising point is that this group of cities includes large, medium and "small" cities. A policy maker taking decisions according to intuitions would unavoidably make mistakes by underestimating the impact of these small cities.

In contrast, when δ is small the only city that consistently appears highly ranked is Delhi. Other than that, no city is at the same time highly central, a key city, involved in a key road and involved in an optimal road addition. The answers provided to a policy maker will then highly depend on what precise question he is interested in.

This difference between low and high interaction levels highlights the importance of correctly evaluating how synergies are created. Determining whether δ is low, medium or high is of crucial importance to policy decisions as this parameter changes the importance of the network structure in determining efforts.

5 Concluding remarks

In this paper, we have shown how the transportation network connecting large cities in India, and in particular its organizational features, can impact their productions' efforts. For that purpose, we considered a simple model of interdependent efforts on networks, and we conducted some calculations based on the calibration of towns' productivity. We determined the town that contributes the most to the current aggregate production effort as a function of the network structure and of the intensity of interaction, and, likewise, we found the existing road with maximal impact. We also considered the possibility of adding some roads to the current network and we found the optimal choice.

The objective of this theoretical exercise is by no way to be predictive, because we have made simplifying assumptions in order to keep exposition clear. The purpose was rather to shed some light on the often underestimated aspect of production and growth that accrues to synergies created by networks. We provide a methodology to account for these synergies. Some simple lessons can be grasped from our exercise.

First, network effects can be of sensitive magnitude. Furthermore, these effects cannot be captured by simply focusing on bilateral relations, because roads generate potentially large indirect effects. Second, the identification of important roads and cities highly depends on the intensity of interactions δ. This advocates for a careful inspection of local interactions. Last, the details of the network matter a great deal, because removing or adding a link might drastically change the results.

This analysis provided herein can be pushed further if one wishes to derive more realistic policy implications. For instance, it could be useful to consider heterogeneous costs of efforts (i.e. $c_i \neq c_j$) and to allow for different road qualities. This later variation of the model would allow for adjacency matrices with elements between 0 and 1 instead of considering binary values only.

References

[1] Ballester, C., A. Calvó-Armengol and Y. Zenou, 2006, Who's who in networks. Wanted: the key player, Econometrica, vol. 74(5), 1403-1417.

[2] Bonacich, P., 1987, Power and centrality: a family of measures, American Journal of Sociology, vol. 92, 1170-1182.

[3] Debreu, G. and I. Herstein, 1953, Nonnegative square matrices, Econometrica, vol. 21, 597-607.

[4] Handbook of Regional and Urban Economics, 2004, Volume 4, Edited by J. Vernon Henderson and Jacques-Franois Thisse, Elsevier.

[5] Wasserman, S. and K. Faust, 1994, Social Network Analysis: Methods and Applications, collection Structural Analysis in the Social Sciences, Cambridge University Press.

6 Appendix

Nash equilibrium profile of the game defined by (2): Individuals have the following utility function:

$$U_i(x_i, x_{-i}; G, \delta) = a_i x_i - \frac{c_i}{2} x_i^2 + \delta \sum_{j=1}^{n} g_{ij} x_i x_j$$

which leads to the first order condition:

$$x_i^* = a_i + \delta \sum_{i=1}^{n} g_{ij} x_j^*$$

In vectorial notations, the Nash equilibrium profile X^* satisfies:

$$(I - \delta G) X^* = A$$

This solution has a unique solution whenever the matrix $(I - \delta G)$ is invertible. The condition $\delta < 1/\mu(G)$ guarantees that the matrix is invertible, and it also guarantees that the inverse matrix is nonnegative, i.e., $M = (I - \delta G)^{-1} \geq 0$. Therefore

$$X^* = MA$$

and $x_i^* \geq 0$ for all i. Developing, we get for city i

$$x_i^* = \sum_{j=1}^{n} m_{ij} a_j$$

The term $m_{ij} a_j$ represents the contribution of town j to town i's effort. The term m_{ij} admits a natural interpretation in terms of network paths. To see this, notice that M can be decomposed as

$$M = \sum_{k=0}^{+\infty} \delta^k G^k$$

where the matrix G^k keeps track of paths of length k on the network.

The equilibrium profile can thus be rewritten as

$$X^* = \sum_{k=0}^{+\infty} \delta^k G^k A$$

Developing, we get for city i

$$x_i^* = \sum_{j=1}^{n} \left(\sum_{k=0}^{+\infty} \delta^k G_{ij}^k \right) a_j$$

Therefore, we can identify m_{ij} as the total number of paths between i and j, where paths of length k are weighted by a factor δ^k.

Figure 8: India National Highway Network Map

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0	0	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1
3	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	0
5	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0
6	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
8	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
9	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
11	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0
12	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0	1	0	0	1
13	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
14	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1
16	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0
18	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
19	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	0	1	0	0	0	0	0	0	0	1	0	1	0	0	1	0	0	0	0	0

Figure 9: The adjacency matrix