Modelling Transportation as a Network Industry

Yuri V. Yevdokimov, Assistant Professor
Departments of Economics and Civil Engineering
University of New Brunswick
P.O. Box 4400, Fredericton, NB E3B 5A3
Tel. (506) 447-3221
Fax (506) 453-4514
E-mail: yuri@unb.ca

Key words: network effects, systems approach, transportation system, social optimum, computer simulation

Abstract

Production and consumption of transportation services are analysed in light of the network effects discussed and developed in the literature on economics of networks. Such an approach is an alternative to conventional transportation economics in general and the existing spatial transportation modelling in particular. The latter treats transportation networks in a cartographical sense in which coordinates and geographical locations matter. In contrast, the proposed framework regards transportation networks as a system of compatible devices to move people and freight in terms of system capacity and capacity utilization.

As well, the systems approach to model transportation is discussed. Basic principles of the approach with respect to transportation are formulated. According to these principles, transportation networks are incorporated in an economic system that has a hierarchical structure. Horizontal and vertical linkages of the transportation network within the economic system are identified. The approach overcomes two major drawbacks of the existing spatial transportation models, namely the failure to incorporate macroeconomic feedbacks, and separation of travel demand from freight demand in the network.

Finally, a computer simulation exercise, based on the designed model, is performed in order to show that if the network effects are not taking into account, social optimum becomes unattainable in principle.

Introduction

Analysis of the transportation economics literature has revealed two basic frameworks in which transportation has been modelled : (i) transportation as a market, and (ii) transportation as a network. The first one is just applied microeconomics and is supported by the majority of transportation economists. The second one recognizes some specifics of transportation, and instead of market uses network as the relevant structure. Three specifications of a transportation network can be identified within this framework: (i) transportation network as infrastructure (usually as public capital); (ii) transportation network as a linear cost function of distance, and (iii) spatial transportation networks. Since the latter comes the closest to the engineering understanding of transportation networks and real world, it is necessary to look into it in more detail.

Usually the spatial transportation models study transport-location interaction in regional and urban economics contexts (for a review of the models, see Button and Gillingwater, 1983, and Torrens, 2000). In this context, the spatial nature of the transportation network has cartographical or topological interpretation in which co-ordinates and geographical locations matter. In general, spatial models involve the development of a set of mathematical equations in which a transportation network is represented by the generalized cost of travel/shipping as a linear or sometimes non-linear function of the distance between different locations, called either origins and destinations or supply and demand areas. Large dimension of these models requires a massive database, and the model's calibration imposes tremendous strain on the available data. The following chart (Figure 1) summarize the above discussed two frameworks in terms of their basic features.

Based on the analysis of the mainstream transportation economics literature, it is possible to draw the following conclusion: *Conventional transportation economics fails to correctly describe* *production and consumption of transportation services because of (i) violations of the competitive market conditions; and (ii) narrow definition of a transportation network.*

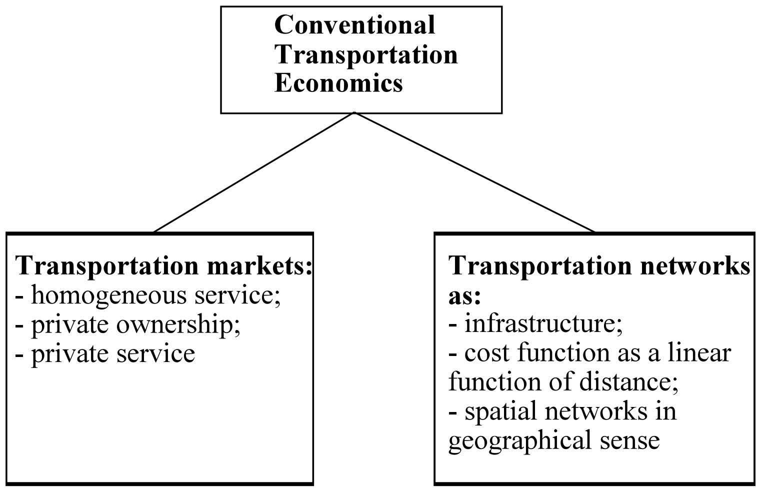


Fig 1. Basic features of conventional transportation economics

The first failure is associated with transportation being a set of heterogeneous services produced and consumed in a multi-purpose transportation network some parts of which are publicly owned. Therefore, strictly speaking the service is not a private one on the one hand, and it is not homogeneous on the other.

The second failure is associated with the literature on economics of networks in which it is argued that the spatial nature of networks should denote the dimension or size of the network components, not geographical locations.

This study takes on the approach discussed in the literature on economics of networks. Although this literature has mainly focussed on information networks, its fundamentals are still valid for transportation.

Basic principles of the approach

The following two principles should be applied to the economic modelling of a transportation system:

- transportation is a network industry;
- transportation is a part of a broader economic system.

With respect to the first one, it is argued in the literature on industrial organization that whenever sunk cost and the degree of uncertainty are very high, the right market structure should be that of a network. Gentzoglanis (2001) claims that *"if, for example, uncertainty is low, the benefits from using the market are much higher than for other strategies. At intermediary levels of uncertainty, integration is the more advantageous strategy, while when uncertainty is at a high level, networks provide a better choice. There is then a relationship between uncertainty (as measured by the amount of sunk costs) and net benefits of each alternative strategy"*

As a matter of fact, transportation is characterized by both, high sunk cost and high uncertainty. For example, Boyer (1998) writes *"... A good example is a railroad embankment. While railroad company must pay to create a level roadway, the result has no scrap value. In order for the land to be used in its next best use, the embankment may have to be removed. It thus has no opportunity cost, and any payment made to create it is a sunk cost... The dominance of transportation expenditures by sunk costs strongly affects the analysis of the sector"* (Boyer, 1998, p. 99) . Regarding uncertainty, Boyer points out: *"[Transportation] fixed facilities investment are also risky because they often are large and lumpy and are made in advance of demand"* (Boyer, 1998, p. 229).

Second principle is associated with the linkages transportation has with other components of an economic system. Transportation does not exist on its own, and therefore its development and functioning depend on the development and functioning of the whole economic system which provides feedbacks and interactions. In turn, as Oum and Dodgson (1997) note, the general failure of the existing spatial transportation modelling is the lack of these feedbacks.

These two principles, transportation as a network industry and as a part of a broader economic system, have two implications for the design of the right economic framework:

- network effects should be directly incorporated into economic analysis of transportation;
- systems analysis should be applied to study transportation.

This paper is mostly dedicated to the first implication, and that is why later the transportation network effects are discussed in detail. However, before looking into these effects, the systems approach should be briefly discussed. The approach is based on the statement that an object under study can be viewed as a system represented by elements whose collective behaviour is essential to the whole, but not relevant by themselves. Applied to transportation, the hierarchical structure shown in Figure 2 should be described while modelling transportation system. With respect to this

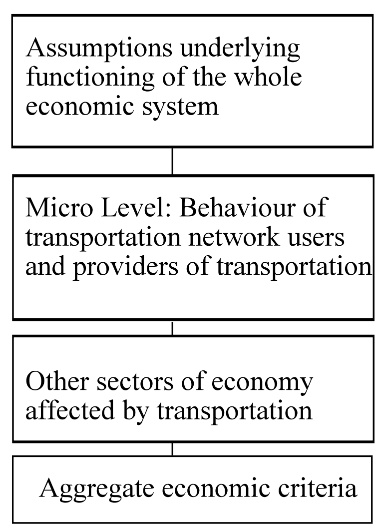


Fig 2. Systems approach to modelling transportation network

structure, if assumptions at the first level are changed, it will affect the economic system in general and the behaviour of transportation providers and users in particular. On the other hand, developments in transportation will affect other components of the economic system, and that is why all these components should be taken in their interaction.

Such an approach implies that there are two types of linkages associated with transportation system: (i) vertical linkages as shown in Figure 2; and (ii) horizontal linkages as it is depicted in Figure 3 below:

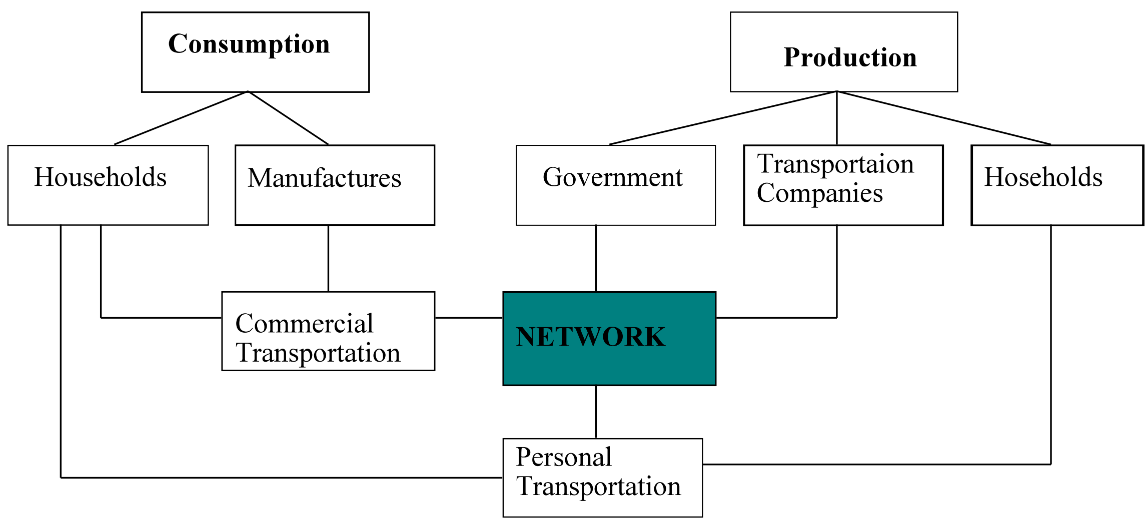


Fig. 3. Horizontal linkages of a transportation system

As noted, this paper is dedicated to the horizontal linkages of a transportation system. The main idea is to show that network effects should be taken into account because otherwise the social optimum becomes unattainable in principle. Therefore, below a simple mathematical model is introduced and then a computer simulation exercise is discussed to illustrate the described approach.

The model

Consumption

The fact that transportation services are consumed in a transportation network results in two additional effects that should be included on consumption side: (i) congestion, and (ii) positive network externalities. It is possible to do so by defining the network user's utility function that directly incorporates these effects.

For example, MacKie-Mason and Varian (1994) start their analysis of congestion in a network with the following quasi-linear utility function:

$$u_i(t_i, Y) + m_i \tag{1}$$

where t_i is user's i use of the network, Y is the network's capacity utilization and m_i is money that the user i has to spend on other goods and services. Furthermore, the network's capacity utilization is defined as

$$Y = \frac{T}{K} \tag{2}$$

where T is the total use of the network and K is the network's capacity. Let us assume that there are N identical users in a transportation network of capacity K. In such a setting, t_i is the desired volume of transportation by user i and T is total traffic in the network.

In addition, according to the Metcalfe's Law (see Katz and Shapiro, 1994), the value of a network is proportional to the number of the network's users N squared because of the positive network externalities. Therefore, positive network externalities enter the utility function through N^2, and complete specification of the utility function becomes

$$u_i(t_i, Y, N^2) + m_i \tag{3}$$

Suppose that p_T is the price of transportation. The utility maximization problem becomes

$$\textit{maximize } u_i(t_i, Y, N^2) + m_i \tag{4}$$

$$\textit{Subject to } m_i + p_T t_i \leq M_i$$

where M_i is total money income earned by the user i and m_i is money spent on goods and services other than transportation. At this point it is necessary to note that Y and N are functions of the volume of transportation t_i:

- capacity utilization Y is an increasing function in t by definition;
- N is an increasing function of t because an increase in number of users N increases user i's transportation t_i and vice versa by assumption of positive network externalities.

Solution to the utility maximization problem (4) is

$$\frac{\partial u_i}{\partial t_i} + \frac{\partial u_i}{\partial Y} \times \frac{\partial Y}{\partial t_i} + \frac{\partial u_i}{\partial N} \times \frac{\partial N}{\partial t_i} = p_T \tag{5}$$

Partial derivative $\frac{\partial u_i}{\partial Y}$ is negative because an increase in capacity utilization Y eventually increases congestion and consequently reduces utility obtained from transportation. In turn, partial derivative $\frac{\partial u_i}{\partial N}$ is positive because the value of the transportation to the user i (utility) depends on the value of the network which is proportional to N^2. Individual Marshalian demand for transportation can be derived from (5) as

$$t_i = f(p_T, Y, N) \tag{6}$$

or

$$t_i = f(p_T, K, N) \tag{7}$$

since $Y = \dfrac{Nt_i}{K}$. Aggregating (7) over N identical users, market demand for transportation is

$$T_D = \sum_{i=1}^{N} t_i = F(p_T, K, N) \tag{8}$$

where T_D is the volume of transportation demanded by all N users of the network.

Production

For simplicity, let us assume that transportation is provided in the transportation network by a single provider, a private transportation company, according to the following production function:

$$T = F(V, L) \tag{9}$$

where V is transportation capital (number of vehicles) and L is transportation labour (operators of vehicles or labour hours of operators). Government provides infrastructure, expressed in terms of the network's capacity K. Therefore, from a standpoint of the economics of networks, where a network is a system of compatible devices and the associated services to carry a specific function, production of transportation services involves: (i) immobile transportation capital (infrastructure); (ii) mobile transportation capital (vehicles), (ii) transportation labour (vehicle operators). From an economic viewpoint, these three are factors of production. Furthermore, fixed capacity K of the network causes congestion in production (negative production externality) which can be captured by a means of the earlier defined network capacity utilization *Y*.

Combining all of the above, the generalized production function for transportation services produced in the network can be written as follows:

$$T = F(Y, K, V, L) \tag{10}$$

or

$$T = F(N, K, V, L) \tag{11}$$

since $Y = \frac{Nt_i}{K}$.

The above described framework can be represented by the following chart:

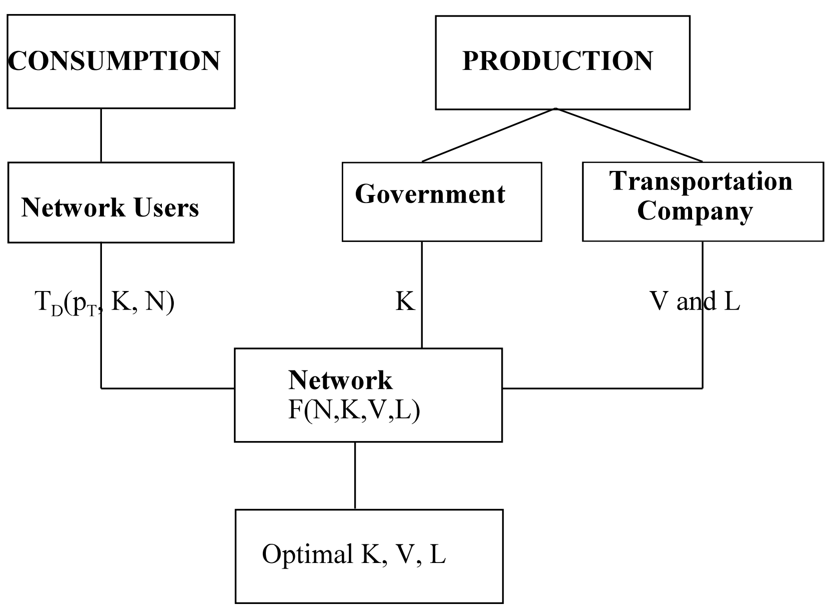


Fig 4. Reduced form of production and consumption of transportation in the transportation network

Technically the problem becomes that of cost minimization:

$$\textit{minimize } C(K,V,L) \tag{12}$$

$$\textit{Subject to } F(N,K,V,L) \geq T_D(p_T,K,N)$$

where *C(K,V,L)* is total expenditure on transportation by the government and transportation company. In this problem, the price of transportation p_T and the number of the network users *N* are given exogenously. In such a way, solution to (12) are optimal values of capacity (infrastructure) *K*, number of vehicles *V* and number of labour hours of vehicle operators *L*. Moreover, this is a social optimum since all externalities in consumption and production were internalized. In order to illustrate the framework, a computer simulation exercise has been performed.

Computer simulation exercise

First, it is necessary to explain the choice of the two key functions: (i) utility function, and (ii) generalized production function , used in the computer simulation. The following functional form of the utility was chosen

$$u(K,N,T) = T - \frac{N}{K}T^2 + N^2 T \tag{13}$$

In this utility function, $\frac{\partial u}{\partial T}$ is positive for small values of *T*; however, it is negative when *T* becomes large. It is consistent with observed behaviour, since the network users enjoy transportation up to a point when the network becomes congested (high values of *T*). The same is true regarding $\frac{\partial u}{\partial N}$. It is positive for small values of N but negative when N becomes very large (over-crowded network). And finally, $\frac{\partial u}{\partial K}$ is strictly positive which is also consistent with the observed behaviour: any

improvements in infrastructure increase the value of the network to its users.

The following form of the generalized production function was chosen:

$$F(N,K,V,L,T) = [1-(\frac{TN}{K}-1)^2]V^{0.5}L^{0.5} \tag{14}$$

The fundamental property of the function is: It is an increasing in T and N at a decreasing rate as long as $TN \leq K$, and a decreasing when $TN > K$. In terms of the transportation network, as long as the network capacity is not reached, the production of transportation services increases with increase in the number of vehicles and labour hours of vehicle operators; however, once the network's capacity is achieved, the production decreases which is a phenomenon known to transportation engineers through flow-density diagram. It turns out, that the utility function and the generalized production function capture the earlier discussed network effects rather well.

Computer simulation was performed from a standpoint of a private producer of transportation, a transportation company, and then from a social standpoint. The results are presented in the following tables:

Table 1. Social optimum

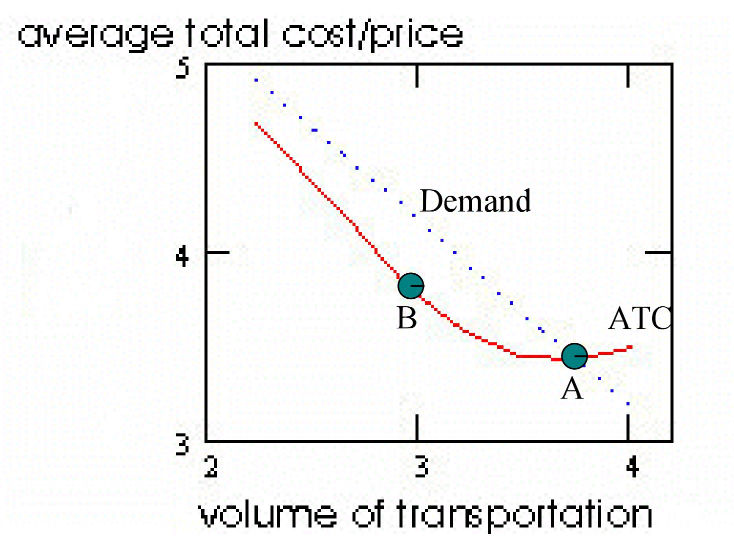
T	**V**	**L**	**ATC**	**Y**
2.236	3.018	3.018	4.678	0.774
2.667	3.062	3.062	4.173	0.870

2.942	3.112	3.112	3.826	0.945
3.214	3.246	3.246	3.576	0.990
3.483	3.493	3.493	3.439	0.998
3.750	3.908	3.908	3.417	0.960
4.015	4.260	4.260	3.493	0.941

Table 2. Private optimum

target T	***real T***	***V***	***L***	***ATC***
2.236	1.553	2.237	2.237	6.100
2.667	2.086	2.668	2.668	4.953
2.942	2.444	2.943	2.943	4.454
3.214	2.804	3.215	3.215	4.076
3.483	3.162	3.484	3.484	3.784
3.75	3.516	3.751	3.750	3.556

In private decision making, neither congestion nor positive network externalities are taken into account. As a result, the private producer consistently under-produces (*real T* versus *target T*) which results in higher average total cost (ATC). This result can be presented graphically as follows: Social optimum is associated with point A while private optimum can be depicted by point like



B. Private producer consistently mis-evaluates the magnitude of the rightward shift of demand over time due to positive network externalities plus it cannot correctly choose the required level of production due to congestion since the network capacity K is not a part of the producer's decision.

Both types of externalities can be internalized by the authority through congestion tax or subsidy. The choice of the right strategy depends on the magnitude of both effects. If congestion over-weighs positive network externality, the authority has to introduce congestion tax. However, the size of the tax depends on the network externalities as well. In such a case the congestion tax, which is a form of user fee, equals the difference between the two externalities. On the other hand, if positive network externalities exceed congestion, then the authority has to subsidize private producer in order to correctly internalize these externalities. Again, the size of the subsidy depends on the difference between the two types of externalities. In this computer exercise, the latter case was obtained which pointed out at a subsidy as the right policy option to achieve the socially desirable outcome.

Conclusion

Although transportation services are produced and consumed in transportation networks, transportation economists have failed to correctly incorporate the network effects. While congestion was given a considerable attention in transportation economic literature, positive network externalities were mostly ignored. However, as it is shown in this paper, the choice of an appropriate public policy in transportation depends on the interplay between two types of externalities: (i) congestion in production and consumption, and (ii) positive consumption externalities, based on Metcalfe's Law. It turns out that if the two are not taken into account, the final outcome becomes sub-optimal.

References

Boyer, K.D. , *Principles of Transportation Economics,* Addison Wesley Longman, 1998

Button, K.J. and Gillingwater, D. , *Transport, Location and Spatial Policy,* Aldershot: Gower, 1983

Gentzoglanis, A., "Networks and Proximity: An Empirical Analysis", *Paper Presented at the ACEA Conference,* September 2001

Katz, M.L. and Shapiro, C., Systems Competition and Network Effects. *Journal of Economic Perspectives*, Vol. 8, 1994, 93-115

MacKie-Mason, J.K. and Hal R. Varian (1994), "Pricing Congestible Network Resources", *University of Michigan Department of Economics Research Paper,* November 1994

Oum, T.H., Dodgson, J.S. and others (ed.), *Transport Economics: Selected Readings,* Harwood Academic Publishers, 1997

Torrens, P.M. , "How Land-Use-Transportation Models Work", *Centre for Advanced Spatial Analysis, University College London*, 2000