A model of the vicious cycle of a bus line

Asaf Bar-Yosef\textsuperscript{a}, Karel Martens\textsuperscript{b}, Itzhak Benenson\textsuperscript{a,*}

\textsuperscript{a}Department of Geography and Human Environment, Tel Aviv University, Israel
\textsuperscript{b}Institute for Management Research, Radboud University Nijmegen, The Netherlands

**Abstract**

It has been frequently noted that in a non-regulated environment the development of public transport service is self-adjusting: Faced with decreasing demand, operators will tend to reduce service to cut costs, resulting in a decrease in the level-of-service, which then triggers a further drop in demand. The opposite may also occur: high demand will induce the operator to increase supply, e.g. through an increase in frequency, which results in a higher level-of-service and a subsequent increase in passenger numbers, triggering another round of service improvements. This paper adds to the literature by presenting an analytic model for analyzing these phenomena that we call vicious and virtuous cycles. Based on field data regarding passengers' variation in willingness-to-wait for a public transport service, we investigate the dynamics of the line service and show how the emergence of a vicious or virtuous cycle depends on the total number of potential passengers, the share of captive riders, and bus capacity. The paper ends with a discussion of the implications of the findings for the planning of public transport services.

1. What is vicious cycle?

When a bus line operator is faced with low demand, he tends to cut expenses instead of improving level-of-service (LOS)\cite{Reinhold2008}. The easiest way to do that is to reduce the frequency of buses. In response, some of the passengers may use the service less or not at all, as they change transportation mode, destination choice, or decide to forego a trip altogether\cite{Castaline1980, Bly1987}. This decrease in the demand for the bus service may cause a further decrease in bus frequency. This vicious cycle continues until only captive passengers, who do not have an alternative mode of travel, continue using the line\cite{Downs1962, Mohring1972}.

The positive, virtuous cycle, is also observed: high demand for a bus line induces the operator to supply more buses and this causes an increase in LOS and perceived satisfaction of passengers. Non-captive passengers react to the improvement, switch from private cars to buses\cite{Liu2010}, leading to a rise in demand\cite{Kingham2001}, which in turn triggers a further increase in LOS.

The above negative and positive cycles are well known among transportation researchers\cite{Mohring1972} and have been investigated both theoretically\cite{Martens2011} and in practice, as in\cite{Xu2010}, who mention in their review of the history of the public transportation network in Beijing that poor public transportation systems may lead to lines' vicious cycle. In contrast, Reinhold and Kearney,\cite{Reinhold2008}, in a study of the Berlin bus network, shows that a virtuous cycle can be induced, even when facing a declining revenue stream from passengers. By deliberately increasing the frequency on a lim-
Mohring (1972) was the first who studied this problem analytically focusing on the level of subsidies necessary for preserving a socially optimal level of public transport frequencies. He demonstrated that urban public transport operations should be subsidized due to economy of scales on the user’s side – higher frequency of buses results in the decrease of total user costs due to reductions in waiting time. This “Mohring effect” was further investigated in many papers (Jansson, 1979; Pedersen, 2003; Jara-Diaz, Gschwender, 2003; Clifton and Rose, 2013). Van Reeven (2008) tried to defect the Mohring effect, but Basso and Jara-Diaz (2010) demonstrated that his result is based on the unrealistic assumption that trip demand does not react to prices, no matter how high they are. Other aspects that influence bus line use, apart from passengers’ willingness-to-wait and fare prices, that have been studied, include congestion (Tirachini and Hensher, 2011), information provision (De Borger and Fosgerau, 2012) and interaction between buses (Ibarra-Rojas and Rios-Solis, 2012; Bartholdi and Eisenstein, 2012) are also considered.

The papers of Mohring and others do not directly investigate the dynamics of the vicious and virtuous cycles, but analyze whether a subsidy would be welfare enhancing and, hence, whether a solid argument can be made in favor of subsidization of public transport. While a subsidy may indeed stop the progress of a vicious cycle, vicious and virtuous cycles can also occur in a subsidized environment if external conditions change. A vicious cycle may occur in case passenger numbers decrease due to, for instance, a decrease in gasoline prices, the opening of a new road, or a decrease in the population. Such an externally induced drop in passenger numbers will not only lead to a decrease in direct operator revenues from ticket sales. Following the Mohring argument, it should also lead to a reduction in the subsidy to the operator, as fewer passengers imply less total waiting time and hence less overall welfare loss. If we assume that the subsidized operator is allowed to adjust public transport frequencies in response to changing demand, than, to preserve the total profit, the operator will reduce the frequency, thereby unintentionally restarting the vicious cycle. These arguments also hold true for the virtuous cycle. If external factors (such as population growth or increasing gasoline prices) cause an increase in the number of bus passengers, not only do revenues from ticket sales go up, but, following the Mohring argument, subsidies should also be increased. Following such an increase in passengers numbers, and irrespective of any changes in the amount of subsidy, the operator may be required to increase the level of service or may voluntarily decide to do so, which may generate an increase in passengers and ultimately a virtuous cycle. The public transport system is an open sub-system of the larger transportation system and Mohring-like subsidization thus cannot prevent vicious and virtuous cycles.

Levinson and Krizek (2008) were the first who tried to investigate the dynamics of vicious and virtuous cycles analytically. They introduced three hypothetical dependencies: one linear, of bus speed as a function of bus waiting time, and two non-linear, of the number of passengers as a function of bus speed, and of waiting time as a function of the number of passengers that use the line. The non-linear dependencies resulted in a qualitative conclusion that the bus line has two stable states – one in which the bus waiting time is very high and the number of passengers is close to zero, and the other, in which the waiting time is close to zero and all potential riders use transit. At the same time, the dependencies they used are purely hypothetical and the authors do not propose any mechanisms that can explain them. As a result, the number of passengers in an equilibrium state is unrealistically high or low (see Appendix B for more details).

In what follows, we focus on passengers’ reaction to public transport frequency exclusively and ignore other factors that may change the demand for public transport trips, such as fare prices, costs of operating a car, or changing land use patterns. We assume all these factors to be constant and investigate ridership as dependent on the mutual interdependency between passengers’ willingness-to-wait and public transport frequencies. We furthermore assume that the public transport operator can increase or decrease service frequency in a continuous way, and that the costs of transporting one additional passenger is lower than the extra revenue generated from additional ticket sales (and/or additional subsidies). In line with the literature we also assume that the bus operator is interested to maximize its profits by attracting more passengers (Van Nes, 2002).

Based on these assumptions and on experimental data on passengers’ willingness-to-wait for a public transport service, we develop a full analytical model of the interaction between the passengers and the operator, and formulate the conditions when the negative (vicious) or positive (virtuous) cycles emerge. The aim of the paper is to provide a systematic understanding of the dynamics of the vicious cycle under various circumstances. While we refer to bus services throughout the paper, the analysis and conclusions are also relevant for other forms of scheduled public transport services.

The paper is organized as follows. In Section 2, we present our assumption regarding the distribution of passengers’ willingness-to-wait, based on a brief overview of surveys into the issue. Then, in Section 3, we analyze the interplay between passengers’ willingness-to-wait and public transport frequency. Based on this, we present the dynamics of the vicious cycle as dependent on total number of potential passengers and share of captive riders (Section 4). Section 5 applies the model to the problems of vehicle size and the introduction of a new line. We end with a discussion in which we elaborate on the relevance of the findings for real-life public transport services and suggest some directions for further refinement of the model (Section 6).

2. Field estimates of passengers’ willingness-to-wait for a bus

Public transport ridership depends on a diversity of factors, such as in-vehicle travel time, walking distance to a bus stop, walkability of the urban environment, public transport fares, as well as the quality and cost of alternative modes
of transport. Among these factors, waiting time is particularly important, as riders view waiting time as much more burdensome than an equivalent amount of time spent in travel (Ceder, 2007). Conventional wisdom holds that average waiting time equals one half of the expected headway (the time gap between two consecutive buses) (Hess et al., 2004). While passengers can reduce their waiting time by synchronizing their arrival at a public transport stop with the service schedule, this strategy is only effective if transit service is reliable. Even in this case some waiting time will often remain, when defined as the difference between passengers’ preferred departure time and the scheduled departure time provided by the service (van Reeven, 2008). Note that this waiting time will be less visible at public transport stops, as passengers may prefer to wait at the origin or a trip (i.e., at home), hence the term ‘hidden waiting time’ (Furth and Muller, 2006; Basso and Jara-Díaz, 2010).

The willingness-to-wait for a bus or other public transport service varies among passengers. This has been shown in various surveys performed over the past two decades that estimated the distribution of the maximal bus waiting time \( \tau \) that passengers are ready to accept in case they do not adjust their arrival to the bus stop to the bus timetable.\(^2\)

Peterson et al. (2006), in a stated preference survey, studied how long bus users are willing to wait for a free transfer between two bus operators. They obtained that only 10% of the respondents indicated to be willing to wait for more than 15 min and only 3% for more than 20 min. Kim and Ceder (2006) asked potential passengers if they would use a planned shuttle service as depending on the time interval between buses. Half of the passengers responded that they definitely or probably accept a 5 or 10 min waiting time, while more than 70% raised doubts that they would wait for 15 min or more. In a city-wide survey conducted in Dublin, 90% of the respondents claimed that they are not willing to wait for more than 20 min, and 70% were not willing to wait for more than 10 min (Caulfield and O’Mahony, 2009). Only 2% claimed they would be willing to wait more than 30 min.

Since stated preference studies do not necessarily accurately reflect people’s revealed preferences, it is worthwhile to compare these findings with a revealed preference survey. In such a survey, Hess et al. (2004) studied the transport behavior of college students who had the choice between two identical bus lines serving the same origin–destination pair, one line served by blue buses and the other by green buses. Students could ride for free on the ‘blue’ bus (with the university paying for the fare on behalf of the student), but had to pay a fare of 0.75 US $ for a ride on the ‘green’ bus. The headway between buses on the blue line was on average 10 min; between green buses, it was 12 min. Hess et al. (2004) measured the waiting time of the students who decided to wait for a blue bus rather than board a green bus when the latter arrived first at the bus stop. They found that the average additional waiting time for these riders was 5.8 min (with an STD of 3.3 min); the median elapsed time was 4.5 min. With a headway of 10 min between the blue buses, these findings are largely in line with the conventional wisdom that actual waiting time equals, on average, one-half of the expected headway. Since virtually all students used the bus line on a regular basis, they will have had a relatively accurate estimate of the expected additional waiting time when deciding not to board an available green bus. The findings can thus be interpreted as students’ actual willingness-to-wait for a bus. Important for the purposes of this paper, is the finding that the distribution of willingness-to-wait time is comparable to those found in stated preference studies (see below). Hence, we conclude that stated preference findings can be used for the analysis of the vicious cycle phenomenon.

Fig. 1 presents the results of the two stated preference surveys of passengers’ willingness-to-wait that are employed in our model. The results have been derived from the studies of Peterson et al. (2006) and Caulfield and O’Mahony (2009). We approximate these empirical distribution densities by the function \( f(\tau) = C\tau^a e^{-b\tau} \), with the values of parameters \( a \) and \( b \) estimated using the non-linear regression method of SPSS19 and \( C \) serving as a normalizing constant for ensuring \( \int_0^{\infty} f(\tau) = 1 \).

The reaction of the passenger population to waiting time and, hence, to bus frequency, is crucial for understanding modal shift between car and bus (Kingham et al., 2001) or, more generally speaking, for understanding bus use. Hence, it is a key element in the model of the vicious cycle and in what follows we perform a deep analytical investigation of this phenomenon. We intentionally ignore other factors that influence the usage of public transport, such as accessibility of destinations, comfort of the service, car travel times, or parking availability and costs. These factors obviously will have an impact on the occurrence of vicious and virtuous cycles in a real-world setting, but do not change the fundamental dynamics of the phenomenon.

3. From willingness-to-wait to bus line dynamics

The model considers a circular bus line of a trip time \( L \). Let the entire population that can be served by the line during this trip be indicated by \( P_{\text{total}} \), and the fraction of captive passengers among them be \( g \), while the fraction of non-captive passengers be \( 1 - g \), i.e. the line population consists of \( gP_{\text{total}} \) Captive passengers and of \( N = (1 - g)P_{\text{total}} \) non-captive passengers. Captive passengers always make the trip, while non-captive passengers decide on using the bus depending on the time \( \tau \) they have to wait at a stop. In line with Mohring’s paper and much of the literature, we assume that passengers do not adjust their arrivals to the bus stop to the timetable (van Reeven 2008). Let us denote as \( N(\tau) \) the number of non-captive passengers whose maximal waiting time is between \( \tau \) and \( \tau + d\tau \), that is, \( N = \int_0^{\infty} N(\tau)d\tau \).

\(^2\) In the remainder of the paper, we will talk about bus services rather than the general category of public transport services, as most of the studies on willingness-to-wait have been carried out for bus services. However, our model also holds for other forms of public transport, although these modes may differ in the extent to which frequencies can be adapted to changes in the level of ridership.
Let us consider the dynamics of the passengers taking the line at a daily time resolution. At a given day \(d\), some non-captive passengers who waited for the bus that day longer than their maximal waiting time \(s\), may decide to stop using the bus line in the future. We assume that this decision is made after several failures and, thus, the fraction of those who left at a day \(d\) is lower than the fraction of those who waited for too long that day. In parallel, we assume that some of the non-captive passengers, who currently do not use the bus, attempt to use it in the hope that the waiting time is reduced in comparison to previous experiences. Some of them may succeed to board several times in a row after waiting less than \(s\) and may thus change their mode back to the bus.

Let \(B(d)\) be the number of buses serving the line at a day \(d\), and \(T_B(d) = L/B(d)\) be the time interval between buses. Let us further assume that every passenger uses the bus once a day and that the daily "leave line" rate at which passengers switch from being a user to becoming a non-user is \(b\) (\(0 < b < 1\)), while the "join line" rate of the opposite switch is \(a\) (\(0 < a < 1\)). Users for which \(s > T_B(d)\) are all served, while users with \(s < T_B(d)\) are served if they arrive to the stop \(s\) minutes or less before the bus arrival. Assuming that the passengers arrive to the bus stop randomly in time, the fraction of the served users among those of \(s < T_B(d)\) is \(s/T_B(d)\).

Let us denote the number of non-captive passengers whose maximal waiting time is \(s\) and who use the bus at a day \(d\) as \(N_u(s, d)\), and those who do not use bus at a day \(d\) as \(N_n(s, d)\).

The dynamics of \(N_u(s, d)\) and \(N_n(s, d)\) can be presented as follows:

If \(\tau < T_B\) then

\[
N_u(\tau, d + 1) = N_u(\tau, d) + aN_n(\tau, d) \left( \frac{\tau}{T_B} - \beta N_u(\tau, d) \right) \left( 1 - \frac{\tau}{T_B} \right)
\]

\[
N_n(\tau, d + 1) = N_n(\tau, d) - aN_u(\tau, d) \left( \frac{\tau}{T_B} + \beta N_u(\tau, d) \right) \left( 1 - \frac{\tau}{T_B} \right)
\]

Otherwise

\[
N_u(\tau, d + 1) = N_u(\tau, d) + aN_n(\tau, d)
\]

\[
N_n(\tau, d + 1) = N_n(\tau, d) - aN_u(\tau, d)
\]

The number \(P(d)\) of passengers served at a day \(d\) is given by:

\[
P(d) = gP_{total} + (1 - g) \int_0^{\infty} N_u(\tau, d) d\tau
\]

To proceed, let us assume that a bus company decides on the bus frequency once in a quarter of a year, and that the leave rate \(\beta\) is essentially higher than a join rate \(\alpha\). That is, passengers are more inclined to stop using a service if the bus does not arrive during their maximal waiting time, than to resume using the service in a hope that waiting time has decreased. Fig. 2 presents the distribution of the passengers who continue to use the bus line (dark grey) versus the initial distribution of the passengers by their maximal waiting time, at \(d = 0\) (light and dark grey together), and after a quarter of a year \((d = 90\) days\), for \(T_B = 15\) min and different values of \(\alpha\) and \(\beta\).

According to Fig. 2, during a three month period the majority of passengers with \(\tau < T_B = 15\) leave the line in case \(\beta/\alpha\) is sufficiently large.
that continue to use the bus after the end of the quarter. In this case, 
\[ t = \frac{1}{4} \] and 
\[ TB_{total} = \frac{L}{B} \] where \( t \) be the time interval between the buses for the line served by \( L \).

For buses needed to carry the current bus riders, \( X \) is potential passengers.

The distribution of willingness-to-wait according to Peterson et al. (2006), assuming all non-captive passengers attempt to use the bus service from day \( d = 0 \) onwards.

Fig. 2. The distribution of \( N(t, \tau, d) \) at day \( d = 0 \) (light grey) and at day \( d = 90 \) (dark grey), for, \( a = 0.005, \beta = 0.05, 0.25 \) and \( 0.5 \) (all per day), \( T_B = 15 \) min and the distribution of willingness-to-wait according to Peterson et al. (2006), assuming all non-captive passengers attempt to use the bus service from day \( d = 0 \) onwards.

Let us, for simplicity, ignore passengers with \( \tau < T_B \) that continue to use the bus after the end of the quarter. In this case, \( \int_0^T N_n(t, \tau, d)d\tau \) can be presented as \( (1 - g)P_{total}\int_{\tau=0}^{\tau=\infty} f(\tau)d\tau \), and the dynamics of the bus users, by quarters \( t \), can be presented as:

\[ P(t + 1) = gP_{total} + (1 - g)P_{total}\int_{\tau=0}^{\tau=\infty} f(\tau)d\tau \]  
(3)

where as defined above, \( P_{total} \) is the overall number of potential passengers, \( g \) is the fraction of captive passengers among them, and \( t \) denotes a quarter of a year.

The full model, which accounts for passengers with \( \tau < T_B(d) \) who continue to use the bus line, is presented in Appendix A, where we demonstrate that the dynamics of (3) and of the full model are qualitatively similar.

For the next step of the analysis, let us denote by \( m \) the optimal number of passengers per bus that makes it maximally profitable for the operator to run the line. The value of \( m \) evidently depends on the maximal capacity of a bus, but will be somewhat lower than that to avoid overcrowded buses (and thus a risk that potential passengers will have to be left behind at a bus stop due to a lack of capacity), but higher than the number of passengers necessary to merely cover operation costs (Meignan et al., 2007).

Let us denote as \( X(t) \) the fraction of all potential passengers using the bus line, \( X(t) = P(t)/P_{total} \) and as \( B(t) \) the number of buses needed to carry the current bus riders, \( B(t) = P(t)/m \). Let \( B_{total} \) be the number of buses needed to carry the total potential of passengers, \( B_{total} = P_{total}/m \), and \( T_{B_{total}} \) be the time interval between the buses for the line served by \( B_{total} \) buses, \( T_{B_{total}} = L/B_{total} \). In this notation

\[ T_{B(t)} = \frac{L}{B(t)} = \frac{Lm}{P(t)} \frac{Lm}{X(t)P_{total}} = \frac{T_{B_{total}}}{X(t)} \]  
(4)

Substituting \( T_{B(t)} \) by \( T_{B_{total}}/X(t) \), Eq. (3) is transformed into the equation of the dynamics of the fraction of bus riders by quarters of a year:

\[ X(t + 1) = F(X(t)) = g + (1 - g)\int_{\tau=0}^{\tau=\infty} f(\tau)d\tau \]  
(5)

Eq. (5) represents the dynamics of a bus line’s vicious cycle through the dynamics of the served fraction of the population of potential passengers.

4. Study of the vicious cycle

The dynamics of (5) are defined by the shape of \( F(X(t)) \) (Holmgren, 1996); let us investigate it as dependent on \( B_{total} \) and \( g \). \( F(X(t)) \) monotonously grows with the growth of \( X(t) \); \( F(X(t)) \) is an integral of a positive function and with the growth of \( X(t) \) the lower limit of the integral decreases.

Equilibriums: Equilibrium fraction of served passengers \( X^* \), if it exists, satisfies the equation \( X^* = F(X^*) \). Equilibrium \( X^* \) is locally stable if \( F'(X^*) < 1 \) and unstable if \( F'(X^*) > 1 \) (Holmgren, 1996). For monotonously growing \( F(X) \), \( F(X) \) is always positive and, thus, the conditions of local stability can be simplified: \( X^* \) is stable if \( F'(X^*) < 1 \), and unstable, if \( F'(X^*) > 1 \).

For a real-world non-monotonous distribution of the willingness-to-wait \( f(\tau) \), as in Fig. 1, the possible number of equilibriums of (5) is one or three. To investigate the stability of these equilibriums, let us note that they are the points of intersection of the curve \( X(t + 1) = F(X(t)) \) and of the straight line \( X(t + 1) = X(t) \). Stability of the equilibriums is thus defined by the tangent to \( F(X(t)) \) at the point of intersection.
In Fig. 3, two qualitatively different situations are represented regarding the equilibriums of (5) as dependent on the total number of potential passengers, in terms of buses \( B_{\text{total}} \) and \( g \).

The model dynamics for a high fraction of captive passengers (high \( g \)) is quite clear-cut. In case the total number of passengers is low, the equilibrium fraction of served passengers is only slightly above \( g \), and the vast majority of riders are captive passengers (Fig. 3a). This is because the bus frequency is not high enough to attract an essential fraction of non-captive passengers. Note that in case of high operating costs, no equilibrium may actually exist in this case, as the number of captive riders may be too low to warrant any public transport service. In contrast, a high number of potential passengers entails, for a high \( g \), an initially high number of captive riders that will make it worthwhile for the operator to establish a sufficiently high bus frequency. This will attract non-captive passengers and as a result cause the development of a virtuous cycle that will ultimately result in an equilibrium fraction of riders that is close to 100% (Fig. 3b).

The model dynamics in case of a low fraction of captive passengers (low \( g \)) depends on the total number of potential passengers in a more complex manner. When \( P_{\text{total}} \) is low, no matter how high the initial number of buses, the operator will follow a vicious cycle and decrease bus frequency until only captive passengers will remain on the bus (Fig. 3c). However, when the number of potential passengers is high, the vast majority of them are not captives, the model (5) has three equilibriums (Fig. 3d). A stable low equilibrium \( X_{\text{low}} \approx 0.21 \), when the frequency of buses is low and only captive passengers use the line, a stable high equilibrium \( X_{\text{high}} \approx 0.86 \), when the frequency of buses is high and the majority of potential passengers use the line, and an intermediate unstable equilibrium \( X_{\text{int}} \approx 0.44 \), separating between the domains of attraction towards the low and high equilibriums. If the initial number of buses is below \( X_{\text{int}} \), a typical vicious cycle starts; if it is above \( X_{\text{int}} \), the system will enter a virtuous cycle.

Fig. 3 presents the process of convergence of the fraction of the bus users for four specific sets of parameters, quarter by quarter, to the equilibriums. The full bifurcation diagram (Fig. 4a) represents the dependence of the equilibrium fraction of passengers, according to (5), on the size of the population of potential passengers (in terms of \( B_{\text{total}} \)) and the share \( g \) of captive riders for the entire interval of \( B_{\text{total}} \) and \( g \). As in Fig. 3, the diagram is constructed for the experimental distribution of willingness-to-wait of Peterson et al. (2006) presented in Fig. 1.

The dependence of the equilibrium fraction of riders on the increasing population size is clarified in Fig. 4b and c. If the majority of the population consists of captive riders (\( g \) is high), as in a “poor area” with a low level of car ownership, then the bus operator will react to an increase in \( B_{\text{total}} \) by increasing bus frequency over time, thereby also inducing more and more passengers to use the bus.
non-captive passengers to change their transportation mode in favor of the bus. Given \( g \), the higher the potential number of passengers, the closer is the equilibrium fraction of riders to 100%. This dynamics are reflected by the cross-section of the diagram for \( g = 0.8 \). For this value of \( g \) and a value of \( B_{total} = 15 \), the equilibrium value \( X_{X/8} \approx 0.92 \), while for a value of \( B_{total} = 35 \) the equilibrium fraction of riders \( X_{X/8} \approx 0.98 \) (Fig. 4b).

In case the majority of the population consists of non-captive passengers, the dynamics of the system are more complex. These dynamics are presented by the case of \( g = 0.2 \) (Fig. 4c). When the overall number of potential passengers is low, only one equilibrium state \( X_{\text{low}} \) is possible, in which virtually only captive passengers are served. With an increase in \( B_{total} \), the equilibrium bus frequency will slowly grow and, thus, the fraction of non-captive riders will slowly grow too. However, until the total number of potential riders is high enough, the growth of the equilibrium fraction of riders is, in itself, insufficient for attracting a substantial number of non-captive riders and initiating a virtuous cycle. The latter becomes possible after \( B_{total} \) (for the values of \( g = 0.2 \) and \( m = 50 \)) passes the lower threshold \( B_{\text{threshold}} \approx 22 \), and the number of potential riders passes \( mB_{\text{threshold}} \approx 1100 \), when two more equilibriums emerged: an unstable equilibrium \( X_{\text{int}} \) and a second stable equilibrium \( X_{\text{high}} \). For the values of \( B_{total} \) above \( B_{\text{threshold}} \), the public transport operator can force the line into the virtuous cycle by instantaneously raising the number of buses above the unstable equilibrium \( X_{\text{int}} \) (the dotted line in Fig. 4c). The situation changes once again when the total population size passes, in terms of number of buses, the higher threshold \( B_{\text{threshold}} \approx 34 \). That is, when the total population of riders exceeds \( mB_{\text{threshold}} \approx 1700 \), the low stable equilibrium \( X_{\text{low}} \) and unstable equilibrium \( X_{\text{int}} \) do not exist and the only equilibrium is \( X_{\text{high}} \). In this case, the line develops according to the virtuous cycle, and in its equilibrium serves almost all potential passengers (Fig. 4c).
Note that within the interval $B_{1}^{\text{threshold}} < B_{\text{total}} < B_{2}^{\text{threshold}}$, the higher is $B_{\text{total}}$, the lower is the number of additional buses that should be added by the transport operator to initiate a virtuous cycle, and the higher the value of the second equilibrium $X_{\text{high}}$ to which the system will converge after that (Fig. 4c). At the same time, if the increase in bus frequency is insufficient and the number of buses added by the operator would remain below $X_{\text{int}}$, the operator will be forced back into the vicious cycle that will return the system to the low equilibrium $X_{\text{low}}$.

The top view (Fig. 5) of the 3D diagram in Fig. 4a is also helpful for understanding the bus line dynamics as dependent on the overall number of potential riders as expressed by $B_{\text{total}}$ and the fraction $g$ of captive passengers. For bus capacity $m = 50$ passengers, the line has one equilibrium as long as the value of $B_{\text{total}}$ remains below $B_{\min} \approx 17.5$, i.e., as long as the total number of potential passengers remains below $mB_{\min} \approx 875$ passengers (left bottom domain in Fig. 5). The number of potential passengers in this case is insufficient to justify a high frequency of buses and, even if the initial bus frequency is high, it will decrease and, in an equilibrium state, the vast majority of the remaining users will be captive passengers who have no alternative means of transportation (Fig. 3c). The equilibrium is also unique if the potential number of passengers is high and the share $g$ of captive passengers is high, above $g_{\max} \approx 0.43$ (right top domain in Fig. 5). In this case, the number of captive passengers is high enough to cause a virtuous cycle irrespective of the initial frequency of buses. If the point in the $(B_{\text{total}}, g)$-space that represents the total number of potential passengers and the share of captive riders is between the corners 1 and 2, then the line has one equilibrium only, in which some fraction of non-captive passengers use the bus. The closer the point to corner 2, the higher is this fraction.

The dynamics of the bus line for the parameters within the “3 equilibriums” domain of the $(B_{\text{total}}, g)$-space is described above.

5. Extensions of the vicious cycle model

5.1. Bus size as a model parameter

One option for increasing bus frequency and, thus, initiating the virtuous cycle, is the use of smaller buses. Putting aside operation costs, the same total capacity of a bus line can be achieved by serving it with a higher number of smaller buses. In model terms, Eq. (4), where the size of a bus is reflected by the parameter $m$, is convenient for analyzing the impact of such an intervention. Just as above, if for a certain population size, the situation of one equilibrium (a line serving only captive passengers) is the only possible one (Fig. 6a), then substitution of the line’s large buses by more smaller buses, with smaller $m$, can turn the dynamics into a situation with two stable equilibriums (Fig. 6b) or even one (Fig. 6c).

The bifurcation diagram for the model (4) for $P_{\text{total}} = 1000$ and varying $m$ and $g$, is presented in Fig. 7.
5.2. One or two lines?

Let us consider establishing a bus service in a new area, assuming, for simplicity, that the spatial structure of the area enables a public transport service consisting of several circular lines only, in order to serve (parts of) the population in a

Fig. 6. Evolution of model equilibriums with a decrease in bus capacity \( m \) for \( P_{\text{total}} = 500 \) and \( g = 0.2 \); (a) \( m = 50 \), (b) \( m = 22 \), and (c) \( m = 15 \).

Fig. 7. (a) Full bifurcation diagram of Eq. (5) for \( P_{\text{total}} = 1000 \), varying \( m \) and \( g \), and distribution of willingness-to-wait according to Peterson et al. (2006); (b) cross-section of the full diagram for low fraction of captive passengers, \( g = 0.2 \); and (c) cross-section of the full diagram for high fraction of captive passengers, \( g = 0.8 \).
reasonable way. Then, the basic question for the public transport operator is whether it is better – i.e., more profitable – to operate only one line that serves half of the population or to establish two lines that will serve twice as much potential passengers but with only half of the number of buses per line?

Let us apply the model of the vicious cycle for the simplest case of an area that is characterized by uniform distributions of captive and non-captive passengers. Given a number of buses B and their size m, we compare one line of length L that serves \( gP_{\text{total}} \) captives and \( (1 - g)P_{\text{total}} \) non-captive passengers by \( B \) buses to two non-related lines of length \( L \), each serving \( P_{\text{total}} \) captive riders and \( (1 - g)P_{\text{total}} \) non-captive passengers with \( B/2 \) buses. Formally, the situation of two lines is identical to the situation of a line which length is \( 2L \) and which is served by the same number \( B \) of buses as the line of a length \( L \).

In the first case, the interval between buses will be \( L/B \), and the potential number of passengers \( P_{\text{total}} \). While in the case of two lines, the interval between buses will be \( 2L/B \), but with a potential number of passengers of \( 2P_{\text{total}} \). Applying (4), the dynamics of the fraction of passengers for each case, by quarters of a year, are given by:

\[
X_1(t + 1) = g + (1 - g) \frac{\int_{0}^{\infty} f(\tau) d\tau}{X_1 + P_{\text{total}}}
\]

\[
X_2(t + 1) = 2g + 2(1 - g) \frac{\int_{0}^{\infty} f(\tau) d\tau}{2X_2 + P_{\text{total}}}
\]

where \( X_1(t) \) and \( X_2(t) \) are the fractions of the bus riders in case of one and two lines, respectively, and \( B_{\text{total}} = P_{\text{total}}/m \).

As can be seen directly from (6) and (7) in case of a majority of captive passengers, i.e. \( g \sim 1 \), two lines are preferable, as \( 2gP_{\text{total}} \) passengers will be served, about twice more than in case of one line, no matter what are the other parameters of the system.

However, in case of a non-captive majority, the situation is different: until the frequency of buses \( 2L/B \) is sufficient to pass the threshold that is necessary for entering the virtuous cycle, it is worth to put all buses into one line and not to serve the second line at all. For the intermediate values of \( g \) the situation becomes more complex: depending on the population density along the line \( P_{\text{total}}/L \), the number of equilibriums for one line of length \( L \) can be different from the number of equilibria of the two-line situation. We delay the study of this case to future papers.

6. Conclusion and discussion

In this paper, we have presented an analytical model for analyzing the vicious and virtuous cycles in public transport based on passengers’ decisions to use a public transport service dependent on waiting time. While based on an analysis of a single line and a number of simplifying assumptions – passenger do not adjust their arrivals to bus stops to the time-table; uniform distribution of captive and non-captive passengers in space and over the day; costs of transporting one additional passenger is lower than the extra revenue generated from additional ticket sales – the model shows that service provision will either enter a vicious or virtuous cycle. The dynamics of the vicious cycle converges to an equilibrium that is characterized by low service frequency and low ridership, while the dynamics of the virtuous cycle converges to an equilibrium of high service frequency and high ridership. The system’s inherent parameters – total number of potential passengers \( P_{\text{total}} \), the share of non-captive passengers \( g \) and bus maximal capacity \( m \) – determine whether the system can converge to vicious or virtuous cycles. For every triple \( (P_{\text{total}}, g, m) \), the model makes it possible to determine if vicious, virtuous or both types of cycles are possible. In case both vicious and virtuous cycles are possible, the model provides the minimal frequency that is necessary to initiate a virtuous cycle. If adequate data are available on the variations in willingness-to-wait among a line’s potential passengers, these estimates can be exploited in reality, for estimating the minimal bus frequency that is necessary to induce a virtuous cycle.

We have analyzed the conditions of the vicious or virtuous cycle based on available field data regarding passengers’ variation in willingness-to-wait for a public transport service. The analysis substantially adds to the discussion by Levinson and Krizek (2008) who used hypothetical dependencies between the frequency of buses and the number of passengers. In future work, we aim to address the issues of non-uniform distribution of passengers over space, variation in trip lengths and reliable real-time travel information provided through, e.g. smart phones.

Despite intentionally simplifying assumptions, we argue that a number of practical lessons can be drawn from the analysis of our abstract model. Let us address two situations that typify a large part of the world: city regions in wealthy countries and city regions in rapidly developing but relatively poor countries.

Typically, city regions in wealthy countries like the US, Japan, or in Western-European countries, are characterized by low shares of captive riders. Total number of potential passengers in these city regions depends highly on the urban structure. In low density sprawling areas, public transport lines, if still existent, will inevitably enter a vicious cycle if subsidies are terminated (Fig. 3a). In denser areas, with a clustering of employment and other activities in a limited number of centers, the possible line dynamics are probably best described by Fig. 3c and both a vicious and a virtuous cycle may be possible:

- Frequencies on many public transport lines will currently be below the level that separates the low and high equilibriums and so are potentially in a vicious cycle. The level of service on these lines will currently be maintained by subsidies. However, as the share of captive riders drops over time due to increasing real incomes and related car ownership, these lines
will keep facing a decreasing ridership and thus subsidies will inevitably have to increase to maintain the level of service. This may not be possible due to budget limitations, forcing a decrease in service. In real-world circumstances with limited budgets, a vicious cycle therefore seems inevitable, even if the public transport service is fully regulated and heavily subsidized. Eventually, these lines will converge to a politically acceptable minimal level of service or will be terminated altogether.

- Some lines may actually be close to the level that separates the low and high equilibrium. For these lines, a temporary increase in subsidy combined with a requirement to increase the bus frequency may lead to a virtuous cycle. If subsidies were maintained long enough, the line, through the virtuous cycle, would converge to a high equilibrium. Note that subsidies could be reduced (substantially) at the moment the high equilibrium is reached, depending on operating costs and fare box revenues, without risking the development of a vicious cycle. This suggests that short-term increases in subsidies may actually lead to a better level-of-service and lower annual subsidies in the long run.

- Few lines may already be above the level that separates the low and high equilibrium, but are part of integrated subsidized tenders. In such cases, the public transport operator has no incentive to increase the level of service on these lines. It may be considered to take these lines out of the tenders and to allow operation on a commercial basis only, to trigger public transport operators to make better use of the potential offered by the virtuous cycle. Other lines may benefit from this split, both from the increase in per line subsidy and, through a network effect, from a total increase in ridership.

Recent societal changes, such as a decreasing car ownership among young adults and a leveling out of car mobility in a number of Western countries, suggest that high-density cities may actually experience an increase in the share of (voluntarily) captive passengers over the coming decade. This can be a trigger for a virtuous cycle on an increasing number of public transport lines. The experiences in Berlin, as reported by Reinhold and Kearney (2008), show that it is indeed possible to ‘engineer’ a virtuous cycle for particular bus lines by increasing service frequency. However, the Berlin bus system is still heavily subsidized and service frequency is determined top-down rather than in direct response to passenger demand. It thus remains unclear whether the virtuous cycle for the main bus lines there is due to sustained subsidies or whether these lines could maintain a high equilibrium without government support.

Overall, our analysis suggests that in wealthy countries it might be attractive, both in terms of ridership and total subsidy needs, to change from area-based tenders of public transport service (Farsi et al., 2007) to a system that makes a distinction between bus lines vis-à-vis the vicious/virtuous cycle, in spite of the possible drawbacks of such a system.

Urban public transport services in the emerging economies are often in a completely different situation. They are usually characterized by a large potential passenger population and a large share of captive riders. However, several developments suggest that these cities may rapidly move into a situation in which a vicious cycle could occur. First, these cities are experiencing a rapid increase in car ownership. Second, operating costs and therefore ticket prices are likely to increase over the coming decade, due to higher demand for quality of vehicles/services and increasing labor costs. Taken together, these developments suggest a decreasing share of captive riders and a lower use of public transport among non-captive riders, which may in turn induce operators to reduce frequencies, with a risk of entering the vicious circle. The challenge for these cities will be to identify impending vicious cycles and prevent them from happening through ‘early-bird’ subsidies and/or strategies to reduce operating costs and increase service frequencies, such as free bus lanes and traffic light priority, in order to return to a path of a virtuous cycle. If successful, this may also reduce the rate with which the captive population will turn into a non-captive one.

A final observation regards the importance of the use of models for studying the possible development of the transport system over time. Currently, virtually all transportation-planning agencies, whether dealing with public and/or private transport, use static and aggregate models to analyze the dynamics of real-world transport systems. The models are static in nature, in that they assume that future travel demand can be forecasted largely independently of future transport infrastructure or transport policies. The models are aggregate in that they do not simulate the behavior of the individual traveler. As such, these models are insufficient for simulating the interrelationships between the users of the transport system, the impacts of policies over time, and the intricate relation between short- and long-term dynamics of the transport system. Static aggregate transport models tend to generate results that satisfy predicted demand, but cannot identify possibilities to change the system through measures which impact becomes only apparent over time. Our study underscores that high-resolution dynamic models may substantially enrich transportation modeling and, in its wake, transportation policies.

Appendix A.

If $\tau < T_B$ then

$$N_u(\tau, d + 1) = N_u(\tau, d) + \alpha N_u(\tau, d) \frac{\tau}{T_B} - \beta N_u(\tau, d) \left(1 - \frac{\tau}{T_B}\right)$$

$$N_a(\tau, d + 1) = N_a(\tau, d) - \alpha N_a(\tau, d) \frac{\tau}{T_B} + \beta N_a(\tau, d) \left(1 - \frac{\tau}{T_B}\right)$$

(A1)
Let:
\[
\alpha \frac{\tau}{T_B} = \epsilon; \beta \left(1 - \frac{\tau}{T_B}\right) = \delta; N_a(\tau, d) = U_t; \text{ and } N_n(\tau, d) = V_t
\]

Eq. (A1) can be thus presented as:
\[
\begin{align*}
U_{t+1} &= U_t(1 - \delta) + \epsilon Y_t \\
V_{t+1} &= \delta U_t + V_t(1 - \epsilon)
\end{align*}
\]  
(A2)

And in a matrix form as:
\[
\begin{pmatrix}
U_{t+1} \\
V_{t+1}
\end{pmatrix} =
\begin{pmatrix}
1 - \delta & \epsilon \\
\delta & 1 - \epsilon
\end{pmatrix}
\begin{pmatrix}
U_t \\
V_t
\end{pmatrix}
\]

To find the eigenvalues of matrix \(\begin{pmatrix} 1 - \delta & \epsilon \\ \delta & 1 - \epsilon \end{pmatrix}\) we have to solve the equation:
\[
\lambda^2 - \lambda(2 - \delta - \epsilon) + (1 - \delta - \epsilon) = 0
\]

Let us denote \((1 - \delta - \epsilon)\) as \(a\); the roots of (3) are, thus
\[
\lambda_{1,2} = \frac{1 + a \pm \sqrt{(1 + a)^2 - 4a}}{2} = \frac{1 + a \pm (1 - a)}{2}
\]

That is
\[
\lambda_1 = 1
\]

and
\[
\lambda_2 = a = 1 - \delta - \epsilon = 1 - \beta + \frac{\tau}{T_B} (\beta - x) < 1
\]

The solution of Eq. (A2) thus converges, in days, to the eigenvector corresponding to the largest eigenvalue \(\lambda_1 = 1\). The components of this eigenvector are as follows:
\[
U^* = \frac{N_0 \epsilon}{\epsilon + \delta}; \quad V^* = \frac{N_0 \delta}{\epsilon + \delta}
\]

That is, for given \(\tau < T_B\), the ratio between the number of potential passengers who use the bus and the number of potential passengers who do not use the bus, converges, in days, to
\[
\frac{N_n(\tau)}{N_{n0}(\tau)} = \frac{U^*(\tau)}{V^*(\tau)} = \frac{\epsilon}{\delta} = \frac{\alpha}{\beta} \frac{T_B - \tau}{T_B}
\]

For \(\tau < T_B\), the equilibrium fraction of the non-captive bus users that are characterized by the willingness-to-wait \(\tau\) is, thus, equal to:
\[
N_n(\tau) = f(\tau) \frac{1}{1 + \frac{\beta}{\alpha} \frac{T_B - \tau}{T_B}}
\]  
(A3)

Fig. 2 in the main text shows the effect of \(\alpha\) and \(\beta\) on the number of users at the equilibrium.

The number of non-captive bus users at \(d = 90\) that are characterized by the willingness-to-wait \(\tau < T_B\) is thus given by:
\[
P(t + 1)_{\tau < T_B} = (1 - g)P_{total} \int_0^{T_B} f(\tau) \frac{1}{1 + \frac{\beta}{\alpha} \frac{T_B - \tau}{T_B}} d\tau
\]

And depends on the analytical expression of \(f(\tau)\).

For example, in case \(f(\tau)\) is uniform on \([0, T_{max}]\)
\[
P(t + 1)_{\tau < T_B} = (1 - g)P_{total} \int_0^{T_B} \frac{C}{1 + \frac{\beta}{\alpha} \frac{T_B - \tau}{T_B}} d\tau
\]

where \(C = 1/T_{max}\) and, analytically,
\[
P(t + 1)_{\tau < T_B} = C(1 - g)P_{total} T_B \left(\frac{\beta}{\alpha}\right).
\]

where \(f(\alpha) = \frac{\alpha^{\alpha \beta - 1} \beta^{\alpha - 1}}{\Gamma(\alpha) \Gamma(\beta)}\)

According to (A3), the full model of a bus line is, thus:
\[
P(t + 1) = gP_{total} + (1 - g)P_{total} \left[\int_0^{T_B(t)} \frac{f(\tau)}{1 + \frac{\beta}{\alpha} \frac{T_B(t) - \tau}{T_B(t)}} d\tau + \int_{T_B(t)}^{\infty} f(\tau) d\tau\right]
\]

(A4)
Or, substituting $P(t)$ by $X(t)P_{total}$ and applying relations (4) in the main text:

$$X(t+1) = g + (1-g) \left[ \int_{0}^{\frac{1}{2}} \frac{f(\tau)}{1 + \frac{g}{\tau} \int_{\frac{1}{2}}^{x} X(t) d\tau} d\tau + \int_{\frac{1}{2}}^{\infty} f(\tau) d\tau \right]$$

Eq. (A5) differs from the model (5) in the main text by the additional (first) term in the square brackets. This addition, however, does not change in any qualitative way the dynamics of the system. Fig. A1 repeats Fig. 3c in the main text for high $B_{total}$, low $g$, $\alpha = 0.005$, and three different values of $\beta$. The black curve represents the same dependency for the simplified model (5) that is investigated in the paper (Fig. 3c).

**Appendix B. Analysis of Levinson and Krizek’s (2008) model**

The illustrative model of the vicious cycle presented by Levinson and Krizek (2008) is based on three heuristic equations that relate between bus speed $v$, number of passengers $P$, and bus waiting time $W$. They do not specify in their analysis which of the equations is chosen for performing the time-step transition and we have arbitrarily chosen their second equation, as presented in (B2) below:

$$v_t = 1.5 - 0.5W_t$$

$$P_{t+1} = \frac{e^{v_t}}{3000 + e^{v_t}}$$

$$W_t = 30 \left( \frac{1}{0.5 + 0.02P_t} \right)$$

Fig. B1. Equilibriums of Levinson and Krizek’s (2008) model: (a) dependence of $P(t+1)$ on $P(t)$ for low values of $P(t)$ and (b) dependence of $P(t+1)$ on $P(t)$ for the entire range of $P(t)$.
To analyze the system we express \( P_{t+1} \) as a function of \( P_t \), substituting (B1) into (B2):

\[
P_{t+1} = 3000 \frac{e^{1.5 - 0.5W_t}}{1 + e^{1.5 - 0.5W_t}}
\]

and further, substituting (B3) into the result:

\[
P_{t+1} = f(P_t) = 3000 \frac{e^{1.5 - \frac{250}{W_t}}}{1 + e^{1.5 - \frac{250}{W_t}}}
\]

Dependency (B4) is presented in Fig. B1b, while Fig. B1a presents a zoom of the coordinate plane for the values of \( P_t \) and \( P_{t+1} \) between 0 and 500. As can be seen, the equations of Levinson and Krizek (2008) produce, for the chosen values of parameters, three equilibriums, but result in non-realistic values of \( P_{\text{low}} \approx 0 \), \( P_{\text{int}} \approx 150 \) and \( P_{\text{high}} \approx 1,200,000 \).

References


