Technische Universität Berlin **CLIMATECON**

Climatecon Working Paper Series

No. 3-2012

Transport costs, urban form and optimal public transit

Felix Creutzig

Feb 28, 2012

Department of Climate Change Economics – Technische Universität Berlin
Room EB 238-240 (EB 4-1) - Straße des 17. Juni 145, 10623 Berlin
www.climatecon.tu-berlin.de

Transport costs, urban form and optimal public transit

Felix Creutzig
Department of Economics of Climate Change
Technical University Berlin, Germany

Abstract

Urban form and transportation infrastructure mutually influence each other. For example, dense Hong Kong boasts a viable and efficient public transit network, whereas many sprawled US cities are best served with automobiles. Here we present a simple model of a monocentric city with public transit and automobiles that explains modal share as a function of urban form, infrastructure investment and marginal transport costs. The contribution to the literature is two-fold. First, adding to urban economic theory, we derive two conditions of optimal public transport infrastructure provision. We also identify a market failure: A private mass transport provider under-invests into public transport infrastructure. Second, adding to the ongoing discussion on urban transport and energy use, we argue that this two-modal model is a useful explanatory framework to investigate empirical observations on urban form, transport energy use and modal share.[1]

1 Introduction

The global discourse on cities and mobility shifted focus in the recent decade. The second half of the last century was dominated by the idea of the "American way of life", imagining unlimited freedom of private motorized mobility. While the implementation of this dream remained illusionary in many places worldwide, its mindset has dominated the discourse among politicians, transport planers and consumers alike. With climate change, the plausible end of

[1]A preliminary version of this paper was presented at the North American Regional Science Conference in Miami, November 2011

cheap oil, the realized limitation of automobility in congested urban areas, and with shifts in life-style the scientific but also public discourse on urban mobility began to shift to questions around sustainability and accessibility, and specifically to urban form and modal choice [Glaeser and Kahn, 2010; Weisz and Steinberger, 2010].

A widely cited paper by Newman and Kenworthy (1989), from hereon NK89, marked the beginning of a vivid discussion on the role of urban form in transport energy consumption. In this international comparison of cities, urban population density is inversely correlated with urban transport energy use. Evidence [Weisz and Steinberger, 2010; Ewing and Cervero, 2010] and modeling [Ewing, 2007] point to considerable mitigation potential in densification. However, the causal relationship between urban form and transport energy use is complicated. Some authors remained skeptical about the value of pursuing compactification [Gordon and Richardson, 1997]. In some US-American studies population density and job density seem to have a surprisingly little effect on vehicle miles travelled (VMT) once controlled for accessibility of destinations and street network design [Ewing and Cervero, 2010]. Mindali et al. (2004) took a closer look at the Newman-Kenworthy data and conclude that the specifics matter: Inner area employment density and concentrated mixed-use design, together with public transit, are more relevant than population density. In addition, studies disagree on the role of residential location self-selection by inhabitants in explaining parts of the correlation between urban form and transport energy use [Cao et al., 2009; Ewing and Cervero, 2010].

Another observation is that public transit activity is negatively correlated with private transport activity [Newman and Kenworthy, 1996]. Public transit is often seen as part of reasonable densification strategies. In particular, transit-oriented development requires sufficient density around a transit corridor to support a rail-based system or high-capacity busses [Cervero, 1998; Bongardt et al., 2010]. A minimum density is seen as a prerequisite for financially viable and environmentally effective public transport [Frank and Pivo, 1994].

While causalities remain somewhat unclear in this literature, there is a disparate discipline - urban economics - which has mostly developed around models explaining the interaction between transport and urban form. The model framework dates back to von Thünen in the early 19th century, who tried to explain the relationship between agricultural product choice, land rent, and transport distance to the central market place [von Thünen, 1826]. It was Alonso, and later Muth and Mills, who transferred this framework to residential location choice, and commuting costs to the central business district (CBD) [Alonso, 1964; Fujita, 1989]. In this model, increasing commuting costs are compensated by decreased land rent, resulting in decreasing population density towards the urban fringe. In these models, lower marginal transport costs imply more urban sprawl.

Mills suggested to expand his model to additional transport modes, a recommendation taken up by some of his peers. Capozza (1973) emphasized the tradeoff between capital and land as scarce input factors for subway versus road infrastructure: in the inner city, land becomes so scarce and valuable that land-intensive road infrastructure is substituted by capital-intensive subway lines. Haring et al. (1976) demonstrate that an additional transport mode reduces the land rent differential between CBD and urban fringe. Anas and Moses not only specify two modes but also explore the role of discrete transport corridors, producing a number of varying urban forms as a function of generalized costs in dense and sparse radial transport networks [Anas and Moses, 1979]. Also in 1979, Kim, a PhD student of Mills, published a two-mode model in which a city with two million inhabitants has sufficient population density to support a subway system, in contrast to a city with only one million inhabitants [Kim, 1979]. Recent articles have made progress in addressing optimal mode choice in presence of income heterogenities and externalities [Borck and Wrede, 2008; Proost and Dender, 2008]. Sasaki [1989] studies the impact of a change in fixed and variable costs on rent levels, welfare and spatial structures in an urban system with two modes. Su and DeSalvo [2008] confirms empirically that urban area contracts with public transit subsidies but expands with an auto subsidy.

This world of models from urban economics remains surprisingly unconnected to the empirical data and correlation studies initiated by NK89. In this paper we want to contribute to initiating engaged communication between these two communities. We do so by building on the literature on urban form and transportation costs with two transport modes. We maintain that the analytical framework of urban economics can provide additional intuition on relevant causal relationships related to urban form and transport energy use, especially if complemented with numerical simulations. In section 2, we introduce a new model, in mode specification comparable to [Sasaki, 1989], but including an additional infrastructure component, in which public transit is fully endogenous to urban form and generalized costs of private transport. In this model, generalized transport costs, urban form and mode choice are nonlinearly related to each other. In section 3, we derive and interpret first-order conditions for public infrastructure provision. We also demonstrate that free-market provision of public transit is socially inefficient. Finally, in section 4 we show that this model can reproduce some relevant results from NK89/96, in particular the relationships between population density and transport energy use, and between transport distance of different modes.

2 A bi-modal city

This section introduces a density and modal share modeling framework - based on the Alonso-Mills-Muth model of a monocentric city with transport costs and housing market.

Definition of important variables

F_c:	Fixed costs of car ownership
m_c:	Marginal costs of car driving (m_c). Includes fuel and insurance costs, as well as time costs, and is, hence, also dependent on the quality and scope of road infrastructure. In DEMOS, m_c is the key variable driving urban form and modal shares
m_p:	Marginal price of public transit charged to patrons per unit distance
m_{infra}:	Marginal costs of public transit per unit distance due to infrastructure provision. In the social planner equilibrium $m_{infra} \equiv m_p$; in the market solution, $m_{infra} \leq m_p$
C:	Infrastructure cost of public transit per unit area
z:	Composite consumer good
s:	Lot size of residence
Y:	Aggregate income
$T(r)$:	Transport costs of commuter living at r
r_p:	Radius of the city area served by public transit
r_c:	Radius of the total city area as served by cars
$R(r)$:	Unit rent costs of commuter living at r
R_a:	Agricultural land rent at r_c
$\rho(r)$:	Population density at r
TT:	Average total transport costs of commuters
RT:	Average rents costs of commuters

2.1 Household location in a monocentric city

We introduce the well-known model framework - a household location theory - following Alonso (1964) [Alonso, 1964] who generalizes the agricultural bid rent theory of von Thünen [von Thünen, 1826]. We specify a closed-city model (the population is constant) and absentee land ownership (in contrast to public ownership of the land) - the so-called CCA model [Fujita, 1989]. The city is characterized as monocentric with a dense radial transport system without congestion. All travel consists of commuters who travel from their residences to the city center where their work is located. The land is featureless, and public goods and externalities are absent. This well-established model is characterized by two agents, the households and absentee landowners.

Economic agent 1: The Households. Households maximize utility $U(z, s)$ where z is a composite consumer good, and s is the lot size of the house. Furthermore, rents are denoted as $R(r)$, and transport costs as $T(r)$, assumed to be monotonically increasing in r. The optimization problem of each household is given by

$$\max_{r,z,s} U(z, s) \; , \qquad \textit{with} \qquad z = Y - T(r) - sR(r) \; .$$

For identical individuals with identical income and utility function, households locate in equilibrium such that utility is equal, i.e. $U(z, s) = \bar{u}$ in equilibrium. If otherwise, some households could improve their utility by moving.

Economic agent 2: The Absentee Landowners. Land belongs to the absentee landowners. They maximize land rent $R(r)$ at any given distance. No housing construction market is specified here. Hence, land rent $R(r)$ is only a function of land demand and not of housing stock. In this model, the market land rent always coincides with the bid rent in equilibrium [Fujita, 1989].

For given utlity $U(z, s) \equiv u$, the log-linear transform of a Cobb-Douglas utility function is given as

$$U(z, s) = \alpha \log z + \beta \log s$$

with $\alpha + \beta = 1$, and $\alpha > 0$, and $\beta > 0$. Under these conditions, it can be shown that rent costs and lot size are given as [Fujita, 1989]:

$$\begin{aligned} R(r, u) = & \quad \alpha^{\alpha/\beta} \beta (Y - T(r))^{1/\beta} e^{-u/\beta} \; , \\ s(r, u) = & \quad \alpha^{-\alpha/\beta} (Y - T(r))^{-\alpha/\beta} e^{u/\beta} \; . \end{aligned} \tag{1}$$

The density profile, or urban form, is then given as $\rho(r, u) = \frac{1}{s(r,u)}$. Furthermore, in a radial symmetric city the city area is specified as a function of distance to the city center as $Ar(r) = 2\pi r dr$. The city boundary is denoted as r_c, corresponding to the outer radius of the car transportation mode (see below).

In this model, utility $U(z,s)$ changes with different transport costs and density profile. In contrast, total population N is kept constant. According to the classification of, e.g., Fujita [Fujita, 1989], fixed population and variable utility corresponds to a closed city model. This allows to solve utility as a function of population and transport costs.

$$\begin{aligned} N &= \int_0^{r_c} Ar(r)\rho(r,u)dr \\ &= \int_0^{r_c} \frac{2\pi r}{s(r,u)} dr \ . \end{aligned} \tag{2}$$

Substituting equation 1 into equation 2 and solving for u, we obtain utility directly as a function of transport costs:

$$u(r,T(r)) = -\beta \ln N[2\pi\alpha^{-\alpha/\beta} \int_0^{r_c} (Y - T(r))^{-\alpha/\beta} 2\pi r dr]^{-1} \ . \tag{3}$$

We can normalize with respect to the population, i.e., without loss of generality, $N = 1$.

Optimal land use. Is the closed-city model with absentee landowner socially optimal? If one uses a Benthamite welfare function and adds individual utilities, the model is generally not optimal. In fact, the discretionary nature of land allocation and asymmetry caused by transport costs produces asymmetrical production possibilities: total welfare can more easily be enhanced by giving additional land to those living farther from the city. As a result, inequality is socially optimal in this model. This observation has been first characterized by Mirrlees (1972), and been analytically treated by Arnott and Riley (1977). As a correlate, in the absence of externalities (summarized, e.g., in [Brueckner, 2000]) and with Cobb-Douglas utility function, social welfare is optimized by subsidizing marginal transport costs [Gusdorf and Hallegatte, 2007]. In a Herbert-Stevens model the social surplus for given identical utility level is maximized. Under such conditions, it is always possible to find an income tax such that the market equilibrium is also efficient [Fujita, 1989].

In the following, we assume the equilibrium land use market of the CCA model as given and investigate social welfare exclusively from optimal public transport provision.

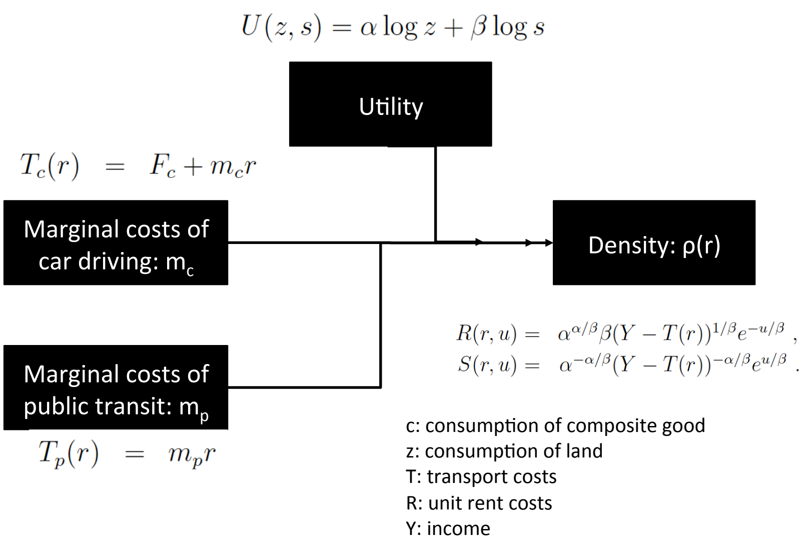

Figure 1: **Structure of a common transport land-rent model with two modes.**

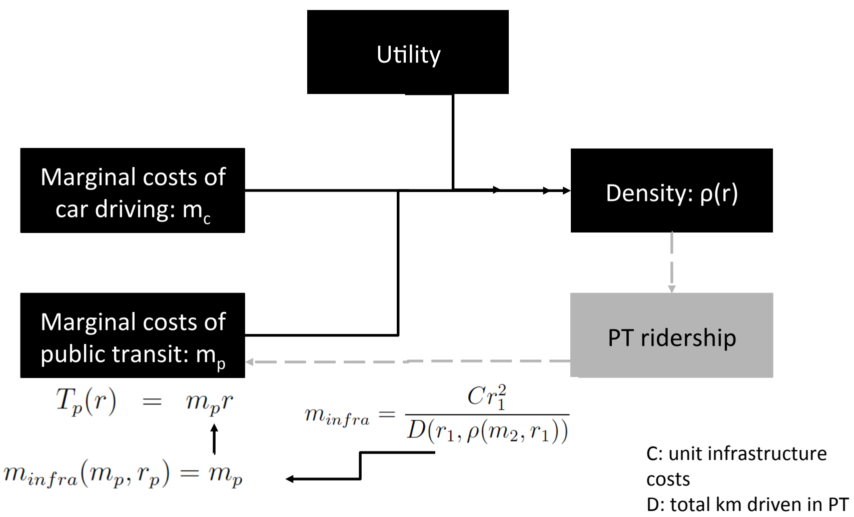

Figure 2: **In this crucial model extension, public transit price is endogenously determined by urban form and ridership.**

2.2 Endogenized costs of public transport

So far, transport $T(r)$ is a function assumed to be increasing in r. We now specify transport costs and introduce two modes, public transport and cars. Each mode is characterized by marginal user costs and fixed capital costs. In this model, public transport (denoted with the subscript p) is marginally more expensive than car transport (denoted with the subscript c) i.e $m_p > m_c > 0$. Clearly, marginal costs are understood here not only as monetary travel prices (usually higher for car transport) but also include time and convenience costs (usually higher for public transport). Public transport is assumed to have no capital costs for users, whereas car users need to invest into vehicles, $F_c > F_p = 0$.

Transport costs $T_i(r)$ are given as follows:

$$T_p(r) = m_p r \tag{4}$$

$$T_c(r) = F_c + m_c r \tag{5}$$

with $m_p > m_c$. In an inner circle $r < r_p$, public transport is more economic than car transport. More precisely, mode choice is given as follows:

$$\begin{aligned} i = p: &\quad 0 < r < r_p \\ i = c: &\quad r_p < r < r_c \; . \end{aligned}$$

Here, r_p denotes the outer radius of the public transport area and r_c denotes the outer radius of car usage, i.e. $R(r_c) = R_a$.

As a central property of our model, we aim to endogenize public transit as a function of the population density profile and m_c (see Figure 2). More precisely, marginal costs of public transit depend on ridership. Ridership itself depends on the proportion of the population living close to the city center for whom public transit is more attractive then car driving, i.e. on urban form - which itself is a function of the marginal costs of both modes of transportation.

To operationalize this interdependence, consider that public transport relies on infrastructure (e.g. a bus or subway system), represented by area unit infrastructure costs C. Infrastructures are a public investment that are here recovered via marginal pricing of public transport which is a function of urban form. Varying marginal costs of car driving (m_c, e.g., fuel pricing including fuel taxes) influence urban form and, hence, the viability of public transport. Overall, the marginal operation costs per passenger and per distance, are given by total service provision per overall ridership.

Let us start with the second component, overall ridership. To calculate this, we need to determine some other variables. The total proportion of

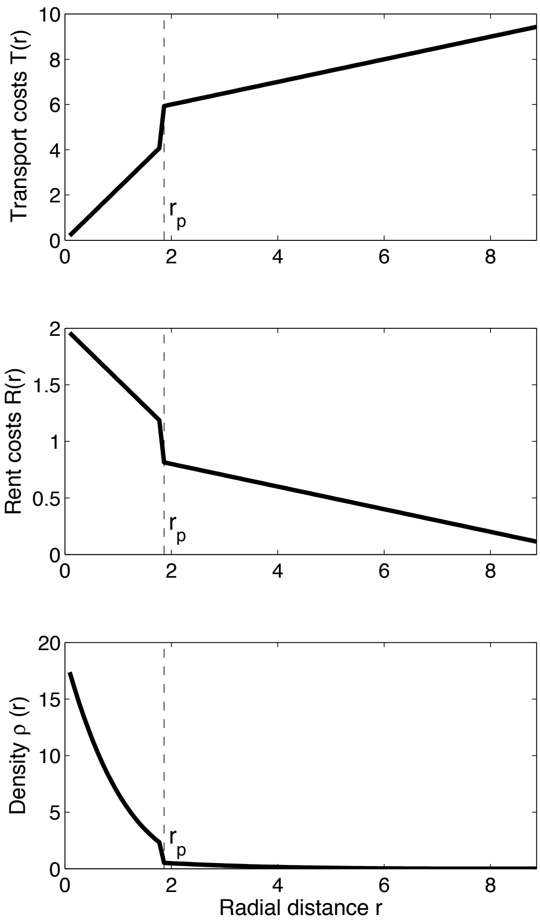

Figure 3: **Characteristics of urban profile as a function of radial distance. A** Transport costs; **B** Rent per unit area; **C** Population density (number of households per unit area). r_p denotes the outer radius of the public transport area. The following parameters were used: $a = 0.8$, $b = 1-a$, $Y = 10$, $C = 1$, $F_c = 5$, $R_a = 0.02$, $m_c = 0.7$.

people living at distance r is given by

$$n(r) = \frac{1}{N} 2\pi r \rho(r) \ .$$

The population density is normalized with the total population N, given in equation 2. For modeling purposes, and without loss of generality, $N = 1$. The total distance traveled is then

$$\begin{aligned} D(r) &= \int_0^r n(r) r dr = \int_0^r \frac{1}{N} 2\pi r^2 \rho(r) dr = \int_0^r \frac{1}{N} \frac{2\pi r^2}{S(r)} dr \\ &= \int_0^r \frac{1}{N} 2\pi r^2 \alpha^{\alpha/\beta} \beta^{-1} (Y - T(r))^{\alpha/\beta} e^{-u/\beta} dr \ . \end{aligned}$$

Now, we can turn our attention to the first term, infrastructure costs. The unit area costs of public transport infrastructure are denoted by C. The total area covered by public transport is πr_p^2. Hence, the total provision costs of public transport within r_p are denoted by $C\pi r_1^2$. Then the marginal costs from supply side are

$$m_{infra} = \frac{C\pi r_p^2}{D(r_p, \rho(m_c, r_p))} \ .$$

The public transport infrastructure term essentially makes the marginal costs of public transport a function of the other variables, in particular of the marginal costs of car transport m_c via urban form $\rho(r)$. In equilibrium then the supply side costs of public transport equal demand side costs, $m_{infra} = m_p$. In Appendix A, we prove that this equality is true for exactly one value of r_p. The recursive and non-linear dependence of m_{infra} and m_p prohibit an analytical characterization of the equilibrium. A numerical treatment is presented in Chapter 3.

For completeness, let us also define the average transport costs TT and average rent costs RT of a city with radius r_c:

$$\begin{aligned} TT &= \int_0^{r_c} n(r) T(r) dr \ , \\ RT &= \int_0^{r_c} n(r) R(r) dr \ . \end{aligned}$$

2.3 Optimal provision of public transport

We are now in position to determine the equations of optimal public transit provision. The social planner maximizes utility $u(m_p, r_p)$ over (m_p, r_p) as given in equation 3 for given m_c under the side constraint $m_{infra}(m_p, r_p) = m_p$:

$$L(m_c) = \max_{m_p, r_p} u(m_p, r_p) + \lambda(m_{infra} - m_p)$$

.

Hence, the first order conditions (FOC) are:

$$\frac{du}{dm_p} = \lambda(1 - \frac{dm_{infra}}{dm_p}) \ , \tag{6}$$

$$\frac{du}{dr_p} = -\lambda \frac{dm_{infra}}{dr_p} \ . \tag{7}$$

Together with the side constraint, these equations characterize optimal public transport provision.

The first order conditions can be interpreted as follows. The first FOC (equation 6) reveals that utility increases with m_p if m_{infra} grows sublinearly with m_p ($\frac{dm_{infra}}{dm_p} < 1$). In other words, if the marginal costs of adding public transport infrastructure are lower than the marginal willingness-to-pay of residents, additional infrastructure provision is welfare enhancing. Utility decreases with m_p if m_{infra} grows super-linearly with m_p ($\frac{dm_{infra}}{dm_p} > 1$).[2]

To interpret the second FOC (equation 7) , consider

$$\frac{dm_{infra}}{dr_p} = \frac{d(\frac{C\pi r_p^2}{D(r_p)})}{dr_p} = \frac{C\pi r(2D(r_p) - rD'(r_p))}{D^2(r_p)} \equiv \kappa(r_p)(2D(r_p) - rD'(r_p)) \ .$$

The first term, $\kappa(r_p)2D(r_p)$, corresponds to the marginal increase in m_{infra} due to increased area coverage. The second term, $-\kappa(r_p)rD'(r_p)$, corresponds to the marginal decrease in m_{infra} due to increased total distance traveled - inhabitants at the edge of the public transport area travel longer distances. The relative magnitude of these two effects decides on the sign of $\frac{dm_{infra}}{dr_p}$. If the costs of additional area provision exceed the the cost reduction effect due to total distance traveled, $\frac{dm_{infra}}{dr_p} > 0$ and the marginal

[2]For fixed r_p, $\frac{dm_{infra}}{dm_p} > 1$ because the infrastructure cost remain constant while the total transport distance traveled in public transit is reduced.

utility with public transport area expansion is negative. Note that for the interpretation of the two FOCs, we implicitly assumed $\lambda > 0$. This can be seen by looking at equation 7. If m_{infra} and hence total transport costs increase with distance, utility must decrease, whereas if m_{infra} decreases with distance, utility must increase. While the derivates can be stated explicitly, the equations cannot analytically be solved for m_p and r_p. Hence, we will retreat to numerical investigation to characterize optimal public transit.

3 Results

The model focusses on the the relationship between transport costs, density and urban form. In the following we present and visualize the model results.

3.1 Urban form displays a kink where car transport substitutes for public transport

Optimal public transport provision can be numerically solved. The resulting city profile can be characterized by transport and rent costs, and density as a function of radial distance to the city center (Figure 3). The transport cost increase with radial distance (top panel). At the private transport radius r_p, the transport cost curve is non-monotonous and displays an step increase. For $r > r_p$, the marginal increase of transport costs is lower than for $r < r_p$, as car transport has lower marginal costs than public transport $m_c < m_p$. Further outward unit rent costs decrease (medium panel). However, as lot sizes become larger, total rent per household increases. Households further outwards derive most of their utility from living amenities, i.e. lot size S, whereas households further inwards have higher proportion of income that can be used for consumption. As an inverse of the plot size at distance r, the population density $\rho(r)$ decreases monotonously with radial distance, and also displays a non-linearity at r_p. As we have now established the model characteristics and understand the dependence of urban and public transport infrastructure on m_c in particular instances, we are now in a position to analyze the impact of varying m_c.

3.2 m_c determine modal regimes

In the model, m_c determines which mode dominates the transport system. There are three different regimes. For low m_c, car transport is the exclusive mode, for intermediate prices, car and public transport coexist, for high m_c,

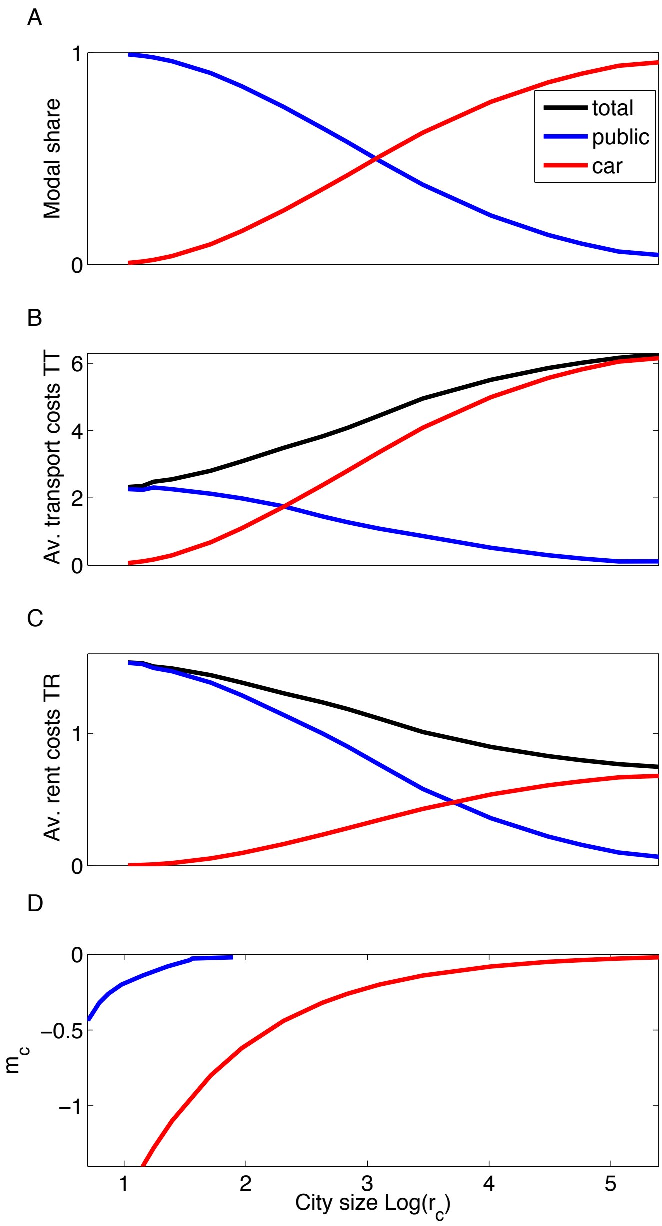

Figure 4: **Two modal regimes as a function of city size, the radius of the city.** Note that city size itself is determined by m_c. The figure should be read from bottom to top, starting with m_c in panel D. **A** Modal share. **B** Transport costs. **C** Rent costs. **D** City size determined by m_c. The following parameters were used: $a = 0.8$, $b = 1 - a$, $Y = 10$, $C = 1$, $F_c = 5$, $R_a = 0.02$.

public transport is the exclusive mode (Figure 4). In the following paragraphs we discuss different parameters.

In Figure 4D the city size r_c is the dependent variable. For high m_c, the city is small in extension. The city grows rapidly in spatial scale with lower m_c. With constant population, the spatial city size is influencing the land rent costs and the average transport costs, as well as the modal share. The city size is an intermediate variable, and the m_c is the causal driver. We discuss the impact of the intermediate variable, city size and urban form, on transport, land rent and modal share. Rent costs are depicted in Figure 4C. In the low m_c regime, the city is large, and rent costs are low because inhabitants can find affordable land further outwards. In contrast, with higher m_c the city size decreases and the rent costs increase. Transport costs are depicted in Figure 4B. Transport costs comprise all marginal costs, including the fixed costs of cars. In the high m_c regime, the transport costs T_1 are caused by public transport use. In the low m_c regime (large city size), all transport costs are caused by car use. In the intermediate regime. In the high m_c regime (small city size), total transport costs are lower than in the low m_c regime. This is because the reduction in total distance traveled compensates for m_c increase, and because public transport has no capital cost component for users, i.e. $F_p = 0$. The modal share is depicted in Figure 4A. For low m_c, and large city size, the transport system relies solely on cars. For intermediate m_c, cars and public transit coexist while for high m_c, commute relies on public transport exclusively.

3.3 Spatial structure induces market failure

Consider a firm providing public transport infrastructure under profit maximization. The firm will chose r_p such that

$$P(r_p) = D(r_p)(m_p(r_p) - m_{infra}(r_p)) \tag{8}$$

is maximized. Equation 8 introduces the market solution. In this market solution, m_p and m_{infra} are not equal for non trivial m_c, i.e. $m_c > 0$. Hence, the market solution is not optimal. Conceptually, this is founded in the asymmetrical production possibility of space. A inner city-dweller is prepared to pay mark-up costs of public transport, essentially as long as total transport costs remain below the fixed costs of car transport.

Let us more closely characterize how the divergence from the social optimum depends on fuel costs. The numerical solution is depicted in Figure 5. Only for high m_c, the average difference in utility becomes relevant. Note that a number of externalities, such as greenhouse gas emissions, congestion, air pollution and imperfect land markets further confound the picture.

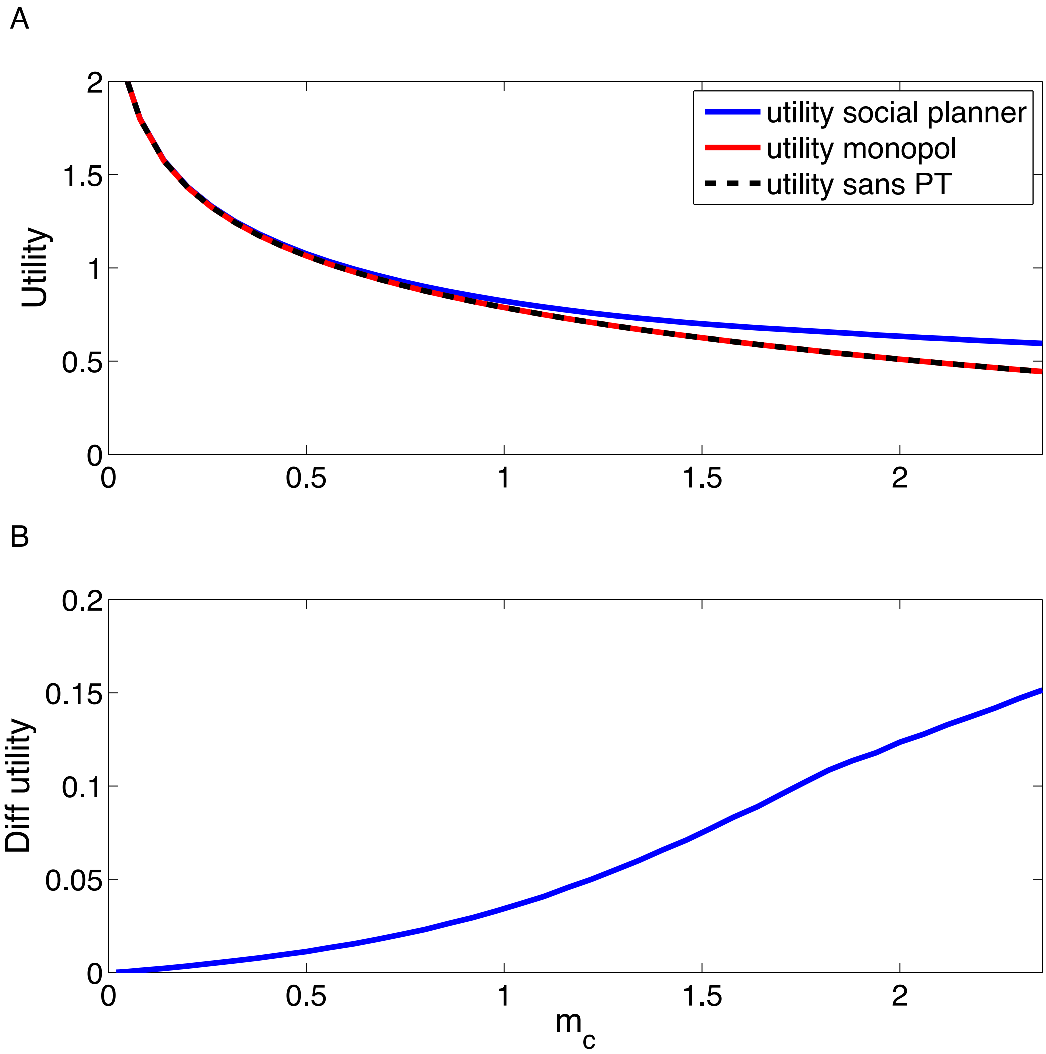

Figure 5: **The market solution reduces average utility level in a high m_c setting.** **A** Utility decreases with m_c as a higher share of income needs to be invested for commuting and smaller plots are used for housing. **B** The difference in utility between social planner infrastructure provision and market infrastructure provision increases with fuel costs. Parameters are as in Figure 4.

4 Urban transport in global cities

Newman and Kenworthy (1989,1996) demonstrated a relationship between urban density and transport energy use on a set of global cities (Fig. 6A, reproduced with data from Kenworthy and Laube [2001]). The correlation is mostly driven by intercontinental differences: US, Australian, European, and Asian cities all cluster together, and the correlation within each cluster is much less significant than the correlation across clusters. Urban economics predicts a clear cut relationship between generalized transport costs, urban density, and resulting transport energy use. A calibration of the model to physical parameters allows to reproduce the inverse relationship between urban density (which is is a function of m_c) and transport energy use (blue line Fig. 6A). While the urban economics framework is highly simplifying in assuming a monocentric city with homogeneous agents, the model still provides an intuition on the observed relationship.

In this light, the low urban density and high transport energy use of US-American cities is based in low m_c, high income and an extensive road transport system which enables rapid unhindered transport in and between cities. A low m_c in the USA corresponds to very low fuel taxes and a well-developed road infrastructure, reducing monetary and time costs of car travel. European cities are historically denser, and provide less space for cars, thus increasing generalized transport costs, in addition to higher fuel taxes. Finally, inhabitants of Asian cities have, in average, lower income, by this raising the relative costs of fuel, and are living in very dense settlements which prohibit high car use.

Another observation of Newman and Kenworthy [1996] is the inverse relationship between transport km of motorized vehicles and of public transit across cities (Fig. 6B). In this case, the model is less able to reproduce observed city data, systematically underestimating public transit km traveled (blue line, Fig 6B). The main reason for this divergence is that the model assumes dense radial coverage of public transit. However, most public transit networks are linear in character, and develop along public transit axes. Obviously, such sparse networks cover longer distances. The arguments of the model still holds for such networks. Density still decreases with distance, but now with higher density public transit axes surrounded with lower density vehicle-served areas. The relative increase in total distance traveled in public transit is determined by the average transit corridor length and corridor thickness. This terms can be collapsed in a linearization factor γ, determining the relative scale of public transit km covered. The general shape of this relationship can then be approximated by the model (red line in Fig. 6B, with $\gamma = 5$). The considerable variance displayed in Fig. 6B requires further investigation, possibly considering income effects, city size, a potential primacy effect (public transit of primary cities might be well subsidized), and other urban form characteristics.

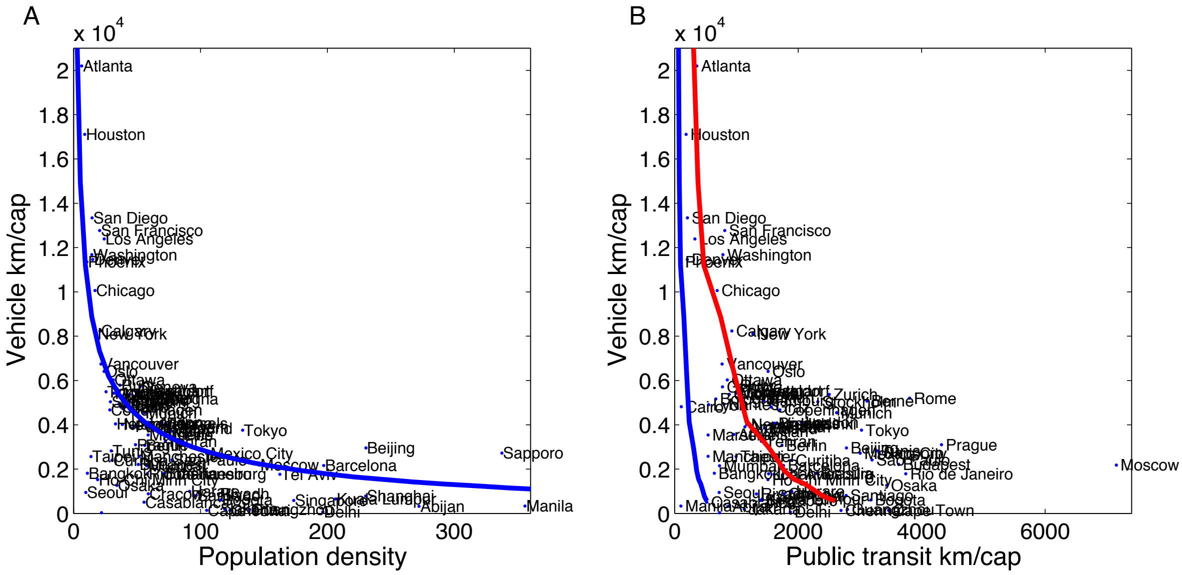

Figure 6: **Urban form and transport demand of global cities. Data are from Kenworthy and Laube [2001]. A:** Transport energy costs fall inversely proportional to urban density. Blue line: model prediction. **B:** Cities with higher car use display lower public transit use and vice versa. Blue line: model prediction. Red line: model with linearized public transit network). The following parameters were used: $a = 0.8$, $b = 1 - a$, $Y = 25$, $C = 1$, $F_c = 5$, $R_a = 0.02$, $\gamma = 5$.

5 Conclusions and discussion

In this paper, we introduce a model of urban form and modal share, building on the Alonso-Muth-Mills framework. In this model, the economic feasibility and the spatial scope of public transit depends on urban form and m_c. As a result, urban form is characterized by two modal areas, an inner city which is served by public transit, and an outer city which is served by car transport. Urban form displays a kink at the border between inner and outer city. Crucially, m_c determines not only car transport but also urban form and modal share.

We derive two first-order conditions, providing necessary conditions for welfare-enhancing investments in public transport infrastructure. The first

first-order condition specifies the marginal cost of infrastructure provision in terms of the willingness-to-pay of potential public transit users in the socially optimal case, giving the micro-economic condition for individual transport users. The second first-order condition specifies the equalization of additional infrastructure costs and additional revenue at the spatial margin, giving the macro-economic budget condition for the social infrastructure planner. We also demonstrate that a monopolistic private agent provides a suboptimal amount of public transport provision. The corresponding decrease in utility becomes particularly relevant for high marginal costs of car driving.

This conceptual framework provides an intuitive explanation of difference in viability of public transport across major world regions and can explain observation on the relationship between urban form and transport energy demand [Newman and Kenworthy, 1989, 1996]. For example, in the US low m_c, available land and urban forms that are mostly unrestricted by historical pre-automobile developments allowed low-density development which makes public transport financially unviable and environmentally ineffective [Chester and Horvath, 2009] with exception of dense coastal city regions. In contrast, in Europa higher m_c, limited land availability and historically denser cities restricted urban sprawl to some degree. As a result, public transport serves larger areas with acceptable subsidy levels. In Hong Kong, very high density and land value capitalization enables very high modal share of public transit and full recovery of infrastructure provision.

Statistical analysis of the same data set [Kenworthy and Laube, 2001] by Souche (2010) revealed that transport costs of private car and public transport and urban density are the statistically relevant factors explaining transport energy use [Souche, 2010]. Urban density seems to more relevant in explaining fuel demand than fuel prices [Souche, 2010; Karathodorou et al., 2010]. The AMM framework predicts that urban density is, ultimately, itself a function of transport costs. To test this hypothesis, intertemporal data on income, fuel prices, road infrastructure and urban form need to be statistically analyzed in a dynamical model.

Clearly, the model framework is simplistic and not suitable for application to real cities. For example, in the model the population size is kept constant, and citizens are homogeneously receiving identical income. Urban travel is characterized one-dimensionally, i.e specifying a mono-centric rather than, e.g., a poly-centric city. Distinct transport corridors have a distinct impact on urban form [Anas and Moses, 1979]. Nonetheless, while a monocentric model is highly unrealistic, the relationship between density and transport distance still holds in more complex real cities [Ewing and Cervero, 2010]. Also, the setting is static, and neither path-dependence of transport infrastructures nor the question of time scales in transport and real estate markets is analyzed (e.g. Gusdorf and Hallegatte [2007] address the importance of inertia in real estate markets in reaction to fuel price shocks). The modal choice is limited to two modes, ignoring numerous slow modes (walking, cycling, e-bikes), neither differentiating between different public modes (bus, tram, subway, regional train) nor between cars (electric cars, trucks).[3] The utility estimation excludes the social value of climate change mitigation, air quality improvement and other common good provisions or co-benefits [Creutzig and He, 2009]. An extended framework can be useful to conceptualize co-benefits and common good provision of urban densification and transport pricing policies. We suggest that the presented framework will be useful to investigate sustainable urban form addressing environmental pollution, energy security and climate change challenges.

A Appendix

Lemma A.1 *For any combination of* $(m_c, F_c, C, u, Y, \alpha)$*, there is a unique* r_p^* *in* $(0\ r_c]$ *such that* $m_p(r_p^*) = m_{infra}(r_p^*)$*.*

We first demonstrate that for at least one value of r_p, $m_p(r_p) = m_{infra}(r_p)$. Second, we show that this equality is true for exactly one value of r_p. To see the first part, observe that $m_p(r_p) \to \infty$ for $r_p \to 0$, and $m_p(r_p) \to m_c$ for $r_p \to \infty$. Also, as $r_p \to 0$, $D(r_p)$ decreases with degree < 3. Hence, $m_{infra}(r_p) \to \infty$ with degree > -1. Finally, as $r_p \to 0$, $m_{infra}(r_p) \to \infty$ ($D(r_p)$ remains finite as the population and the city radius r_c is always finite). Hence, $m_p(r_p) > m_{infra}(r_p)$ for $r_p \to 0$, and $m_p(r_p) < m_{infra}(r_p)$ for $r_p \to \infty$. As both functions are continiuous, they must cross at least once.

We turn to the second part of the proof. The slope of $m_p(r_p)$ changes with $-\frac{1}{r_p^2}$. As $m_{infra}(r_p)$ has degree > -1 for all r_p, the slope of $m_{infra}(r_p)$ has always degree > -2. Hence, the difference $m_p(r_p) - m_{infra}(r_p)$ is monotonously increasing in r_p. We conclude that $m_p(r_p) = m_{infra}(r_p)$ for exactly for one r_p.

[3]The contribution of technology to address externalities such as greenhouse gas emissions was not part of this paper. More than half of required transport emission reduction can reasonable expected to come from advanced technology, but are best addressed on national or regional level [IEA, 2009; Creutzig et al., 2011].

References

Alonso, W. (1964). *Location and Land Use*. Harvard University Press.

Anas, A. and Moses, L. N. (1979). Mode choice, transport structure and urban land use. *Journal of Urban Economics*, 6:228–246.

Arnott, R. and Riley, J. G. (1972). Asymmetrical production possibilities, the social gains from inequality and the optimum town. *The Scandinavian Journal of Economics*, 79(3):301–311.

Bongardt, D., Breithaupt, M., and Creutzig, F. (2010). Beyond the fossil city: Towards low carbon transport and green growth. *GTZ working paper. Presented at the Fifth Regional EST Forum, 23 - 25 August 2010, Bangkok, Thailand.*

Borck, R. and Wrede, M. (2008). Commuting subsidies with two transport modes. *Journal of Urban Economics*, 63(3):841 – 848.

Brueckner, J. K. (2000). Urban sprawl: Diagnosis and remedies. *International Regional Science Reviews*, 23(2):160–171.

Cao, X., Mokhtarian, P. L., and Handy, S. (2009). Examining the impacts of residential self-selection on travel behaviour: A focus on empirical findings. *Transport Reviews*, 29(3):359–395.

Capozza, D. R. (1973). Subways and land use. *Environment and Planning*, 5:555–577.

Cervero, R. (1998). *The transit metropolis: A global inquiry*. Island Press.

Chester, M. and Horvath, A. (2009). Environmental assessment of passenger transportation should include infrastructure and supply chains. *Environmental Research Letters*, 4(2).

Creutzig, F. and He, D. (2009). Climate change mitigation and co-benefits of feasible transport demand policies in beijing. *Transport Research D*, 14:120–131.

Creutzig, F., McGlynn, E., Minx, J., and Edenhofer, O. (2011). Climate policies for road transport revisited (I): evaluation of the current framework. *Energy Policy*, 39(5):2396–2406.

Ewing, R. (2007). Growing cooler: The evidence on urban development and climate change.

Ewing, R. and Cervero, R. (2010). Travel and the built environment – a Meta-Analysis. *Journal of the American Planning Association*, 76(3):265–294.

Frank, L. D. and Pivo, G. (1994). Impacts of mixed use and density on utilization of three modes of travel: Single occupant vehi- cle, transit, and walking. *Transportation Research Record: Journal of the Transportation Research Board*, 1466:44–52.

Fujita, M. (1989). *Urban Economic Theory*. Cambridge University Press, first edition.

Glaeser, E. L. and Kahn, M. E. (2010). The greenness of cities: Carbon dioxide emissions and urban development. *Journal of Urban Economics*, 67(3):404 – 418.

Gordon, P. and Richardson, H. W. (1997). Are compact cities a desirable planning goal? *American Planning Association. Journal of the American Planning Association*, 63(1):95–106.

Gusdorf, F. and Hallegatte, S. (2007). Compact or spread-out cities: Urban planning, taxation, and the vulnerability to transportation shocks. *Energy Policy*, 35:4826–4838.

Haring, J. E., Slobko, T., and Chapman, J. (1976). The impact of alternative transportation systems on urban structure. *Journal of Urban Economics*, 3:14–20.

IEA, I. E. A. (2009). Transport, energy and CO2: moving toward sustainability. Technical report, International Energy Agency, Paris.

Karathodorou, N., Graham, D. J., and Noland, R. B. (2010). Estimating the effect of urban density on fuel demand. *Energy Economics*, 32(1):86 – 92.

Kenworthy, J. and Laube, F. (2001). *The Millennium Cities Database for Sustainable Transport. (CDROM Database)*. International Union (Association) of Public Transport, (UITP), Brussels and Institute for Sustainability and Technology Policy (ISTP), Perth.

Kim, T. J. (1979). Alternative transportation modes in an urban land use model: A general equilibrium approach. *Journal of Urban Economics*, 6:197–215.

Mindali, O., Raveh, A., and Salomon, I. (2004). Urban density and energy consumption: a new look at old statistics. *Transportation Research Part A*, 38:143–162.

Mirrlees, J. A. (1972). The optimum town. *The Swedish Journal of Economics*.

Newman, P. and Kenworthy, J. (1989). Gasoline consumption and cities: A comparison of u. s. cities with a global survey. *Journal of the American Planning Association*, 55(1):24–37.

Newman, P. and Kenworthy, J. (1996). The land use–transport connection : An overview. *Land Use Policy*, 13(1):1–22.

Proost, S. and Dender, K. V. (2008). Optimal urban transport pricing in the presence of congestion, economies of density and costly public funds. *Transportation Research Part A: Policy and Practice*, 42(9):1220 – 1230.

Sasaki, K. (1989). Transportation system change and urban structure in two-transport mode setting. *Journal of Urban Economics*, 25:346–367.

Souche, S. (2010). Measuring the structural determinants of urban travel demand. *Transport Policy*, 17(3):127 – 134.

Su, Q. and DeSalvo, J. S. (2008). The effect of transportation subsidies on urban sprawl. *Journal of Regional Science*, 48(3):567–594.

von Thünen, J. H. (1826). *Der Isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie*. Hamburg.

Weisz, H. and Steinberger, J. (2010). Reducing energy and material flows in cities. *Current Opinion in Environmental Sustainability*, 2(3):185–192.