University Library



A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Navidi, Z;Nagel, K;Winter, S

Title:
Toward identifying the critical mass in spatial two-sided markets

Date:
2020-11-01

Citation:
Navidi, Z., Nagel, K. & Winter, S. (2020). Toward identifying the critical mass in spatial two-sided markets. Environment and Planning B: Urban Analytics and City Science, 47 (9), pp.1704-1724. https://doi.org/10.1177/2399808319842181.

Persistent Link:
https://hdl.handle.net/11343/258740

Towards Identifying the Critical Mass in Spatial Two-sided Markets

Journal Title
XX(X):1–12
©The Author(s) 2016
Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/ToBeAssigned
www.sagepub.com/
SAGE

Abstract

Unlike their non-spatial counterparts, *spatial* multi-sided platforms are matchmaking platforms with an additional layer of complexity: their customers expect to meet in space, not only virtually. This additional challenge will be studied in this paper in the context of a two-sided ride-sharing platform, which serves drivers and passengers. As with any two-sided platform, there is an interdependence between both groups of customers: More drivers are more attractive for passengers, and vice versa. This interdependence creates the old *chicken-and-egg* problem, only that here drivers and passengers need to be matched not for a virtual transaction, but by their ability to meet physically and travel jointly. We argue, and illustrate by simulations, that in spatial multi-sided markets there is not a single critical mass frontier that needs to be reached in order to make the system self-sustained (as in non-spatial markets), and that this frontier is varying from one location to the next, depending on the density and distribution of the demand and supply over space and time. Identification of the critical mass frontier will allow for better evaluation of implementation policies and regulations.

Keywords

Critical mass, Multi-sided platform, spatial critical mass, simulation, ride-sharing

Introduction

The recent advancements in Information and Communication Technologies (ICT) have made flexible and shared modes of transport, such as demand responsive transport, car-sharing or ride-sharing closer to reality. However, there are certain operational issues, specific to these services, that are yet to be fully understood: Reports of many startups failing and some booming are often seen in the news. The key appears to be developing critical mass, which in these transport applications is heavily influenced by their dependency on spatial and temporal properties of demand and supply. *Critical mass* is sufficient amount of demand and supply in a system to make it self-sustaining, i.e., growing without any external investment. ICT can be seen as the enabler, or matchmaker, between local demand and supply in these modes of transport. Economics calls these matchmakers *multi-sided platforms* (Evans and Schmalensee 2016). One important characteristic of these platform is *critical mass frontier*, which consists of all possible situations, in which there are (barely) sufficient members of each group's participants for a system to be self-sustained (Evans and Schmalensee 2016). In other words, the main difference between ICT enabled multi-sided markets and traditional one-sided markets is that multi-sided markets do not only consist of demand and supply sides, but platform providers enter as an intermediary player facilitating the interaction between suppliers and consumers. The facilitation reduces significantly the friction between the market participants (Evans and Schmalensee 2016). For example, passengers find traditional modes of shared transportation by studying schedules and walking to salient pick-up points (e.g., bus stops, train stations). With flexible shared transportation partners, pick-up points and pick-up times have to be individually negotiated. Here matchmaking platforms serve as intermediary matchmakers resolving the complexity of the negotiation process.

The dilemma of multi-sided matchmaker platforms is the network effect: the more users on all sides, the more valuable is the service. Empirical research claims that the network effect is even quadratic, not linear. Applied to a flexible transportation system, this network effect can be illustrated:

- A car-sharing company puts up a vehicle at a fixed cost. Compared to the fix cost, the marginal costs of serving each customer are low, meaning that average costs decrease with increasing numbers of customers. In other words, the more customers ask for a shared car, the more attractive is car-sharing for the company.
- Customers look for access to flexible mobility (here: car-sharing). They do not want to check out each company and private operator for current offers each time they look for a car; instead they look for a single place of information – and this is the case for a multi-sided platform. The platform becomes more attractive with more available cars (car-sharing companies), as the customers will have more versatile options.

Thus, the car-sharing platform provides a service to both market participants: it attracts customers for the companies they would not find on their own, and it provides access to cars a single company would not be able to provide. And this platform could earn money from subscriptions, transaction fees, or advertising.

For flexible shared transportation platforms, the dilemma with this network effect sits in the beginning. A (new) platform has first to reach a certain spatial density of vehicle coverage to be able to offer an attractive service to passengers

at certain locations and times. Customers expect a vehicle nearby when they need one. In practice, this condition requires a heavy upfront investment of a transportation provider, or combined upfront investment of a larger number of providers. On the other hand, this certain density highly depends on the number and distribution of customers, and can vary in space. If there are not sufficient vehicles close to where the customers are, passengers will turn away. Higher customer density may require a higher vehicle density and vice versa. Therefore, to solve the *chicken-and-egg* problem of providing critical mass to both sides of the market the platform has to look at the spatial and temporal components of demand and supply.

Some existing flexible shared transportation platforms have succeeded to reach a critical mass of customers as well as of transport providers. Uber is an example of such a multi-sided platform, with a large number of independent contractors providing the vehicles on one hand, a large number of passengers on the other, and mechanisms (such as surge pricing) to direct vehicles to the current demand. But many other flexible shared transportation platforms have failed in reaching a critical mass.

Thus, the main difference between many multi-sided platforms and flexible shared transportation platforms is that location matters as well, not only the number of drivers (or vehicles) and passengers. For the passengers, the vehicles need to be there at the right time, and for the vehicles large demand elsewhere is irrelevant.

This additional challenge will be studied in this paper. We argue, and illustrate by simulations, that in spatial multi-sided markets, other than in non-spatial markets, there is not a single critical mass frontier that needs to be reached in order to make the system self-sustained, and that this frontier is varying from one location to the next, depending on the density and distribution of the demand and supply over space and time. We call this dynamic location-based frontier, *spatial critical mass frontier*. Identification of the spatial critical mass frontier in the context of flexible transport systems will allow for better evaluation of implementation policies and regulations, which is discussed in detail in Discussion.

We will use a two-sided ride-sharing market for illustration. Plausibly, this market is more attractive to customers when the density of offerings is higher. Conversely, the market is more attractive to suppliers when the density of customers is higher. In consequence, it may happen that a densely populated central area is already "super"-critical, i.e. has sufficient demand for a self-sustained system, while less densely populated areas are still "sub"-critical, i.e. it is impossible to have a self-sustained system in that area. Interestingly, the literature so far concentrates on homogeneous markets. Even a seemingly spatial paper such as by Djavadian and Chow (2017) considers a fixed region, and then considers the question if the region as a whole is viable or not. That paper does not address the question whether a change of the region boundary would change its criticality.

The rest of the paper is organized as follows: First, we explain the relevance of a vehicle sharing platform to a two-sided market and the necessity of developing the concept of a spatial critical mass frontier. Then, we elaborate on the basics of two-sided markets and its differences to a traditional one-sided market. In the next section, we elaborate the conceptual model and frame our expectations, which set the scene for describing the experiments, including the model, simulation configurations and scenarios. Then, we present the results of the experiments. Finally, in the last two sections, we present a thorough discussion and conclusion alongside with the limitations of the work and future plans.

Literature Review

Evans and Schmalensee (2016) discuss the challenges of some multi-sided platforms to reach critical mass. The commonality of these platforms is that their market participants – the 'sides' – are stuck with their location. Multi-sided platforms that have to consider location are a crisply defined subset of multi-sided platforms. News, media, or advice can be provided or sought independent of location, but for example booking a table in a restaurant (the prime example of Evans and Schmalensee (2016)) requires to consider the location of the customers as well as the location of the restaurants. A person wanting to book a table for the evening is interested only in restaurants within an acceptable travel distance. In this category are other services as well, such as platforms matching repair services, or matching travelers for shared or flexible forms of transportation.

Evans and Schmalensee (2016) discuss how a particular restaurant booking platform, seemingly unaware of the spatial aspects of the network effect, struggled to achieve critical mass originally, and only by trial and error found strategies to create critical mass at least in centers. The fundamental empirical principles of geographic information science (Goodchild 2009, 2011; Chrisman 2012) are highlighting that geographical phenomena are heterogeneous, and that near things are more related than distant things (Tobler 1970). These principles should guide the notion of *critical mass* for platforms, i.e., the determination of *critical mass* must become location-dependent.

A remarkable consequence of spatial autocorrelation is the hierarchical organization of space (Batty 2006). Central place theory (Christaller 1933; Batty 2009) postulates that (economic) centers serve surrounding areas depending on the range people would travel for particular goods. These hierarchic centers are forming a lattice, and the regularity has later been linked to fractal nature (Batty and Longley 1994; Jiang and Brandt 2016). Spatial hierarchy emerging from economic considerations impacts on settlement structures, transportation networks (Jiang 2009), and induced mobility demand (Bento et al. 2005). For example, a common challenge in shared transportation is the reallocation of resources that have left the centers: Algorithms still assume (falsely) a homogeneous demand (Fricker and Gast 2016).

The hierarchy and inhomogeneity of the road network has led to characterizations by network centrality measures (Claramunt and Winter 2007). However, network centrality is induced by the inhomogeneous distribution of settlements. Thus network centrality alone cannot explain transportation demand; the population distribution adds to the complexity of urban morphology and derived transport demand (Kazerani and Winter 2009). On the other hand, the

emergence of flexible transportation systems, namely ride sharing, car pooling, car sharing, and demand responsive transport systems, provides a potential unique opportunity to address this complexity. However, their failure in the real world, mostly due to economical reasons (Enoch et al. 2006; Sulopuisto 2016), proves that their market is yet to be fully understood. The concept of critical mass and critical mass frontier (see Theory of Multi-sided Markets and Critical Mass) may be the key to their success, as it guarantees the success in other platforms or phenomena.

Since the development of the two-sided market concept by Rochet and Tirole (2003), numerous studies have explored platform competition in this context (e.g., Armstrong 2006; Caillaud and Jullien 2003; Kodera 2015), and many have investigated the critical mass in mostly technology related platforms (e.g., Grajek and Kretschmer 2012; Dubé et al. 2010). However, a flexible transportation market as an inherently spatial market with many unique characteristics cannot easily adopt or benefit from those studies and requires more specific and to-the-point investigations, which are scarce. For instance, Wang et al. (2016) have looked at pricing in taxi hailing applications by investigating the existence and stability of equilibria in a two sided-market with the taxi hailing application being the matchmaker. Djavadian and Chow (2017) have framed the flexible transportation systems in a two-sided market, and demonstrated the expected direct interdependence of demand and supply. Their work includes a simulation platform showing that there is benefit in investigating flexible transport systems in a two-sided market.

However, none discusses critical mass or provides insight into the critical mass frontier in the context of spatial systems. Identifying critical mass in spatial systems, such as transportation, requires deep understanding of the impact and its significance of urban structure, and demand and supply spatial characteristics. In this work, we aim at bridging this gap by providing an initial study on the effect of basic spatial characteristics of an area in critical mass of a spatial two-sided market.

Theory of Multi-sided Markets and Critical Mass

Here, we explain the traditional and two-sided markets using charts in 1a to 1c. In these charts, the vertical and horizontal axis represent the amount of demand and supply. Assuming the intersection of each pair of arrows is one possible market situation (i.e., the level of demand and supply), the arrows represent the direction of change for respective axis, and their sizes demonstrate the rate of change. Each figure is described in detail in the followings.

Standard economics assumes that average costs increase with increasing supply. In such a market, companies can start with small amounts, and increase production at increasingly higher cost until marginal revenue is equal to marginal costs. More precisely:

- If there is high demand but low supply, prices are high, reducing demand and increasing supply as denoted by the arrows (top left in Fig. 1a).
- If there is low demand but high supply, prices are low, increasing demand and decreasing supply as denoted by the arrows (bottom right in Fig. 1a).
- If both demand and supply are low, prices should be medium, thus somewhat increasing the low demand to medium levels, and somewhat increasing the low supply to medium levels as denoted by the arrows (bottom left in Fig. 1a).
- If both demand and supply are high, prices should also be medium, thus somewhat decreasing the high demand to medium levels, and somewhat decreasing the high supply to medium levels as denoted by the arrows (top right in Fig. 1a).

Overall, the dynamics has only one attractive fixed point, which lies in the center of Fig. 1a, and is more conventionally given by the intersection of the demand and the supply curves.

With economies of scale on the supply side and network effects on the demand side, these cases behave as follows:

- If there is high demand but low supply, prices are high, reducing demand and increasing supply (top left in Fig. 1b).
- If there is low demand but high supply, prices are low, increasing demand and reducing supply (bottom right in Fig. 1b).
- If both demand and supply are low, prices are medium. For economies of scale on the supply side, this will mean that the price is below the supply curve, further decreasing supply. Conversely, for network effects on the demand side, the price is above the demand/willingness-to-pay (w.t.p.) curve, also further reducing demand (bottom left in Fig. 1b).
- If both demand and supply are high, prices are again medium. For economies of scale on the supply side, costs will be lower than prices, thus further increasing supply. Similarly, for network effects on the demand side, the willingness to pay will be above prices, thus further increasing demand (top right in Fig. 1b).

That is, the dynamics has two attractive fixed points, one in the bottom left, one in the top right. A supplying company (roughly) needs to estimate the size of the market (the size of the demand in the upper right corner of Fig. 1b), divide the cost of the supply by the size of the market, and then assess if the market will bear the resulting price. If the product appears to be profitable, and all underlying assessments are correct, then a company can in principle (e.g., barring competition) enforce the outcome by subsidizing the product long enough until demand has followed the large supply. That is, the dynamics has in principle two attractive fixed points, but the supplier by its own actions can render one of them irrelevant.

The two-sided market with economies of scale on both sides has a flow diagram similar to a market with economies of scale on both sides (Fig. 1c). However, the platform provider, even when being a monopolist, is not the supplier, and thus can no longer force the outcome of the dynamics. Clearly, high subsidies and marketing campaigns can reduce the basin of attraction of the bottom left fixed point. Yet, it cannot be rendered irrelevant by the platform provider alone. Evans and Schmalensee (2016) use the term “critical mass

frontier", which can be seen as the line separating the basins of attraction of the two fixed points, when plotting numbers of customers and number of suppliers (Fig. 1c).

Conceptual Model and Expectations

We are interested in two-sided flexible and shared transportation markets: markets where the platform itself does not directly control drivers (supply) and customers (demand). Some systems seem similar, e.g., bicycle sharing systems, but in general are not, since the platform provider and the supplier are often the same. They could, however, be converted into a two-sided market, for example, in a station-based bicycle sharing system, where stations would be provided by persons and institutions, and not by the platform provider itself.

In order to make this concrete for a spatial approach, let us assume that our area of interest is divided into regular cells. It is now plausible to assume that we will have the dynamics similar to Fig. 1c *at each grid cell*. However, we will additionally have an infection process: If one cell is far in the UR corner, it will infect its neighbors because both the high demand and the high supply will radiate into the neighboring cells. Similarly, if a cell is far in the LL corner, this cell will not help its neighbors to become served, and thus effectively inhibit them. Overall, the dynamics becomes quite similar to that of the well-known Ising model (Ising 1925; Chandler 1987), where spins, which are either "up" or "down", try to align to each other, but are also subject to some random noise. There is an elaborate theory of what happens when the noise is larger or smaller; for the paper here, we assume small noise, and thus have a so-called first-order phase transition between "most spins up" and "most spins down". The model can also be used to describe aspects of segregation (Müller et al. 2008); translating this to our two-sided transport market, "up" would correspond to "served", and "down" to not-served.

From this theory, one can come up with predictions. For the following, it is assumed that each individual cell follows the dynamics according to Fig. 1c, and we consider the same plot, but in which all cells' demand and supply values are averaged over the whole system. For *homogeneous* systems (same population density everywhere), one would expect the following:

- In a system of ***infinite* size**, one would expect that the dashed line in Fig. 1c divides the dynamics deterministically into two basins of attraction: If the system starts to the lower left (LL) of that line, it will deterministically go to zero demand and supply (sub-critical); when starting in the upper right (UR), it will deterministically go to high supply and demand (= super-critical = "served").
- In a system of ***finite* size**, one would assume that boundary to become blurred, and the deterministic behavior becomes replaced by probabilities: When starting the system somewhere inside the LL region, one would still expect a non-zero probability to become super-critical (= served); that probability would be 50% at the dashed line, and become smaller with increasing distance from it. Conversely, when starting somewhere inside the UR region, one would

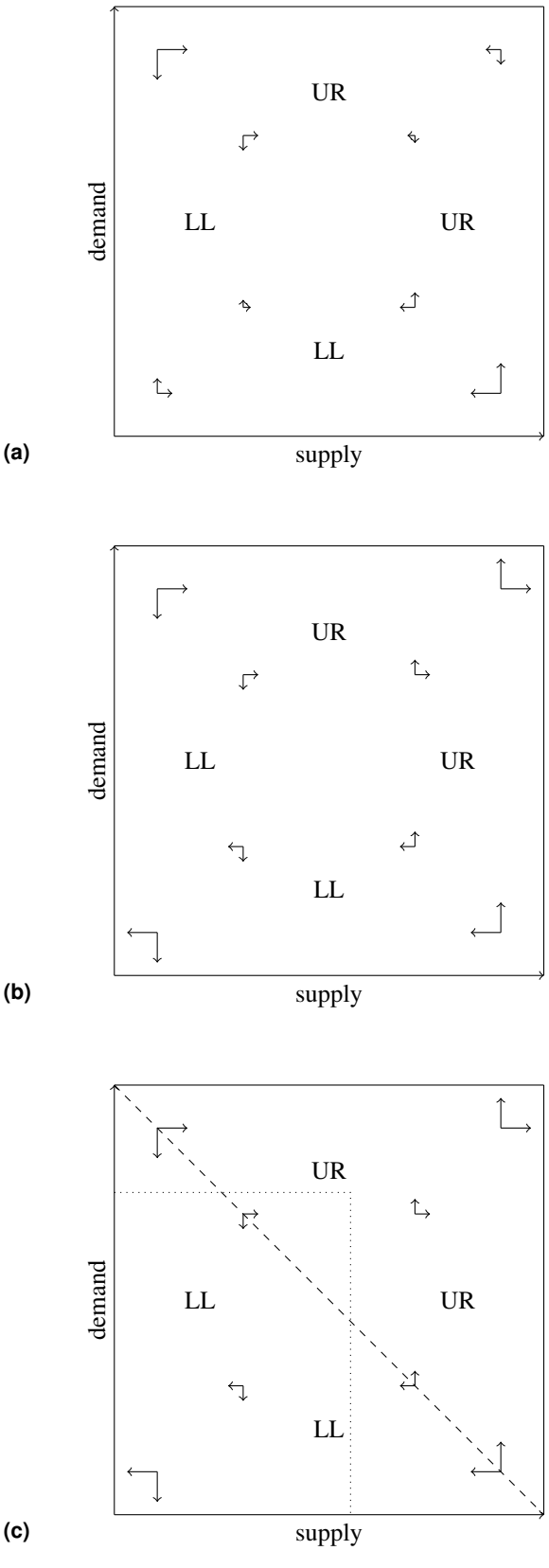


Figure 1. Dynamics for different levels of demand and supply in different types of markets. A larger value along the demand axis means a higher demand; a larger value along the suppy axis means a larger supply. Each (demand,supply) pair corresponds to a point in the 2-dimensional plot. The respective driving forces are denoted by arrows. (a) Regular market. (b) With economies of scale and network effects. (c) Two-sided market with economies of scale and network effects. The dashed line divides UR (upper right) from LL (lower left). The dotted line denotes a possible market where the densities of demand and supply are smaller (e.g., in rural areas) than they maximally could be (e.g., in urban centres).

still expect a non-zero probability to become sub-critical; again, that probability would be 50% at the dashed line, and become smaller with increasing distance from it.

- One would expect that the **transition region (in state space = demand/supply space) becomes more narrow with larger systems**. That is, for small systems, the transition from small to large probability to become super-critical is slow and smooth; for large systems, it is fast and steep; and in the limit of infinite system size, it becomes a deterministic switch. Plots that demonstrate this are the NU and UU plots in Fig. 3, from left to right.
  This also implies that a two-sided market of infinite size could not become super-critical, since the starting point is in the LL corner. Becoming super-critical is thus only possible through some special pathway, typically given by some inhomogeneity in the system (e.g., starting with a specific subset of the population, e.g., only computer-affine people).

In general, however, we expect real-world systems to be inhomogeneous. Fig. 1c shows, by the dotted lines, a possible market where the reachable densities of demand and supply are smaller than what they could be in an urban core. One clearly sees that a transition to the UR area would be more difficult to achieve. With such information, it would be possible to delineate regions that could potentially become super-critical, and others that cannot – the latter being regions where the dotted rectangle does not extend into the UR area at all.

The pathway of an area, consisting of many cells, into super-criticality will, in general, not be a homogeneous transition, but rather some spatial location becoming super-critical, and then **infecting its neighbors**. This also implies that connected areas that can potentially become super-critical are either all together super-critical, or all together they are not. As an illustration, an already served urban core will "infect" all neighboring suburbs as long as they are potentially super-critical. Once the borders of that connected region have been reached, the growth of the service area will stop. From then on, improvements in the cost structure will make additional areas super-critical, and the service area will then grow into those regions. This will somewhat resemble invasion percolation (Wilkinson and Willemsen 1983), except that spatial densities are not distributed randomly, but according to population densities and people's preferences.

Experiments

We will illustrate and test our discussion with simulations, which we will make increasingly more realistic. The simulations are coded in Java language and the important aspects of them are:*

- We will assume a fixed population, which in general will be distributed non-homogeneously in space. Of this fixed population there will be different shares of people interested in the market.
- Initial supply will either be homogeneously or normally distributed (cf. Fig. 2).
- In general, both demand and supply will "drift" with a small rate into the market. Whether they remain in the market will depend on the utility of the service for the customer and on the profit for the supplier.

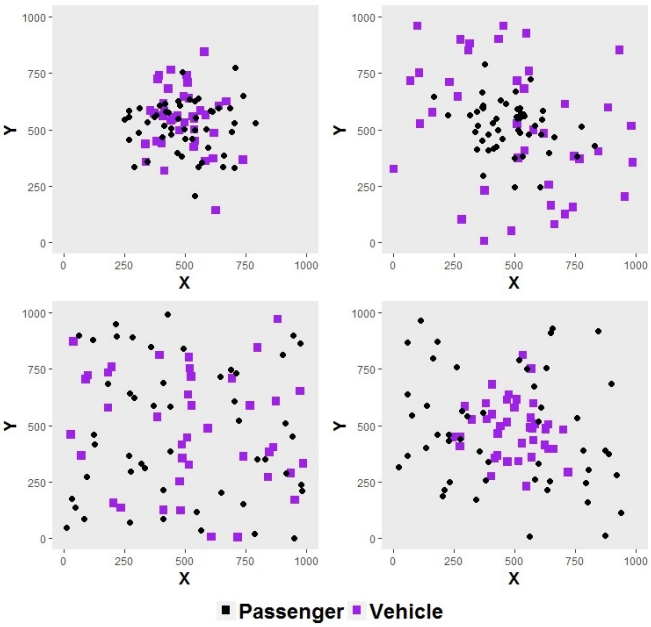


Figure 2. An example of different distributions of vehicles and passengers. Upper left: Both passengers and vehicles are normally distributed (NN). Upper right: Passengers normally distributed, vehicles uniformly (NU). Bottom left: Both passengers and vehicles uniformly distributed (UU). Bottom right: Passengers uniformly distributed, vehicles normally distributed (UN).

Model and Simulation Configurations

The simulation consists of an environment and two types of entities: passenger and vehicle. The environment is the area that contains all the entities in space. A passenger i has three individual characteristics: Number of Neighbours (N_i), Interest (TS_i), and Utility (U_i). N_i is counted for each passenger and is equal to the number of vehicles around i that are within the Acceptance Threshold (AT_i). TS_i can be either 1, meaning Passenger i is interested in being matched with a vehicle, or 0, meaning otherwise. U_i is calculated inversely proportional to the distance of Passenger i to their matched vehicle, if applicable. The vehicle entity j has only one individual characteristics: Utility (V_j), which is equal to the utility of the respective matched passenger, or 0.

The simulations work iteratively. In each scenario, which are explained below, the iterations start after the population generation (described in Section Scenarios). In each iteration, first, a certain number of vehicles are randomly generated within the environment, which are either normally or uniformly distributed (again cf. Fig. 2), and added to the already existing vehicles. This number of additional vehicles decreases by 10% for the next iteration if the average utility of all vehicles is below a certain threshold (VT), and stays the same otherwise.

*Full code is available on

Second, the passengers' interests, with default values of 0, are updated based on the availability of vehicles or their past experience (utility from the previous iteration). More technically, TS_i becomes 1 if either N_i in the same iteration is higher than a defined threshold (the Neighbour-based Interest Threshold $NTST$), or if U_i from the previous iteration, if it exists, is bigger than a Utility Threshold (UT).

The changes in the number of vehicles and interest of passengers replicate their interdependence. If there are more vehicles, more people will get interested, so the chances for a vehicle to be matched and keep the overall utility higher is bigger, which results in more vehicles in the next iteration.

Then, all the vehicles within the acceptance threshold (AT) of each interested passenger are identified and the closest one is matched with that passenger, the passenger's and the vehicle's utilities are updated based on their distance and the matched vehicle is removed from the pool. Finally, all the vehicles with utility below VT are deleted (they lose interest) and the rest stay for the next iteration. Each scenario is run for 100 iterations.

Further, as we assume that in the systems of finite size the system's success is associated with a probability, each 100 iterations of one scenario is repeated 100 times to calculate the probability of the success. Success is defined as at least one passenger having interest. Furthermore, since we are focused only on the concept of critical mass, for the sake of simplicity we compute with Euclidean distances and do not consider network distances. As Hua et al. (2018) have demonstrated recently, the differences between the two distance measures on spatial queries (e.g., on ordering) are not necessarily relevant. Moreover, network distances make a difference when there are either strong physical barriers such as mountains, lakes or rivers in the area of investigation, or singularly fast transportation infrastructure, such as motorways without speed limits. Neither of these occur in the (populated parts of the) study area. Also, our model assumes "success" when a match is within $100\ metres$ (AT threshold); we would expect that instead using a network distance of, say, $150\ metres$, would yield similar results. Note that we are using network-based models for our more realistic simulations (citation is withheld due to anonymous review).

Scenarios

We have first designed scenarios in an artificial environment to show the impact of the spatial distribution of entities, their density and the size of the area on the critical mass frontier in a controlled environment. In these scenarios, the overall population is static and fixed in terms of their location, and in turn density and distribution. However, the demand, i.e., people interested in using the ride-sharing system, is dynamic and uncontrolled, as it depends on the number of available vehicles nearby. In terms of supply, there is a controlled starting point in terms of distribution and density, however, only distribution is controlled in the iterations and density may get higher or lower. Then we have run simulations on a real-world scenario, where not all of these parameters can be controlled any longer.

Artificial Environment The scenarios in this category happen in a square area and include variations of area size, and the density and the distribution of entities. Three area sizes, 1 km^2, 9 km^2, and 25 km^2, and Normal and Uniform distributions have been considered. Combining these parameters comprises the twelve different cases presented in Table 1. According to Conceptual Model and Expectations Section, we expect the system to have different behavior based on its size, therefore, three different sizes are chosen to test the hypothesis. The specific numbers are based on the suburb sizes of Melbourne, Australia, the metropolitan area of the real-world scenario (see below). One square kilometre is close to the smallest suburbs. Nine is considered a medium size suburb and 25 km^2 is among the large suburbs. Although there are bigger suburbs in the metropolitan area, we did not chose to go higher, as the test runs proved that 25 is sufficiently big to test the impact of size.

For each case 100 scenarios are run which includes variation of both entities' densities from 10 per km^2 to 100 per km^2 with an increment of 10 per km^2. The maximum density for the vehicles are 100 per km^2 and in no scenario we have investigated more. In all scenarios the acceptance threshold (AT) and neighbour interest threshold ($NTST$) are the same for all individuals and are set to 100 meter and 5 neighbours, respectively. Moreover, the utility thresholds for people to loose interest and for vehicles to decrease their number for the next iteration are the same and equal to 2.0. In Section Sensitivity Analysis, we explain how changing these thresholds affects the results.

Real-world Scenario: Yarra Ranges, Australia, is a local government area (LGA), located in the outer northeastern suburbs of Melbourne [†]. Yarra Ranges consists of 14 suburbs, one of which includes only national parks and no resident. Figure 4 presents the population density of the residential suburbs at 10% of the total population. Mooroolbark has the highest density and is surrounded by other high density suburbs, namely, Kilsyth, Monterose, and Chirnside Park. In contrast, Upwey-Tecoma with the second highest density is surrounded only by low density areas. This provides an opportunity to study the impact of boundaries and neighbouring suburbs on the transition to super-criticality.

The population generation is based on the census data of 2016 (Australian Bureau of Statistic 2016), which reports the number of people living in an area on a granularity of *Mesh Block*, the smallest area defined by Australian Bureau of Statistics (ABS) for statistical data reporting and analysis [‡]. To generate the population various percentage of the whole population is considered, and sufficient points are created in each mesh block to represent the location of one person. The simulations are run with 5% to 100% of the total population. All other parameters and vehicle densities are the same as the scenarios from artificial environment. There is only one difference in the method for adding vehicles. While in the artificial environment all vehicles are treated as one pool, in the Yarra Ranges scenario, the vehicles are divided into 13 different pools based on the 13 occupied suburbs of Yarra Ranges. At the end of each iteration, the average utility of all vehicles in each suburb is calculated, and if it is lower

[†] Map of Yarra Ranges is available as supplemental material

[‡] http://www.abs.gov.au/websitedbs/censushome.nsf/home/meshblockcounts

Table 1. Artificial environment scenarios

Area Size (km^2)	Population Distribution	Vehicle Distribution	Case Code
1	Normal	Normal	1_NN
1	Normal	Uniform	1_NU
1	Uniform	Uniform	1_NU
1	Uniform	Normal	1_UN
9	Normal	Normal	9_NN
9	Normal	Uniform	9_NU
9	Uniform	Uniform	9_UU
9	Uniform	Normal	9_UN
25	Normal	Normal	25_NN
25	Normal	Uniform	25_NU
25	Uniform	Uniform	25_UU
25	Uniform	Normal	25_UN

than VT, in the next iteration that suburb receives 10% fewer vehicles. However, the passengers are able to look for a vehicle globally, i.e., they do not care if they are in the same suburb as the matched vehicle.

Results

The objective of this work is to investigate how the spatial characteristics of demand and supply affect the critical mass frontier in a spatial two-sided market.

As explained previously, we expect the critical mass frontier to be a blurred area showing the transition from sub-critical to super-critical conditions. To this end, we present the results as heat maps, where the horizontal and vertical axes represent passengers' and initial vehicles' densities in the area respectively, while the colors represent the probability of the system's success and ranges from dark red, representing 100% probability of the system's success, to dark blue, meaning its 0% probability.

Artificial Environment

Figure 3 demonstrates the results for cases in the artificial environment. In all graphs inside this figure, the transition region is visible, which is the area between the dark red and dark blue in each graph.

Let us start with the **UU case** (third row), which is closest to the theoretical situation of Section Conceptual Model and Expectations. One can make the following observations:

- The outcome is probabilistic: For the same initial conditions, different outcomes are possible.
- For very low densities of both passengers and vehicles, the probability to become supercritical is close to zero. Conversely, the probability is close to one for large densities of passengers and vehicles.
- The width of the transition region (from blue to read) becomes smaller with larger system sizes. This is consistent with the theory from Section Conceptual Model and Expectations.

The **NU case** (2nd row) is similar to the already discussed UU case (3rd row). In contrast, **both *N** cases (1st and 4th row) are similar to each other, but different from the two *U cases (2nd and 3rd row). This implies that the initial vehicle distribution (second letter) has a stronger influence on the outcome than the distribution of the population. Initially, concentrating the vehicles in a smaller area (the *N cases) yields a much higher probability of overall success than spreading them out. This implies that whoever wants to make the system a success should concentrate its seed vehicles into small initial service areas rather than spreading them out.

Sensitivity Analysis

In designing the scenarios, we have made a number of assumptions, namely acceptance, neighbour-based interest, and utility thresholds, to be able to demonstrate and discuss the results. Understanding how the variations of these numbers changes the system behaviour is critical, and makes sensitivity analysis a requirement.

We have done the sensitivity analysis on the 9 km^2 area for two reasons: First, according to the results presented in the previous section, too small or too large system may exaggerate or eliminate the impact of parameters on the transition area, thus, it is important to conduct the sensitivity analysis in a medium size area. Second, 9 km^2 is a reasonable approximation of many suburbs in the study area.

The analysis hardly showed any observable change in NN cases when changing the parameters. But unlike the NN cases, some behavioural changes were manifested in the NU and UU cases due to parameters variation, which includes mostly the width (or shape) and the position of the transition area. By width, we mean how fast the systems changes from dark blue to dark red; while, by position, we mean how high the area is in the graph. For example, the transition area in 1_NU case is higher and wider than the one in 25_NU case in Figure 3.

Variations of the utility threshold have proven not to be impactful. In other words, if the threshold increases or decreases by up to 30% the difference in the transition area is minimal, i.e., there is no observed change in its shape and only minimal changes in its position. Since the utility is only a function of distance, variations in acceptable distance thresholds have been investigated as well. Unlike the changes in utility threshold, alterations in the acceptance threshold appear to be critical to a certain extent. While relaxing this threshold from 100 meter to 200 meter results in a much narrower and lower transition area, and changing it up to 400 meter converts the transition area to one deterministic line in all three cases (9_NN, 9_NU, and 9_UU). This sensitivity to changes in acceptance threshold but not to the utility

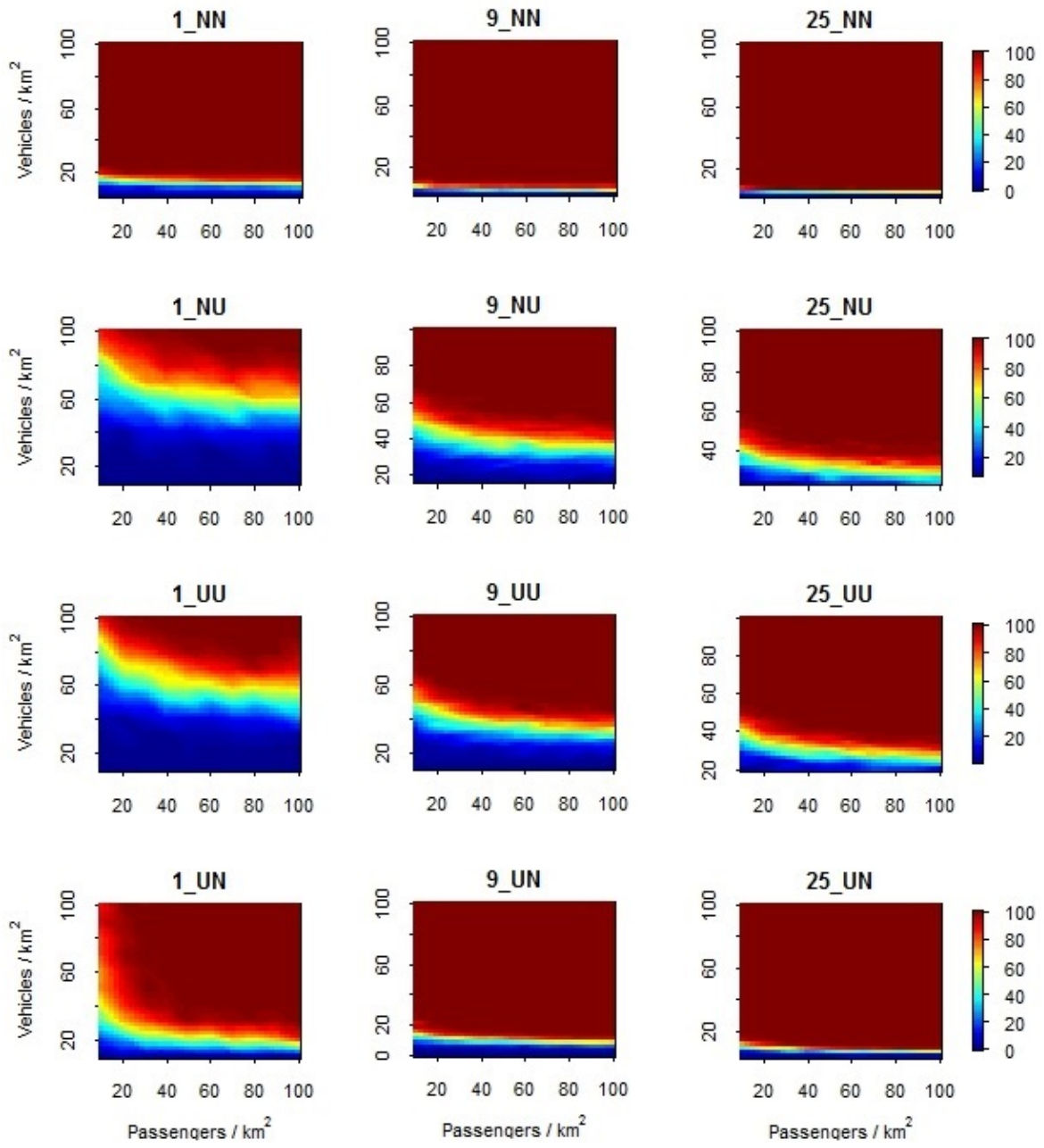


Figure 3. The success probability of systems with different sizes, densities, and distributions. The x-axis denotes the density of potential passengers (boundary condition), while the y denotes the initial density of vehicles.

threshold shows that the system behavior's sensitivity to the distance is discrete. It also makes sense in the real world, as people are indifferent to small changes in their walking distance, e.g., no one would walk 100 meters to the station, but not 120 meters.

The last parameter is the number of neighbours sufficient for a passenger to get interested in being matched with a vehicle ($NTST$). Increasing this threshold (i.e., more neighbours are required, stricter threshold) makes the transition area slightly wider but significantly higher, and vice verse. This means, while people's tendency to use the system may play a crucial role in a faster potential success of the system (lower transition area), it is less critical to the certainty of the system's success (slight changes in the shape).

Overall, the more relaxed any of the thresholds is the wider or the lower the transition area becomes. Furthermore, there is always a threshold at which, regardless of the density, the system is always successful (e.g., $AT = 400meter$) and one at which the system cannot succeed (e.g., $NTST = 10$).[§]

Real-world Scenarios

According to Figure 4, the population distribution in this area is similar to a normal distribution, thus, its heat map is expected to follow the trend from the 25 km^2 cases in Figure 3 and have a more deterministic and narrower transition area than the one in 25_NU case, which is assumed to be the most

[§]The results of all sensitivity analyses will be made available online as supplemental material.

similar theoretical scenario. Figure 5 shows the success heat map of Yarra Ranges scenarios next to the one from 25_NU case, which fulfills the expectation.

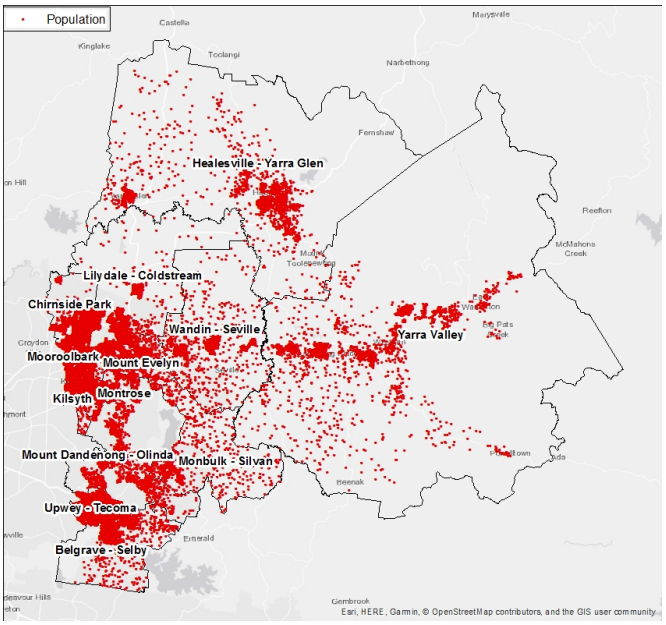


Figure 4. Population distribution of Yarra Ranges demonstrating 10% of the whole population

However, more detailed analysis is required for a better understanding of the system's behaviour towards super-criticality. Thus, we visualise and investigate the super-criticality in a number of suburbs in Yarra Ranges.

We have run a number of Yarra Ranges scenarios up to 1000 iterations to be able to provide a better understanding of the system's behavior in the longer run and allow time for changes to happen. We present the results from Yarra Ranges scenario with initial vehicle density of 60 per km^2 and 10% of the population (these numbers are chosen as an example, simulations of other combinations show the same level of consistency with the theoretical expectations) through iterations in Figure 6. The passengers' utilities are aggregated and illustrated in hexagonal cells.

Mooroolbark, Upwey-Tecoma, and Kilsyth, in this order, with population densities over 100 people per km^2 and area sizes between 8 and 13 km^2, are expected to have relatively similar results, as they all fall in the UR area of their most similar theoretical case, 9_UU. However, Figure 6 reveals that on an aggregated level the passengers utilities in Upwey-Tecoma are significantly lower than in Mooroolbark and Kilsyth. Moreover, it shows that Mooroolbark becomes super-critical in the early iterations (100) and starts infecting the surrounding suburbs¶, while Upwey-Tecoma stays in the same stages of success as the early iterations. Although the process of infecting the neighbouring slows down as it gets closer to the boundary of the high density areas, it does not stop as there is still connections to the neighbouring areas.

This does not mean that there is no one with high utility in Upwey-Tecoma. The system in this suburb behaves as expected from graphs in Figure 3 and contains a number of passengers that are interested and has been matched with vehicles around them. However, since the number is not as high as in Mooroolbark and the passenger's density is relatively high, the average stays low and in the blue spectrum.

Discussion

In this work, first, we investigated the impact of different spatial characteristics of an area on the critical mass frontier in an artificial environment, which allows for a strict control of the parameters. Further, to evaluate the theoretical results, we implemented our system in a real-world scenario.

The results from the artificial environment demonstrated the importance of spatial characteristics of an area in forming the critical mass frontier for a spatial two-sided market. We showed that a densely populated spatial core significantly impacts on the critical mass frontier and helps any system become viable much faster. In all NN cases, the LL areas are significantly smaller than their counterparts in other cases. This means that the UR area is bigger and includes more $\vec{\rho}(x, y)$, resulting in more certainty of the system's success.

The outcome of the real-world scenarios strongly supported the results from the theoretical experiments and further provided insights on the impact of neighbouring areas. In line with our expectations, the results demonstrated that if there are a number of areas in the UR region of their relative graphs, they all become super-critical together and start infecting the neighbouring areas. This process slows down or stops at the border of these areas and may advance again with changes in pricing or other characteristic of the system, which is beyond the scope of this work to illustrate and investigate.

As a consequence of our experiments, one can conceptually delineate a spatial region into areas that are "potentially super-critical" and other ones that are not:

- Areas inside such a potentially super-critical area *and* connected to an urban core can expect to eventually be served (e.g., Kilsyth and Monterose).
- Areas inside such a potentially super-critical area but not connected to an urban core may eventually be served, or not (e.g., Upwey-Tecoma).
- Areas outside the potentially super-critical area will not be served, except when the overall costs for the service become lower.

This has, in fact, important consequences for policy, in particular for regulation. When left to itself, super-critical areas will be served by commercial companies, while sub-critical areas will not. The surplus of the super-critical areas will, depending on the competitive situation, either go to the suppliers or to the customers; and it is in fact quite plausible that it will go to the suppliers since platform markets are not highly competitive, since they are difficult to invade Katz and Shapiro (1985). Sub-critical areas would, in consequence, not be served at all, or only with taxpayer subsidies.

An alternative scenario would be to give out regulatory licenses that would force a licensee to serve the complete market, including the sub-critical areas. Surplus on super-critical areas would thus be used to cross-subsidize

¶An animation of the changes will be made available online as supplemental material.

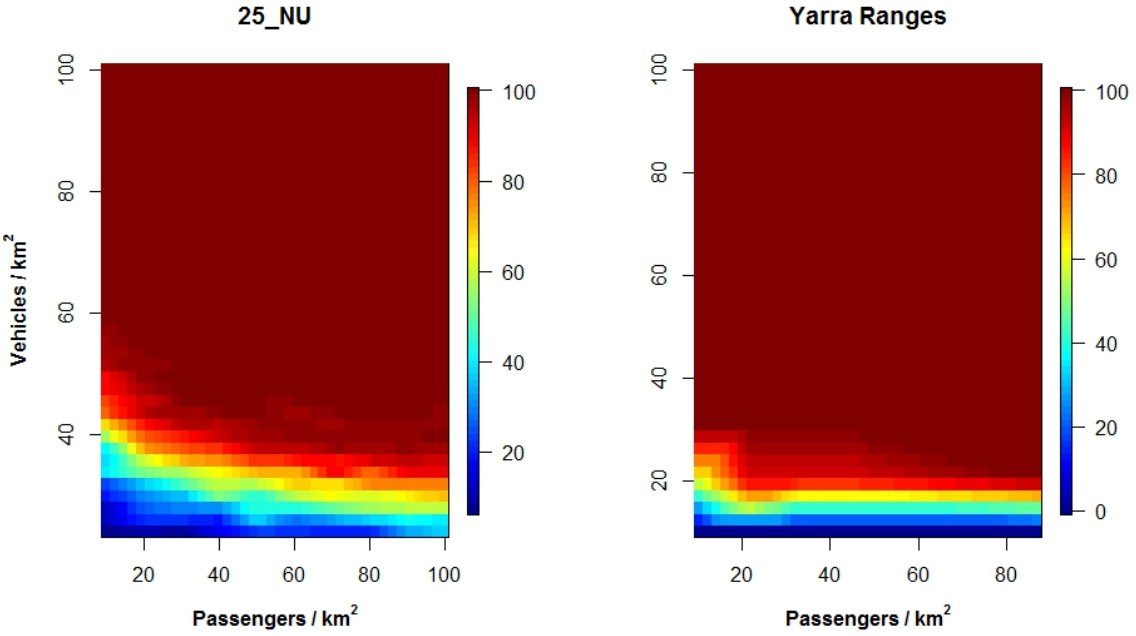


Figure 5. System's success probability in Yarra Ranges compared to 25_NU scenarios

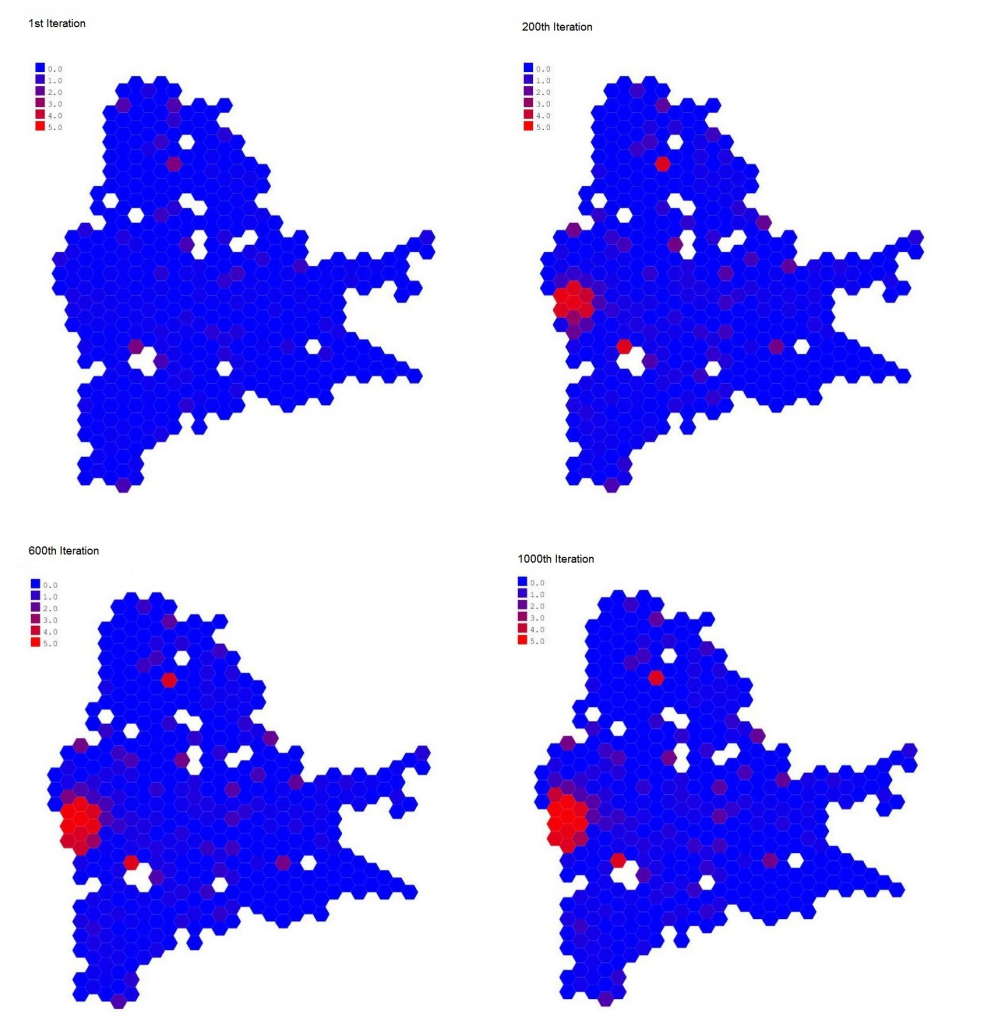


Figure 6. The changes in passengers' utilities through iterations, spreading high utility from the dense center to the surroundings. Also note that spreading eventually stops, confirming that there are areas that cannot be made super-critical even when they have a super-critical neighbor.

sub-critical areas. Clearly, this will make the super-critical areas either more expensive to customers, or less profitable for suppliers; presumably the latter. It would, however, also either decrease the taxpayer burden or improve services to sub-critical areas. Clearly, this is a regulatory intervention. It is, however, an intervention that is quite standard in infrastructure markets (water, electricity, telecommunications), where the connection cost is regulated to be the same for all locations. Thus, a public debate is needed if certain mobility services should be treated in a similar way. This debate is even more necessary since dynamic transport systems bear the promise of being a much better system for sparsely populated areas than the current systems, and thus the question becomes if a society wants to realize these promises or not.

It also becomes clear that in strongly super-critical areas, prices will be such that they are either much higher than cost, or much smaller than the users' willingness to pay. It may thus make sense to use a strategy also used in other areas of infrastructure engineering, which is to enforce a more uniform service provision by giving out licenses. For example, in water, electricity or telecommunication, it is often normal that the connection fee is the same for everybody, no matter where they are located. This effectively spreads revenues from relatively-easy-to serve densely populated areas to more-expensive-to-serve less densely populated areas, and forces suppliers to supply at those prices even in areas where these do not cover costs. In Germany, such schemes are enforced by a certain interpretation of its constitution, which demands "equality of living conditions" across the country. Other countries may have similar norms in place; such schemes may also be seen as a measure to help spread economic and technology gains from the urban centers to other areas. Australia has no comparable legislation.

The mechanism to achieve this would thus be to give out licenses for regions that include urban cores but also rural areas (e.g. for all of Victoria). At the same time, one would demand minimum service standards across the full region. Platform providers could, for example, set incentives for suppliers in rural areas in order to reach them. If the precise details of such a license cannot be determined by government, it could do what it also does in other infrastructure markets: the licenses could be auctioned off.

Conclusion

Critical mass frontier is the border of success and failure in a two-sided market. Its identification prior to implementation is highly crucial for any platform in such market. The main hypothesis of this paper is that unlike in non-spatial markets, e.g., media sharing, news media, and credit card platforms, there are certain platforms, whose success highly depends on the location of their members.

Using simulations, we confirmed the hypothesis by demonstrating that spatial characteristics of an area, namely size, the density and distributions of demand and supply, have significant impact on the critical mass frontier of an implemented spatial platform based on two-sided market. Here, we aim at drawing overall insight on how the spatial characteristics of demand and supply may help or hinder an area achieve super-criticality, not at providing a concrete number as the required number of vehicles or passengers to create a self-sustained two-sided spatial market. Moreover, we illustrated that the critical mass frontier is almost never a deterministic line and there is always specific uncertainty in achieving success.

This work is the first step towards developing a framework for identifying critical mass frontier in spatial two-sided markets. One main shortcoming of this work was the lack of time consideration in matching vehicles and passengers, which makes the results more optimistic than reality. However, since the focus of this work is neither optimizing the dispatching algorithm nor finding specific solutions for specific condition this shortcoming has been neglected.

In future work, it is necessary to make more realistic assumptions, e.g., considering a street network, time dependency of demand and supply, and build a more comprehensive simulation platform that allows for investigating more specific scenarios.

References

Armstrong M (2006) Competition in two-sided markets. *The RAND Journal of Economics* 37(3): 668–691.

Australian Bureau of Statistic (2016) Yarra Ranges (local government area), population profile. Extracted on 08 December 2017.

Batty M (2006) *Hierarchy in Cities and City Systems*, *Methods Series*, volume 3. Berlin: Springer, pp. 143–168.

Batty M (2009) *Cities as Complex Systems: Scaling, Interaction, Networks, Dynamics and Urban Morphologies*. New York, NY: Springer, pp. 1041–1071.

Batty M and Longley P (1994) *Fractal Cities – A Geometry of Form and Function*. London: Academic Press Ltd.

Bento AM, Cropper ML, Mobarak AM and Vinha K (2005) The effects of urban spatial structure on travel demand in the United States. *The Review of Economics and Statistics* 87(3): 466–478. DOI:10.1162/0034653054638292.

Caillaud B and Jullien B (2003) Chicken & egg: Competition among intermediation service providers. *RAND journal of Economics* : 309–328.

Chandler D (1987) *Introduction to modern statistical mechanics*. New York: Oxford University Press. ISBN 9780195042771.

Chrisman N (2012) *A deflationary approach to fundamental principles in GIScience*. gisci-concepts.org: CC BY-NC 3.0, pp. 41–64.

Christaller W (1933) *Die zentralen Orte in Süddeutschland*. Jena: Gustav Fischer.

Claramunt C and Winter S (2007) Structural salience of elements of the city. *Environment and Planning B* 34(6): 1030–1050.

Djavadian S and Chow JY (2017) An agent-based day-to-day adjustment process for modeling 'mobility as a service' with a two-sided flexible transport market. *Transportation Research Part B: Methodological* 104: 36–57.

Dubé JPH, Hitsch GJ and Chintagunta PK (2010) Tipping and concentration in markets with indirect network effects. *Marketing Science* 29(2): 216–249.

Enoch M, Potter S, Parkhurst G and Smith M (2006) Why do demand responsive transport systems fail? In: *85th Annual*

Meeting of the Transportation Research Board. Washington DC, USA.

Evans DS and Schmalensee R (2016) *Matchmakers: The New Economics of Multisided Platforms*. Cambridge, MA: Harvard Business Review Press.

Fricker C and Gast N (2016) Incentives and redistribution in homogeneous bike-sharing systems with stations of finite capacity. *EURO Journal on Transportation and Logistics* 5(3): 261–291. DOI:10.1007/s13676-014-0053-5.

Goodchild MF (2009) What problem? spatial autocorrelation and geographic information science. *Geographical Analysis* 41(4): 411–417.

Goodchild MF (2011) Challenges in geographical information science. *Proceedings of the Royal Society A* 467(2133): 2431–2443.

Grajek M and Kretschmer T (2012) Identifying critical mass in the global cellular telephony market. *International Journal of Industrial Organization* 30(6): 496–507.

Hua H, Xie H and Tanin E (2018) Is euclidean distance really that bad with road networks? In: *11th ACM SIGSPATIAL International Workshop on Computational Transportation Science*. ACM, pp. 11–20.

Ising E (1925) Beitrag zur Theorie des Ferromagnetismus. *Zeitschrift für Physik* 31(1): 253–258.

Jiang B (2009) Street hierarchies: A minority of streets account for a majority of traffic flow. *International Journal of Geographical Information Science* 23(8): 1033–1048.

Jiang B and Brandt S (2016) A fractal perspective on scale in geography. *ISPRS International Journal of Geo-Information* 5(6): Paper 95.

Katz M and Shapiro C (1985) Network externalities, competition, and compatibility. *American Economic Review* 75(3): 424–40.

Kazerani A and Winter S (2009) Modified betweenness centrality for predicting traffic flow. In: Lees B and Laffan S (eds.) *10th International Conference on GeoComputation*.

Kodera T (2015) Discriminatory pricing and spatial competition in two-sided media markets. *The BE Journal of Economic Analysis & Policy* 15(2): 891–926.

Müller K, Schulze C and Stauffer D (2008) Inhomogeneous and self-organized temperature in Schelling-Ising model. *International Journal of Modern Physics C* 19(03): 385–391. DOI:10.1142/S0129183108012200. URL https://www.worldscientific.com/doi/abs/10.1142/S0129183108012200.

Rochet JC and Tirole J (2003) Platform competition in two-sided markets. *Journal of the European Economic Association* 1(4): 990–1029.

Sulopuisto O (2016) Why Helsinki's innovative on-demand bus service failed. URL http://citiscope.org/story/2016/why-helsinkis-innovative-demand-bus-service-failed.

Tobler W (1970) A computer movie simulating urban growth in the detroit region. *Economic Geography* 46(2): 234–240.

Wang X, He F, Yang H and Gao HO (2016) Pricing strategies for a taxi-hailing platform. *Transportation Research Part E: Logistics and Transportation Review* 93: 212–231.

Wilkinson D and Willemsen JF (1983) Invasion percolation: a new form of percolation theory. *J. Phys. A: Math. Gen.* 16: 3365–3376.