Determining the optimal municipality for a new metro stop in Medellín

Alex Saad[a,1]

[a]Department of Mathematics, University of Oxford

April 5, 2019

We analyse the transport network of the Aburrá Valley to determine a proposed location for a new metro stop. We compare data on commutes, demographics and infrastructure from 2017 (and predictions for 2018) with a modified version of a single-constrained gravity model for mobility. The model error is significant in several instances that lead us to conclude that the new station should be built in the Barbosa municipality.

Mobility | Gravity model

The Colombian government has announced funding for a new metro station in the Aburrá Valley, Colombia's second-largest metropolitan area, which is divided into ten municipal zones. There are four municipalities in Medellín with no metro stops, and the new station is intended to be built in one of them.

We use data on demographics and commutes at both the municipality and comuna level, as well as a gravity model of mobility, to determine that the new metro stop should be built in Barbosa, the northernmost municipality.

Recommendation

We recommend that the new metro station should be built in the Barbosa municipality, in its "Urbana Barbosa" comuna. It should attach to the northern end of the metro A-line by connecting with the metro stop "Niquía" in Bello.

Background

The city of Medellín lies within the metropolitan area of the Aburrá Valley, which is split into ten municipalities, displayed in Fig. 1. In the rest of this report we will refer to municipalities by indices 1-10; listed in order of indices, they are Medellín (1), Barbosa (2), Bello (3), Caldas (4), Copacabana (5), Envigado (6), Girardota (7), Itagüí (8), La Estrella (9) and Sabaneta (10). The Medellín municipality is the urban centre. Each of these ten municipalities is further subdivided into "comunas" (neighbourhoods), with a total of 66 comunas across the whole valley.

The Medellín Metro covers large regions of the Aburrá Valley, with its core located within municipality 1. However, there are four municipalities with no metro stops: Barbosa, Caldas, Copacabana and Girardota. We will refer to the set of these municipalities as S_0. It has been decided that the funding for the new metro stop will be allocated to a municipality in S_0, and we attempt to provide some justification for which municipality should be chosen.

The Medellín metro. Medellín's metro system spans 6 of the 10 municipalities in the Aburrá valley across underground, tramway, cable car and bus rapid transit (BRT) systems. The longest metro line is the A-line, connecting Niquía in the Bello municipality in the north of the valley with La Estrella in the south. Much of the rest of the metro network is centered in, or passes through, the central Medellín municipality. A subway map of the metro system is displayed in Fig. 2.

Fig. 1. A map of the Aburrá Valley indicating its partition into municipalities.

Relevant municipality demographics. Here we discuss some of the relevant features of several different municipalities for our analysis. A breakdown of populations in each municipality can be found in Table 1, and a selection of other statistics in Fig. 3.

Medellín and Bello – urban centres. Medellín municipality has the greatest population of all the municipalities, and is the most urbanised with only a 1% rural population. Bello is the second largest city in the Aburrá valley after Medellín, and is similarly

Significance Statement

The gravity model is one of several models of mobility. Commute flows between two locations are modelled by assigning factors to zones simulating their "gravity" e.g. population, employment opportunities, and cost of travel. We use a gravity model to analyse the transport network of the Aburrá Valley and compare the predictions with observed commute data.

[1]E-mail: alexander.saad@st-hildas.ox.ac.uk

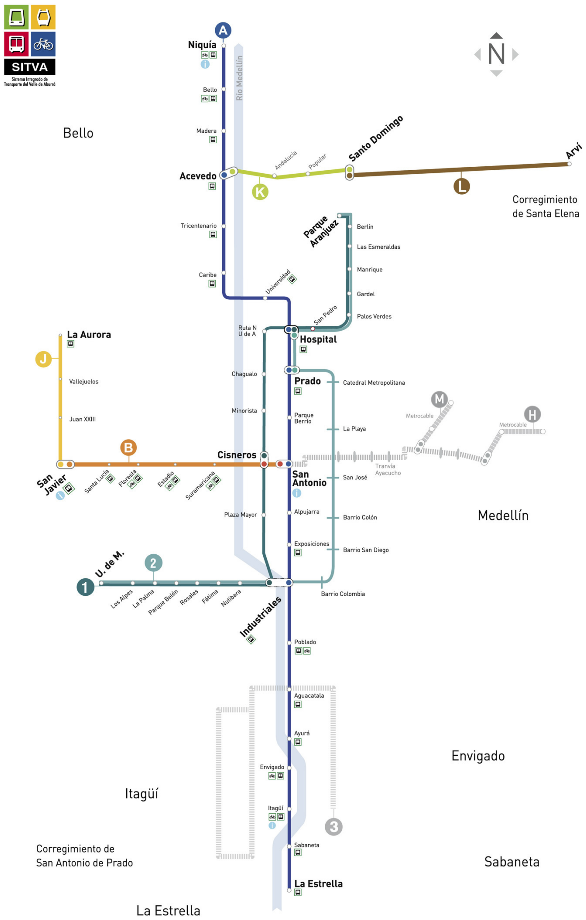


Fig. 2. The Medellín metro system.

Table 1. Population breakdowns for each municipality.

Municipality	Total	Urban	Rural	% Rural
Medellín	2219861	2183557	36304	1
Barbosa	42547	18721	23826	56
Bello	373013	359404	13609	3
Caldas	68157	52632	15525	23
Copacabana	61421	53033	8388	14
Envigado	175337	166742	8595	5
Girardota	42818	25195	17623	41
Itagüí	235567	213297	22270	9
La Estrella	52763	28538	24225	46
Sabaneta	44874	35528	9346	21

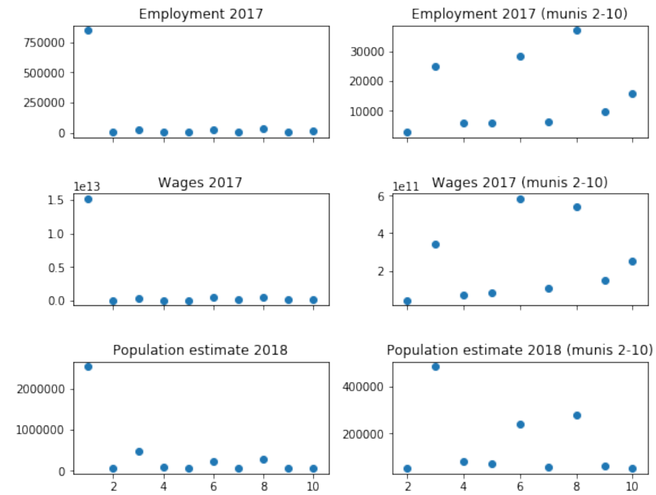


Fig. 3. Demographics for municipalities. The x-axis is the municipality index. Since municipality 1 is so dominant, the right hand column shows the same scatter plots for just the remaining municipalities 2-10.

highly urban. Both municipalities are more economically developed than any of the municipalities in S_0, with significantly higher employment rates and wages.

Barbosa, Girardota and La Estrella – rural areas. The municipalities of Barbosa, La Estrella and Girardota are, in decreasing order, the most rural, with 56% of Barbosa classed as rural. La Estrella is connected to the metro system at the south of the A-line, and has slighly higher wages and employment than any of the municipalities in S_0. This could be indicative of a higher accessibility value in La Estrella due to its closer proximity to the metro system.

Barbosa has the lowest wages of any municipality, and commute data shows that it has high intra-municipality commute rates. This is possibly due to it being an agricultural area with many people working close to where they live. It also has very high commute rates to Bello municipality.

Girardota also has relatively low wages and employment rates and is 41% rural. However, unlike Barbosa, it does not display large commute numbers to any particular municipality.

Methods

Our decision that the new station should be built in Barbosa is based on several factors. First, we construct a municipality-level origin-destination matrix M that indicates that there are significantly more commutes from Barbosa and Copacabana into Bello than any commutes between other municipalities, despite neither Barbosa or Copacabana having a metro station. This provides two good candidates for municipalities in need of a metro station.

Then we compare this with a modification of a single-contrained gravity model for mobility. The gravity model is a reasonably good estimator for commuter numbers generally, but has several significant model errors. We discuss these and attempt to draw useful conclusions from them.

Modified gravity model. We applied a modified version of the singly-constrained gravity model for mobility to estimate the number T_{ij} of commutes from municipality i to municipality

j. The model estimates this quantity via

$$T_{ij} = Z_i P_i E_j e^{-\beta c_{ij}}, \quad [1]$$

where

- P_i is the population of municipality i (more generally this can be any measure of *demand* in the origin zone). We use the predicted 2018 population data in order to most closely fit employment and commuter data;

- E_j is the number of employment opportunities in municipality j (more generally this can be any measure of *supply* in the destination zone). We use the 2017 employment numbers for each municipality;

- c_{ij} is the driving time * to get from municipality i to municipality j (more generally this can be any measure of *cost* from travelling from zone i to zone j). We use the drive time (in seconds) between comunas, measured at 9am on a weekday, and then compute c_{ij} as the average drive time over all comunas within the two municipalities;

- Z_i is a normalisation constant, depending on the municipality;

- β is a parameter that needs to be chosen to best fit the data.

Let C_i denote the number of people in municipality i that commute, so that $0 \leq C_i \leq P_i$. By imposing the condition

$$\sum_j T_{ij} = C_i \quad [2]$$

on the model predictions, we can determine Z_i as

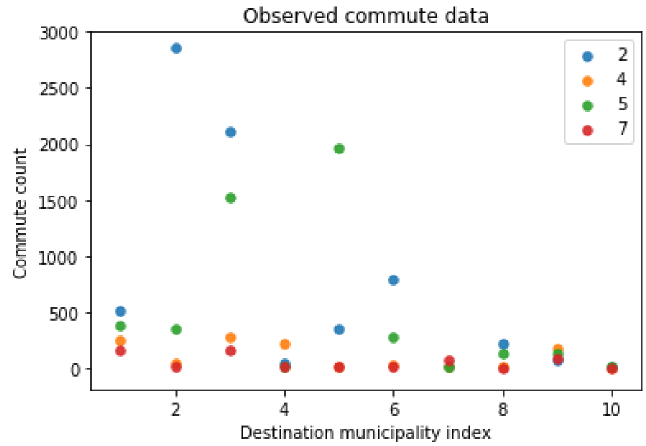
$$Z_i = \frac{C_i}{P_i} \frac{1}{\sum_j E_j e^{-\beta c_{ij}}}.$$

We use Eq. (2) instead of that defined by Piovani et al. [1], where $\sum_j T_{ij} = P_i$, because in this case the fraction C_i/P_i of the population in each municipality that commute to work is very low. The predictions from this "modified gravity model" are much closer to the actual data, whereas using the gravity model in [1] produces large overestimates for T_{ij}.

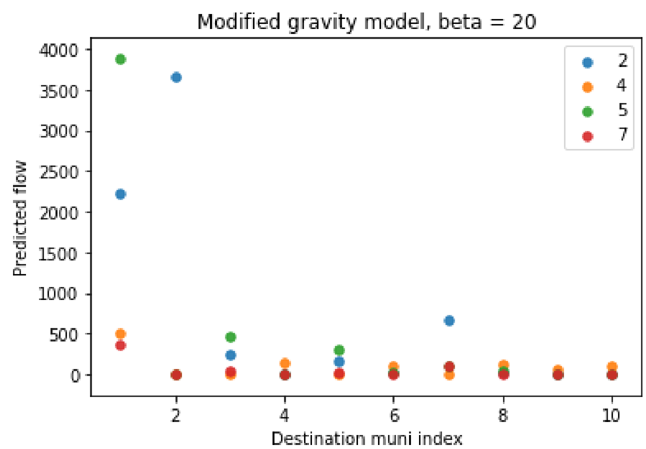
Given this modification in the model, we plot values of T_{ij} as functions of β, which are displayed in Fig. 5. We omit the values T_{ii}, which gradually stabilise to relatively high values as β increases (since the cost c_{ii} is very small in this case), as well as the values T_{i1} which are very large for small values of β. Each curve has a maximum – some of these maxima occur with $\beta < 0$ and then become very small as β increases, but those that have a displayed maximum have it in the region $10 < \beta < 50$.

For our model we choose a value of $\beta = 20$; this prevents some of the values T_{ij} from underpredicting M_{ij}, and seems to most closely fit the observed data. A comparison between the gravity model predictions with $\beta = 20$ and the observed data is given in Fig. 4.

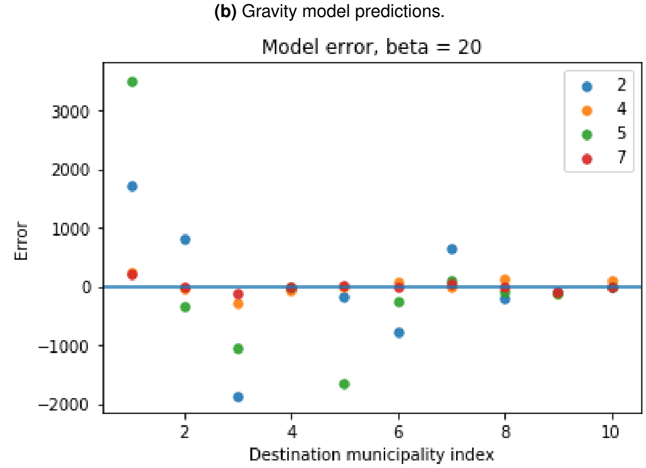
*We chose drive time as a measure of cost as it is the easiest function to implement that also takes into account some of the infrastructure.



(a) Actual commute data for each of the municipalities in S_0 into all other municipalities.

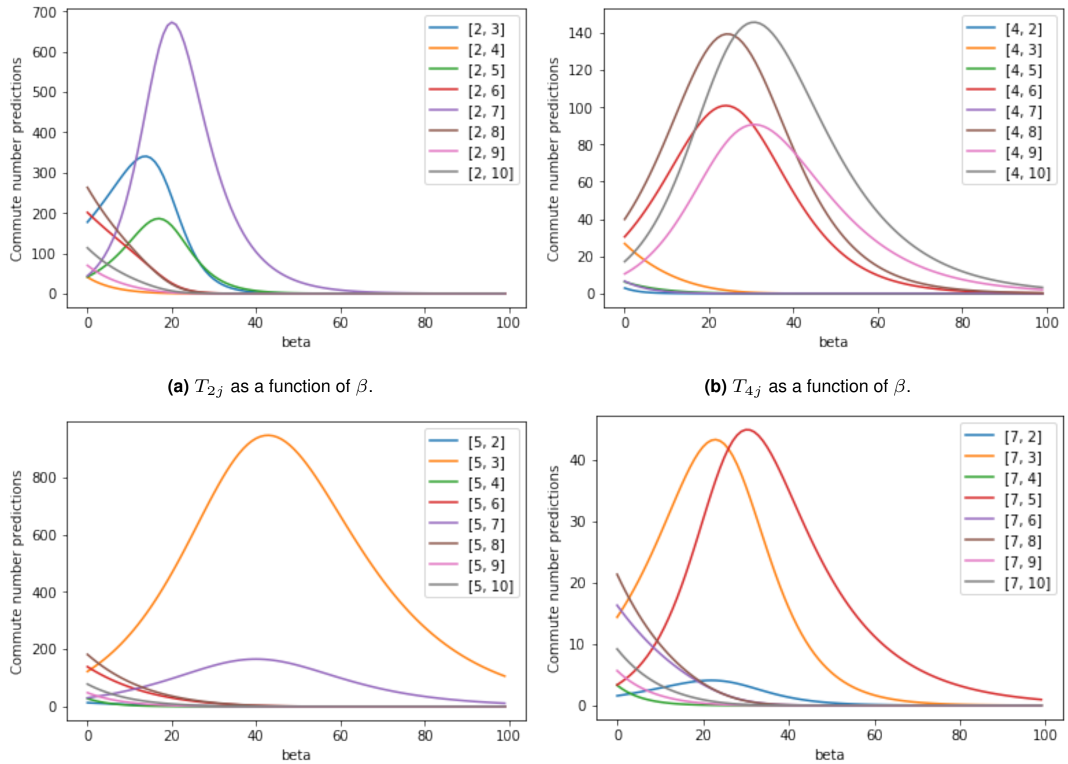


(b) Gravity model predictions.



(c) The model error.

Fig. 4. Plots of the observed number of commutes M_{ij}, gravity model predictions T_{ij} with $\beta = 20$ and model error $T_{ij} - M_{ij}$ with $\beta = 20$.



(a) T_{2j} as a function of β.

(b) T_{4j} as a function of β.

(c) T_{5j} as a function of β.

(d) T_{7j} as a function of β.

Fig. 5. Plots of the variation of T_{ij} as a function of β for $i \in S_0$ and $2 \leq j \leq 10$. Note that we have omitted the plots for T_{i1} and T_{ii} as these are much larger quantities.

Municipality-level origin-destination matrix. From the available data on comuna-level commutes, we construct a comuna-level origin-destination matrix A. This is a 66×66 matrix with integer entries where A_{ij} is the number of commutes from comuna i to comuna j.

By the assignment of each comuna to a municipality, we obtain from this a municipality-level origin-destination matrix M. This is a 10×10 matrix with integer entries where M_{ij} is the number of commutes from municipality i to municipality j. The entries M_{ij} for $i \in S_0 = \{2, 4, 5, 7\}$ and $1 \leq j \leq 10$ are plotted in Fig. 4a.

This commute data covers all methods of transport between areas – in particular, it takes into account the metro system.

Analysis

The gravity model predictions T_{ij} for $i \in S_0$ and $1 \leq j \leq 10$ are displayed in Fig. 4b, and the actual observed commute numbers M_{ij} are shown in Fig. 4a. The model error $T_{ij} - M_{ij}$ is displayed in Fig. 4c. The model error is generally fairly close to zero. However, there are some notable discrepancies between the model and the data, discussed in the following subsections.

Spikes at M_{22} and M_{23}. Fig. 4a shows very high commuter numbers from Barbosa into municipalities 2 (Barbosa) and 3 (Bello): $M_{22} = 2845$ and $M_{23} = 2110$. These values are much larger than any of the other entries except for M_{53} and M_{55}.

The high value of M_{22} indicates that a large number of people living in Barbosa commute completely within Barbosa. This could be explained by the fact that Barbosa is a predominantly rural municipality; residents may be more likely to work agricultural jobs where they stay in the rural local area. Additionally, those working in the urbanised centres may choose to live in a more central municipality equipped with a metro stop to commute into work.

The gravity model with $\beta = 20$ overpredicts a high value of $T_{22} \approx 3655$, giving a model error of ≈ 810. This is likely due to the high value of $\beta = 20$ we have chosen, which largely concentrates commute flows within municipalities.

The value M_{23} is somewhat different, and indicates that the number of commuters making the trip from Barbosa to Bello is still much higher than any other commute rates from municipalities in S_0 (except for those already captured by M_{22}). This is significant as it shows that Barbosa, despite being the most geographically isolated municipality, is the main contributor to commute flows into an urban centre.

Bello had reasonably high employment and wages in 2017 compared to other municipalities, so it is likely to offer more opportunities for those living in Barbosa; although it does not have the highest employment rates or wages (see Table 1) its additionally large population suggests that there may be opportunities there for low-to-moderately paying work that would be an incentive to commute from Barbosa.

Bello is also the closest municipality to Barbosa that is connected to the Medellín metro system. This means that many of these commutes could be made in order to access the metro and travel to other parts of the Aburrá valley.

The gravity model, however, does not predict a significant number of commutes from Barbosa to Bello, with $T_{23} \approx 189$.

By varying the parameter β (displayed as the blue curve in Fig. 5a) we can increase the value of T_{23}, but only to a maximum value of ≈ 340, attained at $\beta \approx 14$. This is significantly less than the actual value $M_{23} = 2110$. This suggests that the opportunities for Barbosa residents in Bello encompass more than its available jobs, and could be due to its metro connection.

Spikes at M_{53} and M_{55}. We observe similar phenomena with very large commute numbers from Copacabana to Bello ($M_{53} = 1527$) and within Copacabana ($M_{55} = 1971$). Copacabana is 14% rural, and had relatively low employment and wages in 2017. It is adjacent to Bello, which could partially account for the higher commute numbers there.

This proximity is detected in part by the gravity model since the cost c_{53} is relatively low and the population of Bello is large, but the model always produces underestimates for this quantity. Indeed, $T_{53} \approx 465$ when $\beta = 20$ and achieves a maximum of ≈ 945 at $\beta \approx 42$ as indicated by the orange curve in Fig. 5c.

This suggests that there could be more reason to commute from Copacabana to Bello than just accessibility of available employment there. This could again be explained by the existence of the metro station in Bello.

No spike at M_{i1} for any $i \in S_0$. The gravity model predicts very high commute flows from Barbosa and Copacabana into Medellín, with $T_{21} \approx 2220$ and $T_{51} \approx 3880$. This is easily explained – Medellín has a very large gravity in the model, and the costs of travelling to Medellín are relatively low from these two municipalities.

Despite this, we do not see large numbers of commutes into Medellín from any municipality in the observed data, and the commute flows from Barbosa and Copacabana into Bello are significantly higher.

Decision to place new station in Barbosa

The fact that we do not see large commute numbers from Barbosa and Copacabana into Medellín could suggest unmet demand in the central municipality. Indeed, if there is already a significant cost in those municipalities commuting to Bello, a more nearby urban centre, then commuters may be less likely to travel to similar jobs further afield in the Medellín municipality. It could also suggest that the increase in opportunities in going from Bello to Medellín may be more relevant to higher-paying jobs; to the relatively poor municipalities Barbosa and Copacabana these opportunities may not be as accessible, and so the benefit of travelling to a more urbanised area not as large.

It seems appropriate to allocate the new metro stop to either Barbosa or Copacabana, given their prolific commute numbers into Bello. However, these two municipalities are also separated by Girardota, which could also be an intermediary option. Determining which municipality to place the stop in therefore requires us to consider several local factors.

Firstly, placing the station in Girardota as an intermediate point between Barbosa and Copacabana does not seem to increase accessibility notably in any municipality. It would change the journey from Barbosa to Bello from a 48 minute drive to a 25 minute drive plus a metro journey, but it seems unlikely that Copacabana residents would travel to Girardota to use such a stop, or indeed a new metro stop in Barbosa; the driving time saved would require travelling further away from Bello to get to the metro, thus potentially making the journey more difficult.

By placing the new station in Copacabana, residents there would reduce the 25 minute drive to Bello to a shorter train ride. However, it would only reduce the 48 minute drive from Barbosa to Bello to a 33 minute drive from Barbosa to Copacabana, followed by a further train journey. Thus Barbosa would stand to benefit little from such an outcome, which would also have the effect of increasing traffic travelling into Copacabana.

Therefore it seems most reasonable to place the new station in Barbosa, in its "Urbana Barbosa" comuna, and connect it to the "Niquía" stop in Bello at the northern end of the A-line. This would reduce the traffic flow from Barbosa into Bello via the Girardota and Copacabana municipalities, potentially shortening their commute times to Bello as a by-product.

This is a feasible modification to make to the network. Since we only propose to extend one of the arms of the metro system by an extra stop, the topology of the network is not changed. The results in Fig. 4a indicate that there will be high numbers of new commuters using the new line from Barbosa to Niquía. If there is additional funding available in the future, a second new stop could be built on this new stretch of line at Copacabana.

(1) Piovani, D., Arcaute, E., Uchoa, G., Wilson, A., Batty, M. (2018). Measuring Accessibility using Gravity and Radiation Models. *Royal Society Open Science*, 5(9): 171668.