Appendix:

Methodology

Data Cleaning, Manipulation, and Creation of the Working Dataset

This analysis starts with the collection of 2016 Census data from Statistics Canada at the Dissemination Area (DA) level; however, the analysis throughout this paper is done at the Traffic Zone level.[1] Traffic zones are the geographic areas in which the Transportation Tomorrow Survey (TTS) is done, similar in size to Das, i.e., 400700 people.[2]

We assign each DA either fully, or partially, to a traffic zone, depending on their intersection.[3] In some cases, DAs are larger than, and cover the entirety of, a traffic zone; in this case, the traffic zone takes on 100 percent of the DA population and accompanying characteristics. In many instances, DAs do not belong to only one traffic zone, and instead overlap with several. In this situation, the share of a DA that overlaps a traffic zone is assigned to said traffic zone.[4]

We then match DA-level Census data to each traffic zone and DA pairing. Census variables that are population counts are multiplied by how much of their DA lies in a given traffic zone.[5] Note that in order to account for population growth, we multiply Census population counts in each DA by the estimated population growth rate between 2015 and 2019 for the CMA in which they reside.[6] Using the population that each DA contributes to a zone, we calculate the total zone population, and then the share of a zone's population that each DA contributes to a traffic zone in order to create zone-level income figures, which are a weighted average of each contributing DA's income data. We collapse data observations by traffic zone, yielding estimated Census data for each traffic zone.

1 Census 2016 data is available for download at https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/details/download-telecharger/comp/page_dl-tc.cfm?Lang=E. This paper uses data found in the data file entitled "Canada, provinces, territories, Census divisions (CDs), Census subdivisions (CSDs) and dissemination areas (DAs) - Ontario only."

2 The Transportation Tomorrow Survey (TTS) is said to be "the largest and most comprehensive travel survey ever conducted in Ontario or perhaps anywhere in North America." It provides rich information about all members of a surveyed household and the travel details on the household over one entire weekday. More information about the TTS can be found at http://dmg.utoronto.ca/transportation-tomorrow-survey.

3 Traffic Zone Boundary, or Shape files, were obtained from the Data Management Group at the University of Toronto – the group responsible for the design and implementation of the TTS. Transportation Tomorrow Survey Traffic Zone Boundary/shape files are publicly available (after creating an account) at http://dmg.utoronto.ca/survey-boundary-files#tts.

4 DA and traffic zone shape files are in different projections, and as a result the re-projection of one layer in order to intersect the two layers in GIS leads to very minor differences in how the borders of polygons line up, which if in the same projection would have otherwise lined up. To counter these projection errors, two adjustments are made after the intersection of DAs and traffic zones. If the area of a DA is 99.5 percent or more within the bounds of a traffic zone, it is rounded up to 100 percent coverage; similarly, if area of a DA is 0.5 percent or less within the bounds of a traffic zone, it is rounded down to 0 percent coverage.

5 In other words, these are variables for which the total composition of a traffic zone can be found simply by adding the share of variable that corresponds to the share of how much a DA is resented by a traffic zone.

6 Population change estimates are calculated annually by Statistics Canada. The data file used to calculate the estimated population growth rates between 2015 and 2019 in this paper can be found at https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1710013601.

We calculate the orthodromic distance, or great-circle distance, between the centre points of 2947 "traffic zones" in the Greater Toronto Area (GTA), which are within 50 kilometers of each other.[7] 1.95 million unique distances result across this area. This file is referred to here as the Master Dataset.

Finding the Agglomeration Coefficient

Using the aforementioned file, we calculate the total surrounding labour force within walking distance – which we assume to be 1 kilometer – for each traffic zone.[8] We further collapse the data, summing population counts by zone, yielding a data set of 2,947 observations, each of which represents a traffic zone, the labour force within 1 kilometer and 50 kilometers, and several other control variables for said zone.

We regress average household income on surrounding labour force within 50 kilometers (in natural logarithmic form). We control for the share of population having obtained a post-secondary education, the share without education (including secondary school), unemployment, and the number of children and non-citizens. Given its log form, the coefficient on the labour force within 50 kilometers variable represents the elasticity of household income with respect to surrounding population; that is, the estimated coefficient on the Size of Labour Force can be interpreted as a percentage figure; the regression results are shown in Table A1. The coefficient of interest, 0.042, implies that a 10 percent increase in the size of the surrounding labour force increases a household's income by 0.42 percent. This estimate is referred to as the Agglomeration Coefficient in the remaining text.[9]

This estimate is within the range found by existing literature from studies around the world that use a similar methodology to that used here. Rosenthal and Strange (2004) survey the existing literature that studies the relationship between city size and productivity, finding that work done across the globe places this elasticity in the range of 0.04-0.11. In Spain, de la Roca and Puga (2017) find an elasticity of 0.0489. In France, Combes et al. (2010) conclude this elasticity to be 0.05. Estimates from the United Kingdom and United States (Rice et al. 2006 and Glaeser and Ressenger 2010, respectively) yield slightly lower estimates, most similar to that of this paper, of about 0.04.

Part 1: The Wider Economic Benefits of the TTC

The core of this analysis relies on data from the 2016 Census, 2016 TTS, and the agglomeration

7 Orthodromic, great-circle, or spherical distance is the distance between two points on a sphere as measured along the surface of the sphere, in this case Earth. This is in contrast to the linear distance which in Euclidean space would take the distance between two points in a straight line, ignoring the curvature of the Earth. The resulting distance would calculate the distance between two points going through the Earth's interior. Instead, the great-circle distance takes into consideration the radius of the Earth and calculates distance as they are if one could travel in a straight line along the surface.

8 The 2016 TTS puts this average at 1.2 kilometers.

9 Previous studies (Combes et al. 2010 and Rosenthal and Strange 2004) use instrumental variables of a region's historical population and geology (that should have no effect on today's income) to determine the later region population count that affects regional agglomeration economies. These studies find that the causality we assume here is the correct interpretation of results.

Table A1: Regression Results

	(1)
Regressors	Average Household Income (Log)
Size of Labour Force (Log)	0.0420*** (0.01)
Share of Population With a Bachelor's Degree or Higher	0.0782*** (0.01)
Share of Population Without a Secondary-school Education	-0.194*** (0.01)
Unemployment Rate	-0.0410*** (0.00)
Number of Children	0.000224*** (0.00)
Number of Constituents Without Canadian Citizenship	-0.000289*** (0.00)
Constant	10.75*** (0.13)
Observations	2,842
R-squared	0.417

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors calculations from sources listed in appendix.

coefficient found as above.[10] In order to determine the wider economic benefits of TTC services in the Toronto area, we consider the population surrounding each traffic zone under two comparative conditions. First, we consider only the population of surrounding traffic zones which are within a zone's average trip distance when travelling by foot, which averages 1.2 kilometers across traffic

10 The coefficient from the agglomeration estimate enters our equation linearly. One can easily prorate the benefits based on reasonable estimates of the upper and lower bound of what the agglomeration coefficient should be. Our 95th percentile lower end of estimated benefits would be 61.9 percent of our estimates and the upper bound of estimated benefits would be 138.2 percent of our estimates.

zones; let this be called the walking population.[11] Second, we consider the population of all traffic zones that are within a zone's average trip distance when taking public transit, which averages 17.3 kilometres; call this the transit population.

Using these two population counts, we calculate the change in effective surrounding population as the transit population minus the walking population, divided by the transit population. We calculate the total wider economic benefit presented to each traffic zone by the TTC's existence as:

*Eqn,(1) WEB = Change in Effective Surrounding Population * Zone's Population * Average Zone Income * Agglomeration Coefficient*

We prorate this benefit by the proportion of each traffic zone's population that takes transit, yielding the wider economic benefits estimate experienced by each zone as a result of the availability of the TTC.[12]

We account for transit users switching to other modes of travel if suddenly no transit system were available by using the findings of several studies which analyze how transit users react—through changes in their primary mode of transport—to public transit strikes. Nguyen-Phuoc et al. (2018) explore the impact of public transit strikes in Melbourne, Australia, which affected services identical to that of the TTC – subway trains, streetcars, and bus services. These strikes completely shut down all transit services across the city. They survey transit users a few months following the strikes, finding that on average, 31 percent of those that used transit had decided to instead drive during the service cancellations. We multiply the share of population in each traffic zone that uses the TTC by 31 percent and calculate the wider economics benefits as in equation 1, where *Change in Effective Surrounding Population* is equivalent to the population accessible by car less the population accessible by walking, divided by the population accessible by car. Average trip distance by car is about 12.5 kilometers across all traffic zones.

Similarly, Fuller et., al. (2019) investigate how transit strikes in Philadelphia impact the use of the City's public bike-share program. They carry out a difference-in-differences methodology to identify the increase in bike-share use in comparison to other US cities (Washington, Boston, and Chicago). They find that bike-share use increased 57 percent during the strikes, equivalent to about 92 new users per 100,000 population. We prorate by 0.00092 the population of each traffic zone, which, on average, estimates four new cyclists in each traffic zone. As above, the wider economic benefits of bike use are calculated using equation 1 where *Change in Effective Surrounding Population* is equivalent to the population accessible by bike less the population accessible by walking, divided by the population accessible by bike. Average trip distance by bike is about 3.5 kilometers. Both bike and car trip distances were collected from the 2016 TTS.

Part 2: How TTC Service Improvements Provide Wider Economic Benefits to the GTA

Part 2a: Service Improvements, New Services, and Line 5 and 6 LRT's

The bulk of new WEBs from the TTC are the result of new services – namely the Eglinton LRT, Finch West LRT, and new Express bus routes – as well as more frequent service for some existing bus services. These service improvements increase

11 Using the 2016 Toronto Transit Survey (TTS), we determine the average distance of trips by foot, car, public transit, and all means of travel, by zone.

12 This is again obtained from the 2016 TTS.

the speed at which commuters can travel. For instance, the Eglinton LRT now allows those who used to take a bus to travel along a similar route at almost twice the speed – going from an average speed of 14 kilometers per hour on the bus to over 25 kilometers per hour on the LRT during peak afternoon commute times.

Wider Economic Benefits are not concerned with how often an individual can access a new group of people, rather, the focus is whether or not there these groups of people are accessible to individuals if they so choose to do so. This means that whether a new service is added during only the peak service hours or operates all day, the WEBs are the same. That said, our analysis is not concerned with whether or not a new route has all-day service or only peak-hour service.

Instead, when multiple service times are scheduled, we consider only the time frame which sees the largest speed increase. In most instances, regular bus service is replaced by either an Express Bus or one of the new LRT lines. That said, the speed of a regular bus service is dependent on roadway congestion, and thus the time of day. In all but four cases, (of 13) the most significant gain in speed happens during afternoon rush hour. Additionally, when a bus service operates several variations of one route, i.e., A, B, C, and D, which travel different distances on the same route, we calculate the average speed of the route across all its variations.

For service changes that are only comprised of the addition of buses to an existing service, i.e., adding more frequent service without changing any other aspect of the service, we only consider changes which result in a headway reduction of five minutes or more. This results in the inclusion of two bus services in Scarborough. For these service improvements, we consider the effective speed of the bus along the entirety of the route. To do so, we create a weighted average in which we multiply the run time of the route by the average route speed and multiply the headway by a speed of zero kilometers per hour. The change in effective speed is realized when the headway, i.e., the time spent travelling zero kilometers per hour, decreases. In all instances, the reduced headways increase the effective route speed by at least one kilometer per hour.

First-order Benefits

First-order benefits are one of the two ways considered in this paper by which investment in transit infrastructure and changes to services create wider economic benefits. The idea here is to determine how much further can a person travel under the improved services or by using the new infrastructure, ceteris paribus, and in turn, how many new people do they now have access to. In other words, how much further can a person travel in the same time as before (which assumes no change in people's value of time) when the effective speed of transit has increased? This further distance equates into greater access to people. Using the agglomeration elasticity, we estimate above, we then calculate the wider economic benefit of this improved access to people.

We create a list of traffic zones for each impacted service, including the forthcoming Line 5 Eglinton and Line 6 Finch West LRT lines, that fall within one kilometer of their route. This yields a dataset of traffic zones that are likely most affected by the service changes for each route. For each service change, we calculate the benefit only for the zones within 1 kilometer of where the route runs. We multiply the average Census commute time in each traffic zone and the initial and new speed for a given route to determine the initial and new travel distance, respectively. We sum the surrounding population within the initial and new travel distance (separately) for each impacted traffic zone – traffic zones within one kilometer of where the given service runs.

Using equation (1), we calculate the estimate of total WEB for each traffic zone.[13] Using these two population figures, we calculate the change in effective surrounding population as the population within the new travel distance less the population within the initial travel distance, divided by the population in the initial travel distance. After this is repeated for each service change, we prorate each WEB by the share of each traffic zone that uses the route impacted by the service change.[14] We then aggregate WEBs by zone. This yields the wider economic benefits accrued to users of TTC routes that are impacted by the service changes of interest. Note that these benefits are only experienced by those who use the impacted services.

Second-order Benefits

Second-order benefits are the second of the two ways considered in this paper by which investment in transit infrastructure and changes to services create wider economic benefits. The idea is to determine how those not using the impacted services, whether transit users or not, benefit from these service changes. While the individuals that are taking trips (referred to as those from travelling zones) benefit from accessing new people, those that remain in their traffic zone (those in stagnant zones) benefit from travellers coming into their area. In other words, investment in transit infrastructure allows for the creation of new connections by enhancing the mobility of those who prior may have been more restricted in their travelling abilities – i.e., transit users.

As in calculating the estimate of first-order benefits, we use the average Census commute time in each traffic zone times the initial and new speed for a given route to determine the initial and new travel distance. We then calculate a prorated population count – the number of transit users in a travelling zone multiplied by the share of people that use the said route. We then aggregate the prorated population of travelling zones that are within the new travel distance but further than the initial travel distance of each stagnant zone. This yields a population count for each stagnant zone of transit users in travelling zones that can now travel to each stagnant zone but couldn't before – the stagnant traffic zone population gain. We calculate the population count that was accessible to each stagnant zone prior to this influx of people from travelling zones – the stagnant traffic zone nearby population.[15]

Using these two population figures, we calculate the change in effective surrounding population as the stagnant traffic zone population gain divided by the stagnant zone nearby population. Using equation (1) we estimate the total WEB for each

13 The equation for the first-order benefits is that we sum over every zone i that sees a service improvement k we calculate the benefit as the agglomeration parameter*total income of all zone residents $15\text{-}64_i$*(change increase in accessible population from service_k)/(existing surrounding population based on zone commuting distance of residents_i)*share of zone residents that use improved transit option_{ki}.

14 This share is estimated using the 2016 TTS. A person is considered to be a user of a transit service if at it is taken as either the first or second link of a trip; further links are not considered as it is unlikely that any person within one kilometer of a service uses two or more other links to reach the said service. This count is then divided by the traffic zone's population as estimated by the TTS.

15 This is the sum of population ages 15-64 in all traffic zones that are within the mean trip distance of each traffic zone. These are residents in the zone *r* that benefit from residents in zones *i* as in the previous equation travelling to zone *r* using improvements in transit option *k*. We calculate this as the sum over *r*, *k*, and *i* using the equation = agglomeration parameter*total income of all zone residents $15\text{-}64_r$*(number of people from zone_i that can access zone_z from service change_k)/(existing surrounding population based on zone commuting distance of residents_z).

stagnant traffic zone, where *Zone's Population and Average Zone Income* is that of the stagnant zone. This yields the second-order wider economic benefits for each traffic zone resulting from the service changes of interest.

Part 2b: Line 1 Automation

When looking to the automation of the Line 1 Subway service, there are two main elements to which attention must be paid. First, the automation allows the train to travel at a faster average speed, decreasing the times it takes to travel the 38.8-kilometer, end-to-end trip from 165 minutes to 153 minutes. Second, because of this improved speed, headway time is also improved; that is, faster service allows for the addition of five new trains onto the track, in turn providing more frequent service. The TTC estimates that the wait time between trains will become 2 minutes, shortening 21 seconds from the old headway of 2 minutes and 21 seconds.

Both the change in route speed and change in headway are calculated as in previous sections; however, because both elements are improved simultaneously, analysis is carried out in order to separate the benefits of faster travel speed and improved headway. First, we determine the old effective speed of the subway using the distance of the route, divided by the sum of the previous headway and end-to-end trip time in hours; i.e., 38.8 ÷ ((2.35 + 165)/60). We perform the same calculation, changing the end-to-end trip time from 165 minutes to 153 minutes; this provides the effective speed if headway was to remain the same. Last, we change both the headway and trip time, replacing them with their improved times, 2 minutes and 153 minutes, respectively. In conjunction, these three effective speeds allow us to separate the WEBs provided from automation.

Using these speeds and the average commute time from the 2016 Census, we calculate the distance that a person from each traffic zone could travel under each of the three effective speeds described above. This yields three possible commute distances for each zone. We sum the population of all surrounding traffic zones that fall within each of these three distances, separately. From here, subtracting aggregated surrounding population under the initial effective speed from the two improved effective speeds provides the new population that a traffic zone would be able to access.

We then divide this newly accessible population under both the initial and improved headway by the total population that was accessible before automation in order to determine the change in effective surrounding population. This is then used in equation (1), as given above, in order to estimate the total wider economic benefits of automation, with, and without, taking the improved headway into consideration. We prorate this benefit by the share of each traffic zone that uses the Line 1 Subway as the benefit only occurs for those that use the transit service.[16] We estimate the second-order benefits of the Line 1 Subway improvements in the same manner as described in Part 2a.

Part 3: The Impact of a Fare Change on Ridership

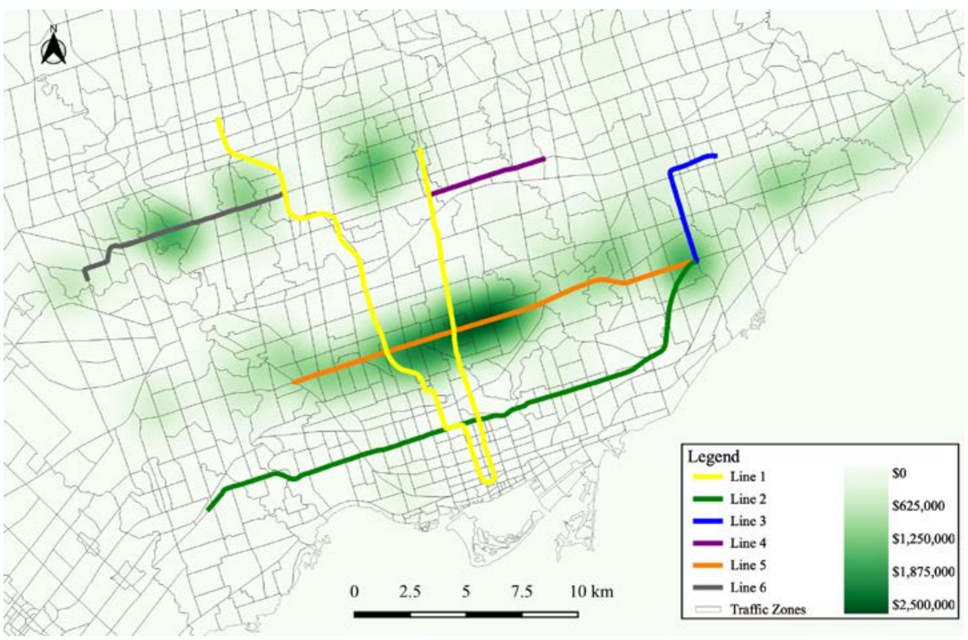
Following existing literature, we use both short-run (6-12 months) and long-run (8 or more years) fare elasticities to determine how demand for the TTC's services would be affected by a fare increase. In order to estimate the total WEBs forgone from the fare increase, we prorate the total WEB of the TTC ($2.7 billion, from Part 1a) by the loss in ridership from a fare change. We assume a 5% fare increase

16 This share is estimated using the 2016 TTS. A person is considered to be a user of the Line 1 Subway if at any time they use it during a trip of any type. This count is then divided by the traffic zone's population as estimated by the TTS.

across all TTC services. We multiply the total WEBs provided by the TTC by 0.988 and 0.944 in each zone to find the lost benefits due to the fare increase in the short- and long-run, respectively. This yields lost estimated wider economic benefits of about $32 million in the short-run and $152 million in the long-run. Assuming a total of 1.69 million fare charges at $3.00 each, the net revenue gain in the year following a 5% fare increase would be approximately $68 million, taking into consideration the short-run loss in ridership that would result from such an increase.

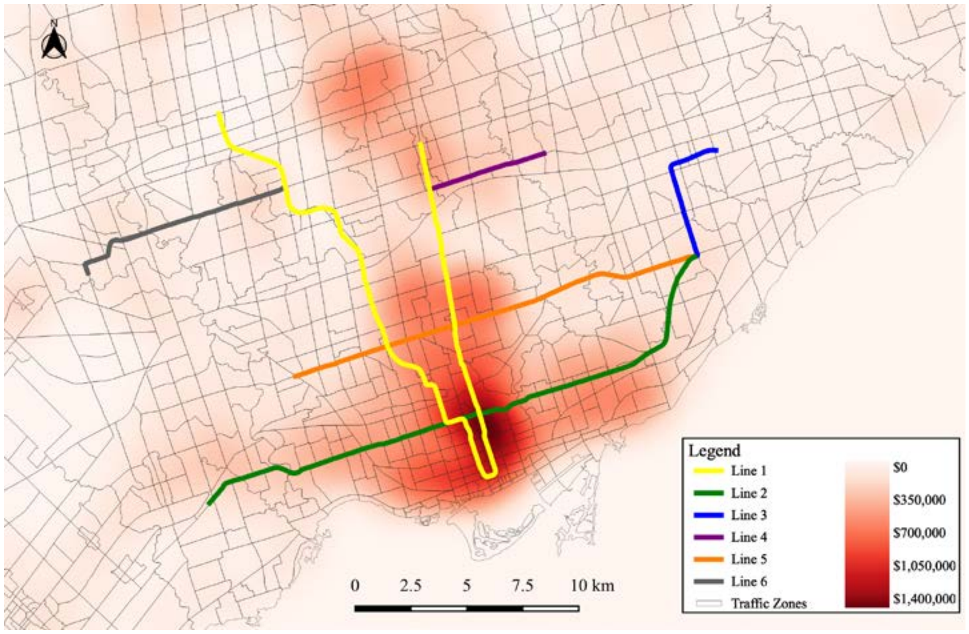
Additional Figures

Figure A1: First-Order Wider Economic Benefits of General Investments in Transit



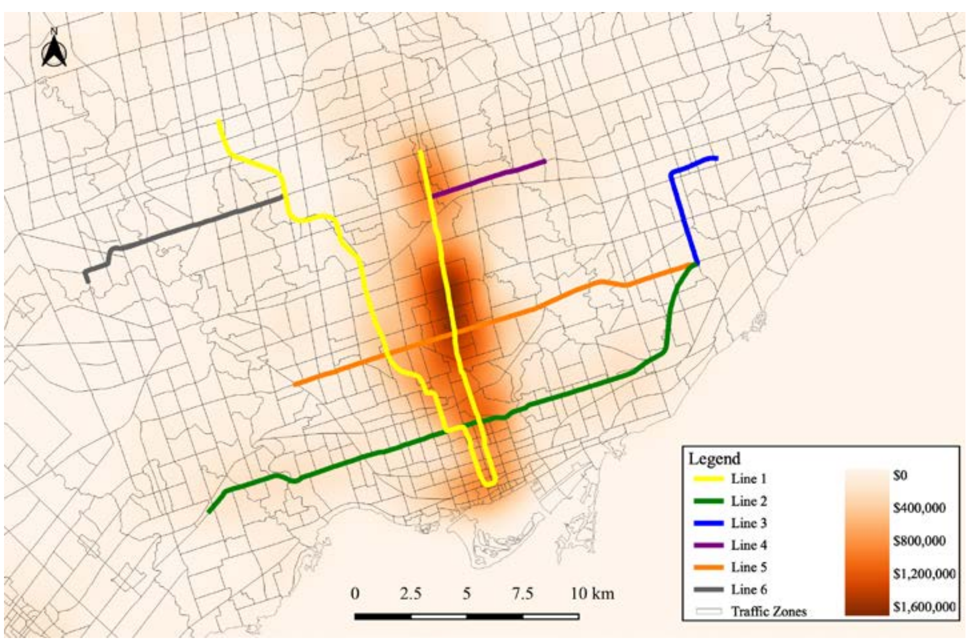
Source: x

Figure A2: Second-Order Wider Economic Benefits of General Investments in Transit



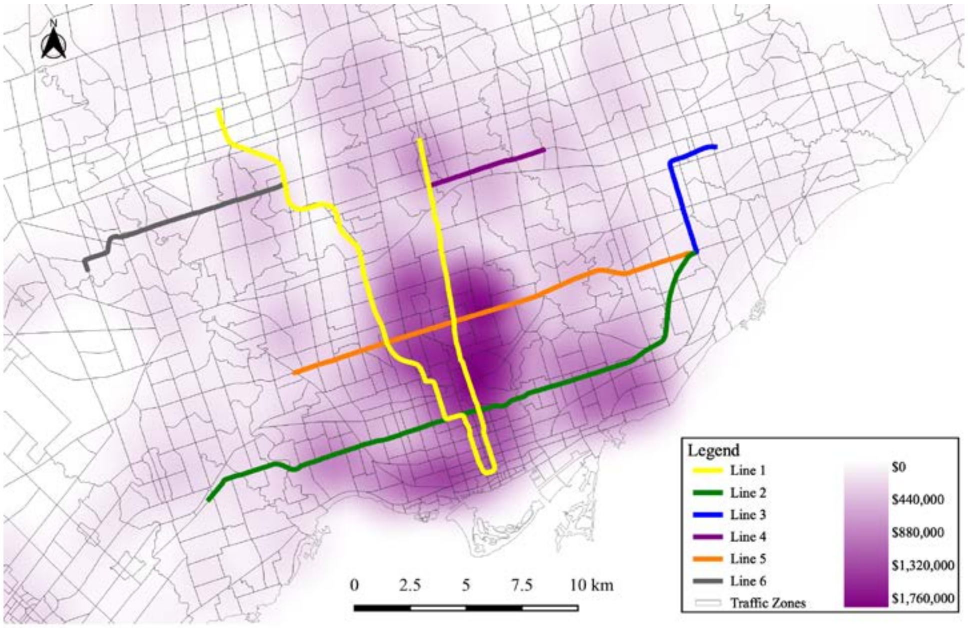
Source: x

Figure A3: First-Order Wider Economic Benefits of Line 1 Automation



Source: x

Figure A4: Second-Order Wider Economic Benefits of Line 1 Automation



Source: x