Online Appendix for "Subways, Strikes, and Slowdowns: The Impacts of Public Transit on Traffic Congestion"

By MICHAEL L. ANDERSON

AI. Mathematical Appendix

Distance to nearest bus line: Suppose that bus lines are placed in a grid pattern of width $2a$, that the population is uniformly distributed across space, and that people cannot walk diagonally across city blocks. Further suppose that bus stops are placed at locations at which bus lines cross (this assumption simplifies notation but can be relaxed with minimal effect on our conclusions). Then accessing the nearest bus stop requires walking U miles north/south and V miles east/west, where U and V are independent and each is distributed uniform $(0, a)$. The random variable $X = U + V$ measures the distance to the nearest bus stop and follows a triangular distribution with density $f(x) = x/a^2$ for $x \in (0, a)$ and density $f(x) = 2/a - x/a^2$ for $x \in (a, 2a)$.

Proof:

First note that $f(u, v) = f(u) \cdot f(v) = \frac{1}{a^2}$.

Thus for $x \in (0, a)$, $P(X < x) = P(U + V < x)$

$$= \int_0^x \int_0^{x-u} \frac{1}{a^2} dv\, du = \frac{x^2}{2a^2}.$$

Therefore $f(x) = x/a^2$ for $x \in (0, a)$.

By symmetry $f(x) = 2/a - x/a^2$ for $x \in (a, 2a)$.

Transit rider consumer surplus: Without loss of generality, we consider only rail commuters. Let A denote the set of commuters that choose to take rail. Let CS_i denote the consumer surplus from using rail for commuter i in the set A. In this context we calculate CS_i as commuter i's willingness to pay for the rail option minus the current fare; alternatively, it is the difference in generalized cost between the rail option and the driving option for commuter i. Rearranging the inequality for the heterogeneous driving delay scenario in Section **Error! Reference source not found.**, we have:

$$CS_i = v_i \left[c(a_d + w_{di} - a_{ri} - w_r) + \left(\frac{m}{s_d} - \frac{m}{s_r} \right) \right] + m(p_d - p_r)$$

Note that the homogeneous driving delay scenario is a special case of the heterogeneous driving delay scenario in which $w_{di} = w_d$ for all commuters. We calculate average consumer surplus for rail commuters as:

$$CS^r = \frac{1}{N_r} \sum_{i \in A} CS_i$$

where N_r denotes the number of rail commuters. Average consumer surplus per mile is CS^r/m.

We calculate consumer surplus for bus passengers in a similar manner but replace rail parameters with bus parameters.

AII. Sensitivity of Model Calibration Results

In this section we test the sensitivity of our calibration results to alternative parameter choices. In our baseline calibration we choose parameter values that tend to lower the cost of transit and increase the cost of driving. These choices are conservative in that they lower the predicted effect of ceasing transit service under the heterogeneous driving delay model. However, in several cases the parameter choices do not have clear implications for the relative cost of driving versus transit. These cases include the delay multiplier (which applies to driving delays, transit wait, and transit access time), the wage multiplier, and trip length.

Table A1 tests the sensitivity of our predictions to reasonable variations in these three parameters. Column (1) reproduces the predictions from our baseline calibration for comparison purposes. In the baseline calibration, the heterogeneous driving delay model predicts a congestion impact from ceasing transit service that is 5.9 times greater than the homogeneous driving delay model's prediction. Columns (2) and (3) test sensitivity to variations in the delay multiplier, which is set at $c = 1.8$ in our baseline calibration (a delay multiplier of 1.8 implies that individuals value time spent waiting in traffic, waiting for transit, or walking to transit at 1.8 times their normal value of time). Under a high delay multiplier ($c = 2.3$), the ratio of the heterogeneous model prediction to the homogeneous model prediction increases to 7.7. Under a low delay multiplier ($c = 1.3$), the ratio decreases to 3.4. Columns (4) and (5) test sensitivity to variations in the value of time. Under a high value of time (60 percent of the average wage), the ratio of the two models' predictions increases to 7.6. Under a low value of time (40 percent of the average wage), the ratio decreases to 4.1. Columns (6) and (7) test sensitivity to variations in trip length. Since long and short trips are less common than average-length trips, we calibrate the model in these two columns so that predicted transit share is half the overall transit share. This reduces the predicted impact of ceasing transit service under either model (homogeneous or heterogeneous), but it has little impact on the ratio of the two predictions (which is the object of interest). On long trips (a 10-mile rail trip or a 7-mile bus trip), the ratio of the two predictions is 6.1. On short trips (a 5-mile rail trip or a 3-mile bus trip), the ratio of the two predictions is 3.6.

Table A2 tests the sensitivity of our predictions to different assumptions about transportation alternatives and sorting behavior. Column (1) reproduces the predictions from our baseline calibration for comparison purposes. Column (2) introduces a mode-specific error term to the utility function. The error term multiplies the value of time applied to the transit commute, on the assumption that some people find time spent walking or inside transit vehicles to be more or less enjoyable than time spent driving. The multiplicative error term ranges from about 0.6 to 1.6 and is positively correlated with transit access time, since it seems likely that people with an intrinsic preference for transit will choose to live closer to transit stops. The ratio of the heterogeneous model's prediction to the homogeneous model's prediction falls to 4.6, but the overall predicted effect on congestion remains high. Column (3) calibrates the homogeneous and heterogeneous models under the assumption that one-third of the bus riders are "captive" riders who do not own cars; this is an extreme case of mode-specific preferences in which some commuters will not choose to drive under any conditions. Under this assumption the predicted impact of ceasing transit service falls in both models because captive riders do not switch to driving. Nevertheless, the ratio of the two models' predictions is nearly unchanged at 6.1. Column (4) calibrates the two models under the assumption that access time and driving delays are negatively correlated (i.e., people living far from transit experience fewer driving delays). This modification captures the possibility that denser areas have better transit access and more congestion. The ratio of the two predictions increases to 6.5. Column (5) calibrates the two models under the assumption that access time and value of time are positively correlated (i.e., wealthy neighborhoods are farther from transit). This modification captures the possibility that low-income individuals choose to live closer to transit. The ratio of the two predictions decreases to 4.6. Column (6) replaces the triangular distribution of bus access times with a smoother gamma distribution. The ratio of the two predictions decreases modestly to 5.4.

Two patterns emerge from the sensitivity analyses in Tables A1 and A2. First, in all cases the heterogeneous driving delay model predicts a much greater increase in congestion from ceasing transit service than the homogeneous driving delay model. The minimum ratio between the two models' predictions is 3.4, and the maximum ratio is 7.7. Second, despite the consistent qualitative finding that a model incorporating heterogeneous driving delays predicts much greater congestion impacts, the magnitude of the predictions varies substantially with different parameter values. Thus it is infeasible to make an accurate quantitative prediction about the effect of ceasing transit service without additional data or a natural experiment.

Table A1: Model Calibration Results Under Different Parameter Values

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Parameter Modification:	None	High Delay Multiplier (c = 2.3)	Low Delay Multiplier (c = 1.3)	High Value of Time (60% avg wage)	Low Value of Time (40% avg wage)	Long Trip (Rail = 10 miles, Bus = 7 miles)	Short Trip (Rail = 5 miles, Bus = 3 miles)
Outcomes							
Effect of Ceasing Transit on Average Delay (Homogeneous Driving Delay Model)	6.3%	6.4%	6.4%	6.3%	6.3%	3.1%	3.1%
Effect of Ceasing Transit on Average Delay (Heterogeneous Driving Delay Model)	37.1%	49.3%	21.5%	47.8%	25.8%	18.9%	11.2%
Ratio of Heterogeneous Model Effect to Homogeneous Model Effect	5.9	7.7	3.4	7.6	4.1	6.1	3.6
Calibration Parameters (Heterogeneous Model)							
Share of Population within 2 miles of Rail Line	30%	30%	27%	32%	25%	15%	30%
Average Bus Line Spacing in Residential Areas	0.5 miles	0.5 miles	0.5 miles	0.3 miles	0.7 miles	0.6 miles	0.6 miles

Notes: Average delay is chosen to match Parry and Small (2009). Calibration parameter values are the values necessary to equate predicted ridership with observed ridership in the model with heterogeneous driving delays. The effect of ceasing transit under the homogeneous model is smaller in columns (6) and (7) than other columns because observed ridership for long/short trips is assumed to be half the observed ridership for the average trip.

Table A2: Model Calibration Results Under Different Parameter Values

	(1)	(2)	(3)	(4)	(5)	(6)
Parameter Modification:	None	Mode-specific error in utility function	1/3 of bus riders are "captive"	Access time and delays negatively correlated	Access time and value of time positively correlated	Bus access time distributed gamma
<u>Outcomes</u>						
Effect of Ceasing Transit on Average Delay (Homogeneous Driving Delay Model)	6.3%	6.1%	4.7%	6.3%	6.2%	6.2%
Effect of Ceasing Transit on Average Delay (Heterogeneous Driving Delay Model)	37.1%	28.2%	28.6%	41.0%	28.5%	33.7%
Ratio of Heterogeneous Model Effect to Homogeneous Model Effect	5.9	4.6	6.1	6.5	4.6	5.4
<u>Calibration Parameters (Heterogeneous Model)</u>						
Share of Population within 2 miles of Rail Line	30%	9%	30%	28%	22%	30%
Average Bus Line Spacing in Residential Areas	0.5 miles	0.9 miles	0.5 miles	0.5 miles	0.7 miles	0.5 miles

Notes: Average delay is chosen to match Parry and Small (2009). Calibration parameter values are the values necessary to equate predicted ridership with observed ridership in the model with heterogeneous driving delays. The effect of ceasing transit under the homogeneous model is smaller in column (2) than other columns because the "captive" riders do not switch to driving when transit service ceases.

Table A3: Effect of Strike on Traffic Flows over Entire Day

Dependent Variable: Hourly Traffic Flow per Lane

	(1)	(2)	(3)	(4)	(5)	(6)
Strike	-1.9	-13.3	14.3	14.0	-11.1	-5.0
	(6.0)	(8.1)	(7.3)	(13.0)	(9.5)	(6.3)
Date	0.04	-0.16	0.09	-0.03	0.20	0.08
	(0.22)	(0.27)	(0.44)	(0.71)	(0.47)	(0.18)
Date*Strike	-0.85	-1.10	-0.65	-1.15	-0.95	-0.71
	(0.43)	(0.51)	(0.52)	(0.81)	(0.63)	(0.47)
Average Hourly Flow Pre-Strike	1,016	1,133	990	1,004	1,059	987
Freeways	All	101	105	110 & 710	10	Other
Parallel Transit Line		Red Line	Green Line	Blue Line	Rapid 720	
Sample Size	463,848	41,819	80,290	49,433	40,673	251,633

Notes: Each column represents a separate VMT-weighted regression, with weights equal to (length of highway covered by detector i)×(average pre-strike traffic flow over detector i). The observation is the detector-hour, and the sample is limited to weekdays within 28 days of the strike's beginning. Parentheses contain clustered standard errors that are robust to within-day and within-detector serial correlation. All regressions include day-of-week and detector fixed effects.