

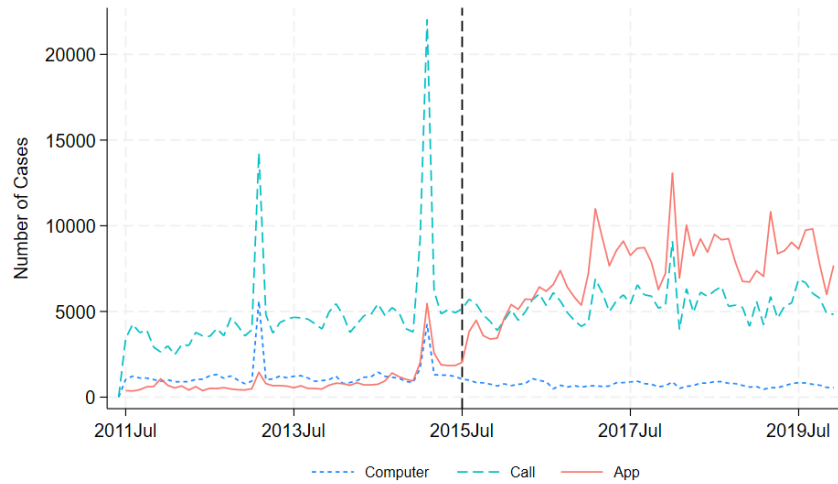
WA1. Web Appendix

WA1.1. Data Cleaning and Model Free Evidence

Table WA1 Data Cleaning Flow Summary

Step	Description	Obs. Dropped	Obs. Remaining
0	Initial sample	—	1,666,441
1	Drop by channel (channels other than phone, website, and app)	317,525	1,348,916
2	Drop missing location data (missing longitude or latitude)	32,491	1,316,425
3	Drop ambiguous department categories ('General' and 'Info')	89,735	1,226,690
4	Drop complaints with no app use	144,274	1,082,416
5	Drop rare departments and rarely reported case types	2,193	1,080,223
6	Drop observations with missing census data	9,734	1,070,489

Figure WA1 The Evolution of Number of Cases by Source Across the Policy Change



Notes: Full sample of data. Before and after the policy change, comparing the number of complaints submitted by channel of submission. 2013 and 2015 were years of heavy snowfall in Boston.

Table WA2 Overview of Control Municipalities Checked for Analysis

Municipality	Included?	Reason	Link
Houston	Yes	Meets analysis criteria	houstontx.gov
Miami-Dade County	Yes	Meets analysis criteria	gis-mdc.opendata.arcgis.com
New York City	Yes	Meets analysis criteria	data.cityofnewyork.us
Oakland	Yes	Meets analysis criteria	data.oaklandca.gov
San Francisco	Yes	Meets analysis criteria	data.sfgov.org
Kansas City	No	No mobile channel	data.kcmo.org
Austin	No	No department data	data.austintexas.gov
Baltimore	No	Does not distinguish website vs. app	data.baltimorecity.gov
Baton Rouge	No	Insufficient pre-treatment coverage	data.brla.gov
Chicago	No	Insufficient pre-treatment coverage	data.cityofchicago.org
Dallas	No	Insufficient pre-treatment coverage	dallas.opendata.com
Denver	No	Major change happened after 2015	opendata-geospatialdenver.hub.arcgis.com
Gainesville	No	Insufficient pre-treatment coverage	data.cityofgainesville.org
Las Vegas	No	No geo-coordinates and closed date	opendataportal-lasvegas.opendata.arcgis.com
Los Angeles	No	Insufficient pre-treatment coverage	data.lacity.org
Louisville	No	No closed date and channel data	data.louisvilleky.gov
Memphis	No	Insufficient pre-treatment coverage	data.memphistn.gov
Milwaukee	No	Insufficient pre-treatment coverage	data.milwaukee.gov
Minneapolis	No	No channel data	opendataphilly.org
Nashville	No	Insufficient pre-treatment coverage	data.nashville.gov
New Orleans	No	No channel data	data.nola.gov
Philadelphia	No	No channel data	opendataphilly.org
Phoenix	No	Insufficient pre-treatment coverage	phoenix.opendata.com
Pittsburgh	No	No closed date data	data.wprdc.org
Sacramento	No	No channel data	data.cityofsacramento.org
San Antonio	No	Insufficient pre-treatment coverage	data.sanantonio.gov
San Diego	No	Insufficient pre-treatment coverage	data.sandiego.gov
Washington, DC	No	No channel data	opendata.dc.gov

WA1.2. Employee's Organizational Incentives

In the face of any major policy change involving technology, it is natural to ask whether this was accompanied by organizational change. We interviewed city employees to find out. These interviews emphasized that the aim of the city of Boston in launching this omnichannel approach was to resolve complaints quickly; there was no organizational shift. We went out to confirm this using quantitative data from within the Boston City organization in the form of a parallel ‘worker app’. As detailed on the Boston 311 website, the city worker app “allows City Workers to create, close, and re-assign cases in real time from the field. Aside from making things more efficient, managers and supervisors now have a better sense of what’s happening on the ground. The app is now mainly used by supervisors in the field.” This verifies that each employee was indeed being monitored for how quickly they resolved these complaints, which again suggests that incentives between the city and workers were aligned.

We investigated the data and report the resolution time and status for cases logged internally by city service employees from the city worker app in Figure WA2. The data show that the percentages of resolved cases and cases resolved on time for complaints that originated with workers remained largely unchanged around the 2015 policy shock. The resolution time for worker-generated complaints also remained stable (though with more fluctuations compared to the other two metrics) around the policy change threshold. This confirms that there was no substantial change in city workers’ incentives to resolve complaint cases caused by the policy shock or any organizational streamlining that provide an alternative explanation.

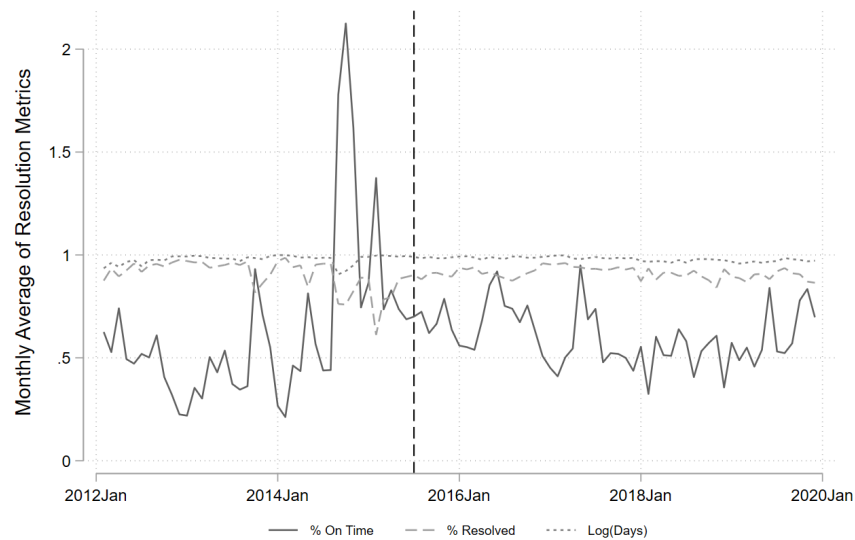
Although this data does not reveal how cases are allocated among employees, the volume of cases generated by field workers is likely to be correlated with the number of workers present at each location to address resident-reported complaints. In Table WA3 Column (1), we regress the number of employee-generated cases on the treatment indicator and its interaction with median household income of a census block group. We did not find any significant relationship between the shock and the number of employee generated cases across neighborhoods of different levels of income.

Table WA3 Number of Employee-logged Cases

	(1)
	Volume
Std. Income	-0.009 (0.035)
After Omnichannel $\times$ Std. Income	-0.033 (0.021)
Fixed Effects	Yes
Observations	89742
R-Squared	0.11
Mean Dep Var	1.76

Notes: The estimated model includes day-of-week fixed effects, year-month fixed effects, census block group fixed effects, department fixed effects, and reason fixed effects. Standard errors clustered at the census block group and month level. Standard errors clustered at the census block group level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Dependent variable is the number of cases logged by city employees in the field.

Figure WA2 Cases Logged by City Workers



WA1.3. Robustness of the Main Effect

WA1.3.1. Flexible Controls for Target Time Because target times may vary systematically across channels and regions and correlate with actual resolution times, omitting them could introduce bias. We therefore still include target resolution time in all subsequent Boston-based analyses. Table WA4 further tests robustness using linear and more flexible higher-order (third- and fifth-order polynomial) controls. The results are consistent across specifications.

Table WA4 The Effect of the Shift to Omnichannel on Apps: Different Controls for Target Resolution Time

	(1)	(2)	(3)
	Log(Days)	Log(Days)	Log(Days)
App=1 $\times$ After Omnichannel	-0.463*** (0.032)	-0.470*** (0.033)	-0.447*** (0.034)
App=1	0.369*** (0.026)	0.385*** (0.026)	0.362*** (0.026)
Target Time Function	Linear	3 rd Order Polynomial	5 th Order Polynomial
Fixed Effects	Yes	Yes	Yes
Observations	1018836	1018836	1018836
R-Squared	0.264	0.315	0.320
Mean Dep Var	1.189	1.189	1.189

Notes: All specifications include day-of-week fixed effects, year-month fixed effects, census block group fixed effects, and department fixed effects. Standard errors clustered at the census block group and month level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Dependent variable is the log of days it took to resolve the complaint. The first four columns use the entire sample. Column (5) compares app-submitted complaints with call-submitted complaints. Column (6) compares app-submitted complaints with computer-submitted complaints.

WA1.3.2. Alternative Dependent Variable (On-time) We check the robustness of our finding to an alternative dependent variable - whether a case was resolved ‘on time’ - in Table WA5. The ‘on time’ status of a case takes 1 if the resolution date of a case was prior to its target resolution date, and 0 if a case was not resolved or if it was resolved after the target resolution date. Since the target resolution date of a case was assigned by the public service representative and reflects the inherent difficulty of resolving the case, the ‘on time’ status of a case captures the efficiency of complaint resolution by measuring whether the case was resolved later than the average expected resolution time given the nature of the case. In Table WA5 in the appendix, we replace the dependent variable in Table 2 to whether a case was labeled as resolved ‘on time’. The results confirm our previous finding that the omnichannel integration improves the efficiency

of complaint resolution for cases reported through the mobile app, compared to other non-mobile channels.

Because the ‘on time’ indicator was determined by comparing the actual to the target resolution time, in addition to replicating the patterns in Table 2, we also include two additional columns in Table WA5 to check the robustness of this specification to the measurement of the target resolution time. In particular, in Column (5) of Table WA5, we add an additional control variable for the target resolution time. This helps us to control for the influence of the endogenously decided target time on on-time resolution rate. Adding this additional control variable does not affect our substantive findings. Given that the on-time resolution rate is high in the data, we also assess the robustness of our results against the possibility that target resolution times may have been set conservatively, potentially inflating the resolution rates. We calculate the mean and standard deviation of the actual resolution times for each complaint type by year and define a ‘maximum required’ threshold for a given case type as the sample mean resolution time plus two standard deviations. We then compute, for each complaint type, the proportion of cases where the target time exceeds this ‘maximum required’ threshold. In Column (6) of Table WA5, we exclude all case types where over 20% of cases had target resolution times exceeding the ‘maximum required’ time, as the on-time resolution rates may have been overestimated for these types (this applies to 5 complaint types). Our findings remain robust whether these complaint types are included or excluded.

Table WA5 Robustness to looking at the dependent variable capturing whether complaints were resolved on

	time							
	All						App v Call	App v Computer
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
App=1 × After Omnichannel	0.086*** (0.005)	0.077*** (0.005)	0.080*** (0.005)	0.079*** (0.005)	0.061*** (0.005)	0.099*** (0.005)	0.083*** (0.005)	0.017*** (0.006)
App=1	-0.066*** (0.005)	-0.060*** (0.005)	-0.094*** (0.005)	-0.092*** (0.005)	-0.048*** (0.005)	-0.057*** (0.004)	-0.095*** (0.005)	-0.084*** (0.007)
Computer=1				0.009*** (0.003)				
Target Time					-0.000* (0.000)			
Day of Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
YearMonth FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Census FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Department FE	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1070586	1070581	1070581	1070581	911956	760701	970366	541985
R-Squared	0.084	0.094	0.183	0.183	0.184	0.206	0.176	0.162
Mean Dep Var	0.816	0.816	0.816	0.816	0.787	0.764	0.815	0.840

Notes: Standard errors clustered at the census block group level. * p<0.05, ** p<0.01, *** p<0.001. Dependent variable is whether the city evaluated the request as being on time. Linear probability model estimates. The first six columns use the entire sample. Column (7) compares app-submitted complaints with call-submitted complaints. Column (8) compares app-submitted complaints with computer-submitted complaints.

WA1.4. Decomposing the Main Effect

Our analysis so far shows that the omnichannel integration decreased the average case resolution time in Boston, and that this improvement came primarily from cases reported through the mobile app channel. In this section, we investigate what contributed to this increase in resolution speed of the app channel relative to other reporting channels. Although the upgrade to the integrated 311 system included backend improvements for all channels, the key change specific to the app was the expansion of the types of complaints it accepts. This was part of the broader effort to standardize the complaint types across all channels. We therefore stratify the complaint types based on whether it could be easily reported via app prior to the omnichannel integration. In particular, we define nine complaint types that were never or rarely reported (fewer than 50 cases from 2011 to July 2015) through the app channel prior to the integration as ‘expanded types’. These types were ‘expanded’ to the app channel in that all of them experienced substantial increases in the volume of app reports after the omnichannel integration. We define the remaining six complaint types as the ‘existing types,’ meaning that they could already be reported conveniently through the app prior to the omnichannel integration. Both types can be easily reported through non-app channels both before and after the integration.

We first investigate whether the integration led to channel substitution and resulted in a substantially lower number of cases being reported through non-app channels. Figure WA3 plots the average daily reports across complaint types for app and non-app channels. The figure shows that the number of non-app reports remained remarkably stable before and after the omnichannel integration, whereas the number app reports increased substantially after integration. This pattern suggests that the integration increased the absolute number of cases, likely by making it easier for residents to use the app to report cases that would have not been reported at all without the integration. It did not seem to significantly reduce the volume of cases reported through non-app channels.

In Figure WA4, we further stratify the number of cases based on case types. The left columns of the two panels of Figure WA4 show that, for non-app channels, while the total number of

cases remained stable, the composition of cases changed over time, with the proportion of existing types decreasing and the proportion of expanded types increasing. The right columns of the two panels show that the volume of both types of cases increased significantly for the app channel. In particular, the number of expanded types increased from nearly zero before the integration to around 150 cases per day after the integration, reaching a level comparable to the volume of expanded type complaints through non-app channels.

Figures WA3 and WA4 show that the omnichannel integration changed how people report complaints by increasing the number of cases submitted through the mobile channel, and that this increase was contributed by both existing and expanded case types. We next investigate how this change in case volume and composition contributed to the overall effect of the omnichannel integration.

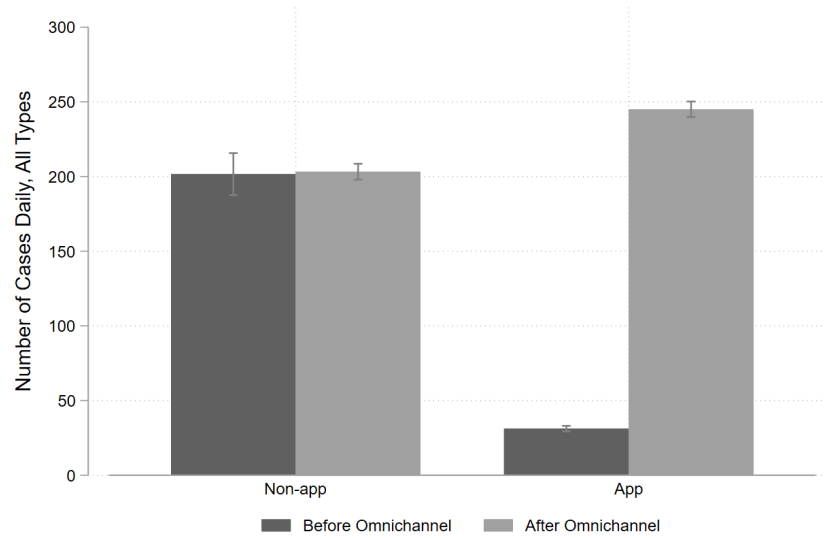
In Columns (1) and (2) of Table WA6, we compare the main effect identified in Section 4.1 (Column (1)) to a specification with fixed effects that control for complaint type (Column (2)). Adding type fixed effects enables us to account for baseline differences in resolution speed across case types, thereby distinguishing the effect due to improved app efficiency from the effect due to changes in composition of case types reported via the app. Results in Column (2) reveal that the omnichannel integration did improve efficiency across complaint types, as the coefficient for the interaction term, $App=1 \times After\ Omnichannel$, is still negative and statistically significant even after controlling for fixed effects for case types. However, this coefficient is substantially smaller than that in Column (1) without the type fixed effects. This suggests that the change in composition of case types after the omnichannel integration also contributed to the overall effect.

To better understand the role of changes in case composition, we include a dummy variable in Column (3) of Table WA6 indicating whether a complaint case belongs to one of the expanded types. The estimate for the term $Expanded\ Types = 1$ is negative and statistically significant, indicating that these complaints were on average easier to resolve compared to the existing types. In particular, the app channel seemed to be much more efficient in resolving these complaints: the

coefficient for $Expanded\ Types = 1 \times App = 1$ is negative, statistically significant, and large in size. In addition, Column (3) also shows that integration led to a small but significant increase in the app's efficiency in complaint resolution (-0.100), but this increase was not significantly different across complaint types. We caution our readers that because expanded types were rarely reported via app prior to the integration, the coefficient of the interaction term $Expanded\ Types = 1 \times App = 1$ is identified based on the very small sample of expanded type complaints that did manage to go through the app channel. These complaints might be inherently different even from other (non-app) complaints of the same type. We should therefore interpret the coefficient of the interaction term with caution.

Combining the evidence from Figures WA3, WA4, and Table WA6, we conclude that the effect identified in Section 4.1 can be attributed to three key factors. First, there was of course a composition effect. The integration made it easier for people to submit complaint cases and led to an influx of cases to the app channel that were inherently easier to resolve (the expanded types). However, there were two other additional effects which illuminate the benefits of the omnichannel approach taken. One is that the integration had an efficiency effect. As shown in Table WA6 Columns (2) and (3), it resulted in a significant improvement in the app's ability to resolve complaints on average. The other factor is that the integration also enhanced allocation efficiency. The expanded case types were particularly well-suited for the app channel but were previously difficult to submit through it. As the volume of these complaints grew, the omnichannel integration ensured they were processed through a more suitable and efficient channel for these specific case types, thereby improving overall resolution efficiency. Importantly, the effect is unlikely to be an artifact of channel switching that left the non-app channels with more difficult-to-resolve cases and therefore decreased the average resolution speed of these channels, as the proportion of the easier-to-resolve case types (the expanded types) also increased for the non-app channel. In fact, the results in Column (5) of Table A5 show that the non-app channel in Boston also experienced an increase in average resolution speed compared to other municipalities (the coefficient estimate for the term $Boston = 1 \times After\ Omnichannel$ is negative and statistically significant).

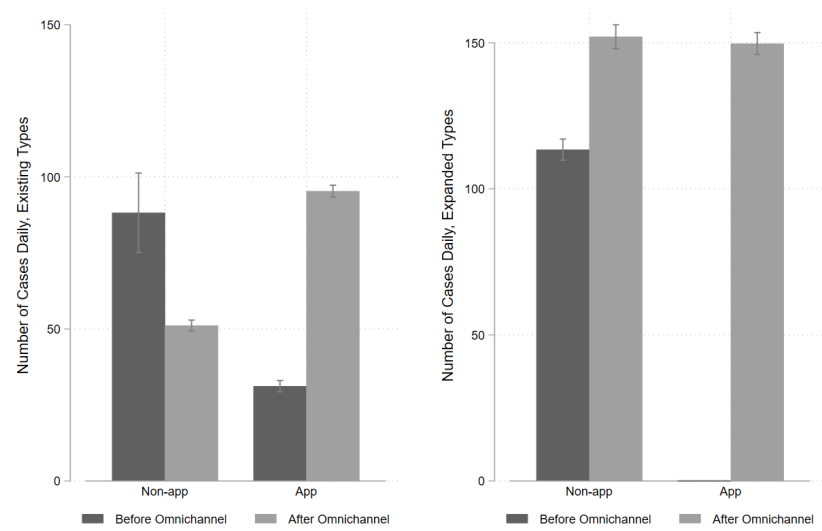
Figure WA3 Daily Number of Cases by Channel



Notes: This figure plots the average daily reports across complaint types for app and non-app channels.

Finally, in the last four columns of Table WA6, we investigate how the treatment effect changed over time. Columns (4)–(6) re-estimate our previous model by comparing pre-treatment data with data up to the end of 2015 (six months after the integration), data up to the end of 2016 (1.5 years after the integration), and data from 2019 alone. In Column (7) of Table WA6, we further include a variable that indicates the number of months away from omnichannel integration and its interactions with $App = 1$ and *After Omnichannel* in the regression. The coefficients for *Months from Omnichannel* and *After Omnichannel* $\times$ *Months from Integration* were dropped due to multicollinearity. The coefficient for the three way interaction term, $App = 1 \times After Omnichannel \times Months from Integration$, is positive but not statistically different from zero. The results show that the omnichannel integration has a long-term effect that lasts over four years after the treatment, and that the effect sustained over time.

Figure WA4 Daily Number of Cases by Case Types



Notes: This figure plots the average daily reports across complaint types for app and non-app channels by case types.

Table WA6 Decomposing the Effect

	Overall Effect	Controlling Type FE	Stratify by Case Types	Up to End 2015	Up to End 2016	Year 2019	Effect Over Time
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Log(Days)	Log(Days)	Log(Days)	Log(Days)	Log(Days)	Log(Days)	Log(Days)
App=1 × After Omnichannel	-0.480*** (0.032)	-0.153** (0.064)	-0.100** (0.038)	-0.664*** (0.035)	-0.611*** (0.022)	-0.524*** (0.033)	-0.444*** (0.074)
Expanded Types=1			-0.548*** (0.062)				
Expanded Types=1 × App=1			-0.525*** (0.143)				
Expanded Types=1 × App=1 × After Omnichannel			0.173 (0.158)				
App=1 × After Omnichannel × Months from Integration							0.006 (0.004)
Target Time	0.001*** (0.000)	0.000*** (0.000)	0.001*** (0.000)	0.000*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)
Day of Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
YearMonth FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Census FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Department FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Channel FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lower Order Interactions	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Type FE	No	Yes	No	No	No	No	No
Observations	1018836	1018836	1018836	373893	509132	494734	1018836
R-Squared	0.265	0.364	0.290	0.200	0.247	0.228	0.265
Mean Dep Var	1.189	1.189	1.189	1.596	1.432	1.399	1.189

Notes: Standard errors clustered at the census block group and month level. * p<0.05, ** p<0.01, *** p<0.001. Dependent variable is the log of days it took to resolve the complaint. Expanded types represent complaint types that were rarely reported via app through the omnichannel integration, including Abandoned Bicycle, Code Enforcement, Enforcement & Abandoned Vehicles, Environmental Services, Housing, Needle Program, Park Maintenance & Safety, Sanitation, and Trees.

WA1.5. Robustness of SES Results

In Tables WA7 and WA8, we examine heterogeneous treatment effects across income levels by comparing Boston with other control municipalities discussed in Section 4.1.2. Table WA7 reports results for each income quartile, using a synthetic differences-in-differences approach. Table WA8 reports results based on a classical differences-in-differences approach. The results show that all groups in Boston experienced a decrease in resolution time after the integration relative to control municipalities, but this improvement in efficiency is the largest for highest-income areas.

Table WA7 Stratification by Income Quartiles with Control Municipalities, SDID

	(1)	(2)	(3)	(4)
	Q1 Income	Q2 Income	Q3 Income	Q4 Income
Treatment Effect	-0.340*** (0.026)	-0.318*** (0.023)	-0.350*** (0.025)	-0.361*** (0.024)
Fixed Effects	Yes	Yes	Yes	Yes
Observations	81,324	81,864	79,812	77,436

Notes: All specifications include year-month fixed effects and city fixed effects. Jack-knife standard errors reported. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Dependent variable is the log of days it took to resolve the complaint. Dependent variable is the log of days it took to resolve the complaint. Income is standardized within each county. The table compares all cases from Boston to cases from a set of five control regions including New York City, San Francisco, Houston, Miami-Dade County, and Oakland.

Table WA8 Stratification by Income Quartiles with Control Municipalities, DID

	(1)	(2)	(3)	(4)
	Q1 Income	Q2 Income	Q3 Income	Q4 Income
Boston=1 × After Omnichannel	-0.408*** (0.038)	-0.371*** (0.041)	-0.447*** (0.036)	-0.469*** (0.038)
Fixed Effects	Yes	Yes	Yes	Yes
Observations	2163392	2251446	1954454	1726580
R-Squared	0.471	0.468	0.452	0.444
Mean Dep Var	1.404	1.408	1.392	1.354

Notes: All specifications include day-of-week fixed effects, year-month fixed effects, city fixed effects, department fixed effects, and reporting channel fixed effects. Standard errors clustered at the census block group and month level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Dependent variable is the log of days it took to resolve the complaint. Dependent variable is the log of days it took to resolve the complaint. Income is standardized within each county. The table compares all cases from Boston to cases from a set of five control regions including New York City, San Francisco, Houston, Miami-Dade County, and Oakland.

In Table WA9, we check the robustness of our findings of the heterogeneous effect across income to an alternative measure of complaint resolution efficiency. Table WA9 reports the result where the dependent variable is whether a complaint case was resolved ‘on time’. The ‘on time’ status of a case therefore captures both whether a complaint case was resolved, and (if resolved) whether it was

resolved later than the average expected resolution time given the nature of the case. Reassuringly, our findings are robust to this alternative dependent variable.

Table WA9 Differences in the Benefits of the App Across Income Quartiles: Robustness to a Dependent variable Capturing whether Complaints were Resolved on Time

	(1)	(2)	(3)	(4)
	Q1 Income	Q2 Income	Q3 Income	Q4 Income
App=1 $\times$ After Omnichannel	0.039*** (0.009)	0.068*** (0.014)	0.097*** (0.008)	0.089*** (0.010)
App=1	-0.052*** (0.009)	-0.071*** (0.013)	-0.110*** (0.008)	-0.114*** (0.009)
Fixed Effects	Yes	Yes	Yes	Yes
Observations	223082	218020	264646	364736
R-Squared	0.195	0.194	0.172	0.174
Mean Dep Var	0.776	0.805	0.833	0.835

Notes: All specifications include day-of-week fixed effects, year-month fixed effects, census block group fixed effects, and department fixed effects. Standard errors clustered at the census block group and month level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Dependent variable is whether or not the complaint was resolved 'on time' given the deadline allocated by the system. Full sample used including complaints from apps, calls and computers. Each column represents the results for a census block whose average household income in the data was in that quartile of income across census group blocks.

We run a series of robustness checks for our findings in Table 4 and Table A6. We report the results of our robustness checks in Table WA10.

Because in the data there were more people using apps to report cases during weekends compared to weekdays relative to the other channels, we first check whether our results are robust to this difference in usage pattern across day of week. In Column (1) of Table WA10, we exclude reports from weekends and re-estimate the three-way interaction model as specified in Column (1) of Table A6. The exclusion of weekends does not affect our substantive findings.

We next check the robustness of our results to complaints potentially reported by participants in the sharing economy. Uber, Lyft, and Airbnb became more popular during this period. Airbnbs are a potential source of complaints in our data, and in fact in 2020, a special category of complaint was created just for them.¹ It is possible that the rise of Airbnbs was more focused in richer neighborhoods as they may be more desirable for tourists. So, in Column (2) of Table WA10, we

¹ <https://www.boston.gov/departments/inspectional-services/short-term-rentals>

drop all complaints related to residence permits from our analysis which are the complaints most likely to be associated with Airbnbs. Our results are robust to this exclusion.

Ridesharing drivers may be more likely to report complaints related to their driving via app since their jobs are app-based. This could also be a concern for rich neighborhoods which are more likely to use ridesharing, which might skew our results. In Column (3) of Table WA10, we drop all complaints related to highway maintenance, which by their nature are more likely to concern drivers that use ridesharing apps. We find that excluding these cases does not have any substantial impact on our results. Overall, these checks suggest that the rise of the sharing economy and reports from sharing economy participants are not likely to be the key driver of our results.

As we are studying Boston, which can have snowy winters, we also check the robustness of our results to snow-related complaints. We first plot the number of cases over time in Figure WA1 and find that there were two major increases in the number of cases in February 2013 and February 2015, which were driven by the unusual snowfall in these months. As rich neighborhoods are more likely to hire private snowplow drivers, snow cleaning requests from such neighborhoods may be resolved faster purely due to more accessible supplementary private services. As a result, the resolution speed of such cases may not reflect the efficiency of the Boston 311 service. To address this concern, we drop all snow-related service requests in Column (4) of Table WA10 and re-estimate the model. Reassuringly, dropping snow-related complaints does not affect our previous findings.

In Columns (5) and (6) of Table WA10, we check whether changes in political focus and priorities of the city could affect the efficiency of complaint resolution. To test this, we include a series of fixed effects for each political district in Boston in Column (5) to control for the differences in political priorities across districts. We further exclude complaints submitted the month before a City Council election (Octobers of 2013, 2015, 2017, and 2019) in our analysis in Column (6). Our results are robust to these checks.

Another possibility is that errors in the app or the lower cost of the digital process could have led to more duplicate requests being submitted. In Column (7) of Table WA10, we check the robustness

of our results to the exclusion of cases that were marked as ‘duplicates’ in the system. Our results are also unaffected by the exclusion of such cases.

In Column (8) of Table WA10, we check whether our results are mainly driven by areas with college campuses. Colleges have larger populations of young, educated smartphone users and university-funded complementary support systems that would help resolve issues faster, so cases from universities and surrounding areas may skew our results. We identify and exclude complaints submitted from areas where the population of 18 - 29 years old exceeds 50%, and report the results in Column (8) of Table WA10. Our results are robust to the exclusion of these cases.

Finally, linguistic nuances and communication styles might also vary across citizens with different socioeconomic backgrounds. These disparities in the effectiveness of communication might also affect our finding. Because our data do not include the full text of each complaint, we analyze the complaint title and compute several linguistic measures to proxy for features of the underlying communication. More specifically, we measure conciseness as the length of the complaint title, informativeness as the number of unique words with substantial meaning (obtained by removing common stop words and punctuation) divided by the title length, readability using the Flesch-Kincaid Readability Score (Flesch 1948), and sentiment based on the algorithm developed in the *sentimentr* R package (Fast et al. 2023). In Column (10) of Table WA10, we replicate our analysis while including all four linguistic features as control variables. Our findings are robust to the inclusion of these additional controls.

Table WA10 Robustness Checks for Whether or not the Effect was Driven by Socioeconomic Status

	Not Weekend	No AirBnB	No Highways	No Snow	Electoral Fixed Effects	No Election Months	No Duplicates	Not Student-Heavy	Linguistic Controls
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Log(Days)	Log(Days)	Log(Days)	Log(Days)	Log(Days)	Log(Days)	Log(Days)	Log(Days)	Log(Days)
App=1 × After Omnichannel × Std. Income	-0.053*** (0.014)	-0.050*** (0.012)	-0.055*** (0.016)	-0.050*** (0.014)	-0.049*** (0.013)	-0.050*** (0.013)	-0.048*** (0.012)	-0.049*** (0.014)	-0.045*** (0.013)
App=1 × After Omnichannel	-0.457*** (0.039)	-0.446*** (0.031)	-0.614*** (0.036)	-0.486*** (0.047)	-0.446*** (0.031)	-0.441*** (0.031)	-0.509*** (0.028)	-0.449*** (0.031)	-0.406*** (0.016)
App=1 × Std. Income	0.046*** (0.011)	0.040*** (0.010)	0.041*** (0.014)	0.043*** (0.011)	0.039*** (0.010)	0.039*** (0.010)	0.039*** (0.010)	0.040*** (0.010)	0.040*** (0.011)
After Omnichannel × Std. Income	-0.008 (0.010)	-0.010 (0.009)	-0.009 (0.010)	-0.010 (0.009)	-0.010 (0.009)	-0.009 (0.009)	-0.011 (0.009)	-0.010 (0.009)	-0.010 (0.008)
App=1	0.340*** (0.035)	0.358*** (0.026)	0.506*** (0.034)	0.407*** (0.049)	0.358*** (0.026)	0.352*** (0.026)	0.408*** (0.023)	0.360*** (0.026)	0.294*** (0.016)
Std. Income	-0.006 (0.013)	-0.010 (0.012)	-0.010 (0.013)	-0.011 (0.012)	-0.009 (0.012)	-0.011 (0.013)	-0.009 (0.012)	-0.011 (0.013)	-0.009 (0.010)
Target Time	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)
Day of Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
YearMonth FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Census FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Department FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City Council District	No	No	No	No	Yes	No	No	No	No
Linguistic Controls	No	No	No	No	No	No	No	No	Yes
Observations	843437	1018836	926252	951287	1016125	977559	973072	917095	1018507
R-Squared	0.267	0.264	0.288	0.266	0.264	0.264	0.276	0.259	0.279
Mean Dep Var	1.192	1.189	1.187	1.166	1.186	1.193	1.191	1.190	1.188

Notes: Standard errors clustered at the census block group and month level. * p<0.05, ** p<0.01, *** p<0.001. Dependent variable is the log of days it took to resolve the complaint. Column (1) is a sample that excludes observations of complaints submitted at the weekend. Column (2) excludes complaints related to residential occupancy. Column (3) excludes complaints about the conditions of highways. Column (4) excludes complaints relating to the Snow. Column (5) includes the full sample but includes fixed effects for the political district. Column (6) excludes complaints submitted the month before an election for the city council, which happens every two years. Column (7) excludes complaints that were marked as duplicates in the system. Column (8) excludes census block groups where the population of 18-29 year olds exceeds 50 percent in an attempt to remove the influence of student populations. Column (9) controls for perceived severity of each case, measured by the target resolution time set by city employees. Column (10) controls for four linguistic features (conciseness, informativeness, readability, and sentiment) of each complaint as defined in Section 4.3.

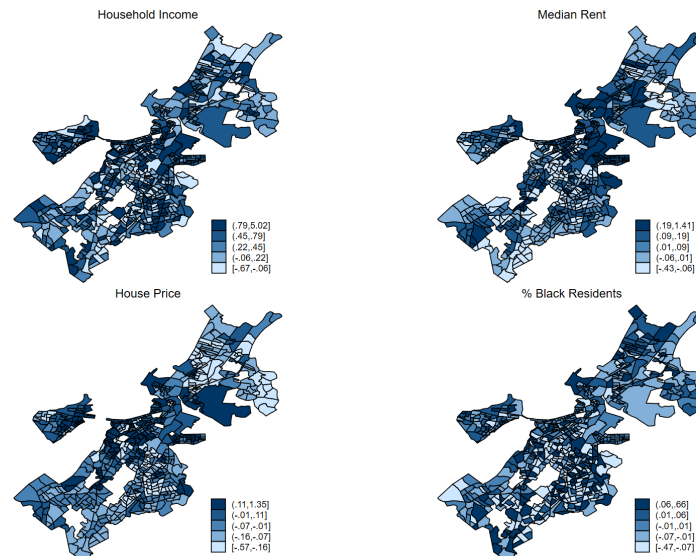
WA1.6. Supplementary Information for the Mechanism

Table WA11 Robustness to Population Density Controls

	Full Sample (1) Log(Days)	Enforcement Cases (2) Log(Days)
App=1 × After Omnichannel × Std. Income	-0.049*** (0.012)	-0.192*** (0.035)
App=1 × After Omnichannel × Census Block Group Population	0.000 (0.000)	0.000* (0.000)
App=1 × After Omnichannel	-0.459*** (0.042)	0.020 (0.072)
Target Time	0.001*** (0.000)	0.120*** (0.003)
Fixed Effects	Yes	Yes
Observations	1018836	279929
R-Squared	0.26	0.58
Mean Dep Var	1.19	0.46

Notes: All specifications include day-of-week fixed effects, year-month fixed effects, census block group fixed effects, department fixed effects, and lower-order interaction terms. Standard errors clustered at the census block group and month level. * p<0.05, ** p<0.01, *** p<0.001. Dependent variable is whether or not the complaint was resolved 'on time' given the deadline allocated by the system. Full sample used including complaints from apps, calls and computers. Each column represents the results for a census block whose average household income in the data was in that quartile of income across census group blocks.

Figure WA5 Visualizing Socioeconomic Indicators of Gentrification in Boston



Notes: This figure visualizes the socioeconomic indicators of gentrification in the city of Boston. The first panel reports proportional change in household income for each census block group from 2011 to 2019. The second panel reports proportional change in median rent for each census block group from 2010 to 2016. The third panel reports proportional change in median home value for each census block group from 2010 to 2016. The fourth figure reports proportional change in the percentage of Black population for each census block group from 2011 to 2019.

References

- Fast, M. J., B. M. E. Kurtz, and M. J. Barr (2023). *sentimentr: Calculate Textual Sentiment*. R package version 3.2.1.
- Flesch, R. (1948). A new readability yardstick. *Journal of applied psychology* 32(3), 221.