Social Networks and the Demand for News

Lisa M. George
Hunter College and the Graduate Center, CUNY*

Christian Peukert
Institute for Strategy, Technology and Organization, LMU Munich†

November 21, 2013

** Highly Preliminary**

Abstract

The well-documented relationship between group size and consumption in media markets is typically attributed to the incentive for firms to target products at large groups. Using data on monthly local and non-local visits to news outlets online, we document a positive relationship between group population in local markets and consumption of national media online. This pattern of preference externalities cannot be explained by standard supply-side arguments. We offer evidence that local social networks increase with group size, and that the richer social networks made possible by larger groups with shared tastes can explain the relationship between local group size and consumption of national media. Social networks matter in the market for news.

Keywords: News, Social Media, Media, Internet, Digitization
JEL No.:

*Lisa M. George, Department of Economics, Hunter College, 695 Park Ave., New York, NY 10065. Email: lisa.george@hunter.cuny.edu.

†Christian Peukert, Institute for Strategy, Technology and Organization, Ludwig-Maximilians-University Munich, Schackstr. 4, 80539 München, Germany. Email: c.peukert@lmu.de.

1 Introduction

The importance of shared preferences in media consumption is well documented across news and entertainment. Larger markets see higher newspaper sales, greater radio listening, and more television viewing. "Preference externalities" are especially important for minority groups with distinct tastes, and a larger black or Hispanic population in a market, all else equal, is linked to higher news consumption and greater political participation among individuals in these groups.(George and Waldfogel, 2003; Waldfogel, 2003, 2004; Oberholzer-Gee and Waldfogel, 2009).

The theoretical mechanism driving the relationship between group size and media consumption is understood to arise from supply-side incentives in the framework of spatial model: when some costs are fixed, more people with shared preferences brings forth more varieties targeted to those tastes, which in turn increases consumption. In cases where variety does not increase with group size, more people with common tastes induce firms to position products to satisfy those tastes, or perhaps to increase product quality (George and Waldfogel, 2003; Berry and Waldfogel, 2010; Wang and Waterman, 2011). In all cases, higher participation follows.

However despite consistent evidence that larger groups with shared preferences have a higher tendency to consume media products, the link between supply and consumption remains in some ways incomplete. It might be the case that demand-side factors contribute to or even dominate the relationship. For example, markets with more blacks have papers with a larger share of news content targeted at minority tastes and more radio stations playing songs preferred by black listeners. But rather than products driving consumption, it might be that a larger minority population fosters more extensive social networks which raise the utility of consuming news and other cultural or political media products.

Why does the distinction matter? Supply-driven models grounded in a spatial framework hinge on fixed production costs that limit entry in small markets. Spatial

models also naturally apply to local markets in the sense that products are targeted at tastes of a defined group. An important consequence of what might be called "social" externalities in media consumption is that a larger population with shared preferences might raise consumption for products that *do not* exhibit increasing returns to scale in production. Perhaps more important, with social network externalities the relationship between group size and consumption can operate for national products, not solely local ones. The important determinant of preference externalities would not then be the presence of *local product markets*, but rather the presence of *local social networks*. Because social networks are more likely to be tied to the local environment than product markets, the scope for shared preferences to drive consumption is significantly expanded.

A deeper theoretical understanding of social networks in media demand is also important for policy in the context of remedies. Lower news consumption is associated with reduced political participation, and this link drives policy interest in the industrial organization of media markets. If consumption and engagement result from low levels of targeted content, then preference externalities can be remedied with familiar supply-side interventions. If, instead, reduced media consumption is driven largely by social networks, then supply-side interventions have less traction.

This paper studies the role of social networks in news consumption on the internet. The study proceeds in two parts. The first section studies the relationship between group size and digital news consumption. We find that local news consumption by black individuals is positively related to the black population in a market and negatively related to the white population. White readership also depends on group size, though to a lesser extent. The results closely correspond to the pattern documented in George and Waldfogel (2003) which used zipcode-level daily newspaper circulation to identify the effect of shared preferences on readership. We next build upon the baseline result and study local and non-local news separately. We find a strong relationship between

a larger minority population and minority consumption in local news, as would be expected with supply side models. But we also find that black readership of national products is higher in markets with more blacks, which cannot be explained by targeting or entry models.

With this evidence that shared tastes matter for consumption for non-local products, the second part of the paper evaluates whether the larger social networks made possible by a larger population with shared preferences might drive higher news consumption. To do this, we first study overall social media usage by black and white households as well as news visits originating from social media sites. We find that social media use increases with group population for black households, suggesting that larger markets foster stronger social networks. When we account for social networks in news visits, we find that the direct effect of group size on news consumption shrinks. Social networks provide a mechanism for the operation of preference externalities in news consumption.

In addition to the now substantial literature on preference externalities cited above, this paper relates to a growing literature in marketing on word-of-mouth advertising and its role in consumer purchase [Dina Mayzlin 2011, Catherine Tucker 2013]. From an econometric standpoint, the research shares challenges with the literature on peer effects [xx], though the theoretical mechanism is more akin to indirect network effects than peer effects traditionally conceived [cite]. We also contribute to a small literature on digital news markets.[cite].

The paper proceeds as follows. Section 2 describes the data. Section 3 outlines the identification strategy and presents results for local and non-local news readership. Section 4 presents results for social media. Section 5 concludes.

Table 1: News Consumption Sample Statistics

	Observations	Mean	S.D.	5th Pct	95th Pct
News Visit Probability	469,224	0.66	0.47	0.00	1.00
News Visits	469,224	12.52	41.64	0.00	56.00
Local News Visits	469,224	1.35	8.16	0.00	5.00
Non-Local News Visits	469,224	11.16	38.63	0.00	49.00
Black Household	39,102	0.31	0.46	0.00	1.00
High Income Household	39,102	0.16	0.37	0.00	1.00
Household Size	39,029	3.18	1.52	1.00	6.00
Black MSA Pop	322	0.09	0.22	0.00	0.43
White MSA Pop	322	0.46	0.63	0.07	1.82

2 Data

Our analysis of news consumption is based on monthly visits to local and non-local news outlets by 39,102 black and non-Hispanic white households across 322 MSA's in 2011. The working data are constructed from click-level data from the ComScore Web Behavior Database. Local and non-local news outlets are identified from Burrelle's Media Directory, Bulldog Media directory, and the Newspaper Association of America web site as well as sites appearing on Google News. News visits are classified as local to a user if MSA from which a site receives the most visits matches the MSA of the user. The classification process matches that used in George and Hogendorn (2013), which provides additional detail on construction of the working data. We exclude from the analysis households residing outside of MSA's as well as websites with more than 500,000 unique visits per month, largely platforms with a very small share of news visits.[1] The household sample includes a small number of demographics: race, age of oldest household member, income categories, household size and zipcode residence. We supplement with zipcode demographics for some specifications.

Table 1 summarizes the working sample. For the full sample, the probability of a

[1]Excluded domains are aol.com, google.com, msn.com, yahoo.com and youtube.com.

news visit each month is .66, with an average of 12.52 visits per month to all sites and 1.35 visits to local sites. Black households comprise about % of the sample across 322 markets.

For the second portion of our analysis we add to the working data measures of social network use, especially monthly visits to social media outlets and time spent at these sites. We also calculate for each household the number and share of news visits referred from social media. [2]

3 News Consumption and Group Size

Our goal in this section is to estimate the relationship between group size and readership in digital news markets, focusing on black and non-hispanic white households. [3]

As a first step, it is useful to consider monthly household news visits by population quartile, shown in table 2. The left column shows total news consumption for all households. Monthly visits to online news outlets increases with market size, consistent with the pattern found with aggregate newspaper circulation data. The four subsequent columns show local and non-local news consumption by black and white households. News visits for white households generally increase with market size for both groups, though the increase is not in all cases monotonic.

The consumption pattern for black households is more mixed, suggesting that overall population is not as closely linked to consumption.

To study more formally the relationship between group size and consumption, we

[2] The raw data include a referral field which allows us to tag visits referred from social media sites. We also construct the referral variable from lagged visits, but the measures are highly correlated so we use the referral field.

[3] Following the literature on preference externalities, we rely on racial classifications for our analyses not to study race *per se*, but because of well-documented evidence that preferences for media products differ markedly across racial groups in measurable ways. Population by race thus offers a clear and practical measure of shared tastes in a market. In our data, users self-report racial categories as black, white, or asian so we choose this over more nuanced or diverse categories.

Table 2: News Consumption by Market Quartile

	All Visits	Local (W)	Local (B)	Non-local (W)	Non-local (B)
1st Quartile	12.106	1.549	0.563	11.871	7.634
2nd Quartile	12.184	1.767	0.581	12.285	7.206
3rd Quartile	12.554	1.647	0.521	12.684	7.619
4th Quartile	13.248	1.861	0.693	13.646	9.291
Total	12.520	1.696	0.600	12.564	8.065

estimate the following:

$$V^W_{i,m,t} = \alpha_0 + \alpha_1 W_m + \alpha_2 B_m + \gamma' C_{i,t} + \epsilon^W_{i,m,t}, \tag{1}$$

$$V^B_{i,m,t} = \beta_0 + \beta_1 W_m + \beta_2 B_m + \gamma' C_{i,t} + \epsilon^B_{i,m,t}, \tag{2}$$

where $V^W_{i,m,t}$ measures visits to news outlet of a white household i, residing in MSA m, in month t, $V^B_{i,m,t}$ is similarly defined for black households i. W_m is the number of whites in the MSA, and B_m is the number of blacks in the MSA (estimated in millions).

Because some households do not make any news visits in some months, we also estimate the model with $V^W_{i,m,t}$ and $V^B_{i,m,t}$ defined as the probability of a visit to a news outlet in a second step.

The vector $C_{i,t}$ includes household-specific control variables and a time trend implemented with monthly fixed effects. The constants α_0 and β_0 reflect each group's average propensity to consume news online. If readership increases with shared tastes, we expect a positive relationship between "own" group population and readership. In George and Waldfogel (2003) the effect of a larger white population on white circulation was small compared to the black population effect on black circulation, suggesting diminishing returns to larger population. The effect of individuals with different tastes on readership is ambiguous, but should be less than the own effect. More precisely, we expect $\alpha_1 > \alpha_2$ and $\beta_2 > \beta_1$, also $\alpha_1 > 0$ and $\beta_2 > 0$. [4]

[4]Results in George and Waldfogel (2003) indicated $\beta_1 < 0$, with more whites in a market reducing

Table 3 present estimates of equations (1) and (2). Columns (1) and (2) estimate the model using news visits (transformed as log(visits+1)) and columns (3) and (4) with probability of a news visit. In both cases, shared tastes matter for minority readership: a larger black population is associated with higher black readership. An additional million blacks increases total news visits by 20% and the probability of a news visit by 6 percentage points. An additional million whites in the market is associated with an increase in white visits to news outlets by 5%. The population effect in the linear probability model is positive but not statistically significant, suggesting that the effect of an additional person diminishes with larger populations.

Results in the tables also suggest that cross-effects matter: an increase in the white population reduces news visits and the likelihood of a news visit by blacks, though the estimate is smaller and noisier than with own effects. An additional million white residents reduces black news visits by 4% and the probability of a news visit by .016 percentage points. The size of the black population does not appear to affect white readership.

It is worth noting that the pattern of coefficients is very similar to that found in George and Waldfogel (2003) using zipcode-level newspaper circulation. While the coefficients are not directly comparable, the effects measured here are comparable in magnitude to those measured with aggregate purchasing data.

With this evidence that preference externalities operate in markets for online news, we turn to an investigation of theoretical mechanisms. If shared preferences increase consumption by bringing forth in a market more products suited to those preferences, the link between group size and consumption should center on local products, or perhaps products that are produced nationally but whose availability is dictated by local demand. [5] With supply-side mechanism, we would not expect the size of the local

per capita readership among blacks. However because that study used zipcode aggregates rather than individual data, the focus was on relative effects

[5]For example, during the 1990's the *New York Times* expanded first into markets with a large college-educated population.

Table 3: Does Group Size Matter in Digital News Consumption?

	(1) Black Visits	(2) White Visits	(3) Black Pr Visit	(4) White Pr Visit
Black MSA Pop	0.212***	-0.0586	0.0598***	-0.00687
	(0.0405)	(0.0367)	(0.0105)	(0.00709)
White MSA Pop	-0.0421	0.0453***	-0.0162**	0.00404
	(0.0276)	(0.0139)	(0.00652)	(0.00269)
Constant	0.957***	1.185***	0.496***	0.566***
	(0.0413)	(0.0482)	(0.0135)	(0.0134)
Observations	146352	321996	146352	321996
Adjusted R^2	0.021	0.014	0.015	0.011
Mean Dep. Var.	1.258	1.577	0.612	0.684
Mean Black MSA Pop	0.553	0.362	0.553	0.362
Mean White MSA Pop	1.566	1.326	1.566	1.326

Dependent variable in columns 1-2 is log transform of all news visits. Dependent variable in column 3-4 is an indicator for at least one monthly household news visit. All specifications include month fixed effects and controls for household size, income and age, not shown. Standard errors clustered by MSA: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

group population to affect consumption of national products targeted to national preferences. We can investigate this by studying local and national news site visits online. The targeting of national media should not be related to the size of a population in any market, and national media can be accessed freely across markets online.

We re-estimate equations (1) and (2) separately for local and non-local news visits. Results are shown in table 4 for (log transform) news visits and in 5 for linear probability estimates. As expected from a supply-driven model, group size matters for consumption in local media markets: a larger black population has a positive effect on local media consumption among blacks and a larger white population has a positive effect on local news consumption among whites. Black readership declines with a larger white population, consistent with findings in the literature. The coefficient patterns with linear probability models are similar to those with the number of visits. But the tables also show that shared preferences matter for consumption of non-local media

Table 4: Local Preferences and Non-Local Media (News Visits)

	(1) White Local	(2) Black Local	(3) White Nonlocal	(4) Black Nonlocal
Black MSA Pop	-0.0334	0.0902***	-0.0456	0.200***
	(0.0475)	(0.0154)	(0.0323)	(0.0409)
White MSA Pop	0.0350**	-0.0200**	0.0379***	-0.0378
	(0.0166)	(0.00873)	(0.0137)	(0.0282)
Constant	0.175***	0.113***	1.143***	0.930***
	(0.0227)	(0.0141)	(0.0472)	(0.0407)
Observations	321996	146352	321996	146352
Adjusted R^2	0.004	0.010	0.015	0.021
Mean dependent variable	0.332	0.176	1.501	1.217
Mean MSA black population	0.362	0.553	0.362	0.553
Mean MSA white population	1.326	1.566	1.326	1.566

Dependent variable is log transform of visits to local and non-local news outlets. All specifications include month fixed effects and controls for household size, income and age, not shown. Standard errors clustered by MSA: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Local Preferences and Non-Local Media (Linear Probability Model)

	(1) White Local	(2) Black Local	(3) White Nonlocal	(4) Black Nonlocal
White MSA Pop	0.0220**	-0.0125**	0.00365	-0.0155**
	(0.00989)	(0.00600)	(0.00257)	(0.00676)
Black MSA Pop	-0.0129	0.0691***	-0.00553	0.0588***
	(0.0266)	(0.0103)	(0.00643)	(0.0107)
Constant	0.128***	0.0925***	0.559***	0.490***
	(0.0120)	(0.00912)	(0.0133)	(0.0134)
Observations	321996	146352	321996	146352
Adjusted R^2	0.006	0.013	0.012	0.015
Mean dependent variable	0.227	0.144	0.674	0.605
Mean MSA black population	0.362	0.553	0.362	0.553
Mean MSA white population	1.326	1.566	1.326	1.566

Dependent variable is probability of a visit to a local or non-local news outlet. All specifications include month fixed effects and controls for household size, income and age, not shown. Standard errors clustered by MSA: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

products, which cannot be explained by a supply-driven model. The magnitudes are also large for national media: a one standard deviation increase in black population (.22) increases black visits to non-local media outlets by more than 4%, more than double the comparable increase in local news visits. With the linear probability models, the effects are closer to parity, with a one standard deviation increase in black population raising the likelihood of a local or non-local news visit by about .015.

As in the overall tables, a larger white population is also associated with higher white readership, but the effects are smaller and in some cases not different from zero at standard significance levels. This is consistent with diminishing returns to benefits from shared tastes.

4 Social Networks and the Demand for News

The results above indicate that a larger population with shared tastes in a market can increase consumption of non-local products. This suggests that demand rather than supply effects play a role in the observed relationship. We turn now to evidence on whether larger markets foster deeper social networks, and that it is these networks that drive the demand for news.

To tackle this question, we investigate whether a larger group population is associated with more intense use of social media sites on the internet. As a first step,we estimate a social media analog of equations (1) and (2) with visits to social media sites as a dependent variable. A better measure would perhaps be the size of social networks, which we do not observe. (We do observe minutes per month spent at social media sites, but this is highly correlated with visits so we adopt the straightforward visit measure.) Results are shown in table 6. The pattern and magnitude of coefficients is similar to that in 4, with an increase in the market population of blacks of one standard deviation (.22m) raising social media visits by about 4%. A larger white population is associated with fewer visits to online social media. While still small, population explains a larger

Table 6: Do Shared Preferences Enhance Social Networks?

	(1) Black Visits	(2) White Visits
Black MSA Pop	0.188***	-0.0436**
	(0.0336)	(0.0185)
White MSA Pop	-0.0455***	-0.00500
	(0.0175)	(0.00994)
Constant	2.474***	2.476***
	(0.0536)	(0.0588)
Observations	146352	321996
Adjusted R^2	0.034	0.042
Mean Dep. Var.	2.664	2.487
Mean Black MSA Pop	0.553	0.362
Mean White MSA Pop	1.566	1.326

Dependent variable is log transform visits to social media outlets per month. All specifications include month fixed effects and controls for household size, income and age, not shown. Standard errors clustered by MSA: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

share of the variation (about double) in social media visits compared to news visits.

One implication of this result is that larger groups with shared preferences offer more opportunity for members to foster social relationships that might, in part, stimulate interest in current affairs and culture. [6] To evaluate whether social networks play a role in the relationship between group size and media consumption, we want to estimate the news consumption regressions in 4 with social media use as a dependent variable. With a control for social media use, the population coefficients will then reflect the supply-side drivers of preference externalities with demand-side drivers absorbed into the social media coefficients. Because population also drives social media use, the best approach is a two equation model in a seemingly-unrealated-regression. For this approach we want instruments that are correlated with social networks and social media use but not correlated with news consumption. Following the literature [Dora L. Costa and Matthew E. Kahn, The American Economic Review, Vol. 97, No. 4 (Sep., 2007), pp.

[6] cite peer effects literature

Table 7: Shared Preferences and Social Networks in News Consumption

	(1) Black Nonlocal	(2) Black Local	(3) White Nonlocal	(4) White Local
Black MSA Pop	0.0783***	0.157***	-0.0272	-0.0287
	(0.0146)	(0.0364)	(0.0470)	(0.0309)
White MSA Pop	-0.0174**	-0.0282	0.0332**	0.0332**
	(0.00830)	(0.0263)	(0.0165)	(0.0130)
HH Visits to Social Media	0.00174***	0.00637***	0.00297***	0.00804***
	(0.000139)	(0.000182)	(0.000115)	(0.000168)
Constant	0.0665***	0.758***	0.0803***	0.886***
	(0.0148)	(0.0372)	(0.0223)	(0.0436)
Observations	146352	146352	321996	321996
Adjusted R^2	0.040	0.085	0.040	0.091
Mean Dep. Var.	0.176	1.217	0.332	1.501
Mean Black MSA Pop	0.553	0.553	0.362	0.362
Mean White MSA Pop	1.566	1.566	1.326	1.326

Dependent variable is log transform visits to local and non-local media outlets per month. All specifications include month fixed effects and controls for household size, income and age, not shown. Standard errors clustered by MSA: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

1467-1487], we use zipcode level measures of urban density to predict social networks and social media use. Because supply-side incentives for news media to enter markets center on market-level population, zipcode density measures should be uncorrelated with msa group population.

Table 7 reports results for this estimation. Including social media usage in the regression reduces the estimated effect of black population on black readership of non-local media by about 25%, from .2 to .15. Adding the control has a smaller effect on local news readership, a reduction from .09 to .08. This is consistent with a view that entry and product positioning play a dominant role in local media markets, but social networks are more important for national consumption. The effect of additional whites in a market on white readership, a small effect even in table 4, is reduced by a small amount in table 7. Social media has a bigger effect in the non-local than local news

specification.

It should be noted that in these cross-sectional regressions, social media usage is likely picking up some unobserved user characteristics such as time spent online, occupation, etc. Because we are not interested in directly measuring the relationship between social media usage and news reading, unobserved heterogeneity of this sort is not a fundamental problem with this specification. [Robustness check].

[News Referrals]

5 Conclusions

We show in this paper that preference externalities operate for national as well as local media products. This result challenges the standard theory that entry and targeting in local markets drives the effect of shared preferences on consumption in media markets. We show that social networks can explain the relationship between group population and consumption for non-local media, and that including measures of social media reduces the effect of black population on black readership by about 25%, while altering the estimated effect on local media consumption to a much lesser degree. However this proxy does not fully explain the presence of preference externalities in national media, so more work is needed to understand "who benefits whom" in these markets.

References

Berry, S., and Waldfogel, J. (2010). "Product Quality and Market Size." *The Journal of Industrial Economics*, *58*(1), 1–31.

George, L., and Hogendorn, C. (2013). "Local News Online: Aggregators, Geo-Targeting and the Market for Local News." *Working Paper*.

George, L., and Waldfogel, J. (2003). "Who Affects Whom in Daily Newspager Markets?" *Journal of Political Economy*, *111*(4), 765–784.

Oberholzer-Gee, F., and Waldfogel, J. (2009). "Media Markets and Localism: Does Local News en Español Boost Hispanic Voter Turnout?" *American Economic Review*, *99*(5), 2120–2128.

Waldfogel, J. (2003). "Preference Externalities: An Empirical Study of Who Benefits Whom in Differentiated-Product Markets." *RAND Journal of Economics*, *34*(3), 557–568.

Waldfogel, J. (2004). "Who Benefits Whom in Local Television Markets?" *Brookings-Wharton Papers on Urban Affairs*, 257–284.

Wang, X., and Waterman, D. (2011). "Market Size, Preference Externalities, and the Availability of Foreign Language Radio Programming in the United States." *Journal of Media Economics*, *24*, 111–131.