Diverse News Media Decreases Misperceptions about Immigration*

Hanno Hilbig† and Sascha Riaz‡

April 1, 2019

Abstract

Immigration is one of the most salient political issues in Europe and the United States. Perceptions of immigrants shape preferences for redistribution, and likely contributed to the surge of populist parties and candidates. Recent research has shown that misperceptions about the relative size of the immigrant population are widespread. We investigate the news media as a source of those misperceptions: Using survey data in conjunction with a comprehensive data set on the geographic coverage of all German newspapers, we examine how diversity in the local market for news shapes beliefs about the relative size of the immigrant population. Our identification strategy exploits exogenous variation in the number of newspapers available in small-scale geographic units as a result of partially overlapping delivery areas. We present causal evidence that access to a greater number of local news sources decreases misperceptions.

*We thank Merlin Schaeffer and Ruud Koopmans for sharing their data. We thank Pia Raffler, Michael Hiscox, Kosuke Imai, Michael Zoorob, Sun Young Park and Anselm Hager for helpful comments.

†Department of Government, Harvard University, hhilbig@g.harvard.edu

‡Department of Government, Harvard University, riaz@g.harvard.edu

1 Introduction

In a large number of countries, immigration has become a defining policy issue. The New York Times observes that immigration "is reshaping societies around the globe" (Porter and Russell, 2018). A poignant example is one of the core promises of Donald Trump's 2016 campaign: building a wall between Mexico and the United States to limit migration from Latin America to the United States. Similarly, right-wing populist parties have surged across Europe, partly as a response to increasing rates of immigration in recent years. However, recent research suggests that anti-immigrant sentiments are based on perceptions which are far from the truth. Constituents frequently have biased beliefs about the proportion and the characteristics of immigrants. In the US and in Europe, Natives overestimate the proportion of immigrants, both at the national and local level(Alesina, Miano and Stantcheva, 2018; Sides and Citrin, 2007*a*; Hopkins, Sides and Citrin, 2018). Such misperceptions can have far-reaching consequences: In Europe, the perceived share of immigrants has been shown to be a better predictor of anti-immigrant sentiments and social cohesion than the true proportion of immigrants (Koopmans and Schaeffer, 2016; Hooghe and Vroome, 2015). Alesina, Miano and Stantcheva (2018) argue that "the political debate about immigration takes place in a world of misinformation". As biased perceptions have become increasingly salient, a logical next step is to investigate (1) how institutions increase or decrease individual bias and (2) how bias can be mitigated.

In this paper, we examine the connection between misperceptions and an institution that is central to the functioning of democracy: the news media. There is little work that directly examines how societal institutions affect the accuracy of individual perceptions[1]. We ask whether a more diverse local media landscape, as measured by total number of available outlets, affects misperceptions about immigrants. We assume that local news is one of the main sources of information about the local proportion of immigrants. Based on previous research, we also assume that local media outlets likely slant in one way or another (see Gentzkow and Shapiro, 2010; Mullainathan and Shleifer, 2005). We posit that, even when local media reporting on immigrants is biased, greater diversity of outlets can lead to more accurate perceptions when individuals aggregate information. This can happen in two ways: First, individuals can directly aggregate information from a variety of different sources, such that slanting is partially 'neutralized' through aggregation. Second, aggregation of infor-

[1]Previous studies have looked at individual correlates of misperceptions, e.g. Hopkins, Sides and Citrin (2018), Sides and Citrin (2007*b*) and Sides and Citrin (2007*a*). In contrast, we explicitly focus on a contextual factor, diversity of the local news media.

mation may happen through diffusion of information in social networks: In contexts where there is a greater diversity of local news outlets, readers are exposed to more heterogeneous information. When individuals exchange information with each other, aggregation of information can happen in an indirect manner, even when individuals only read one news source.

This study focuses on Germany, a country that has experienced a considerable amount of labor migration since the 1950s, as well as, more recently, the influx of over 1.5 million refugees between 2015 and 2017. In 2011, the combined share of first and second generation immigrants in Germany was about 18.9%, putting it in the top quartile of European countries with the largest foreign-born population.[2] To measure misperceptions about immigrants, we rely on a large-scale survey that was conducted in 2011 (see Koopmans and Schaeffer, 2016). About 5,200 survey respondents explicitly reported their perceptions of the relative size of the immigrant population in their local neighborhood. In conjunction with census data on the true share of immigrants in those areas, this allows us to measure misperceptions at the local level.

To measure media diversity, we combine the survey data with a comprehensive data set which documents the coverage areas of all local newspapers in Germany. Our empirical strategy exploits the fact that partial overlap between the delivery areas of newspapers is common, resulting in discontinuous changes in access to local news between two adjacent zip code regions. Conditional on sufficient geographical closeness, we assume that treatment assignment can be considered as-if random, similar to the design in studies such as Larreguy, Marshall and Snyder (2014) or Ansolabehere, Snowberg and Snyder (2006). We operationalize media diversity as a binary treatment: We consider individuals 'treated' if they live in regions with local local newspaper monopolies: This means that individuals only have access to one local newspaper, while neighboring survey respondents can obtain information from two or more newspapers. We implement a matching algorithm with a geographic distance caliper to ensure individuals with differing treatment statuses are physically close as well as similar with respect to a variety of observable characteristics.

We demonstrate that local media market diversity reduces individual-level misperceptions of the size of the immigrant population: In monopolistic media markets, misperceptions are between three and four percentage points higher than in segmented media markets. The

[2]Specifically, this number refers to individuals with a 'migration background' as defined by the German Federal Statistical Office: a person has a 'migration background' if that person or at least one parent did not possess German citizenship at birth.

results hold both for the distance matching specification, as well as for a more conventional OLS specification that includes small-level geographic fixed effects. In addition to our main result, we document a number of descriptive facts. First, we show that survey respondents overestimate the proportion of immigrants in their local area by a factor of about 1.8, echoing previous studies that examine the extent of misperceptions[3]. Second, the perceived proportion of immigrants is positively correlated with the true share of immigrants, indicating that changes in the true number of immigrants are, to some degree, reflected in the perceptions of individuals.

Our findings contribute to the literature in a number of ways: We present the first causal estimates of the relationship between diversity of the local news media and the degree to which perceptions of immigrants are subject to bias. In doing so, we underline the relevance of local media as an important source of information for constituents. We confirm prior findings that show that, even in times of the internet and social media, traditional news outlets shape perceptions, discussions and behavior around salient national political issues (see e.g. King, Schneer and White, 2017; Olson, 2018; Gentzkow, Shapiro and Sinkinson, 2011). Second, we demonstrate that the provision of information can induce changes in individual perceptions. When individuals obtain their information from a more diverse set of sources, misperceptions are less pronounced. This finding speaks to experimental work such as Hopkins, Sides and Citrin (2018) or Lawrence and Sides (2014), who show that information treatments can, to some degree, decrease misperceptions. Consequently, our work demonstrates that misinformation and biased perceptions are not set in stone: Rather, a healthy and diverse media landscape can counter false beliefs.

The rest of the paper proceeds as follows: we first situate our research in the burgeoning literature on misperceptions about immigrant populations. We then describe our data sources, the operationalization of our outcome of interest, and the implementation of our identification strategy through a matching algorithm. Finally, we present our main results together with a large number of robustness checks across different model specifications.

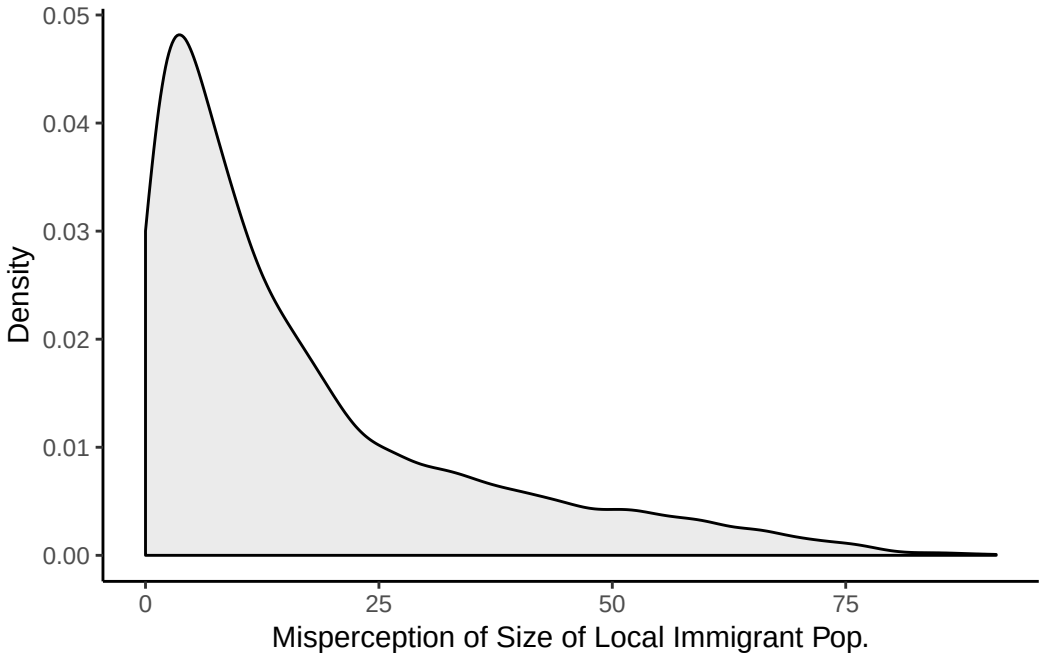
[3]Alesina, Miano and Stantcheva (2018) show that German natives overestimate the number of immigrants by a factor of two. It should, however, be noted that Alesina, Miano and Stantcheva (2018) ask respondents to estimate the number of immigrants at the national level, while our data asks for the immigrants on the local level.

2 Related Research

The fact that individuals tend to overestimate the number of out-group members is now well-established. This holds both in the context of immigrants (Alesina, Miano and Stantcheva, 2018; Herda, 2010; Koopmans and Schaeffer, 2016; Sides and Citrin, 2007*b*), as well as for the minority populations more generally (Hopkins, Sides and Citrin, 2018). It is natural to ask how perceptions affect political preferences and attitudes. Alesina, Miano and Stantcheva (2018) conduct large-scale survey experiments in five European countries and the U.S., concluding that support for redistribution is correlated with the perceived *composition* of immigrants rather than the perceived share of immigrants per se. This is consistent with previous research by Hainmueller and Hiscox (2010) emphasizing the salience of ethnocentric and sociotropic concerns in natives' attitudes towards migrants. Survey experiments in which participants are treated with accurate information about the true size of the immigrant population have yielded mixed results: Lawrence and Sides (2014), Hopkins, Sides and Citrin (2018), and Sides and Citrin (2007*b*) provide evidence that correcting misperceptions has little impact on attitudes toward immigration in the American context, even though it does reduce the perceived size of the immigrant population. Alesina, Miano and Stantcheva (2018) find that providing natives with information about the hard work immigrants commonly perform does increase support for redistribution whereas information on the true size and origin of the immigrant population does not. Facchini, Margalit and Nakata (2016) and Grigorieff, Roth and Ubfal (2016) on the other hand find that information treatments can significantly shift attitudes in support for a more open immigration policy in a variety of countries including Russia and Japan.

To our knowledge, only one study has analyzed the relationship between the media and misperceptions about immigration: Herda (2010) uses data from the European Social Survey to explain why individuals hold biased beliefs about the size of the immigrant population. Herda regresses misperceptions on self-reported measures of radio, newspaper and TV consumption. His results are ambiguous: Media consumption is associated with misperceptions, but the direction of the effect depends on the type of media that respondents consume. While this is an important first stab at the problem, Herda is not able to make causal claims: He cannot rule out confounding or reverse causality, since respondents likely self-select into different types of media consumption (Freddi, 2017).

Figure 1: Distribution of outcome variable 'misperception': absolute difference between the perceived and true proportion of immigrants in the local environment of individual i living in zip code region j.



3 Data and Empirical Strategy

We draw on two data sources: The *Ethnic Diversity and Collective Action Survey* (EDCAS), and a comprehensive data set of local newspaper coverage areas that we obtained from the media market research company *Zeitungsmarktfoschung Gesellschaft* (ZMG)[4] With an effective sample size of around 5,200, EDCAS is, to our knowledge, the largest and most comprehensive survey that measures individuals' perceptions of the size of the local immigrant population in a single European country. In addition to standard socio-economic and demographic background characteristics, EDCAS includes a large number of survey items related to attitudes towards immigrants, social capital and social cohesion as well as trust.

3.1 Outcome: Misperceptions

Our primary outcome of interest are individuals' misperceptions about the relative size of the local immigrant population. We measure this by combining an EDCAS item that asks respondent to estimate the share of immigrants in their local neighborhood with data from

[4]ZMG is a part of the "Federation of German Newspaper Publishers" (BdZV), and is tasked with providing information to companies that want to advertise in German local newspapers. More information can be found at `https://www.zmg.de`.

the 2011 German Census. The survey explicitly asks respondents to give an estimate of the *combined* share of first to second generation immigrants. Accordingly, we apply the same definition for our measure of the true share of immigrants in a given locality. We aggregate the census data, which is provided at the $1\,\text{km}^2$ grid-level, up to the zip code level (each grid is matched to the closest zip-centroid). Assuming that the ratio of first to second generation immigrants is largely constant within counties, we use official data from the German Federal Statistical Office on the true proportion of immigrants at the county-level in combination with the census data to estimate the true % combined share of first and second generation immigrants at the zip code level. The census data by itself only contains information about the share of first-generation immigrants in a given grid area. We operationalize misperceptions of individual i living in zip code region j as the absolute difference between the perceived and true proportion of immigrants in the local environment:

$$Y_{i,j} = |\text{Immigrants}_{i,j}^{\text{Perceived}} - \text{Immigrants}_{j}^{\text{True}}|$$

In Figure 1, we show that misperceptions are widespread: The average individual in our sample overestimates the true proportion of immigrants by about 17 percentage points, a large magnitude given that the average size of the immigrant population across zip codes is merely 13.38% with a standard deviation of 8.31%. Figure 2 displays the perceived share of immigrants conditional on the true share in each zip code area. Individuals consistently overestimate the true proportion of immigrants, although misperceptions seem less severe when the true proportion of immigrants exceeds 40%.

3.2 Treatment: Media Market Diversity

We are interested in the effect of the structure of the local media market on misperceptions. We define a binary treatment variable $T_{i,j}$, where $T_{i,j} = 1$ if individual i lives in a zip code region j in which only a single local newspaper is available and $T_{i,j} = 0$ otherwise. This binary operationalization of the treatment variable $T_{i,j} \in \{0, 1\}$ accurately captures the relationship between the number of available outlets and the competitiveness of the local market for news. To show this, we calculate the Herfindahl index H_j for each zip code region j in Germany. This common metric of market competitiveness is defined as:

$$H_j = \sum_{i=1}^{N} s_{i,j}^2$$

where N is the total number of local newspapers available in zip code j and $s_{i,j}$ is the market

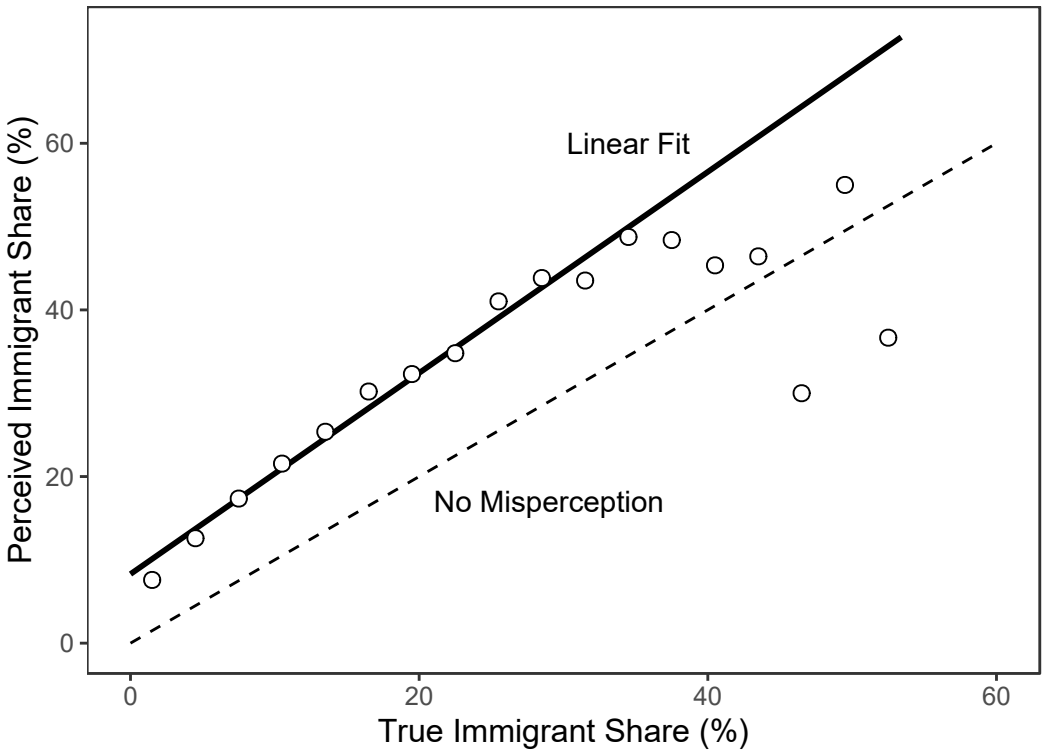


Figure 2: Perceived vs. actual share of immigrants. The x-axis shows the true immigrant share in each zip code area, and the y-axis shows the average perceived share of immigrants for respondents in the respective zip code areas. The binned scatter plot shows average perceptions conditional on levels of true immigrant shares. The dashed line illustrates a scenario in which perceptions are unbiased, i.e. the true share on average equals the perceived share. The solid line displays predicted values from a linear regression model fitted to the EDCAS data.

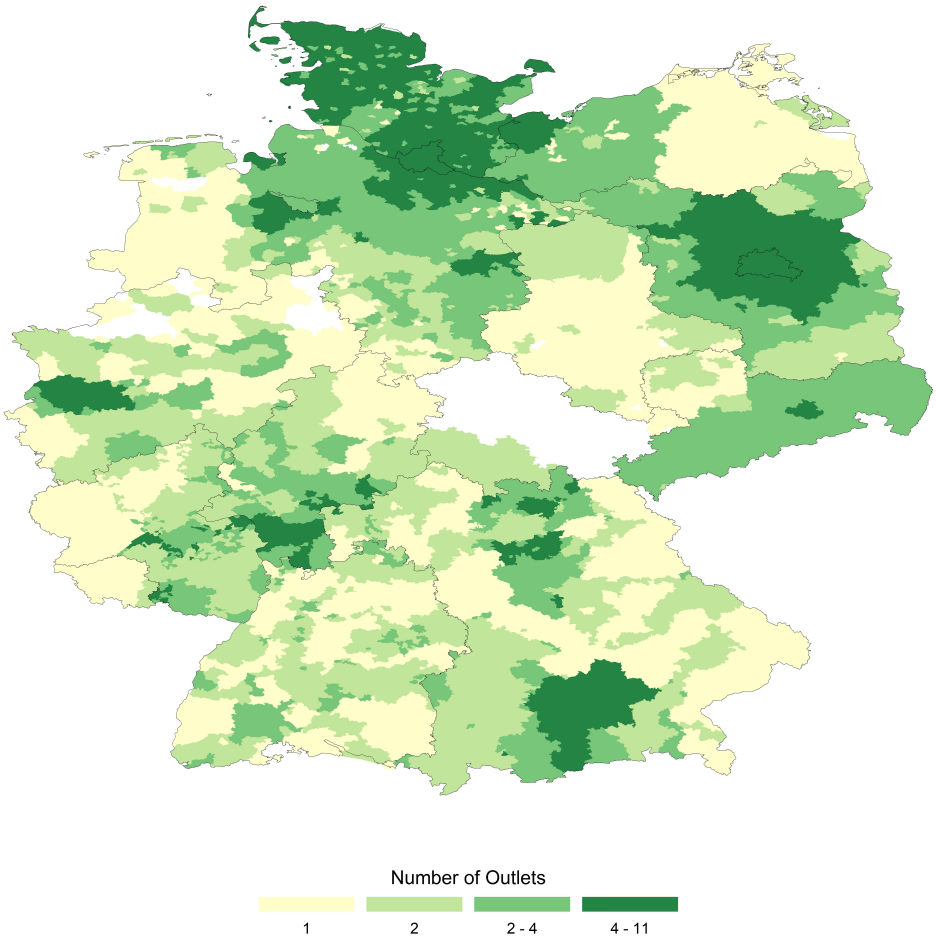
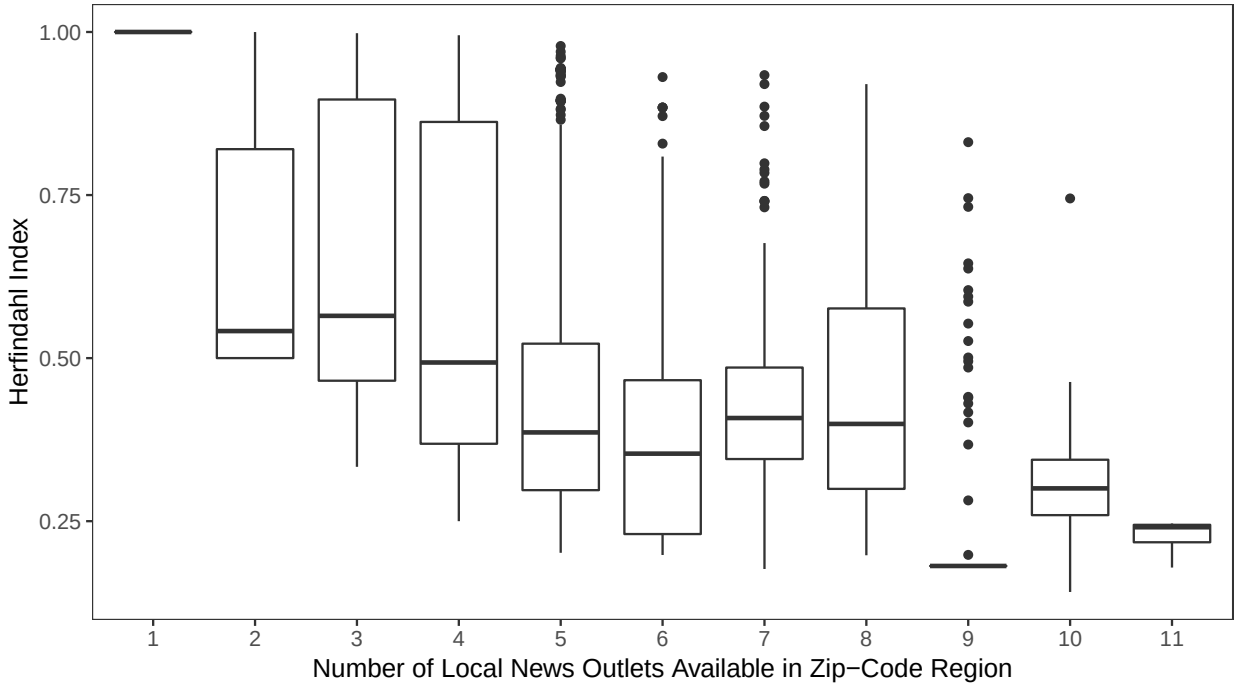


Figure 3: Number of local news outlets available in each zip code region in Germany. darker shaded areas indicate that individuals can choose between a greater number of local news outlets. White areas indicate missing data.

Figure 4: Relationship between the number of local newspapers available in a zip code region and the Herfindahl index of media market competition H_j. Each box shows the 25^{th}, 50^{th}, and 75^{th} percentile of H_j. Whiskers extend to 1.5 times the interquartile range.



share of newspaper i in regional market j. The index ranges from $\frac{1}{N}$ to 1, where $H_j = 1$ indicates a perfectly monopolistic market whereas $H_j = \frac{1}{N}$ describes a situation of perfect competition in which all media outlets have an equal share in the local market. As illustrated in Figure 4, non-monopolistic markets are much more competitive than monopolistic ones. However, competitiveness as measured by the Herfindahl index increases only modestly beyond two local newspapers. Figure 3 illustrates the structure of local news media markets across Germany, showing the count of outlets available in each zip code region. In 31.1 % of Germany's zip code regions, the market for local news is monopolistic.

3.3 Geographic Matching

To identify the causal effect of local newspaper monopolies on misperceptions, we exploit exogenous geographic variation in the coverage of newspaper markets. Specifically, we compare individuals in zip code regions in which only a single newspaper is available to individuals that are similar with respect to a variety of observable characteristics but are situated in neighboring zip code area in which the market for local news is segmented. We implement this identification strategy through a matching algorithm:

1. For each treated unit situated in a zip code region under a local newspaper monopoly $T_{i,j} = 1$, we identify all control units available in the same state within a radius of z kilometers. We conduct robustness checks by letting z vary between ten and fifty kilometers.

2. Among the set of potential control units, we choose the M units that are closest to treated unit i in terms of Mahalanobis distance on individual level and zip code level covariates[5]. As recommend by Imbens and Rubin (2015, p. 451), we vary the maximum number of control units that are matched to each treated unit $M \in \{1, 2, 3, 4\}$ as an additional robustness check.

We include first and second generation migration background, gender, age, employment status, and education on the individual level and population density and the unemployment rate at the zip code level as covariates in this step. We perform matching as pre-processing step of our dataset, discarding all treated units for which no match within a radius of z kilometers could be found. We also remove all unused control units from the dataset. We match with replacement, i.e. one control unit can be matched to multiple treated units We stress that this does not artificially inflate our sample size by duplicating control observations. We hence prioritize unbiasedness over variance reduction. The core identification assumption underlying our analysis is that in the dataset pre-processed as described above, treatment assignment is independent of potential outcomes:

$$T_{i,j} \perp\!\!\!\perp Y_{i,j}(t) \tag{1}$$

In Figure A.1, we present balance on observables before and after matching and at varying distance cutoffs z. Our matching algorithm considerably improves balance on observables, particularly with respect to zip code level characteristics. For all models, we use robust standard errors clustered at the zip code level (Abadie et al., 2017).

4 Results

In Table 1 from OLS specification, using the matched data set that we constructed using the matching algorithm described in Section 3.3. We find that local media market diversity reduces individual-level misperceptions: For individuals in monopolistic newspaper markets,

[5]The Mahalanobis distance $d(\vec{x_{i,j}}, \vec{x_{k,l}})$ between two observations is defined as:

$$d(\vec{x}_{i,j}, \vec{x}_{k,l}) = \sqrt{(\vec{x}_{i,j} - \vec{x}_{k,l}) S^{-1} (\vec{x}_{i,j} - \vec{x}_{k,l})}$$

where S is the covariance matrix of the covariates we balance treated and control observations on.

misperceptions are about three to four percentage points greater than for those in segmented markets, depending on the specification. The estimates of the average treatment effect on the treated (ATT) are significant at conventional levels of statistical significance.

Our main results are based on a choice of distance caliper of $z = 30$ km. We match each treated unit to a single optimal control unit, such that $M = 1$. Since we face a bias-variance trade-off, the choice of z and M is not straightforward: While the total sample size increases in z and M (see Figure A.3 in the SI), the quality of the average match for each treated unit (in terms of Mahalanobis and geographic distance) decreases. To verify that our findings are not strongly dependent on the parameters of the matching procedure, we implement an extensive sensitivity analysis. In Figure A.2 in the SI, we vary the distance caliper z as well as the maximum number of control units M. In general, we observe that our results are robust to different combinations of M and z. Given that smaller distance calipers result in a lower number of observations, we observe that estimates for smaller z are somewhat less precise. However, both the direction and magnitude of the observed effect remains relatively constant for different combinations of z and M.

Across all 72 model specifications estimated in the sensitivity analysis (see Figure A.2), the mean ATT estimate is 3.16, which corresponds to a 0.18 standard-deviation increase in the dependent variable.

As a final robustness check, we present results from reduced form regressions in which we approximate our matching algorithm through a within-county comparison of monopolistic vs. segmented media markets at the zip code level. Table A.1 in the appendix show the reduced form regression results that are consistent with the results presented in Figure A.2.

Table 1: Post-matching Regressions. Standard errors are clustered at the zip-code level.

	DV: Misperception		
	(1)	(2)	(3)
Media market monopoly	3.739**	4.242***	3.165**
	(1.756)	(1.524)	(1.559)
Individual-level Covariates	No	Yes	Yes
Zip-level Covariates	No	No	Yes
Distance caliper	30 km	30 km	30 km
Max. no of. controls per treated	1	1	1
N	564	564	564
R^2	0.010	0.139	0.169

***Significant at** Significant at the 5 percent level. the 1 percent level.*Significant at the 10 percent level.

5 Conclusion

Misperceptions about immigrants are widespread and associated with xenophobic attitudes at the individual level. This paper has explored one mechanism through which such misperceptions can be addressed. We have presented causal evidence that media market competition at the local level has a strong negative effect on misperceptions. While each individual media outlet may slant its reporting in one way or another, individuals in non-monopolistic newspaper markets can aggregate information from a variety of sources and thereby improve the accuracy of their beliefs. Our results point to the societal and political pathologies of monopolistic newspaper markets: not only are local media monopolies detrimental for consumers, they also lead to aggravated misperceptions about vulnerable minority groups. Media market monopolies are not only economically inefficient but entail substantial negative political externalities.

References

Abadie, Alberto, Susan Athey, Guido W Imbens and Jeffrey Wooldridge. 2017. "When should you adjust standard errors for clustering?".

Alesina, Alberto, Armando Miano and Stefanie Stantcheva. 2018. Immigration and Redistribution. Technical Report w24733 National Bureau of Economic Research Cambridge, MA: .

Ansolabehere, Stephen, Erik C. Snowberg and James M. Snyder. 2006. "Television and the Incumbency Advantage in U.S. Elections." *Legislative Studies Quarterly* 31(4):469–490.

Facchini, Giovanni, Yotam Margalit and Hiroyuki Nakata. 2016. "Countering public opposition to immigration: The impact of information campaigns.".

Freddi, Eleonora. 2017. "Do people avoid morally relevant information? Evidence from the refugee crisis.".

Gentzkow, Matthew and Jesse M. Shapiro. 2010. "What Drives Media Slant? Evidence From U.S. Daily Newspapers." *Econometrica* 78(1):35–71.

Gentzkow, Matthew, Jesse M. Shapiro and Michael Sinkinson. 2011. "The Effect of Newspaper Entry and Exit on Electoral Politics." *American Economic Review* 101(7):2980–3018.

Grigorieff, Alexis, Christopher Roth and Diego Ubfal. 2016. "Does information change attitudes towards immigrants? Evidence from survey experiments." *IZA Institute of Labor Economics, discussion paper series No* 10419.

Hainmueller, Jens and Michael J Hiscox. 2010. "Attitudes toward highly skilled and low-skilled immigration: Evidence from a survey experiment." *American political science review* 104(1):61–84.

Herda, Daniel. 2010. "How Many Immigrants?Foreign-Born Population Innumeracy in Europe." *Public Opinion Quarterly* 74(4):674–695.

Hooghe, Marc and Thomas de Vroome. 2015. "The perception of ethnic diversity and anti-immigrant sentiments: a multilevel analysis of local communities in Belgium." *Ethnic and Racial Studies* 38(1):38–56.

Hopkins, Daniel J., John Sides and Jack Citrin. 2018. The Muted Consequences of Correct Information About Immigration. SSRN Scholarly Paper ID 2798622 Social Science Research Network Rochester, NY: .

Imbens, Guido W. and Donald B. Rubin. 2015. *Causal Inference in Statistics, Social, and Biomedical Sciences*. Cambridge University Press. Google-Books-ID: Bf1tBwAAQBAJ.

King, Gary, Benjamin Schneer and Ariel White. 2017. "How the news media activate public expression and influence national agendas." *Science* 358(6364):776–780.

Koopmans, Ruud and Merlin Schaeffer. 2016. "Statistical and Perceived Diversity and Their Impacts on Neighborhood Social Cohesion in Germany, France and the Netherlands." *Social Indicators Research* 125(3):853–883.

Larreguy, Horacio A, John Marshall and Jr. Snyder, James M. 2014. Revealing Malfeasance: How Local Media Facilitates Electoral Sanctioning of Mayors in Mexico. Working Paper 20697 National Bureau of Economic Research.

Lawrence, Eric D and John Sides. 2014. "The consequences of political innumeracy." *Research & Politics* 1(2):2053168014545414.

Mullainathan, Sendhil and Andrei Shleifer. 2005. "The Market for News." *American Economic Review* 95(4):1031–1053.

Olson, Michael P. 2018. "The Print Media and the American Party System: Evidence from the 2016 US Presidential Election." *Quarterly Journal of Political Science* 13(4):405–426.

Porter, Eduardo and Karl Russell. 2018. "Migrants Are on the Rise Around the World, and Myths About Them Are Shaping Attitudes." *The New York Times* .

Sides, John and Jack Citrin. 2007*a*. "European Opinion about Immigration: The Role of Identities, Interests and Information." *British Journal of Political Science* 37(3):477–504.

Sides, John and Jack Citrin. 2007*b*. How large the huddled masses? The causes and consequences of public misperceptions about immigrant populations. In *annual meeting of the Midwest Political Science Association, Chicago, IL*.

A Tables & Figures

Table A.1

	DV: Misperception	
	Model 1	**Model 2**
Media market monopoly	3.559***	2.488*
	(0.954)	(1.387)
State FE	Yes	No
County FE	No	Yes
Covars	Yes	Yes
Outcome mean	17.36	17.36
N	4082	4082
R-squared	0.156	0.172

***p < .01; **p < .05; *p < .1

Figure A.1: Covariate balance before and after one-to-one matching as pre-processing with replacement at varying distance cutoffs. Points show standardized difference in means between treated and control groups in sample with 95% confidence intervals.

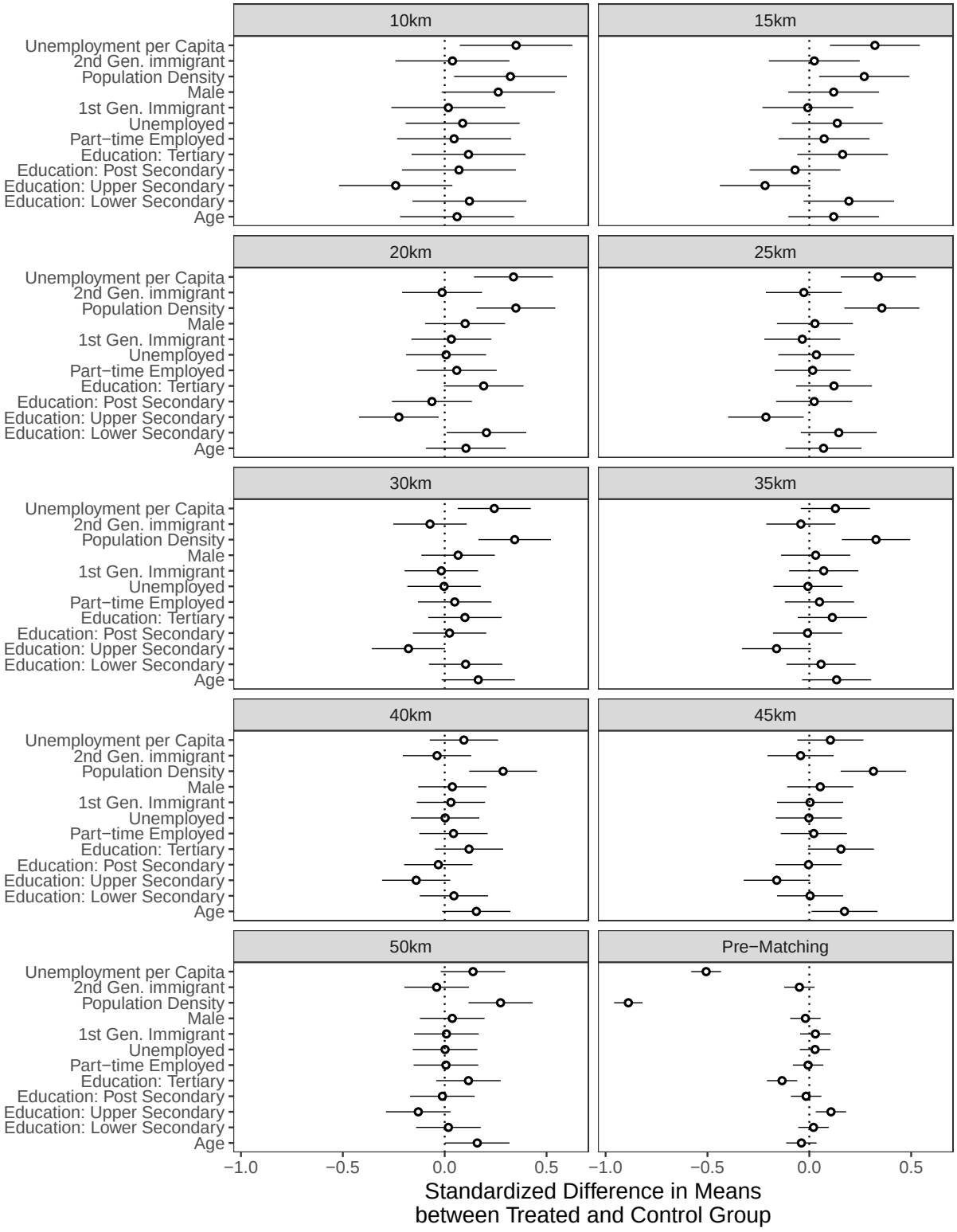


Figure A.2: ATT estimates and 95% confidence intervals after one-to-many matching with replacement based on geographic and Mahalanobis distance. The plots shows the ATT at varying treated-control maximum distance thresholds. M denotes the maximum number of control units matched to each treated unit. Standard errors are clustered by zip code. Information on the total sample size in each matched dataset is provided in Figure A.3 in the appendix.

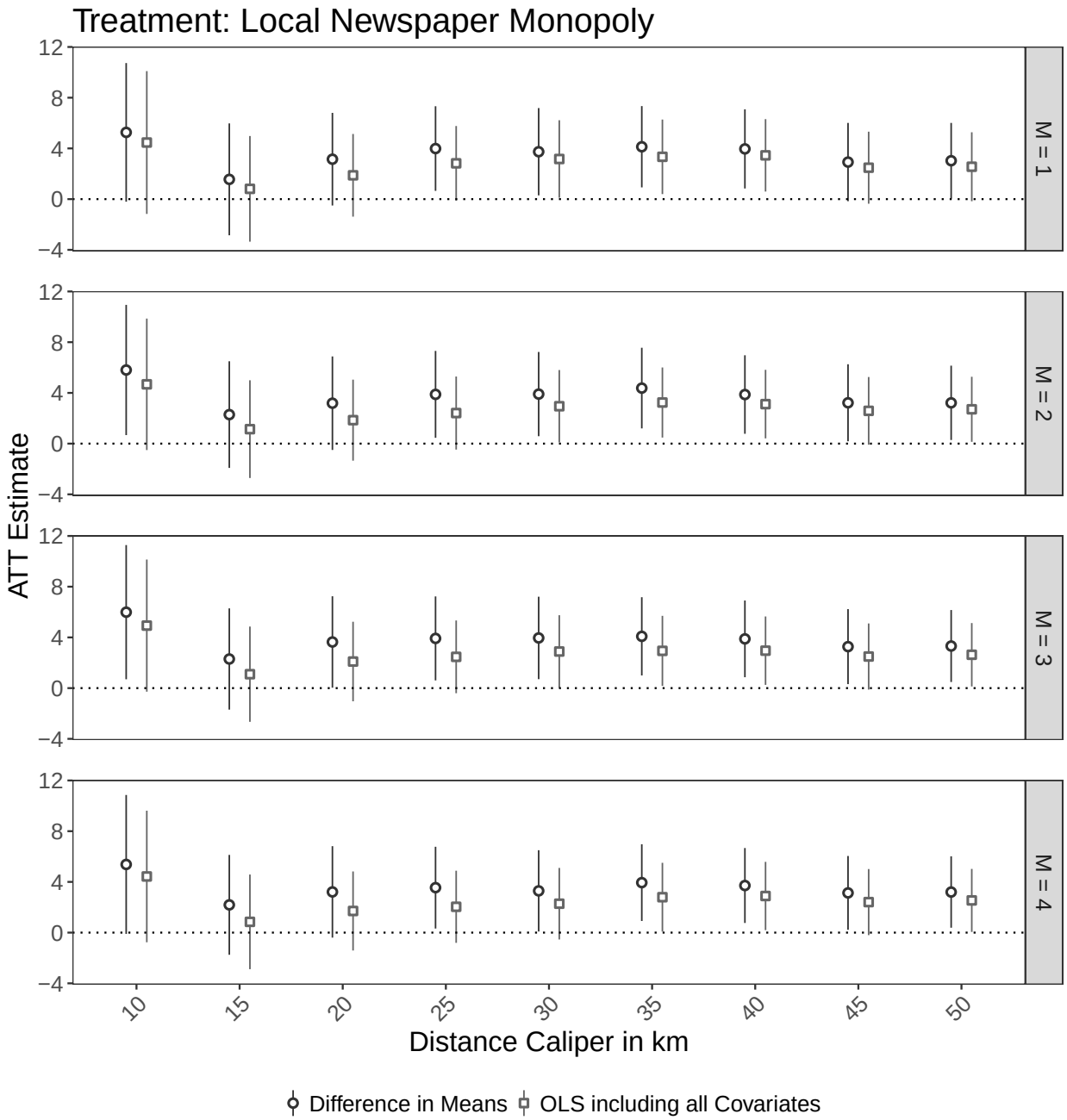


Figure A.3: Total sample size after matching at varying maximum distance thresholds and maximum number of control units matched to each treated unit.