THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITE DE TOULOUSE

Délivré par l'Université Toulouse Capitole

École doctorale : **Sciences Economiques-Toulouse School of Economics**

Présentée et soutenue par

HENNIG Jakob

le 11 Juillet 2019

Three essays in Applied Microeconomics

Discipline : **Sciences Economiques**

Unité de recherche : **TSE-R (UMR CNRS 5314 – INRA 1415)**

Directeur de thèse : Monsieur Paul SEABRIGHT, Professeur, TSE

JURY

Rapporteurs Madame Alice MESNARD, Professeur, City University of London
Monsieur Kaivan MUNSHI, Professeur, University of Cambridge

Suffragants Monsieur Matteo BOBBA, Professeur Assistant, TSE
Monsieur Paul SEABRIGHT, Professeur, TSE

Three essays in Applied Microeconomics

Jakob Hennig

PhD Thesis

Toulouse School of Economics
Université Toulouse Capitole

Acknowledgments

I am greatly indebted to various people, without whom this thesis would not have been possible. First among them is my advisor Paul Seabright. I have often admired his ability to ask the right questions of my work, and steering it into more interesting directions. His high standards have always pushed me to do better, and his support has helped me overcome my frequent setbacks and doubts along the way.

The same is also true of many other professors at the Toulouse School of Economics. Among these I would like to mention Milo Bianchi, Matteo Bobba, Sylvain Chabé-Ferret, Alexandre de Cornière, François Poinas, and Karine Van der Straeten. They all provided helpful comments on my work, and the way I think about research was shaped significantly in workshops and seminars with their participation. Patrick Fève, who was the head of the doctoral school and for whom I was a teaching assistant for many years, has extended me a helpful hand at many times.

I owe thanks to many friends, especially my fellow doctoral students at the TSE. They are too many to enumerate here, but they have contributed to this thesis immeasurably, by discussing my work with me and by helping me through difficult times.

In addition to my years in Toulouse, I have visited Georgetown University and benefited greatly from discussions with researchers there. I am especially grateful to Anna Maria Mayda, and also to Luigi Minale at UC3M.

My family, and especially my parents, deserve my greatest gratitude. They have always provided me and my siblings with an environment that encouraged curiosity and creativity, and where my first supporters when this led to my decision to pursue a PhD.

Lastly, a special thank you must go to my fiancée Esther Muriithi. I owe much of my happiness and sanity during these last years to her.

Contents

Introduction - The Impact of Immigration and Digitalization on Labour Markets and Social Relations

This dissertation consists of three chapters studying global trends in the world of labour markets and social relations. Two of the chapters use the refugee immigration wave in Germany, 2014-2016, as a natural experiment. The first paper takes a local perspective, investigating the areas around refugee shelters in Berlin. These areas have experienced a decline in perceived neighbourhood quality, as represented by real estate prices and by ratings on an online review site. At the same time, there was increased support for right-wing parties. In the second, I show the effect of aggregate refugee immigration on labour market regions in Germany. Employment and GDP have benefited from the immigration wave, due to the additional demand generated by refugees. The third paper is a cautionary look at the digitalization of labour relations - my coauthor Milena Petrova and I use data from an online labour market to show how the reputation mechanism can fail to incentivize seller effort when the returns to reputation fluctuate.

The dissertation is mainly empirical in nature. I use a variety of different data sources - administrative and aggregated data, as well as rich datasets from online platforms. Common to all three chapters is the attempt to find novel natural experiments in observational data. In the chapter on local labour markets, I exploit local policy differences and infrastructure constraints to construct new instrumental variables, creating exogenous variation in local immigration. The other chapter on immigration is equally careful to establish the exogeneity of refugee shelter locations. And in the third chapter, we show that it is seasonal variation in the value of reputation, rather than for example seller characteristics and selection, that systematically induce undesirable outcomes.

The abstracts of the three papers are the following:

Refugee shelters, Neighbourhood Quality and Electoral Outcomes in Germany

Does refugee immigration affect the perceived quality of neighbourhood amenities, and is this an important factor for the opposition to immigration? In this paper, I use real estate listings and online reviews of businesses and public places to demonstrate neighbourhood change due to the establishment of a refugee shelter. The setting is Berlin during the European refugee crisis of 2015. Local authorities had to scramble to find suitable locations for refugee shelters; I show that these locations did not differ from control locations in terms of real estate prices, online reviews or political outcomes before immigration. When a shelter was established, rental prices and ratings for existing places declined in the immediate vicinity. Additionally, I show that the anti-immigrant German AfD party received a higher share of the vote near the refugee shelters. However, these effects are relatively small, and the measured decline in perceived neighbourhood quality explains at most a third of the effect of shelters on voting outcomes (an alternative mechanism could be the increased salience of the continent-wide crisis in the affected areas). The effect on rental prices and ratings is also very local, while the effect on voting outcomes is significant even at a greater distance.

The Demand Effect of Immigration: Refugees and Employment in Germany

The 2014-2015 refugee crisis led to a sharp and unexpected increase in immigration to Germany, particularly from Syria, Iraq, and Afghanistan. While as of 2017 few of the immigrants have found employment, refugees receive social security payments and access local services in the areas where they settle. The arrival of refugee immigrants in a community therefore represents an unexpected demand shock. I exploit the fact that settlement of refugees in Germany was managed to create novel instruments for the arrival of refugees in a given region. Firstly, German federal states used different rules for the internal allocation of refugees, and secondly, the availability of vacant housing was an important factor. I find that immigration increased native employment growth by about 6 percent annually between 2014 and 2017; these gains are just as high for non-university educated workers and occur especially in the service sectors most likely to benefit from the demand shock. Despite this, I find some evidence that refugee immigration has led to increased electoral support for the anti-immigration AfD party.

Reputational Incentives under Heterogeneous Demand Fluctuations

Using data from an online home services marketplace, we study reputational incentives in an environment with heterogeneous seller demand fluctuations. Our strategy exploits the fact that some professionals work in highly seasonal occupations, while others do not. We show that sellers facing an upcoming demand downturn become less likely to receive a positive review, and we argue that this is because they value their reputations less and exert lower effort. Robustness checks rule out several other explanations for this phenomenon, e.g. reverse causality, selection of sellers into jobs or differences in the cost of seller effort. Neither the economic literature nor the platforms relying on reputation systems have sufficiently considered seller heterogeneity and the optimal way to account for it in the information given to the buyer and in seller incentives.

Introduction - L'impact de l'immigration et de la numérisation sur les marchés du travail et les relations sociales

Cette thèse est composé de trois chapitres consacrés à l'étude des tendances mondiales dans les marchés du travail et les relations sociales. Les deux premiers chapitres se basent sur la vague d'immigration de réfugiés en Allemagne, 2014-2016, comme expérience naturelle. Le premier article adopte une perspective locale et examine les zones autour des abris de réfugiés à Berlin. Ces zones ont connu une baisse de la qualité perçue du quartier, représentée par les prix de l'immobilier et par les notations sur un site d'évaluation en ligne. A cette même période, le soutien aux partis de droite a augmenté. Dans le deuxième article, je montre l'effet de l'immigration de réfugiés sur les marché du travail en Allemagne. L'emploi et le PIB ont profité de la vague d'immigration, en raison de la demande supplémentaire générée par les réfugiés. Le troisième article, coécrit avec Milena Petrova, porte un regard d'avertissement sur la numérisation des relations de travail. Grâce aux données provenant d'un marché du travail en ligne, nous montrons que le mécanisme de la réputation n'incite pas toujours les vendeurs à fournir l'effort nécessaire, en particulier lorsque leurs rendements fluctue.

Cette thèse est principalement de nature empirique. J'utilise une variété de sources de données - des données administratives et agrégées, ainsi que de riches ensembles de données provenant de plates-formes en ligne. Chacun des trois chapitres permet de chercher des nouvelles expériences naturelles dans les données d'observation. Dans le chapitre sur les marchés du travail locaux, j'utilise les différences de politique locale et les contraintes d'infrastructure pour construire de nouvelles variables instrumentales, créant une variation exogène de l'immigration locale. Le premier chapitre sur l'immigration est tout aussi prudent pour établir l'exogénéité des lieux d'abris pour réfugiés. Dans le troisième chapitre, nous montrons que ce sont les variations saisonnières de la valeur de la réputation, plutôt que par exemple les caractéristiques et la sélection des vendeurs, qui induisent systématiquement des résultats indésirables.

Les résumés des trois articles sont les suivants:

Abris de réfugiés, qualité du quartier et résultats électoraux en Allemagne

L'immigration de réfugiés affecte-t-elle la qualité perçue des commodités du quartier et s'agit-il d'un facteur important pour l'opposition à l'immigration? Dans cet article, j'utilise des listes de biens immobiliers et des évaluations en ligne d'entreprises et de lieux publics pour mettre en évidence les changements de quartier dus à la création d'un abri pour réfugiés. Ce papier se focalise sur la situation de Berlin lors de la crise des réfugiés en Europe en 2015. Les autorités locales ont dû se démener pour trouver des emplacements convenables pour les abris de réfugiés; Je montre que ces emplacements ne diffèrent pas des emplacements de contrôle en termes de prix de l'immobilier, d'évaluations en ligne ou de résultats politiques avant l'immigration. Lors de la mise en place d'un abri, les prix de location et les évaluations des lieux existants ont baissé dans les environs immédiats. De plus, je montre que le parti politique allemand anti-immigrés AfD a obtenu une plus grande part du vote aux alentours des abris de réfugiés. Cependant, ces effets sont relativement faibles et le déclin mesuré de la qualité perçue du quartier explique au plus un tiers de l'effet des abris sur les résultats du vote (un mécanisme alternatif pourrait être la visibilité accrue de la crise à l'échelle du continent dans les zones touchées). L'effet sur les prix de location et les évaluations est également très local, tandis que l'effet sur les résultats du vote est significatif même à grande distance.

L’effet de l’immigration sur la demande: réfugiés et emploi en Allemagne

La crise des réfugiés de 2014-2015 a entraîné une augmentation forte et inattendue de l’immigration en Allemagne, en particulier en provenance de Syrie, d’Irak et d’Afghanistan. Bien que, en 2017, peu d’immigrés aient trouvé un emploi, les réfugiés reçoivent des paiements de sécurité sociale et ont accès aux services locaux dans les zones où ils s’installent. L’arrivée d’immigrants réfugiés dans une communauté représente donc un choc inattendu de la demande. L’installation de réfugiés en Allemagne a été contrôlée par les autorités, ce qui me permet de créer de nouveaux instruments statistiques pour l’arrivée de réfugiés dans une région donnée. Premièrement, chaque état fédéré allemand utilisait des règles différentes pour la distribution interne des réfugiés et, deuxièmement, la disponibilité de logements vacants était un facteur important. Dans cet article, je montre que l’immigration a accru la croissance de l’emploi des autochtones d’environ 6% par an entre 2014 et 2017; ces gains étant tout aussi importants pour les travailleurs non universitaires et se produisant en particulier dans les secteurs des services les plus susceptibles de bénéficier du choc de la demande. Malgré cela, cet article fournit des preuves que l’immigration de réfugiés a conduit à un soutien électoral accru au parti AfD anti-immigration.

Incitation de la réputation en cas de fluctuations hétérogène de la demande

En utilisant des données provenant d’un marché de services à domicile en ligne, nous étudions l’importance de la réputation dans un environnement caractérisé par des fluctuations hétérogènes de la demande des vendeurs. Notre stratégie exploite le fait que seule une partie des professionnels occupent des emplois très saisonniers. Nous montrons que les vendeurs confrontés à un ralentissement de la demande imminent ont moins de chances de recevoir un avis positif. Nous affirmons que cela est dû au fait qu’ils attachent moins d’importance à leur réputation et exercent un effort plus faible. Les contrôles de robustesse excluent plusieurs autres explications de ce phénomène, par exemple celle de la causalité inverse, de la sélection des vendeurs dans des emplois particuliers, ou de différences dans le coût associé aux efforts du vendeur. Ni la littérature économique ni les plates-formes reposant sur des systèmes de réputation n'ont suffisamment pris en compte le rôle de l'hétérogénéité des vendeurs ainsi que le moyen optimal de la transmettre dans les informations fournies aux acheteurs et vendeurs.

Chapter 1

Refugee shelters, Neighbourhood Quality and Electoral Outcomes in Germany

Does refugee immigration affect the perceived quality of neighbourhood amenities, and is this an important factor for the opposition to immigration? In this paper, I use real estate listings and online reviews of businesses and public places to demonstrate neighbourhood change due to the establishment of a refugee shelter. The setting is Berlin during the European refugee crisis of 2015. Local authorities had to scramble to find suitable locations for refugee shelters; I show that these locations did not differ from control locations in terms of real estate prices, online reviews or political outcomes before immigration. When a shelter was established, rental prices and ratings for existing places declined in the immediate vicinity. Additionally, I show that the anti-immigrant German AfD party received a higher share of the vote near the refugee shelters. However, these effects are relatively small, and the measured decline in perceived neighbourhood quality explains at most a third of the effect of shelters on voting outcomes (an alternative mechanism could be the increased salience of the continent-wide crisis in the affected areas). The effect on rental prices and ratings is also very local, while the effect on voting outcomes is significant even at a greater distance.

1.1 Introduction

Immigration has been one of the central issues in recent European and US electoral campaigns. The European refugee crisis of 2015 in particular has received strong attention, and in Germany, it is widely understood to be the main driver behind the rise of the right-wing AfD party.[1] What explains this opposition to immigration?

One possibility is that voters oppose immigration out of economic self-interest, for example if their job security or wages are threatened. This view is supported by a literature showing negative effects of low-skilled immigration on wages and employment of low-skilled natives (e.g. Card (2001), Borjas (2003), Borjas and Monras (2017)); at the same time, it is especially low-skilled individuals who oppose immigration (Scheve and Slaughter (2001), Mayda (2006), Facchini and Mayda (2009)). On the other hand, Facchini et al. (2013) show that low-skilled natives are more hostile to immigration even in a context where immigrants are high-skilled, and Rozo and Vargas (2018) show that only culturally more distant international refugees provoke a backlash, whereas internally displaced persons help the left-wing, more pro-immigration party.

Therefore, anxiety over cultural change, weakened social norms and declining quality of local amenities may be the more important channels. Using survey data, Card et al. (2012) show that such concerns for local amenities are more predictive of opposition to immigration than labor market concerns (see also Dustmann and Preston (2007), Hainmueller and Hopkins (2014) and Hainmueller et al. (2015)). There are however few studies on how the quality of amenities changes due to immigration or ethnic heterogeneity, and even fewer which investigate whether this is an important mechanism driving voting outcomes.

In this paper, I investigate these questions in the context of the refugee crisis of 2014-2016 in Berlin. I use real estate data from the largest German listings website, immobilienscout24.de, to show how rental prices are affected by the presence of a refugee shelter. Additionally, I use place-ratings from the website Foursquare to show changes in how these neighbourhoods are perceived. The majority of Foursquare places are businesses such as restaurants and shops, but public places such as parks and metro stations are also included. Rental prices decline by 3%, and a rating given to an existing place becomes

[1] A post-election survey of AfD voters revealed that 92% of respondents thought that "the party (AfD) mainly exists to change the refugee policies with its initiatives", while 97% said that they feared a loss of German culture due to immigration (Infratest/dimap (2017))

20% less likely to be positive when a refugee shelter opens nearby. These phenomena point to a perceived decline in the quality of the neighbourhood. There is some evidence for increased creation of new businesses near refugee shelters, especially Turkish and Middle Eastern restaurants, which is consistent with neighbourhood changes which cater to the new inhabitants. Additionally, I show that in the electoral districts closest to refugee shelters, the percentage of voters voting for the AfD party increased by 10%. This is true both for the 2016 Berlin Senate Elections and the 2017 Federal Elections.

I also discuss the relevance of the decline in rental prices and ratings as a channel for the voting outcomes. Using these as intermediate outcomes changes the quantitative importance of proximity to a refugee shelter only moderately. I demonstrate that this method is likely biased in the direction of overestimating the importance of the rental price and ratings channel. Furthermore, areas with many public venues, such as parks and squares, experienced an especially strong decline in ratings when a shelter was opened nearby. However, these areas did not see a larger increase in right-wing voting. Similarly, the effect of shelters on rental prices and ratings is limited to very nearby places, while their effect on voting has a much wider radius.

These findings underline that the effect of shelters on right-wing voting mainly works through channels other than the decline in neighbourhood quality reflected in rental prices and ratings. One such alternative channel could be that the Europe-wide crisis is more visible in areas near a refugee shelter, increasing the salience of this issue in the minds of voters.

My main results come from difference-in-differences specifications. I use a complete panel of shelter locations, capacities and dates of operation to define treatment variables. To address concerns that the locations of refugee shelters are endogenously determined, I show that these locations did not differ from non-treated areas in a large variety of characteristics. This includes levels and trends of real estate prices, political outcomes, ratings and the types of local businesses. In addition, the results are robust to the use of an instrument, namely the availability of infrastructure to house refugees (in public schools).

The previous literature has shown that immigration and social heterogeneity can have a negative impact on the quality of local amenities, by lowering the willingness to engage in the community, contribute to public goods and sanction antisocial behaviour (see e.g. Alesina et al. (1999), Alesina and La Ferrara (2000), Miguel and Gugerty (2005), Dahlberg et al. (2012), and Algan et al. (2016)). I argue that the consequences of such changes would affect not only the quality of housing, but also the restaurants, shops and public places present in the Foursquare database.

Housing outcomes can be affected in various ways by immigration. The demand from immigrants can drive up prices in larger geographic areas (see e.g. Saiz (2007), Ottaviano et al. (2012)); but there is also some evidence that more locally, prices can decline due to natives valuing the area less (Accetturo et al. (2014) and Sá (2014)). This can lead to the out-migration of natives and residential segregation, such as in the case of 'white flight' from US urban centers (see Boustan (2010), Boustan et al. (2010)). The negative price effects I show are unlikely to be partially offset by an opposing demand effect, given that they are very locally constrained, and can therefore be interpreted more easily as a signal of a decrease in the subjective quality of local amenities.

With respect to political outcomes, the larger literature has often found that immigration, particularly of individuals with low skills or strong cultural differences, increases the electoral success of parties opposed to immigration (Barone et al. (2016), Halla et al. (2017), Harmon (2017)). This effect may be limited to low-skilled voters (Mayda et al. (2018)). Dustmann et al. (2016) and Gerdes and Wadensjö (2008) show an increase in right-wing votes in Denmark caused by refugee allocation, while Dinas et al. (2016) show the same for the Greek islands which house refugees. On the other hand, Steinmayr (2016) finds a negative impact of refugees on right-wing votes in Austria, noting that direct exposure can lead to decreased prejudice (the contact hypothesis).

Typically, the unit of observation in these papers are larger areas, e.g. counties or municipalities, rather than neighbourhoods around conspicuous immigrant housing, but Otto and Steinhardt (2014) find the same effect for neighbourhoods in Hamburg.

The wider economic consequences of refugee immigration have also received special attention in recent years. Akgündüz et al. (2015), Del Carpio and Wagner (2015), Borjas and Monras (2017), and Hennig (2018) focus on labour market changes, while Alix-Garcia et al. (2018) and Altindag et al. (2018) show positive effects on prices and production.

To this literature I add an investigation of a new and related outcome, namely how businesses and other amenities in the neighbourhood change due to the establishment of a refugee shelter. While the ratings are indeed negatively affected by the establishment of a shelter, I conclude that the contemporaneous increase in support for right-wing parties is not mainly a consequence of the neighbourhood decline, meaning that other mechanisms musr also be at work.

The paper is organized as follows: section 2 describes the data and the historical context. Section 3 provides estimation results both on electoral outcomes and on the establishment and perceived quality of venues. Section 4 explores whether the decline in real estate prices and ratings is a channel for the electoral impact. Section 5 discusses the robustness of different specifications, including the IV specification, while section 6 concludes.

1.2 Data and historical background

1.2.1 The refugee crisis in Berlin

During the 2014-2016 European Refugee crisis, more than 70 Thousand initial applications for asylum were made in Berlin. This represents roughly 2% of Berlin's population of 3.6 Million, the highest per-capita figure of any German Federal State except for the small city state of Bremen[2]. Because of its dense development and increasingly tight housing market, this meant that Berlin is the place where the refugee inflow was most acutely experienced by the local population.

The Berlin Office for Refugee Affairs (Landesamt für Flüchtlingsangelegenheiten, LAF) provides information on all shelters in operation, including their exact location, capacity and occupancy. I have obtained this list at two different points in time, once in September 2016 and once in January 2018. There is considerable overlap between the two lists, but some shelters have closed while others have opened between the dates.

Additionally, some shelters – especially among those located in gymnasiums of public schools – had already been closed in the summer of 2016. From these lists, I construct a panel of 177 shelters, which I believe to be mostly complete. 96 of those shelters where still in operation in January of 2018. I have also collected their opening and (where applicable) closing dates, to be able to pinpoint the exact month when a neighbourhood would be treated.

According to the LAF, resources where so strained during the refugee crisis that shelters were opened where it was possible, without consideration of political or social consequences. This was communicated to the media and also confirmed to me via email. I largely validate this claim in section 5 (Robustness and validity), showing that the eventual

[2]BAMF, Das Bundesamt in Zahlen 2014-2016, Asyl

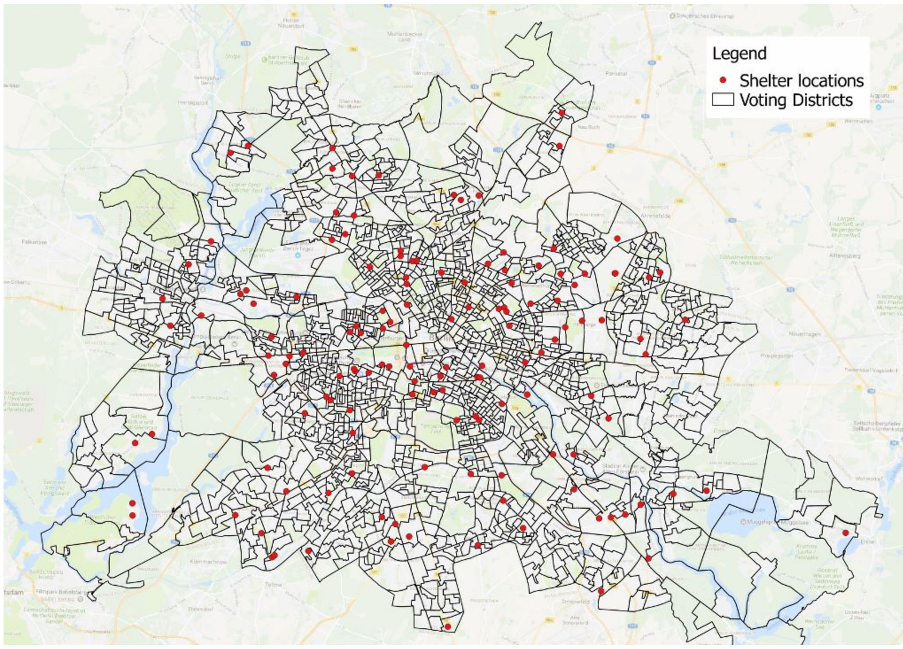


Figure 1.1: Location of refuge shelters in Berlin

shelter locations do not differ from others either in terms of previous results for right-wing parties or in the composition, type and quality of local places.

The majority of shelters were established in pre-existing buildings. Frequently, these were gymnasiums of public schools (48 instances in my list), and I could confirm this for at least 89 other shelters, for example in unused administrative buildings. At least 19 shelters were temporary structures.

The fact that the premises of public schools where often used as locations for a shelter enables me to use the proximity of such a school as an instrument for the eventual proximity of a shelter (see section 4, robustness and validity). However, while the results are robust to the use of this instrument, the locations of schools actually show less balanced characteristics before the crisis than the true locations of shelters, which is why I prefer to simply use these true locations as treated areas in a differences-in-differences setting.

As can be seen in Figure 1.1, the shelters are distributed across Berlin; however, the more densely developed centre of the city has more shelters. The population living in the centre consequently is more likely to live close to a refugee shelter (Fig. 1.2).

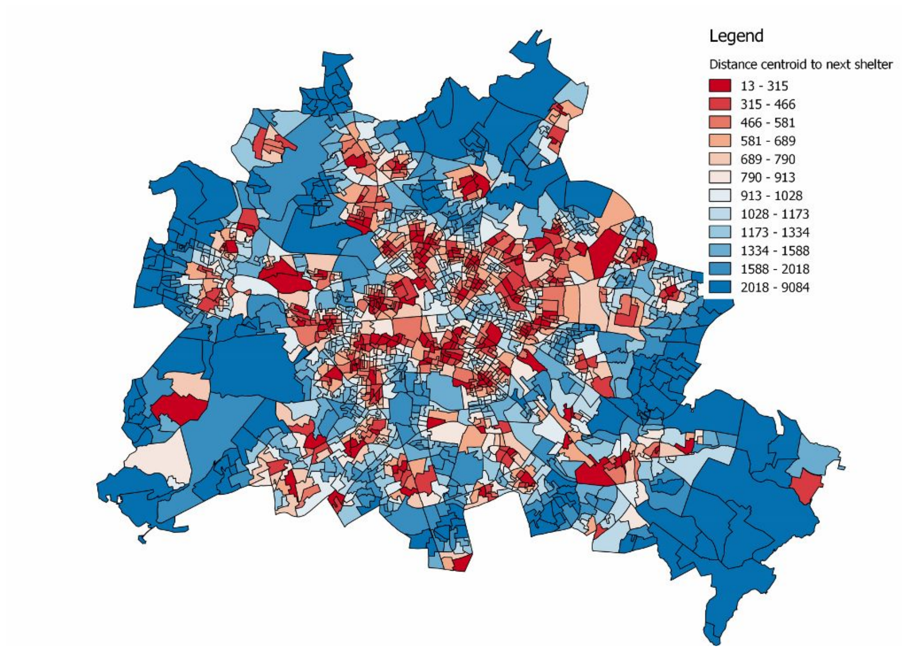


Figure 1.2: Distance of voting tracts to nearest shelter

1.2.2 Independent variables

This distance to a shelter will be my main treatment, either as a continuous variable or as a dummy indicating that a shelter is close. In my preferred specifications, I simply create a treatment dummy T_{it}^{d}, indicating whether there is an open shelter within d meters of the observation. For real estate listings and place ratings, I use the exact address to calculate the distance; in the case of voting precincts, I use the centroid of their area.

During 2016, 1% (371 of 40,080 observations) of real estate listings are within 100m of a shelter. 4.5% (1,832) are within 200m, and 25% (9,942) within 500m. Of the voting precincts, 1% (18 of 1,470) are within 100m, 4% (61) are within 200m, and 23.5% (346) within 500m.

Shelters vary considerably in size. Their capacity ranges from 30 to 1200 places, with a mean of 300.[3] Additionally, some locations are close to several shelters, and distance is a continuous variable, with close proximity presumably having a higher impact.

I therefore create a treatment variable that takes both distance and capacity of all nearby shelters into account. It is defined as the sum of the capacities (in hundred beds) of the

[3]The largest shelter, at the former Tempelhof airport, temporarily had an even larger capacity of more than 4000 places.

three closest shelters, each divided by the square root of the shelter's distance to a place, or

$$T_{it} = \sum_{j=1}^{5} \frac{c_{jt}}{\sqrt{d_{ij}}}. \tag{1.1}$$

Shelters $j = 1, 2, 3$ have capacities of c_j, and are located at a distance of d_{ij} from place i (or the centroid of voting precinct i). [4]

The mean treatment value (when observations are the voting precincts) is .25 during the year of 2016, with a standard deviation of .14. 5% of voting districts have a treatment value of .55 or higher.

I discuss in section 5 how treated and untreated precincts did not differ from each other in trends or levels before the crisis (see e.g. the balance table 1.18).

1.2.3 Elections

My main unit of observation for election outcomes is the voting precinct, the districts served by one polling station. This is the basic unit at which votes are counted. There are 1779 such precincts in Berlin, serving on average 1343 eligible voters (see Table 1.1). The state election supervisor for Berlin publishes party vote totals and percentages at this level for every election held in Berlin, be they senatorial (state), federal or European elections.

The AfD party received 14.2% of the vote in the Berlin senatorial elections of 2016, and 12% of the Berlin vote in the 2017 federal elections. During previous elections, it received a much lower percentage of the vote - 4.9% in the Federal Elections of 2013 and 7.9% in the European elections of 2014 - and before 2013, it did not exist. It was also founded with a very different platform from the one it would adopt after the refugee crisis, namely one with a much larger focus on the European debt crisis and the Euro, rather than immigration.

This must be kept in mind when we use difference-in-difference specifications to estimate our effects of interest: earlier votes for the AfD do not necessarily capture the full potential of anti-immigrant votes before the refugee crisis.

[4] I only use the 3 closest shelters because the minimum distance of a place to the fourth-closest shelter is greater than 500m, a distance at which the impact of a shelter on a neighbourhood (beyond the impact of 3 closer shelters) is small.

Table 1.1: Voting tracts in Berlin, Federal Elections 2017

N: 1779 (in 2017)	Mean	Median	St.D.
Eligible voters	1,343	1285	386
Votes	698	681	195
Voting for AfD (%)	13.5	12.8	6.5
Area (km2)	.5	.2	1.2
Distance from centre (km)	8.7	8.1	4.5
Distance to nearest refugee shelter (km)	1.4	1.2	.98

As can be seen in the map Figure 1.3, the outer voting precincts, and particularly those in the east, voted more strongly for the AfD party than in the centre of the city. This is the reason I control for distance from the centre as well as district, interacted with the treatment dummy for the years 2016-2017, in my main specifications.

The median area of voting tracts is 190t m2, (mean: 500t m2), so that a representative district, if it were square, would have all points within roughly 250m of its centroid. Therefore, I define such precincts where the distance from the shelter to the centroid is 250m or less as treatment districts.

1.2.4 Real estate data

The real estate listings website immobilienscout24.de is the largest such service in Germany. Real estate agencies as well as private landlords use it to make their offers conveniently searchable for prospective buyers and tenants. Beyond currently available listings and accompanying detailed exposés, the website includes an "atlas" where clients can retrieve information about past listings in the vicinity of a given address, to form an impression of price developments in the area.

I have scraped this information from the atlas, since in contrast to the current listings, it includes listings going back to 2010. The available variables are the exact address, price, size, and number of rooms as well as the real estate agency handling the listing.

I was able to webscrape and geolocate roughly 250,000 individual ratings. The majority of these (about 200,000) are for apartment rentals, with the rest being house rentals or purchase listings. Berlin has experienced a real estate boom during the period under

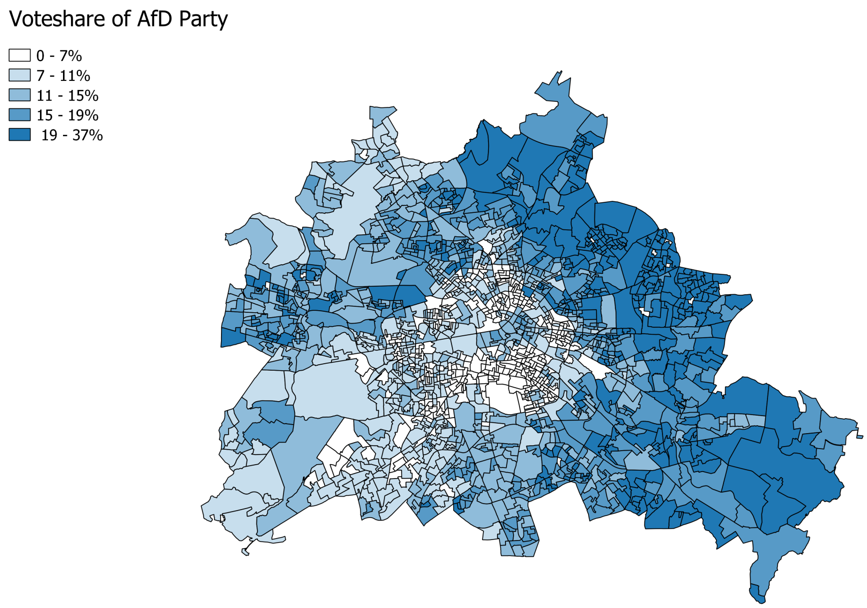


Figure 1.3: Success of AfD party in Berlin, 2017 federal elections

study, which is reflected in my data by a 66% increase in the mean rental price from 6.76 Euro per square metre in 2010 to 11.23 Euro in 2018.

1.2.5 Ratings and Places data

I use data from the local search website Foursquare to construct additional variables. Foursquare makes its data available through an API, free of charge for non-commercial applications. It includes a large database of "venues" - geocoded places such as restaurants and other businesses, but also e.g. parks, streets and bus stops. Users can create these venues and give "tips" on them, which include a rating ("like", "dislike" and "meh") and a text review. For greater clarity, I refer to venues and tips as places and ratings.

The creation of places as well as the ratings come with a time stamp, making it possible to create a panel of many characteristics of areas over time - for example the number and type of businesses as well as average ratings of these businesses.

For Berlin, there are roughly 66 thousand unique places in the database, which have together received 81 thousand ratings (16.5 thousand places have received at least one rating). For each place, one or more categories are given - "Office" and "Café" are the

Number of Venues	68,902
- at least 1 rating	16,252
- at least 5 ratings	4,815
- Office	3200
- Café	1967
- Residential	1785
- Bakery	1743
- Bus Stop	1425
...	...
- Italian	1023
- Doner	380
- Turkish	264
- Middle Eastern	175

Table 1.2: Venues on Foursquare

Number of Tips	80,166
- 'liked'	.49
- 'disliked'	.04
- 'meh'	.06
- none	.4
- English	.56
- German	.36
- Russian	.03
- Turkish	.02

Table 1.3: Tips on Foursquare

most common, but there are also numerous places in categories related to immigration from the refugee origin countries - “Doner”, “Falafel” and “Middle Eastern”, for example.

The first set of variables I constructed from the Foursquare data are counts of newly created places by month and voting precint, as well their as counts in specific categories. I also construct the mean rating and mean price category of these places. In the absence of small-scale census data (the census is only broken down to the 12 districts of Berlin), these variables allow us to assess whether or not refugee shelter locations can truly be seen as similar to locations without shelters.

The aggregate variables on the level of the voting tract can also be used as outcome variables. They give us an indication of how the composition of neighbourhoods changes, e.g. if there is increased creation of new businesses of certain types, and if so if these differ in price and perceived quality from those created not in the proximity of a shelter.

It has to be kept in mind that this is not an official and complete business register. If business creation on Foursquare increases or declines in a certain area, this could be due to the fact that Foursquare users and developers give this area greater or lower attention.

Secondly, I consider if ratings given to places in a certain area are more or less likely to be positive or negative, what language is used in the text etc. This I can do for ratings of all places or only of places in specific categories. Note that there are almost as many places as there are ratings; the majority of places never receives any rating. My estimates on the probability of receiving a positive rating, which will include place fixed effects, are estimates off the minority of places which have received multiple ratings.

In addition to creating these outcome variables, I also use the Foursquare data to create covariates for the analysis of electoral outcomes. For example, the presence of a refugee shelter can have a different impact in areas where there is already a strong immigrant presence. Official population data for Berlin is only available at the coarse level of the district; using this finely grained data enables me to e.g. observe restaurants with immigrant cuisine and use it as an indicator for immigrant presence. I also use the average price ratings on foursquare as an indicator of the wealth of an area.

In Appendix 1.C I show correlations of these Foursquare measures of wealth and foreign population with the official data available on the level of the 12 district.

1.3 Results

1.3.1 Real estate prices and listings

Prices: I first study the impact of refugee shelters on real estate prices. The observations are individual rental listings (apartments and houses)[5]. I regress the price per square metre on treatment and controls as follows:

$$y_{ijt} = \beta T_{it} + \gamma X_{it} + FE_t + FE_j + \epsilon_{ijt} \tag{1.2}$$

where i is the listing and j the voting tract. Since I have precise information on the operation dates of shelters, I can use a monthly panel where the treatment variable T_{tj} varies from month to month, when new shelters are opened. T_{ti} is defined first, for clarity of interpretation, as a dummy taking the value of 1 after a shelter opens within 100m of a place i.

The geographic fixed effects are on the level of the voting precinct, and time fixed effects are for individual months. I also include linear time trends for the voting precincts

[5]purchase listings are similarly impacted, but they are much fewer.

as a robustness check, but since they do not change the results but slow down calculations considerably, I omit them in my main specifications.

The controls included at basement are only size, a dummy for if the listing is for a house (rather than an apartment - the large majority), and the number of rooms (as a categorical variable).

The results show a decline by about 37 cents for those rentals very close (within 100m) of an operating refugee shelter (see table 1.4). That is about 3% of the average rent.

Table 1.4: Nearby refugee shelter and rental prices - treatment dummy

Outcome variable	Rental price (€ per m^2)		
Shelter within 100m	-0.360**	-0.372**	-0.375**
	(0.152)	(0.150)	(0.146)
FE voting precinct	✓	✓	✓
FE month × year (time)	✓	✓	✓
Distance from center × time		✓	✓
Linear trends by voting precinct			✓
R squared	0.201	0.829	0.833
N. of observations	208,390	208,390	208,390

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***. Standard errors clustered at place and time level.

Number of listings and repeated listings: If a refugee shelter increases the number of new listings in a neighborhood, it would be an indication that previous residents are moving away to avoid the immigrants. To study this, it is necessary to aggregate the number of listings by a geographical unit. I use the voting precincts for any such aggregation in this paper.

I regress the number of immoscout24.de listings on a treatment variable as follows:

$$N_{it} = \beta(close_i \times post_t) + \Gamma X_i \times post_t + FE_t + FE_i + \epsilon_{it} \tag{1.3}$$

N_{it} is the number of listings in voting precinct i in year t. The treatment variable is the interaction of a dummy $close_i$ (the closest shelter location is within 200m of the voting precinct centroid) with the post-crisis dummy $post_t$ (any year after 2015). Geographic controls X_i are interacted with the post-crisis dummy to avoid picking up any mechanical effect, by which denser precincts closer to the centre are more likely to be treated and also have different real estate dynamics beyond the impact of shelters.

The results can be seen in table 1.5. We note that the number of listings increases by roughly 2 per year after a shelter opened nearby (specification (1). The average number of listings per year is 16).

One limitation of my data is that I do not have the date at which a listings was withdrawn, so that I cannot observe how long an apartment stays on the market. However, some apartments are quickly re-listed on the platform - real estate agents could do this if they haven't found a new tenant and want to update the exposee e.g. with pictures to give it a more prominent position on the website. Apartments that remain unoccupied for a longer time are more likely to be re-listed in this way.

For the purposes of this study, I define listings for an apartment that has been listed less than 6 months before as re-listings. Since these are long-term rentals (rather than sublets or holiday rentals), I assume that these apartments have remained unoccupied in between the original listing and the re-listing.

When I subtract such re-listings from the number of listings per voting precincts, I obtain the number of new listings. If I regress these on our treatment variable, we see that it increased considerably less than the total number of listings (table 1.5, (2)). The difference is (mechanically) made up by re-listings (3). The effect on new listings is 1.1 relative to an average of 13, while the effect on re-listings is 1.1 relative to an average of 2.5 per year per voting precinct.

Taken together, this indicates that there are indeed slightly more new listings on the website after a shelter opened nearby. However, the effect on re-listings, and therefore on the likelyhood that an apartment remains vacant for a longer time, is more important in relative terms.

Table 1.5: Nearby refugee shelter and number of real estate listing

Outcome variable	Number of listings		
	(1) Total	(2) New listings	(3) Re-listings
Shelter within 100m	2.204***	1.088**	1.105***
	(0.444)	(0.322)	(0.271)
Distance from center × time	✓	✓	✓
FE voting precinct	✓	✓	✓
FE month × year (time)	✓	✓	✓
R squared	0.568	0.629	0.265
N. of observations	9,880	9,880	9,880

Significant at: p < 0.1: *; p < 0.05: **; p < 0.01: ***. Standard errors clustered at place and time level.

1.3.2 Public amenities and their Ratings

Next I investigate the impact of shelters on place ratings. Observations are individual 'tips' - ratings with a review/description text - of which there are 81,122 in Berlin during the years 2010 to 2017. I use a linear probability model taking the form

$$y_{ijt} = \beta T_{tj} + FE_t + FE_j + \epsilon_{ijt} \tag{1.4}$$

where j is an individual venue, e.g. a business. Using fixed effects on the level of the venue, the coefficient of interest β is only identified from those venues that exist before and after the establishment of refugee shelters. Since I have precise information on the operation dates of shelters, I can use a monthly panel where the treatment variable T_{tj} varies from month to month, when new shelters are opened. T_{tj} is defined first, for clarity of interpretation, as a dummy taking the value of 1 after a shelter opens within 200m of a place j. My preferred specification is however the one defined as above in (1.1), taking into account the distance and capacity of all nearby shelters.

I also use voting tract fixed effects rather than venue FE in some specifications, to allow for an effect on the establishment of new venues around refugee shelters, which could have systematically different ratings and reviews.

The outcome variables y_{ijt} are dummies, e.g. for whether or not a rating was positive or whether it was in German. For some regressions, I restrict the sample to certain places of special interest to us.

Table 1.6: Nearby refugee shelter and ratings of existing places, all places - treatment dummy

Outcome variable	(1) 'liked'	(2) 'disliked'	(3) 'meh'
Shelter within 200m	-0.12***	0.01	-0.02
	(0.044)	(0.018)	(0.028)
FE place	✓	✓	✓
FE month × year	✓	✓	✓
R squared	0.31	0.08	0.07
N. of observations	80,166	80,166	80,166

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***. Standard errors clustered at place and time level.

The outcomes can be seen in tables 1.6 for the treatment dummy, and 1.7 for the continuous treatment variable. The results indicate roughly the same effects - as can be seen in table 1.7, ratings for a place are less likely to be positive after refugee shelters have opened in the vicinity of the venue. The likelihood of the rating being explicitly negative is unchanged, and an ambivalent rating ('meh') becomes more likely. The decline of positive ratings by 1.4 percentage points is not large when we consider that about 46% of tips on the platform are positive. The treatment variable is roughly 1 on average, and the 90th percentile is around 1.9. A 'highly treated' place would therefore be around 3 percentage points less likely to receive a high rating (we will see later that places very close to a shelter – within 100m – see a much larger impact). Use of the treatment dummy variable suggest somewhat smaller impacts (table 1.6).

I will use my continuous treatment variable in the remaining regressions in this subsection, since it gives a more complete picture of the intensity of refugee housing in an

Table 1.7: Nearby refugee shelter and ratings of existing places, all places - continuous treatment

Outcome variable	(1) 'liked'	(2) 'disliked'	(3) 'meh'
Shelters capacity/distance	-0.014**	0.001	0.008*
	(0.007)	(0.004)	(0.005)
FE place	✓	✓	✓
FE month × year	✓	✓	✓
R squared	0.31	0.08	0.07
N. of observations	80166	80166	80166

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***. Standard errors clustered at place and time level.

area. When the results of regressions with the simpler treatment dummy differ from those presented here, I will note it in the discussion (those results are available on request).

The main coefficient does not change greatly when we restrict the sample to only those venues which have received 10 or more ratings or when we include the geographic controls introduced in the section on election outcomes (Table 1.8, specifications (1) and (2) respectively).

This decline on ratings could in principle only affect certain types of places. If we restrict our sample to public places such as parks and roads, or to Turkish and Middle Eastern restaurants, we see that the former are impacted at a similar magnitude as all places, while the latter are potentially less impacted (the coefficients are less precisely measured due to the smaller sample). If only German language ratings are considered, we again find no difference to the overall coefficient (these regressions on subsamples are reported in Table 1.9).

Newly established venues and their composition: We are also interested in the question of whether areas around refugee shelters experience changes in business activity. New businesses could open to cater to refugee demand, while at the same time, the area could become less attractive for other businesses. I study the establishment of new places on Foursquare in the vicinity of shelters, keeping in mind that this can only be an in-

Table 1.8: Nearby refugee shelter and ratings of existing places, all places

Model	(1) Restricted	(2) Geo. controls
	'liked'	
Shelters capacity/distance	-0.017**	-0.013*
	(0.008)	(0.007)
Distance to centre× (2016 or 2017)		0.00
		(0.00)
FE venue	✓	✓
FE Year	✓	✓
R squared adjusted	0.29	0.31
N	51,614	80,166

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***. Standard errors clustered at district level in (3).

Table 1.9: Nearby refugee shelter and ratings of existing places, subsamples

Subsample	(1) All places	(2) Public	(3) MEastern	(4) German review
Outcome variable		'liked'		
Shelters capacity/distance	-0.014**	-0.012	-0.002	-0.015
	(0.007)	(0.016)	(0.031)	(0.014)
FE voting tract	✓	✓	✓	✓
FE Year	✓	✓	✓	✓
R squared	0.31	0.35	0.27	0.35
N. of observations	80,166	13,113	3,169	26,585

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***. Standard errors clustered at district level in (3).

dication of true business activity - not all businesses are recorded, and the activity of Foursquare users could itself be subject to changes in the neighbourhood.

It appears that there is a slight increase in the number of newly established venues driven by nearby shelters (table 1.10, (1)), as well as in the establishment of new Middle Eastern and Turkish restaurants (2). On average, 1.56 new places are created in a voting precinct each month, and .008 Middle Eastern and Turkish restaurants. So while the coefficient of .06 on the creation of all new places is not large relative to the base rate, the creation rate of the relevant ethnic restaurants of .002 is relatively high (and this is the value for the average treatment, not the areas at the 90th percentile of our treatment variables).

It should be noted however that these effects on the creation of new businesses are only weakly significant, and are insignificant when we use the 'dummy' treatment variable.

The average rating of newly established businesses is not affected (3), and neither is the average price level (4).

It would be interesting to investigate this further, since in the longer run we do expect a shelter to affect the categories of businesses created more than this limited evidence suggests.

Table 1.10: Nearby refugee shelter and newly established places

Outcome variable	(1) all venues	(2) Turkish & MEastern	(3) rating	(4) price
Shelters capacity/distance	0.063*	0.002*	-0.039	0.021
	(0.032)	(0.001)	(0.065)	(0.014)
FE voting tract	✓	✓	✓	✓
FE Year	✓	✓	✓	✓
R squared	0.814	0.077	0.195	0.124
N. of observations	11,161	11,161	3,067	5,867

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***. Standard errors clustered at district level in (3).

1.3.3 Elections

For ease of interpretation, I will again first show the results using the simple treatment dummy, which is 1 the electoral precinct has any shelter within 200m from its centroid. In following The preferred model I use is a simple diff-in-diff model of the form

$$right\_voteshare_{it} = \beta(close_i \times post_t) + \Gamma X_i \times post_t + FE_t + FE_i + FE_d \times post_t + \epsilon_{it} \tag{1.5}$$

where i is the voting precinct and t denotes specific elections - -the Berlin senatorial elections of 2006, 2011 and 2016, as well as the federal elections of 2009, 2013 and 2017. $close_i$ is a dummy taking the value of 1 if the centre of the polling district is within 250 metres of a refugee shelter, and the dummy $post_t$ is one for the elections of 2016 and 2017. FE_i is the fixed effect on the polling station, while FE_t is the election fixed effect. I also use geographic variables X_i interacted with the post-dummy; these include the precinct's area and distance to the centre of Berlin. I prefer to include these due to the geographic concentration of treated areas near the centre of the city (which also have a smaller area).[6] As discussed before, each electoral district has around 1500 voters.

The results table are presented in Table 1.11). Specification (1) does not include any geographic controls other than the fixed effects. My preferred specifications (2) and (3) include distance to the centre of Berlin as the only such control. They differ in so far as (3) estimates only within-district effects by including an additional FE for each district after the refugee crisis; these additional controls would be important if e.g. there are differences in the impact between Western and Eastern districts. However, they do not change our main coefficient.

These specifications indicate that the voteshare of right-wing parties increased by 1.2 additional percentage points in those voting tracts where a refugee shelter is nearby. This is about 10% of the median AfD voteshare of 12.8%.

This magnitude is confirmed a model using the natural logarithm rather than the level of the outcome variable (model 4). The shelters increase the voteshare by about 10%.

[6]The previous level of right-wing support does not fully capture the potential for increase in radical votes in 2016 and 2017. The highest increase happened in the poorer and less densely populated outer areas of Berlin. Since many of the shelters are in the more densely populated inner residential areas, and since the outer polling tracts are larger in area, the outer polling stations also happen to be further away from shelters on average (see Figures 1 and 2 for visual evidence of this geographic correlation).

Table 1.11: Nearby refugee shelter and right-wing vote share, treatment dummy

Model	(1)	(2)	(3)	(4)
Outcome		right-wing voteshare		ln(rw voteshare)
Independent variables				
Shelter within 200m	0.006**	0.012***	0.012***	0.102***
	(0.002)	(0.001)	(0.001)	(0.019)
Distance to centre× (2016 or 2017)		0.006***	0.005***	0.016
		(0.001)	(0.001)	(0.009)
FE voting precinct	✓	✓	✓	✓
FE election	✓	✓	✓	✓
FE district × (2016 or 2017)			✓	
R squared (adjusted)	0.795	0.827	0.869	0.857
N. of observations	10,958	10,958	10,958	10,857

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***. Standard errors clustered at district and year levels.

Interestingly, the geographic variables are less significant and have smaller importance relative to the treatment coefficient compared with specifications (2) and (3).[7]

The results when using the continuous variable indicate a stronger effect of the presence of shelters on voting outcomes (table 1.12), of 5 percentage points (model 1) or 3 points (model 2, where differences between districts are controlled for). We will see later that the effect of shelters on voters is present at a larger distance than 200m; it is possible that the continuous treatment variable captures e.g. the effect of some large shelters within a distance of 200-1000m of a voting precinct, something that the simple dummy does not capture.

We will see later (in section 4) that these result are in their sign robust to a number of different specifications, including the use of proximity of a public school as an instrument. I will however argue that this OLS specification is preferable; most importantly, the 'treated' areas do not have different political outcomes to untreated areas before the refugee crisis.

1.3.4 Heterogeneity of electoral impact

We expect the impact on electoral outcomes to be different from one voting tract to the other, depending on local characteristics. For example, we might think that having a shelter nearby has a stronger impact on right-wing voting in voting tracts of low density, simply because the shelter and refugees would be more visible in such an environment. I define density as eligible voters per 100m^2, so that it has a mean of 0.79.

A voting tract where there is already a large foreign (and especially Middle Eastern) presence might also be differently impacted, although we can imagine arguments for an effect in either direction. Since census data on the level of voting tracts is not available, I use the Foursquare data to find a proxy for this variable, namely the number of Turkish and Middle Eastern restaurants. On average, there are .55 such establishments in a voting tract, but more than half of them do not have such a restaurant at all.

Lastly, voters in more wealthy areas could also react differently. As a (certainly imperfect) proxy for prosperity, I take the average price category of Foursquare places in the

[7] I will discuss this in more detail in section 4 (Robustness), but the distance from the centre likely picks up an effect that is not accounted for by the additive fixed effects model; the logarithmic specification, in contrast, possibly accounts for this effect by assuming that the location and time fixed effects interact in a multiplicative fashion on the voteshare.

Table 1.12: Nearby refugee shelter and right-wing vote share, continuous treatment variable

Model	(1)	(2)
Outcome	right-wing voteshare	
Independent variables		
Shelters capacity/distance	0.051***	0.030***
	(0.007)	(0.006)
Distance to centre× (2016 or 2017)	0.006***	0.005***
	(0.001)	(0.001)
FE voting precinct	✓	✓
FE election	✓	✓
FE district × (2016 or 2017)		✓
R squared (adjusted)	0.829	0.869
N. of observations	10,958	10,958

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***. Standard errors clustered at district and year levels.

voting tract.

I interact these three variables - density, presence of Middle Eastern restaurants, and average price category of places - with our continuous treatment variable accounting for nearby shelters. The results from this regression can be seen in table 1.13. Density has indeed the expected effect (column 2) - where a new shelter would be more visible (lower density), it has a much stronger impact on right-wing voting. Voting tracts with higher presence of Middle Eastern or Turkish restaurants actually saw a lower impact, probably due to the fact that voters in these tracts are more used to, and more sympathetic towards, foreign immigration. Lastly, the price level as given by our proxy variable also has a negative interaction with the treatment variable in affecting rigt-wing voting (this last interaction is not significant when the treatment dummy is used).

These heterogeneous effects appear to be quite sizeable - recall that density as defined here is roughly .8 on average, and the number of ME restaurants .55. The coefficients on the interactions are therefore large relative to the direct effect. This has important implications for where to locate refugee shelters, if policy makers aim to mitigate the impact of shelters on an electoral backlash.

1.3.5 Type of shelter

The LAF distinguishes two main types of shelters, the Gemeinschaftsunterkunft (community shelter, GU) and the Notunterkunft (emergency shelter, NU). Presumably, refugees will be visible in the vicinity of both GU and NU, but the NU are more clearly a reminder of a crisis situation, and may be seen as a sign of its mismanagement. If these types of shelters have no different impact on our outcomes, it may be seen as a sign that voters simply object to the presence of refugees, while a stronger impact of NU would indicate that at least part of the political backlash is due to the perceived chaotic circumstances rather than simply immigration alone.

I investigate if the impact of a NU shelter is different from other shelters (mostly GU) by interacting the dummy for a nearby shelter with another dummy that indicates a NU. The negative impact on ratings is much higher for NU, which is what we would expect. However, the impact on electoral outcomes may even be smaller (the coefficient is not significant). This is unexpected, but perhaps voters in the vicinity of emergency shelters were reassured that they would only be temporary.

Table 1.13: Nearby refugee shelter and local characteristics, electoral outcomes

Outcome variable	right-wing voteshare			
	(1)	(2)	(3)	(4)
Shelters capacity/distance	0.051***	0.083***	0.064***	0.085***
	(0.007)	(0.012)	(0.009)	(0.010)
Shelters capacity/distance × Density		-0.039***		
		(0.007)		
Shelters capacity/distance × ME rest.			-0.030***	
			(0.004)	
Shelters capacity/distance × price level				-0.028***
				(0.006)
Distance to centre× (2016 or 2017)	0.006***	0.006***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)
FE voting tract	✓	✓	✓	✓
FE Year	✓	✓	✓	✓
R squared	0.829	0.832	0.834	0.830
N. of observations	10,958	10,958	10,958	10,151

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***. Standard errors clustered at district level in (3).

Table 1.14: The impact of emergency shelters (NU) vs. community shelters

Outcome	(1) Right-wing voteshare	(2) Negative ratings
Independent variables		
Shelter within 200m	0.015***	-0.09**
	(0.005)	(0.04)
Shelter within 200m × shelter is NU	-0.019	-0.04**
	(0.012)	(0.02)
FE year	✓	✓
FE district	✓	
FE venue		✓
R squared	0.83	0.18
N. of observations	13,283	160097

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***. Standard errors clustered at district level in (3).

1.4 Is the change in perceived neighbourhood quality a mechanism for voting outcomes?

The previous section established the impact of refugee shelters on real estate prices, the ratings and types of places in the neighbourhood and on contemporaneous electoral outcomes. However, this does not tell us how much the decline in perceived neighbourhood quality, as documented in the real estate and foursquare data, contributed to the electoral backlash against refugee immigration. In fact, I argue in this section that the changes documented in the ratings data are only a moderately important channel for the voting impact.

Mediation analysis: The first strategy is to use the intermediate outcome as a right-hand side variable in a regression of final outcomes on the treatment status and covariates. If this lessens the coefficient of treatment, the difference can, under strong assumptions, be interpreted as the part of the total effect explained by the changes in the intermediate outcome. Formally, the procedure is to estimate the system of equations

$$Y_{it} = \beta^t T_{it} + \Gamma X_{it} + \epsilon_{1_{it}} \tag{1.6}$$

$$M_{it} = \beta^m T_{it} + \Gamma X_{it} + \epsilon_{2_{it}} \tag{1.7}$$

$$Y_{it} = \beta_1^d T_{it} + \beta_1^{id} M_{it} + \Gamma_1 X_{it} + \epsilon_{3_{it}} \tag{1.8}$$

our outcome of interest Y_{it} is $right\_voteshare_{it}$, while the intermediate outcome (or mediator) M_{it} is either the average rental price $rprice_{it}$ in voting precinct i, or the proportion of ratings which are positive $liked_{it}$. The covariates X_{it} include fixed effects as in 1.5, as well as the appropriate geographic controls. β^t would then be interpreted as the total effect of shelters on voting, and β^d as the direct effect (that is, the effect not accounted for by our observed channel). The product $\beta^{id} \times \beta^m$ would be the indirect effect, mediated by the intermediate outcome $liked_{it}$.

This procedure, known as mediation analysis in statistics (see Heckman and Pinto (2015), Imai et al. (2010)), is prone to bias.[8] It requires that

[8] Despite these problems, mediation analysis has been applied in a number of papers recently, particularly in the economics of education (see e.g. Heckman et al. (2013), Oreopoulos et al. (2017)), but also in the more closely related literature on trade, labour markets and voting outcomes (Dippel et al. (2017))

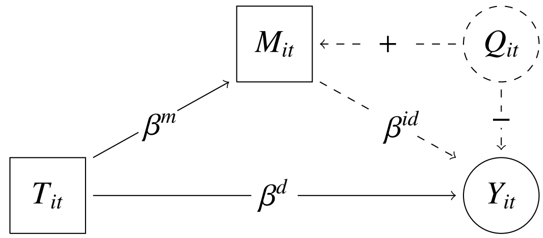


Figure 1.4: Shelters treatment T_{it} affecting voting outcomes Y_{it} directly, and through ratings M_{it} as intermediate outcome; other, unobserved shocks to quality Q_{it}.

$$\begin{aligned} \{Y_{it}, M_{it}\} \perp\!\!\!\perp T_{it}|X_{it} = x \quad \text{and} \\ Y_{it} \perp\!\!\!\perp M_{it}|T_{it} = t, X_{it} = x \end{aligned} \tag{1.9}$$

The first assumption simply states that there are no confounding unobserved variables affecting both treatment status and either the intermediate or final outcome variable. It is the same assumption needed for our estimation of total treatment effects above. The second assumption requires that there are no unobserved variables affecting both the outcome Y_{it} and the intermediate variable M_{it}, for the coefficients of 1.8 to be identified.

Despite my inclusion of fixed effects and other controls, this condition is unlikely to be met in our application. In particular, we can expect that there are unobserved shocks to the quality of a neighbourhood, affecting simultaneously real estate prices, the ratings of foursquare venues and the success of right-wing parties.

However, it may be possible to make additional assumptions on the direction of the influence of this unobserved shock on Y_{it} and M_{it}. If we assume that the unrelated and unobserved decline of neighbourhood quality would lower the ratings of businesses, and at the same time increase the electoral success of right-wing parties, we can show in which direction the coefficients β^d and β^{id} in 1.8 are likely biased. A regression of right-wing electoral success Y_{it} on real estate prices and foursquare ratings M_{it},

$$Y_{it} = \beta M_{it} + \Gamma X_{it} + FE_i + FE_t + \epsilon_{it}$$

finds a negative coefficient β, suggesting that this is indeed the direction of correlation (see table 1.15).

Table 1.15: Correlation of neighbourhood quality measures and voting outcomes

Outcome	right-wing voteshare	
Mean rental price	-0.046	
	(0.025)	
Average rating		-0.016
		.009
FE year	✓	✓
FE precinct	✓	✓
R squared	0.873	0.838
N	9,878	7,509

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***. Standard errors clustered at precinct and year level.

If we assume that equation 1.8 is primarily misspecified by the omission of the unobserved shock to neighbourhood quality Q_{it}, the true relationship between treatment, ratings and voting outcomes would be

$$Y_{it} = \beta_2^d T_{it} + \beta_2^{id} M_{it} + \beta^q Q_{it} + \Gamma_2 X_{it} + \epsilon_{3_{it}} \tag{1.10}$$

while the true specification for determining real estate prices and ratings is

$$M_{it} = \beta^m T_{it} + \beta^{qm} Q_{it} + \Gamma X_{it} + \epsilon_{2_{it}}. \tag{1.11}$$

The omitted variable bias resulting from the omission of Q_{it} in our specification 1.8 relates the estimated coefficients β_1^{id} and β_1^d to the true β_1^{id} and β_1^d; it is

$$\beta_1^{id} = \beta_2^{id} + \beta^q \frac{Cov(M,Q) \times Var(T)}{Var(M) \times Var(T) - Cov(M,T)^2} \tag{1.12}$$

and

$$\beta_1^d = \beta_2^d - \beta^q \frac{Cov(M,Q) \times Cov(T,M)}{Var(M) \times Var(T) - Cov(M,T)^2} \tag{1.13}$$

Derivations of these terms will be provided in the Appendix 1.A. If we assume that $Cov(M,Q)$ is positive – a positive shock to unobserved neighbourhood quality increases real estate prices and ratings – and that β^q is negative – the same shock reduces right-wing voting – this tells us that β_1^{id} is biased downward (away from zero). So is β_1^d, since $Cov(T,M)$, the effect of shelters on average ratings, is also negative.

Estimating specifications 1.6 and 1.8, we find that the treatment effect β_1^d of a nearby shelter on voting outcomes is virtually unchanged when we include the average ratings in the voting precinct as a regressor. The effect of including real estate prices is more sizeable (see table 1.16). The effects of the both real estate prices and average ratings, β_1^{id}, have the expected negative sign, but it is small and insignificant.

The decrease of the treatment coefficient from 1.2 percentage points to .8 percentage points when the mean rental price is included suggests that potentially, a third of the effect of the establishment of the shelter on voting outcomes is due to the mechanism of neighbourhood decline. But as we have seen, β_1^d and β_1^{id} are likely to be biased downward in these regressions, so that this may overstate the importance of this channel. It provides an upper bound, meaning that most of the treatment effect is due to unmeasured mechanisms.

In the following, I present additional pieces of evidence, strengthening this conclusion.

Distance effect: Treatment effects diminish with distance. A shelter within 100m of the centre of a polling area, or within 100m of a venue, should have a larger impact than one within 500m. To investigate this, I use a different treatment definition, one that only takes the distance of the nearest shelter into account. I define four brackets, according to the proximity of the next shelter. In bracket 1, observations are within 100m of the next shelter, and in brackets 2, 3, and 4 the nearest shelter is within 100m-200m, 200m-500m and 500m-1000m respectively. This defines four different treatment groups. During 2016, 39% of voting precincts, 49% of listings, and 47% of ratings now fall into one of the treatment groups.

The results of a regression where these treatment dummies are interacted with post-crisis dummies (as before) are presented in table 1.17. We see the expected declining

Table 1.16: The impact of shelters on voting, RE prices and ratings as intermediate variable

Outcome	right-wing voteshare		
Shelters T_{it}	0.012***	0.008***	0.011***
	(0.001)	(0.001)	(0.001)
Mean rental price		-.0007	
		(0.0004)	
Average rating			-0.015
			(0.009)
FE year	✓	✓	✓
FE precinct	✓	✓	✓
R squared	0.836	0.872	0.838
N	7,509	7,509	7,509

Significant at: p < 0.1: *; p < 0.05: **; p < 0.01: ***. Standard errors clustered at precinct and year level.

impact for voting outcomes, real estate prices, and negative tips. It is notable however, that the impact on the measures of neighbourhood quality falls much faster in distance than the impact on voting outcomes. The stark difference tells us that voters within 500m of the nearest shelter are still more likely to cast their vote for the right-wing party, while rental prices in their block do not decline and businesses in their own close vicinity do not experience a decline in ratings.

Table 1.17: The impact of shelters, decreasing with distance

Outcome	(1) RW voteshare	(2) rental price	(3) Positive rating
Independent variables			
Shelter within 100m	0.012***	-0.37**	- 0.12**
	(0.002)	(0.15)	(0.05)
Shelter within 100m-200m	0.013***	-0.05	-0.017
	(0.004)	(0.11)	(0.022)
Shelter within 200m-500m	0.009***	0.03	0.023
	(0.003)	(0.07)	(0.025)
Shelter within 500m-1000m	0.005**	0.01	-0.023
	(0.002)	(0.06)	(0.018)
FE year	✓	✓	✓
FE district	✓	✓	
FE venue			✓
R squared	0.83	0.83	0.18
N. of observations	13,283	208,390	80,166

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***. Standard errors clustered at district level in (3).

This set of results may not be conclusive, but it suggests that the impact of shelters on right-wing voting works through a channel other than the ratings decline discussed above. With the data at hand, we can not investigate all these possible channels, but one candidate is that a nearby shelter simply increases the visibility and salience of the refugee crisis to voters, leading them to vote for the anti-immigrant party even if there is no perceived decline in the quality of their neighbourhood.

1.5 Robustness and validity

1.5.1 Common trends and balance between treated and non-treated areas

An important concern is whether the placement of refugee shelters was influenced by political considerations - they could for example be located in areas where support for right-wing parties is weaker. Table 1.18 provides an overview of the differences between areas close to the eventual locations of refugee shelters (treated) and those further away. In terms of previous political outcomes, there was no difference between treatment and control groups. I test this not only for our outcome variable, the right-wing voteshare, but also for the share of the Green party (the most pro-immigration large party in Germany) and for the share of nonvoters (since many of the AfD voters came from this group).

Treated areas also did not have different average ratings before, and the share of Turkish or Middle Eastern restaurants relative to all places was also the same. On the other hand, there were more Foursquare places in the treated areas, probably because they tend to be closer to the centre of the city.[9]

This balance between treated and non-treated areas could however hide diverging trends. We could find that ratings in areas where shelters would eventually open were higher at the very beginning of our panel, and falling up to the time when shelters were opened. In that case, the treatment effect could simply be the continuation of that trend.

To address this concern, I will demonstrate that my cross-sectional treatment variables, capturing eventual proximity to refugee shelters, do not determine either ratings or voting outcomes at any time prior to the refugee crisis. The regression specification is

$$y_{i,t} = \alpha_t \left(treatment_i \times \mathbf{TD}_t\right) + \Gamma X_i + FE_i + FE_t + \epsilon_{i,t} \tag{1.14}$$

where the outcome variable $y_{i,t}$ is either the probability of a positive rating or the right-wing voteshare. $\mathbf{TD}_t$is a period dummy (quarterly for ratings, yearly for election out-

[9] Similar comparisons for other variables are available upon request.

Table 1.18: Balance table, treated and untreated voting precincts

	Non-treated		Treated		T-test
Variable	N	Mean/SE	N	Mean/SE	Difference
Rightwing vote share	7675	0.038	343	0.035	0.003
		(0.002)		(0.003)	
Greens vote share	7675	0.177	343	0.193	-0.016
		(0.001)		(0.007)	
Nonvoters	7675	0.414	343	0.407	0.006
		(0.001)		(0.006)	
Mean rental price (apartments)	6745	7.729	299	7.638	0.092
		(0.032)		(0.093)	
- price development (2011-14)	6496	1.574	294	1.603	-0.029
		(0.065)		(0.070)	
Average rating	7094	1.286	325	1.272	0.013
		(0.003)		(0.012)	
Ethnic restaurants	7675	0.014	343	0.017	-0.003
		(0.000)		(0.001)	

Notes: The covariates area and distance to center center are included in all estimation regressions. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

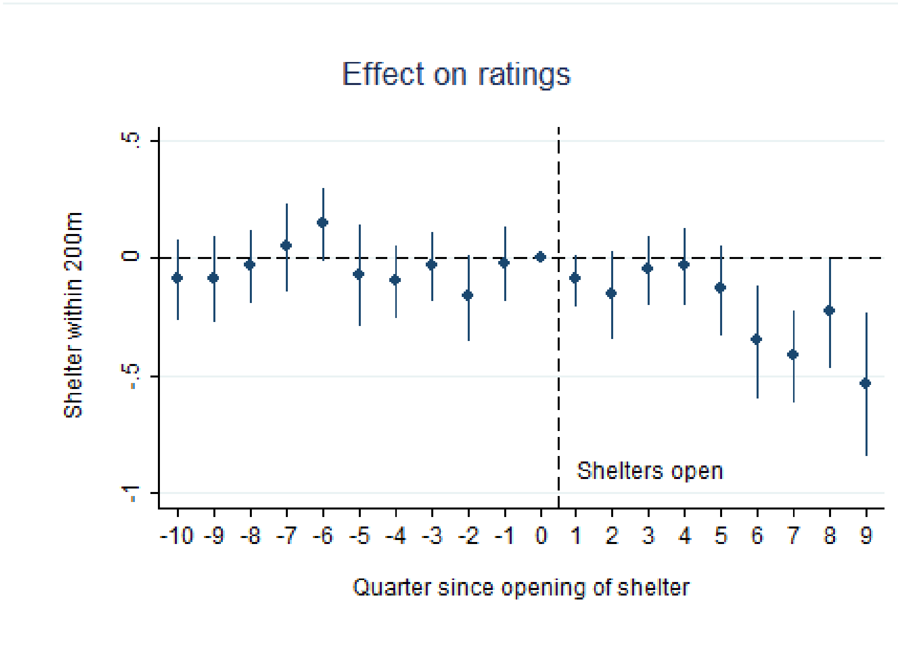


Figure 1.5: Ratings and proximity to shelter location, before and after opening

comes), and the treatment dummy is simply indicating the eventual opening of a shelter within 200m.

Figure 1.5 shows the coefficients from this regression (for positive ratings). As can be seen, the locations of shelters did not follow a different trend before the refugee crisis. It also becomes apparent that the effect of the shelter only appeared a year after opening.

Figure 1.6 shows the same coefficients for all federal and state elections since 2006. It becomes evident that there was no rising support for right-wing parties in these locations before the crisis (the regressions results and coefficient plots for other political outcomes, as well as for other ratings outcomes and treatment definitions, are available on request).

1.5.2 Places and ratings

Shelters were not opened in such voting tracts where tips where generally better or worse before the refugee crisis (table 1.19). I have also investigated the composition of venues in tracts which would eventually host a shelter along several dimensions; on most measures, these are not different from others. This is notably true for the number of Turkish or Middle Eastern restaurants in the area relative to the total number of venues (reported in

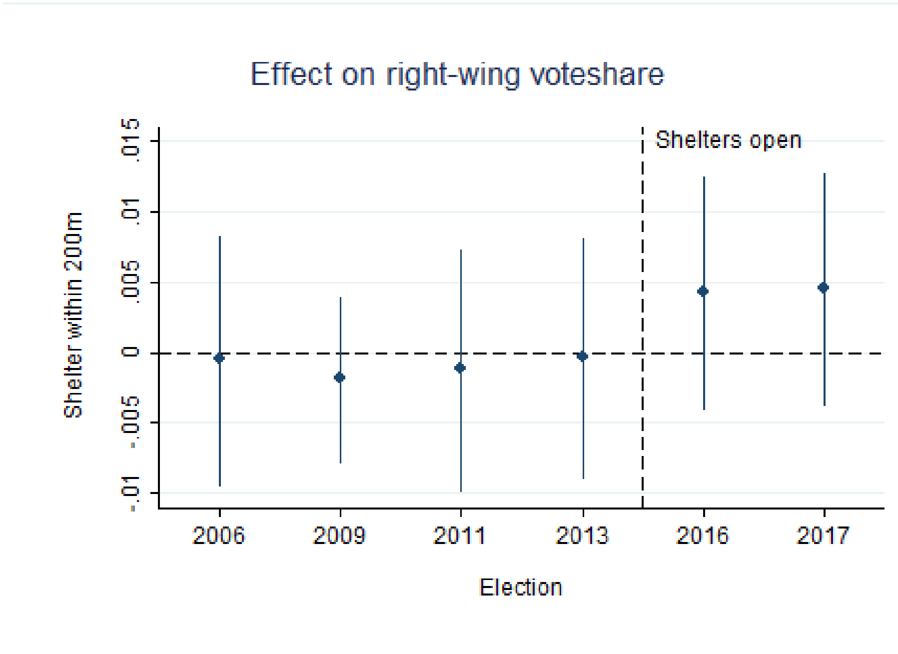


Figure 1.6: Voting outcomes and proximity to shelter location, before and after opening

table 1.20). It appears therefore not true that the foreign population of a voting tract was a consideration when the locations of shelters were decided.

The average ratings of places, their average price and other characteristics are also not predictive of whether or not a shelter would be opened nearby (tables upon request). An exception is the number of total venues in a voting tract: this is significantly higher in such voting tracts which would be near a shelter, even when we control, as always, for the area and distance to the centre. This points to the fact that a general abundance of facilities meant a place for shelter could more likely be found. Conversely, it was not the case that shelters were established mostly 'out of view', in areas of low density (again, regression tables are available upon request). A look at the maps in section 2 also confirms this.

Instrumental variable estimation: Another strategy to address the potential endogeneity of shelter locations would be to instrument them with the infrastructure commonly used for them. Since school grounds, school sports facilities etc. were often used in Berlin, the distance to the next public school is one such candidate. I investigate this instrument in Appendix 1.B.

My main results are quite robust to the use of this instrument, depending on treatment definitions. However, as I show, the instrumented treatment locations are actually less

Table 1.19: Nearby refugee shelter and tips in previous years

Outcome variable				negative rating			
Elections	(1) 2011	(2) 2012	(3) 2013	(4) 2014	(5) 2015	(6) 2016	(6) 2016
Nearby shelter	-0.057	-0.195	-0.054	0.955*	1.968	2.428**	2.849
	(0.062)	(0.589)	(0.116)	(0.526)	(1.528)	(0.951)	(1.885)
Distance to centre	-0.002	-0.060	-0.030***	-0.189***	-0.445***	-0.360***	-0.533***
	(0.005)	(0.046)	(0.009)	(0.040)	(0.122)	(0.073)	(0.144)
Area	0.002	0.050	0.035**	0.183***	0.518**	0.458***	0.576**
	(0.008)	(0.078)	(0.015)	(0.069)	(0.211)	(0.120)	(0.243)
FE district	✓	✓	✓	✓	✓	✓	✓
N. of observations 871	999	1206	1074	984	985	926	

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***. Standard errors clustered at district level in (3).

Table 1.20: Nearby refugee shelter and Turkish and Middle Eastern restaurants in previous years

Outcome variable				negative rating			
Elections	(1) 2011	(2) 2012	(3) 2013	(4) 2014	(5) 2015	(6) 2016	(6) 2016
Nearby shelter	-0.005	-0.010	0.002	-0.008	0.004	-0.005	-0.015
	(0.009)	(0.010)	(0.008)	(0.010)	(0.015)	(0.013)	(0.017)
Distance to centre	-0.001*	-0.001**	-0.001**	-0.001	-0.000	-0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Area	0.001	-0.002	-0.001	-0.001	-0.002	-0.000	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
FE district	✓	✓	✓	✓	✓	✓	✓
N. of observations 1546	1623	1609	1420	1221	1301	1113	

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***. Standard errors clustered at district level in (3).

balanced with non-treatment locations then the true shelter locations. For this reason, I prefer to use the OLS estimations.

1.5.3 Identity of users

An objection to my use of Foursquare ratings as an outcome variable is that it could be that a refugee shelter simply attracts a different set of Foursquare users to an area. These users could be of a type that gives worse ratings, even though their experience is no less positive than that of users who frequented the area before.

Ideally, I would use user ID fixed effects interacted with place ID fixed effects, to only capture users giving repeated ratings to the same place. This is however impossible: there are 27 thousand different user IDs for 81 thousand ratings, in addition to the 16.3 place IDs in the ratings data. Including user fixed effects completely absorbs the remaining variation, making any treatment effect insignificant.

I can therefore only investigate if available user characteristics are impacted by the establishment of shelters. These characteristics are user gender and the language of the review text. The rationale for this is that if a different group of users were giving the reviews after the establishment of a shelter, they would likely differ in language and gender. The immigrants themselves for example would be more likely to be male and non-german.

Regressions of dummies for reviews in different languages - German, English, Turkish, and Arabic - on our treatment and control variables from the main specification show that this is not the case (table 1.21). As before, the 10% most treated areas have a value of 1.8 during 2017. Recall also that, in the whole dataset, the percentages of reviews in different languages are: 56% English, 36% German, 2% Turkish, and .1% Arabic.

There is also no effect on user gender (table 1.22). At baseline, 60% of reviews are by male users, and 30% by female users (the rest are by users who do not identify their gender). These percentages are not significantly affected by the treatment variable.

I conclude from this evidence that the users giving ratings are not discernibly different after the establishment of a shelter.

Table 1.21: Distance to refugee shelter and review language

Outcome variable	(1) German	(2) English	(3) Turkish	(4) Arabic
Independent variables				
Shelters capacity/distance	-0.001	-0.007	-0.002	0.001
	(0.009)	(0.008)	(0.002)	(0.001)
FE place	✓	✓	✓	✓
FE month	✓	✓	✓	✓
R squared	0.185	0.152	0.093	0.055
N. of observations	80,166	80,166	80,166	80,166

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***. Standard errors clustered at place and month level.

Table 1.22: Distance to refugee shelter and review language

Outcome variable	(1) Male	(2) Female
Independent variables		
Shelters capacity/distance	-0.004	0.010
	(0.009)	(0.008)
FE place	✓	✓
FE month	✓	✓
R squared	0.077	0.066
N. of observations	80,166	80,166

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***. Standard errors clustered at place and month level.

1.6 Conclusions

This study suggests that a nearby refugee shelter decreases the perceived quality of local amenities such as restaurants. At the same time, the anti immigration AfD party has benefited from the presence of shelters in elections after the refugee crisis. It appears however that these are two independent outcomes, so that another mechanisms than the perceived decline of neighbourhood quality is responsible for the effect on votes.

There is limited evidence for increased creation of businesses in the vicinity of refugee shelters, and of Turkish and Middle Eastern restaurants in particular. This may go hand in hand with increased unobserved business closures; it would be interesting to return to this question in the future, and perhaps with different data.

It must be kept in mind that the Foursquare data represents subjective opinions about local venues, and that these opinions come from a selected group of individuals. However, taken together with e.g. the survey evidence from Greek islands in Hangartner et al. (2017), these results indicate that emergency housing of refugees is seen as a negative for the community.

The response to refugees may be different in less crisis-laden circumstances. The fact that some of the negative impacts found in this study are more severe for the more makeshift emergency shelters is an indication of this. This is of interest for the future management of refugee housing, as is the finding that more visible shelters, and shelters in areas without a previous presence of co-ethnic inhabitants, appear to provoke a stronger electoral backlash.

Bibliography

Accetturo, A., Manaresi, F., Mocetti, S., and Olivieri, E. (2014). Don't stand so close to me: the urban impact of immigration. *Regional Science and Urban Economics*, 45:45–56.

Akgündüz, Y., Van den Berg, M., and Hassink, W. (2015). The impact of refugee crises on host labor markets: the case of the syrian refugee crisis in turkey.

Alesina, A., Baqir, R., and Easterly, W. (1999). Public goods and ethnic divisions. *The Quarterly Journal of Economics*, 114(4):1243–1284.

Alesina, A. and La Ferrara, E. (2000). Participation in heterogeneous communities. *The quarterly journal of economics*, 115(3):847–904.

Algan, Y., Hémet, C., and Laitin, D. D. (2016). The social effects of ethnic diversity at the local level: A natural experiment with exogenous residential allocation. *Journal of Political Economy*, 124(3):696–733.

Alix-Garcia, J., Walker, S., Bartlett, A., Onder, H., and Sanghi, A. (2018). Do refugee camps help or hurt hosts? the case of kakuma, kenya. *Journal of Development Economics*, 130:66–83.

Altindag, O., Bakis, O., and Rozo, S. (2018). Blessing or burden? the impact of refugees on businesses and the informal economy.

Barone, G., D'Ignazio, A., de Blasio, G., and Naticchioni, P. (2016). Mr. rossi, mr. hu and politics. the role of immigration in shaping natives' voting behavior. *Journal of Public Economics*, 136:1–13.

Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *The quarterly journal of economics*, 118(4):1335–1374.

Borjas, G. J. and Monras, J. (2017). The labour market consequences of refugee supply shocks. *Economic Policy*, 32(91):361–413.

Boustan, L. P. (2010). Was postwar suburbanization "white flight"? evidence from the black migration. *The Quarterly Journal of Economics*, 125(1):417–443.

Boustan, L. P., Fishback, P. V., and Kantor, S. (2010). The effect of internal migration on local labor markets: American cities during the great depression. *Journal of Labor Economics*, 28(4):719–746.

Card, D. (2001). Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics*, 19(1):22–64.

Card, D., Dustmann, C., and Preston, I. (2012). Immigration, wages, and compositional amenities. *Journal of the European Economic Association*, 10(1):78–119.

Dahlberg, M., Edmark, K., and Lundqvist, H. (2012). Ethnic diversity and preferences for redistribution. *Journal of Political Economy*, 120(1):41–76.

Del Carpio, X. V. and Wagner, M. (2015). *The impact of Syrians refugees on the Turkish labor market*. The World Bank.

Dinas, E., Matakos, K., Xefteris, D., and Hangartner, D. (2016). Waking up the golden dawn: Does exposure to the refugee crisis increase support for extreme-right parties? Technical report, Working paper. https://sites. google. com/site/kostasmatakos/research.

Dippel, C., Gold, R., Heblich, S., and Pinto, R. (2017). Instrumental variables and causal mechanisms: Unpacking the effect of trade on workers and voters. Technical report, National Bureau of Economic Research.

Dustmann, C. and Preston, I. P. (2007). Racial and economic factors in attitudes to immigration. *The BE Journal of Economic Analysis & Policy*, 7(1).

Dustmann, C., Vasiljeva, K., and Damm, A. P. (2016). Refugee migration and electoral outcomes. *CReAM DP*, 19:16.

Facchini, G. and Mayda, A. M. (2009). Does the welfare state affect individual attitudes toward immigrants? evidence across countries. *The review of economics and statistics*, 91(2):295–314.

Facchini, G., Mayda, A. M., and Mendola, M. (2013). What drives individual attitudes towards immigration in s outh a frica? *Review of International Economics*, 21(2):326–341.

Gerdes, C. and Wadensjö, E. (2008). The impact of immigration on election outcomes in danish municipalities. Technical report, IZA Discussion Papers.

Hainmueller, J., Hiscox, M. J., and Margalit, Y. (2015). Do concerns about labor market competition shape attitudes toward immigration? new evidence. *Journal of International Economics*, 97(1):193–207.

Hainmueller, J. and Hopkins, D. J. (2014). Public attitudes toward immigration. *Annual Review of Political Science*, 17.

Halla, M., Wagner, A. F., and Zweimüller, J. (2017). Immigration and voting for the far right. *Journal of the European Economic Association*, 15(6):1341–1385.

Hangartner, D., Dinas, E., Marbach, M., Matakos, K., and Xefteris, D. (2017). Does exposure to the refugee crisis make natives more hostile?

Harmon, N. A. (2017). Immigration, ethnic diversity, and political outcomes: Evidence from denmark. *The Scandinavian Journal of Economics*.

Heckman, J., Pinto, R., and Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*, 103(6):2052–86.

Heckman, J. J. and Pinto, R. (2015). Econometric mediation analyses: Identifying the sources of treatment effects from experimentally estimated production technologies with unmeasured and mismeasured inputs. *Econometric reviews*, 34(1-2):6–31.

Hennig, J. (2018). The demand effect of immigration: Refugees and employment in germany. Link.

Imai, K., Keele, L., and Yamamoto, T. (2010). Identification, inference and sensitivity analysis for causal mediation effects. *Statistical science*, pages 51–71.

Infratest/dimap (2017). Survey of AfD voters for the ARD (in German), 24 Sept. 2017. Link. [Online; accessed 11/3/2018].

Mayda, A. M. (2006). Who is against immigration? a cross-country investigation of individual attitudes toward immigrants. *The review of Economics and Statistics*, 88(3):510–530.

Mayda, A. M., Peri, G., and Steingress, W. (2018). The political impact of immigration: Evidence from the united states. Technical report, National Bureau of Economic Research.

Miguel, E. and Gugerty, M. K. (2005). Ethnic diversity, social sanctions, and public goods in kenya. *Journal of public Economics*, 89(11-12):2325–2368.

Oreopoulos, P., Brown, R. S., and Lavecchia, A. M. (2017). Pathways to education: An integrated approach to helping at-risk high school students. *Journal of Political Economy*, 125(4):947–984.

Ottaviano, G. I., Peri, G., et al. (2012). The effects of immigration on us wages and rents: A general equilibrium approach. *Migration impact assessment: new horizons. Cheltenham: Edward Elgar Publishing Limited*, pages 107–46.

Otto, A. H. and Steinhardt, M. F. (2014). Immigration and election outcomes – evidence from city districts in hamburg. *Regional Science and Urban Economics*, 45:67–79.

Rozo, S. and Vargas, J. F. (2018). Brothers or invaders? how crisis-driven migrants shape voting behavior. Available at SSRN. [Version: October 17, 2018, accessed November 3, 2018].

Sá, F. (2014). Immigration and house prices in the uk. *The Economic Journal*, 125(587):1393–1424.

Saiz, A. (2007). Immigration and housing rents in american cities. *Journal of urban Economics*, 61(2):345–371.

Scheve, K. F. and Slaughter, M. J. (2001). Labor market competition and individual preferences over immigration policy. *Review of Economics and Statistics*, 83(1):133–145.

Steinmayr, A. (2016). Exposure to refugees and voting for the far-right:(unexpected) results from austria. *Browser Download This Paper*.

Appendix

1.A Bias in indirect effect regression

Assume the true relationship between voting outcomes Y_{it}, proximity of shelters T_{it} and ratings M_{it} is

$$Y_{it} = \beta_2^d T_{it} + \beta_2^{id} M_{it} + \beta^q Q_{it} + \epsilon_{3_{it}}$$

Since we do not observe the shock to neighbourhood quality Q_{it}, we estimate

$$Y_{it} = \beta_1^d T_{it} + \beta_1^{id} M_{it} + \epsilon_{3_{it}}$$

Let **X** be the matrix of independent variables [T,M]. The estimated coefficients $\hat{\beta}$ are:

$$\hat{\beta} = \begin{bmatrix} \hat{\beta_1^d} \\ \hat{\beta_1^{id}} \end{bmatrix} = \frac{\mathbf{X'}Y}{\mathbf{X'X}}$$

which is

$$\hat{\beta_1^d} = \frac{Var(M) \times Cov(T,Y) - Cov(T,M) \times Cov(M,Y)}{Var(T) \times Var(M) - Cov(T,M)^2}$$

and

$$\hat{\beta_1^{id}} = \frac{Var(T) \times Cov(M,Y) - Cov(T,M) \times Cov(T,Y)}{Var(T) \times Var(M) - Cov(T,M)^2}$$

Developing these expressions by substituting the true specification for Y we have for $\hat{\beta_1^{id}}$:

$$\begin{aligned}
\hat{\beta_1^{id}} = &\quad \frac{Var(T) \times Cov(M, \beta_2^d T + \beta_2^{id} M + \beta^q Q + \epsilon_3) - Cov(T,M) \times Cov(T, \beta_2^d T + \beta_2^{id} M + \beta^q Q + \epsilon_3)}{Var(T) \times Var(M) - Cov(T,M)^2} \\
= &\quad \frac{\beta_2^{id}\Big(Var(T) \times Var(M) - Cov(T,M)^2\Big) + \beta_2^d\Big(Var(T) \times Cov(T,M) - Var(T) \times Cov(T,M)\Big)}{Var(T) \times Var(M) - Cov(T,M)^2} \\
&\quad + \frac{\beta^q\Big(Var(T) \times Cov(M,Q) - Cov(T,M) \times Cov(T,Q)\Big) + V(T) \times C(M,\epsilon_3) - C(T,M) \times C(T,\epsilon_3)}{Var(T) \times Var(M) - Cov(T,M)^2} \\
= &\quad \beta_2^{id} + \beta^q \frac{Var(T) \times Cov(M,Q) - Cov(T,M) \times Cov(T,Q)}{Var(T) \times Var(M) - Cov(T,M)^2}
\end{aligned}$$

Since $Cov(T,Q) = 0$ (the unobserved shocks are independent of treatment), this simplifies to

$$\hat{\beta_1^{id}} = \beta_2^{id} + \beta^q \frac{Var(T) \times Cov(M,Q)}{Var(T) \times Var(M) - Cov(T,M)^2}.$$

By the same logic,

$$\beta_1^d = \beta_2^d - \beta^q \frac{Cov(M,Q) \times Cov(T,M)}{Var(M) \times Var(T) - Cov(M,T)^2}.$$

1.B Instrumental variable estimation:

An appealing strategy to address the possible endogeneity of shelter location is to instrument the location of shelters with the infrastructure commonly used for them. Since school grounds, school sports facilities etc. were often used in Berlin, the distance to the next public school is highly predictive of the distance to the next shelter during the refugee crisis. Via this distance channel, it also predicts the treatment variable (shelters capacity/distance - see table 1.B.1). Equally, having a public school nearby increases the likelyhood of having a nearby shelter - this second specification is a much worse fit however, since there are many more public school locations in Berlin than shelter locations. For the IV regressions, I therefore prefer the continuous treatment variable, rather than the nearby shelter dummy.

Table 1.B.1: Public schools and refugee shelters, first stage

Outcome	(1) Shelter Capacity/Distance	(2) Nearby shelter dummy
Independent variables		
Distance to school	-.038***	
	(0.009)	
Nearby school dummy		0.014*
		(0.008)
Distance to centre	0.0	-0.003
	(0.006)	(0.002)
FE district	✓	✓
FE Year	✓	✓
R squared	0.79	0.015
N. of observations	18,947	18,947

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***. Standard errors clustered at district level in (3).

When I instrument the shelter capacity/distance by the distance to the closest school in the panel model (equation 1.5), the coefficient is close to the one obtained in the OLS model (table 1.B.2). The IV performs much better when continuous distance to shelter is used than when we use a dummy, because the standard error is quite large. However, the direction of the coefficient is the same. This gives us additional confidence that shelter locations where not chosen according to the potential for political discontent.

The results on venue ratings are also robust to this instrumental variable strategy (tables upon request).

However, public school location turn out to be more predictive of political outcomes before the refugee crisis than the actual shelter location (table 1.B.3). The IV specification thus fails a placebo test which the OLS specification largely passes.

This circumstance and our knowledge of the historical and institutional situation – especially the information given by the responsible agency that shelters were opened wherever it was possible, without any political considerations – means that I prefer the OLS specification over the IV.

Table 1.B.2: Distance to refugee shelter and right-wing vote share, OLS and IV

Model	(1) OLS	(2) IV	(3) OLS	(4) IV
Independent variables				
Shelters capacity/distance	0.051***	0.035***		
	(0.001)	(0.004)		
Nearby shelter× (2016 or 2017)			0.012***	0.060
			(0.003)	(0.079)
Distance to centre× (2016 or 2017)	0.007***	0.007***	0.005***	0.005***
	(0.000)	(0.000)	(0.000)	(0.000)
FE polling station	✓	✓	✓	✓
FE election	✓	✓	✓	✓
N. of observations	13283	13283	13283	13283

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***. Standard errors clustered at district level in (3).

Table 1.B.3: Nearby school in previous elections

Outcome variable	right-wing voteshare					
Elections	(1) 2006	(2) 2009	(3) 2011	(4) 2013	(5) 2016	(6) 2017
Distance to school	-0.015***	-0.002	-0.005	-0.009***	-0.025***	-0.014*
	(0.003)	(0.002)	(0.004)	(0.002)	(0.009)	(0.007)
Area	0.002**	-0.001	-0.001	-0.000	-0.000	-0.001
	(0.001)	(0.000)	(0.001)	(0.000)	(0.002)	(0.001)
Distance to centre	0.002***	0.001***	0.002***	0.002***	0.010***	0.007***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
FE district	✓	✓	✓	✓	✓	✓
N. of observations	1,591	1,652	1,596	1,617	1,779	1,779

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***. Standard errors clustered at district level in (3).

1.C Restaurant categories and prices on Foursquare

As discussed in section 2, I construct variables of foreign presence and of wealth in a voting precincts from the Foursquare data.

As an indication of the presence of foreigners in the area, I use the presence of Turkish and Middle Eastern restaurants. The scatterplot Figure 1.C.1 shows how this measure correlates, on the level of city districts where official data is avaiable, how the number of such restaurants is correlated with the percentage of inhabitants who are foreigners.

Figure 1.C.2 shows how on the same level, average real estate prices (per square meter) correlate with the average price ratings received by foursquare places. These correlations are positive as we would expect. It is strong (.82) in the case of the ethnic restaurants, but a bit weaker in the case of price ratings (.59).

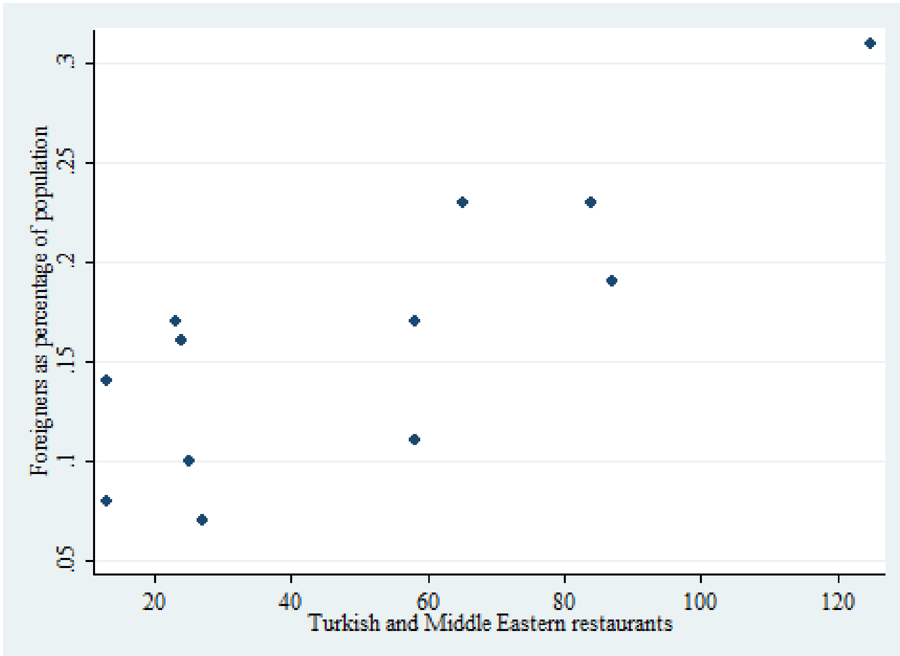


Figure 1.C.1: Foreign population and foreign restaurants on Foursquare by district

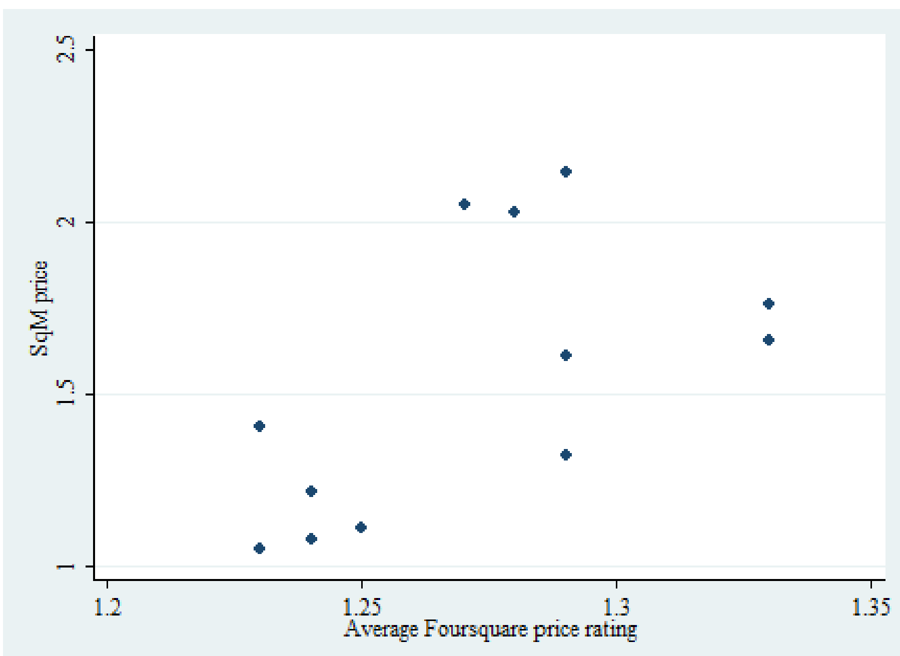


Figure 1.C.2: Real estate prices and Foursquare price ratings by district

Chapter 2

The Demand Effect of Immigration: Refugees and Employment in Germany

The 2014-2015 refugee crisis led to a sharp and unexpected increase in immigration to Germany, particularly from Syria, Iraq, and Afghanistan. While few of the immigrants have found employment, refugees receive social security payments and access local services in the areas where they settle. The arrival of refugee immigrants in a community therefore represents an unexpected demand shock. I exploit the fact that settlement of refugees in Germany was managed to create novel instruments for the arrival of refugees in a given region. Firstly, German federal states used different rules for the internal allocation of refugees, and secondly, the availability of vacant housing was an important factor. I find that immigration increased native employment growth by about 6 percent annually between 2014 and 2017; these gains are just as high for non-university educated workers and occur especially in the service sectors most likely to benefit from the demand shock. Despite this, I find some evidence that refugee immigration has led to increased electoral support for the anti-immigration AfD party.

2.1 Introduction

The arrival in Germany of war refugees from Syria, Iraq, Afghanistan and other countries has received considerable attention in recent years. Given their high numbers – more than 150,000 in 2014, and around one million in 2015 – these immigrants represent a challenge for the German administration and larger society. This challenge has been met at times with a welcoming attitude, at others with concerns about crime, integration and the impact of these new arrivals on local labor markets.

By the end of 2016, only about 50,000 (or 5%) of these immigrants had found a job in Germany, suggesting that displacement effects on native workers, if any, are limited (the number has grown to above 200,000 by the end of 2017). However, this makes it possible to investigate in isolation the impact that immigration can have on local demand, for example in services such as retail and restaurants. In 2016 and 2017, the German government spent almost 20 billion Euros annually, or 0.5% of GDP, on stipends for refugees as well as related expenses such as teaching, construction of shelters and public administration.

At least in the short run, the management of the immigration wave has therefore acted as an unintended employment stimulus for German nationals. The German Employment Agency and the associated research institute IAB reported in 2016 that ca. 50.000 new jobs directly related to the refugee influx have been generated (Arbeitsmarkt, 2016).

This paper investigates these effects of the refugee crisis on local employment, economic development and political outcomes. It exploits a natural experiment arising from the fact that the refugee crisis was partially managed. Refugees were not allowed to settle wherever they wanted, but were distributed across Germany to the 16 Federal States, then to administrative districts and finally to communes and municipalities. Local population size was the most important determinant of how many refugees were to settle in a given area, but certain federal states also considered tax income, surface area, and foreign population when allocating refugees. This paper exploits these differences in dispersal policies to create an instrument for refugee settlement. In practice, the availability of suitable housing for refugees also evolved as an important determinant directing the flow of immigrants. I use this constraint as a second instrument.

I demonstrate that both instruments affect the number of immigrants settling in a given region significantly, and only during the years of the refugee crisis. Furthermore, immigration as well as other outcomes such as employment and GDP growth are not affected by these variables before the refugee crisis.

I find that the immigration wave had a significant positive impact on employment of native workers. Employment growth was boosted by roughly 6 percent annually (0.1 percentage points) over the period of the immigration wave. Workers without a university education benefited strongly, consistent with the finding that most of the effect was felt in service sectors such as construction, retail, hospitality and some public services.

Additionally, and despite these short-term benefits for local employment from refugee immigration, I find limited evidence that areas receiving more refugees voted more strongly for the anti-immigrant AfD party in the 2017 federal elections.

The literature studying the impact of immigration on local employment has mostly focused on how the labour supply shock leads to the displacement of native workers. Card (1990) finds no significant effect on total local employment, while Card (2001) and Borjas (2003) find a moderate negative effect. Ottaviano and Peri (2012) find a positive effect on native wages, arguing that immigrants are imperfect substitutes for native employees, who move to more productive occupations. However, they find immigration to be detrimental to previous immigrants. Borjas and Monras (2017) investigate several refugee waves from the 1960s to the 1990s, coming to the conclusion that they can hurt low-skilled native employees, but also benefit complimentary workers. Clemens and Hunt (2017) try to reconcile the findings from several studies that look at these same refugee supply shocks, and confirm that the negative impacts are rather small.

Moving to the German context, Dustmann et al. (2017) demonstrate the negative employment and wage impact of cross-country commuting along the Czech border. Since the commuting workers do not consume in Germany but rather at the (Czech) location of their homes, this natural experiment provides a compliment to the one I study in this paper, isolating a labour supply shock without increased demand from immigrants. In my case, there is a very limited labour supply shock but strong demand from immigrants due to government assistance, leading to a positive 'stimulus' effect on certain sectors.

This positive effect of immigration on local service sectors is also present in Hong and McLaren (2015), and in Cengiz and Tekguc (2017) who study the immigration of Syrian refugees to Turkey. In their paper, the demand generated by immigrants is coming directly from the immigrant workers themselves, rather than from Government assistance as during the refugee crisis.

Refugee immigrations and the dispersal policies used to house immigrants have been exploited in several recent papers. Glitz (2012) and Foged and Peri (2016) both also investigate the impact on native employment. They look at immigration waves during the 1990 - ethnic Germans coming to Germany from the former Soviet republic in the case of Glitz (2012), and refugees entering Denmark from a variety of origin countries in the case of Foged and Peri (2016). These papers argue that the dispersal policies were independent from labor market conditions and that the resulting distribution can therefore be taken as quasi-random. Glitz shows a moderate negative impact on native employment, while Foged and Peri find that immigration pushed native workers into more cognitively and communicatively demanding jobs, ultimately benefiting them (Peri and Sparber (2009) makes a similar argument while investigating Mexican immigration to the US).

In contrast to these papers, I show that despite the intention to disperse refugees equally during the 2014-2015 crisis, the distribution of refugees across Germany ex-post is quite strongly influenced by pull-factors of immigration. My use of the instrumental variable approach discussed above, which exploits the management of refugee flows more explicitly, therefore provides a methodological alternative. I also contrast my approach with the commonly used strategy of instrumenting immigration with the size of pre-existing communities.

Besides Cengiz and Tekguc (2017), there are a few other working papers studying the effect of the current refugee crisis. For Germany, Gehrsitz and Ungerer (2017) present an analysis of immigration and a number of relevant outcomes (including employment), noting that the frantic management of the crisis introduced randomness in the distribution of refugees. Steinmayr (2016) studies voting outcomes in Austria, using the availability of mass accommodations as an instrument; this strategy has similarities to my use of housing variables. Further afield, Ceritoglu et al. (2017) study labor market outcomes in Turkey, using geographical distance to Syria to differentiate between treated and untreated districts.

The organization of the paper is as follows: section 2 explains the institutional background as well as data sources, and gives summaries of the raw data; section 3 and 4 present the empirical specifications and results respectively. Section 5 provides robustness checks, while section 6 concludes. The appendices provide a theoretical model motivating the investigation of different sectors of employment as well as an exploration of the network instrument as a source of variation. Another appendix investigates how voting outcomes during the 2017 federal election were influenced by the immigration wave.

2.2 Institutions and Data

2.2.1 Data

I use mainly German administrative data, published on the level of administrative districts. The 402 administrative districts (295 rural and 107 urban) do not always form labor markets; often, an urban district and its surrounding rural district should be seen as one market instead. The Federal Office for Building and Regional Planning provides a key by which we can aggregate the districts into labor markets which are defined by commuting patterns[1]. I use this definition to form my basic regional unit of observations, the labor market region (LMR).

The data on immigration by origin country is available from the foreigners registry, which includes the number of foreign nationals, male and female, by nationality and type of residence permit, for each of the administrative districts. Especially during the refugee crisis 2014-16, there was a lag in registrations because the responsible Foreigners Offices were often overwhelmed. However, the numbers from the end of 2016 likely present a realistic picture, and so the total difference in the register between 2014 and 2016 can be taken as total immigration over this period.

Over these three years, the number of residents from the 8 most important non-European refugee origin countries grew by 1.065 Million. These countries - often represented together in German statistics and official announcements - are Syria, Iraq, Iran, Afghanistan and Pakistan as well as Somalia, Eritrea and Nigeria. The number of Syrians alone increased by 580 Thousand, while that of Afghanis is second with 185 Thousand and Iraqis third with 140 Thousand. These three countries together therefore represent 905 thousand immigrants. I will focus on these in my analyses, since I only have detailed employment numbers for these three countries and since they are a relatively homogeneous group with respect to education and cultural background (more so than when e.g. the African or Balkan immigrants were also taken into account). There was also an important immigration from the Balkan countries of Albania, Serbia, Macedonia and Kosovo around the same time, but these immigrants are unlikely to be granted asylum status and unlikely to stay in Germany.

As can be seen in figure 2.2.1, the increase of the registered population from the three main origin countries during 2014-2016 dwarfs immigration during the preceding years.

[1] For a reference, see Raumabgrenzungen (2015).

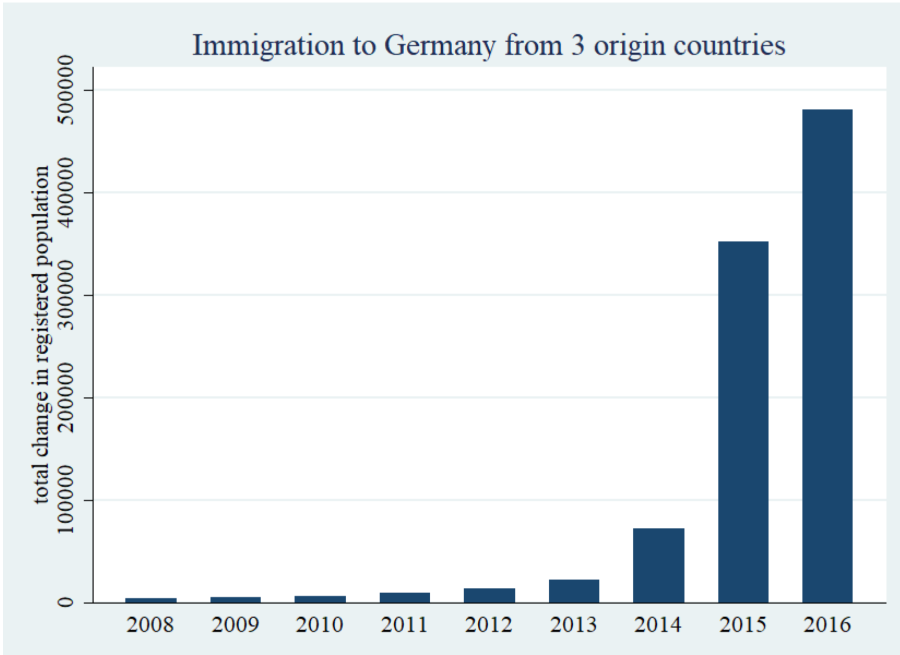


Figure 2.2.1:

The highest increase occurs in 2016, and not in 2015 - this is because of the backlog in registrations, since the foreigners offices were overwhelmed by the immigration wave in 2015. The numbers shown here refer to foreigners registered at their place of residence, rather than the number of foreigners entering Germany, which is why the increase appears only with a delay.

Distribution keys

The federal states of Germany decide how to distribute refugees to administrative districts. When doing so, most only consider the population of each district, but North Rhine-Westphalia (NRW) and Brandenburg take the area of a district into account, and Hesse the share of foreigners in the population; Bavaria tends to assign a higher share of asylum seekers to urban rather than to rural districts, relative to population. These distribution keys were created in the late 1990s or early 2000s, so they cannot have anticipated the current conditions.[2]

[2]The ultimate source of the different distribution keys are the directives issued by the federal states. An overview is given in Geis and Orth (2016).

This gives rise to a situation where two urban or rural districts, which are otherwise similar in population, area, economic development etc., are allocated different numbers of asylum applicants, because their federal states use different distribution keys. A Bavarian city would receive more than one located in Baden-Württemberg, which uses a simple population-based key, while a city in Hesse or NRW would receive fewer. The refugees have hard incentives to remain in the location where they were first assigned, since they receive a stipend from the local authorities which is conditional on their place of residence.

On average, a district in NRW, Bavaria, Hesse or Brandenburg is supposed to receive 9.1% more or fewer asylum seekers than they would with a hypothetical distribution key that considers only population. For the year of 2015, this often means a difference of several hundred immigrants. Figure 1 shows a map of German districts coloured according to this relative difference,

$$key_i = \frac{N_i^{key} - N_i^{pop}}{N_i^{pop}} \tag{2.1}$$

where N_i^{key} is the number of asylum seekers a district i is supposed to accommodate according to the key used in its federal state, and N_i^{pop} is the number it would be supposed to accommodate if a pure population key was used. The quantity key_i, interacted with time dummies for the years of the refugee immigration, is the main instrument I use in my analyses. [3] If we multiply this difference $N_i^{key} - N_i^{pop}$ with the number of refugees arriving in the federal state j of which i is part, we get the number of excess refugees that the district receives due to the idiosyncratic distribution key,

$$excess\_ref_i = (N_i^{key} - N_i^{pop}) \times N_j$$

N_j is the actual number of refugees arriving in the federal state.

This difference is zero in most states, where a pure population key is used. The city states of Berlin, Hamburg and Bremen do not have the same administrative subdivisions. The variation for an econometric analysis comes from the 193 districts in NRW, Bavaria, Hesse and Brandenburg.

[3] the map has not been updated; in an updated map, the districts of the state of Brandenburg would also appear coloured

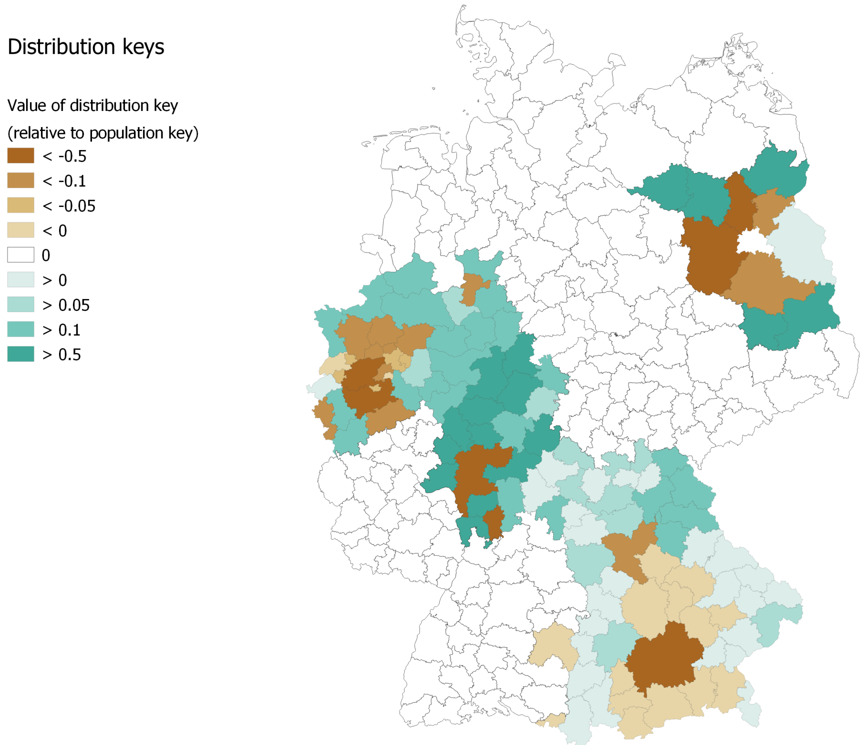


Figure 2.2.2: Labour market regions with non-zero distribution key difference, relative to population key

2.2.2 Actual distribution of refugees and housing availability

The true distribution of refugees ex-post does not show equal numbers of refugees per inhabitants within states, even when this was intended according to distribution key. In my data, the range of the population increase is from zero in one labour market region (Osterode) to 3.5 per hundred inhabitants (Gießen). The mean is 1.2 immigrants per hundred inhabitants, while most LMR receive between .8 and 1.5 , as shown in the histogram Figure 2.2.3. Larger LMR tend to receive more refugees even on a per capita basis, and so do more prosperous ones. There is also a clear regional pattern, as can be seen in the map Figure 2.2.4. The north-west has generally accommodated more immigrants than particularly the former Eastern German states. This is partially due to the distribution key we discussed earlier, but it also points to political factors playing a role.

A strong predictor of immigration is the size of the pre-existing community of the same nationality. Figure 2.2.5 shows in a scatter plot the positive and linear relationship

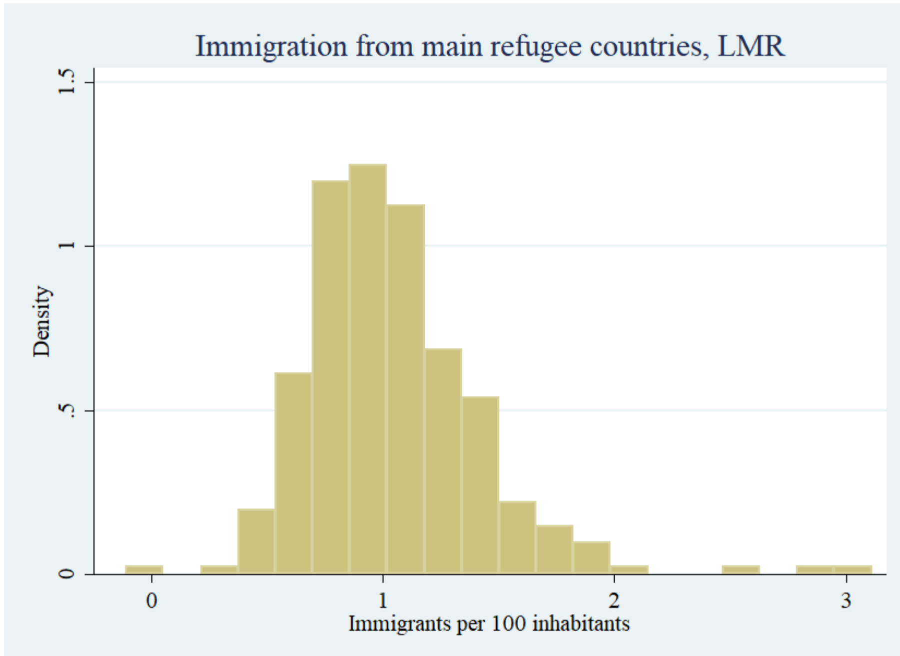


Figure 2.2.3:

between log community size in 2013 and log immigration 2014-2016, for the three main nationalities. The well-known "pull"-factors of immigration clearly play a role in the distribution of refugees in Germany.

I use the availability of suitable housing for refugees, in particular the number of empty apartments in communal ownership, as another exogenous source of regional variation in the number of immigrants.

The difficulty of housing refugees is often anecdotally cited as one of the main determinants of where they were eventually accommodated. An early investigation by the German newspaper Die Zeit, for example, notes that deviations from the prescriptive distribution keys are common, due to pull factors but also to the availability of empty apartments[4]. The Financial Times noted on September 19, 2017: "Refugees tend to be sent to regions with plentiful housing. But these are also places with higher unemployment and lower-quality jobs", corroborating the mechanism I want to exploit but also addressing the main identification challenge of using this instrument.

During the immigration crisis, some municipalities argued that they could accommodate no more newcomers due to the lack of facilities, while others actually volunteered

[4]http://www.zeit.de/gesellschaft/zeitgeschehen/2015-08/fluechtlinge-verteilung-quote

Refugee immigration per inhabitant in German labor market regions

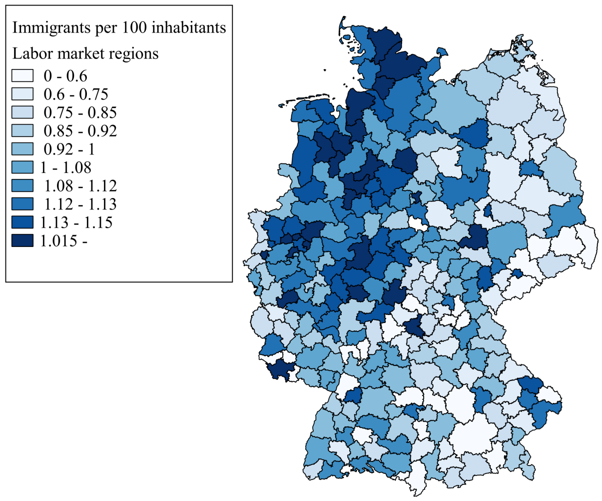


Figure 2.2.4:

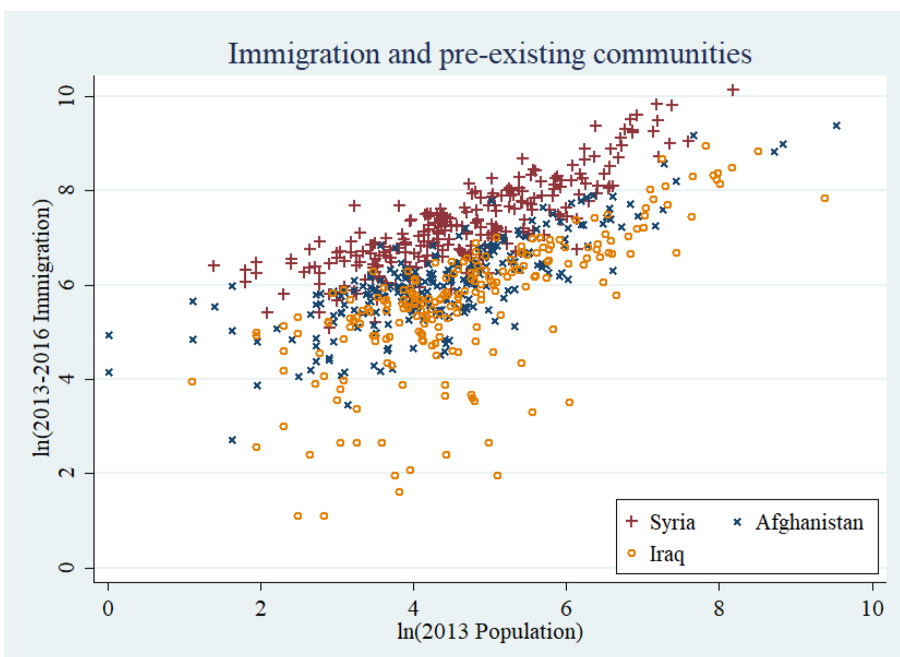


Figure 2.2.5:

to take more than their share (based on population and distribution keys alone), because they had e.g. municipality owned mass accommodation facilities that could be used. Many municipalities built refugee shelters, especially the city states of Berlin and Hamburg, while others appropriated e.g. school gyms and event halls for emergency accommodation. In the longer run however, the refugees are supposed to be housed in regular apartments, and by the end of 2016 this was mostly achieved. It is no surprise then that the number of vacant apartments, and of vacant apartments in public ownership especially, strongly influenced the destination of immigrants.

The data on apartments is taken from the latest census, in 2011 (this also avoids the possibility of communities strategically building, buying or selling publicly owned housing to influence the number of refugees arriving). Of course, larger LMR have more empty apartments (the correlation coefficient between population and number of empty apartments is 0.81); however, since there is a lot of heterogeneity in municipal policies with respect to public ownership of housing assets, the correlation between population and community owned empty apartments is only weak (correlation coefficient of 0.25). On average, there are 7000 empty apartments per LMR, and 685 empty apartments in communal ownership (at the time of the 2011 census). From these numbers, I construct the main cross-sectional part of my second instrument, the number of empty apartments per person:

$$eapp_i = \frac{empty\_apartments_i^{2011census}}{total\_population_i^{2011census}}$$

Before the refugee crisis, neither the number of empty apartments nor of empty apartments in communal ownership impacts immigration from the three most important refugee countries (or from the other countries). During the crisis, the impact is quite significant. We will discuss this when we discuss our main results and their robustness in sections 2.4 and 2.6.

The mean of $eapp_i$ is .025 with a standard deviation of .014. That is, there is roughly one empty apartment for 40 people. The distribution of this variable can be seen in figure 2.2.6; it shows quite a bit of variety, having many values that can be said to be significantly above average.

It must be noted that these variables show some geographical correlation across Germany. In general, the Federal States in former Eastern Germany have more empty apartments per person (the pattern is even stronger with commune owned apartments). This is

Figure 2.2.6:

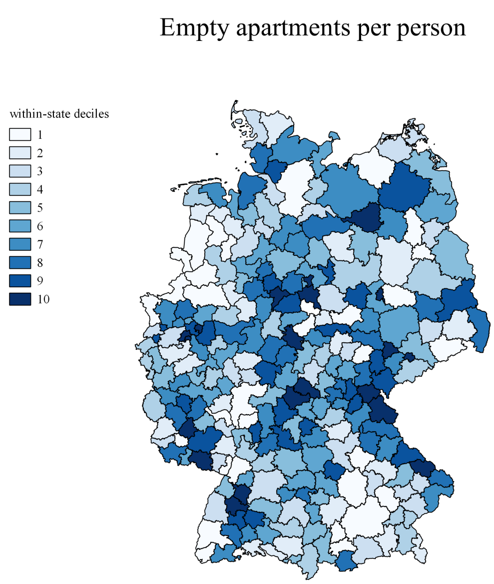


Figure 2.2.7: Labour market regions by within-state decile of empty apartments per person

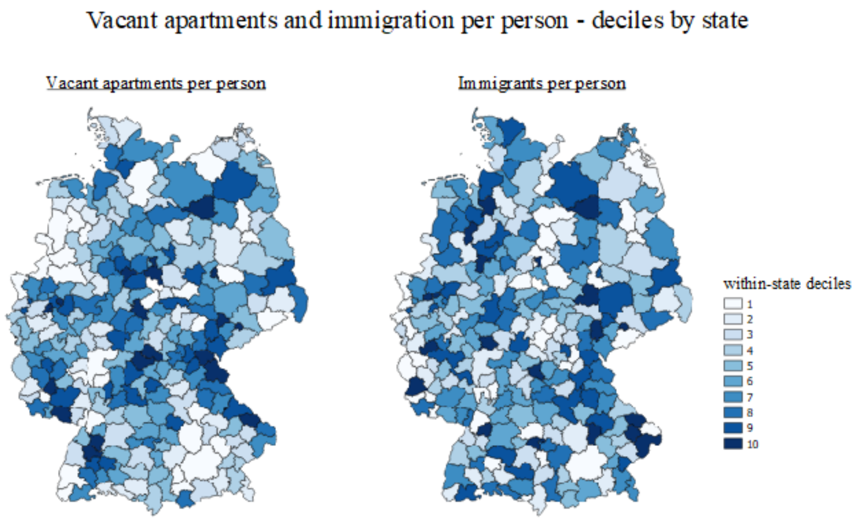


Figure 2.2.8:

one of the reasons why I use state-by-year fixed effects for all of my specifications; this means that the effect I show is not the simple effect of empty apartments per person, but of empty apartments relative to other LMR in the same federal state.

This is important for another reason: the federal states differ in their treatment of refugees. Not only do they receive different numbers of refugees per person according to the federal distribution key, but many policies relevant for labor market access, support and integration of refugees are decided on the level of the state. For example, states decide who is responsible for managing the stipend to refugees - communes or municipalities, or the state itself. Since the police is also a state responsibility, everything relating to policy enforcement also varies from state to state. Thus, some states enforce repatriation of immigrants whose asylum request has been denied, while others do not.

The map Figure 2.2.8 shows empty apartments per person and immigrants per person to the LMR next to each other. However, it shows within-state deciles of both variables rather than absolute numbers; this removes the state fixed effects and therefore provides a better graphical representation of what will be our main first-stage regression. As you can see from the map, LMR which have many empty apartments relative to the state average also tend to receive more refugee immigrants (again, relative to the state average) during the years of 2014-16.

As we have discussed, the size of the pre-existing same origin community has often been used to instrument for the arrival of new immigrants. Appendix 2.A of this paper provides an exploration of this possible instrument, arguing that it is difficult to use in the context considered here. I show that indeed, the network size influences the number of new immigrants, but less so during the crisis years of 2014-2016 than before. On the other hand, regions with strong such networks followed different trends e.g. in unemployment before the refugee wave. In particular, the interactions of a variable indicating the network size with year dummies positively correlate with the change in unemployment, in some years significantly so.

Here, I also use data on the gender composition of immigrants, as well as what type of residence permit they hold (as of Dec. 2016). The impact of the size of the pre-existing community on the share of newcomers who are female is not significant. However, when we look at the share who hold the residence permit which is most tightly linked with asylum - the permit for reasons of international law, humanitarian or political considerations, as opposed to any other permit or a suspension of deportation - it is slightly higher where the network is stronger. This suggests that either there is a selection effect by which "true refugees" rather than economic immigrants tend to seek out stronger networks communities, or that this community actually helps them obtain the more favorable kind of residence permit, allowing them to expect a more secure stay in Germany. Both could confound an analysis relying on this instrument.

2.2.3 Government spending on refugees

Government expenditure on benefits from the asylum seekers act in Germany were 9.5 Billion Euro in 2016, up from 1.5 Billion in 2013 (2014: 2.4 Billion, 2015: 5.3 Billion). The act pays for a stipend and e.g. housing expenses and other basic needs, all of which directly benefit the asylum seekers. It is administered locally, meaning on the level of the district or (in some states) communes; however, since 2015, the federal government has refunded local administrations for these expenses (the totals on expenses as well as breakdowns into varying types come from a report of the German Ministry of Finance (BMF 2017) and a study by the FiFo Institute of the University of Cologne (FiFo Institute 2016)). After an asylum request has been granted, refugees benefit from social security transfers like the native population. In 2016, such transfers to the newly arrived refugees totaled 1.7 Billion Euro; this amount is expected to grow as asylum requests are being processed, while payments through the asylum seekers act would fall.

Table 2.2.1: Federal Spending 2016

Spending due to Asylum policies	Euros, Billion
Total federal spending	21.7
- Refunds to local governments (for payments from refugee seekers act)	9.5
- Combating causes of displacement**	7.1
- Registration, administration and accommodation in temp. shelters	1.4
- Integration efforts (incl. language courses)	2.1
- Social security after asylum request	1.7

Source: Federal Ministry of Finance, 2017

**International spending

Spending through the refugee seekers act is usually accounted as district government spending, even if it is refunded by the federal government. However, local government expenditures have not yet been widely released for the year of 2016. It is available only until 2014 in all states; some states independently publish more recent figures. For other states, it is not yet feasible to directly show how the immigration has influenced local government spending. Therefore, in this version of the paper, I have to rely on (invariant) individual spending, i.e. how much districts are required to spend on individual refugees by law, and assume that we can take these figures as true expenditure in 2014-16.

Expenditure on each individual amount to about 1000 Euros per month and person, whether through social security transfers or through the asylum seekers act; the FiFo Institute report estimates that total expenses per asylum seeker and month are 1063 Euro (including costs of education, which are born by the state rather than the federal government), while they are 1023 Euro for persons benefiting from regular social security (SGB II), such as refugees who's request has been granted. These numbers tally well with the total costs reported by the Ministry of Finance; if we assume that about a thousand euros were payed per month on each of 900 thousand refugees, it would be a total of 10.8 Billion Euros; the totals provided by the Ministry of Finance add up to roughly 10 Billion, 8 Billion through the refugee seekers act + 1.7 Billion through social security. The difference can likely be explained by the fact that not all refugees started receiving benefits in January 2016; a few would only arrive in 2016, while others had arrived in late 2015 but not handed in their asylum requests.

Other expenses associated with the refugees include funds spent for registration, administration and accommodation in temporary shelters as well as efforts to facilitate integration, which includes the language classes and cultural training ('welcome classes'). These together accounts for 3.5 Billion Euros in 2016 (see Table 2.2.1). Since some of these costs (e.g. the integration classes) are paid directly by the federal government rather than paid by local governments and refunded, it will be even more difficult to obtain exact numbers on where this money was spent. The assumption that the local share of this spending is simply the local share of the refugees is therefore harder to substantiate.

The second largest part of total spending does not have a direct benefit on the German economy at all: funds for combating the causes of displacement are more similar to foreign aid. They include e.g. payments to improve security and livelihoods in Syria and for border enforcement in Turkey.

The data available at the moment is not ideal, but the legal requirements as well as our analysis of the states that do provide local breakdowns on spending suggest that within states, where the money is spent follows the pattern of migration closely. For rough estimates of how the demand increase affected employment, I will therefore assume in the following that a refugee brings an increase in spending on housing, consumption, health and education of around 1000 Euros a month. To this we can add the costs not covered by the refugee seekers act or by social security, namely the 3.5 Billion Euros spend on teaching and public administration as well as temporary shelters. Since in the total, this amounts to a quarter of the total domestic expenditure, this translated to 250 Euros per person and month.

2.2.4 Employment of natives and immigrants across sectors

The data on employment by nationality is provided by the German Employment Agency. I use yearly data going back to 2007 and up until 2017. In my main specifications, I use one-year leads of employment changes as outcome variables, owing to the fact that hiring requires some time. 2017 employment changes are thus regressed on 2016 immigration etc. Additionally, I have data on employment by nationality for each sector for 4 years, namely for 2007, 2010, 2013 and 2016. I use these for the explorations of specific sectors. As we will see, the impact in these sectors was felt more instantaneously, so that the effects are significant already in 2016. [5] The available data enables me to use panel data estimation in all my main specifications.

[5] The same breakdown for 2017 was not yet available at the time of writing this version.

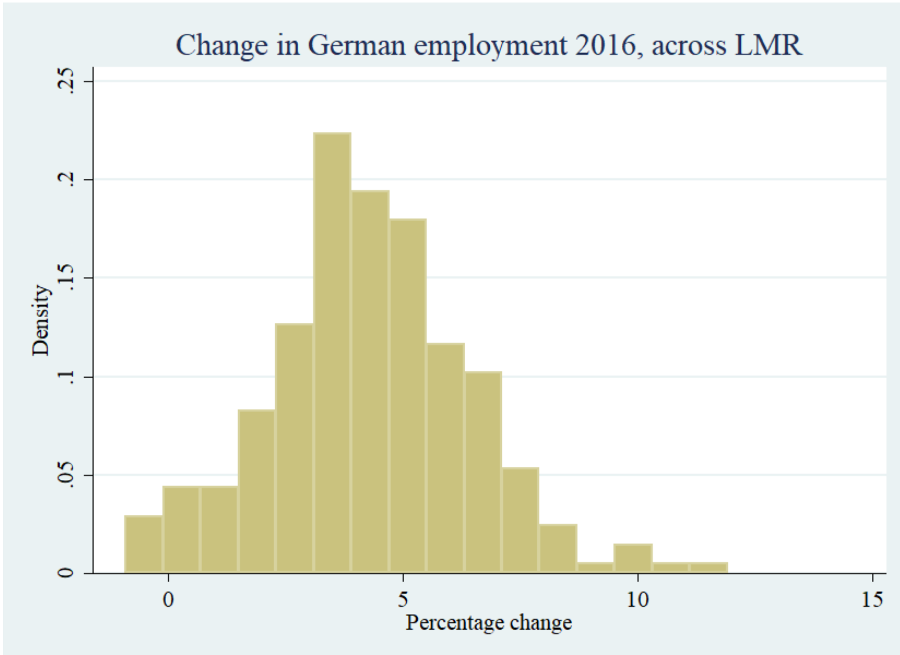


Figure 2.2.9:

The employment situation in Germany steadily improved during the 10 years under study. But across LMR, there is a lot of variation in employment change (figure 2.2.9). Over the three years of 2013 and 2016, the mean employment change was by 4.3 percent, with a minimum of -.89 and a maximum of 11.9 percent.

I use employment numbers for Germans and for Syrians, Iraqis and Afghanis, either separately or aggregated. I have a partition of these employment numbers into 17 sectors, of which I define several as likely to benefit directly from the demand generated by refugees: retail and hospitality. Three more sectors could benefit from expenditure related to the management of immigration, namely construction, education and other public administration.[6]

The most important sector to employ workers from the three origin countries is hospitality/restaurants. On average, 21% of these immigrant employees in a labor market region work in this sector. This is followed by manufacturing (14%), healthcare (12%) and commerce (10% - these relative numbers are from 2013, before the immigration wave,

[6]It should be noted that for some sectors in some LMR, the employment data provided to me was censored/anonymized. This was the case when there were only two or fewer firms active in that sector/LMR. These sectors simply do not appear in my empirical regressions.

Figure 2.2.10: C: Manufacturing, G: Commerce, I: Hospitality, health: Healthcare

but do not change much over time). There is however a lot of heterogeneity across labor market regions in these shares. Figure 2.2.10 shows how the share of employees in these main sectors is distributed across LMRs.

While, as we have noted, the size of the population and the total number of workers from the three origin countries strongly influence immigration during the refugee crisis, these employment shares seem unrelated with new arrivals. For example, a region where most Syrian workers are employed in manufacturing does not attract more or fewer immigrants than one where most are in hospitality or healthcare, conditional on the total number of Syrians. These employment shares also do not affect selection to the (very limited) extend that we can investigate it (that is, they do not affect the share of immigrants who are women or the share who have their asylum request approved). The regressions supporting this are shown in the appendix.

2.3 Empirical strategy

2.3.1 Instrumental variable strategy - distribution keys, housing, immigration and employment

The main difficulties in studying the impact of the refugee immigration on labor markets are the possibilities of reverse causality and joint determination: the processes driving employment changes could also determine where the immigrants settle, and of course, the expectation of positive labor market developments is a pull-factor in itself. This section details our strategy to address these issues.

The timing of the refugee crisis follows the Syrian and Iraqi civil wars, and other Middle Eastern and African conflicts. It can therefore be said that it is exogenous to local labor market conditions in Germany. However, simply using the arrival of refugees in Germany as an instrument for local refugee immigration would not suffice, as it does not give us any variation between German labor markets.

We are interested in the impact of immigration on native employment changes, and therefore ultimately in a specification such as:

$$\delta L_{i,t}^{DE} = \beta\, I_{it} + \Gamma X_{it} + FE_i + FE_{t,r} + \epsilon_{i,t} \tag{2.2}$$

$\delta L_{i,t}^{DE}$ is our outcome variable, the percentage change in German employment growth in labor market region i at time t, overall or in specific sectors. The main exploratory variable is I_{it}, the number of new immigrations in the region per person (the difference between the old and new population size, normalized by the population size of the region).

The obvious problem would be that there are other changes over time that affect employment trends, which could be correlated with immigration because they also fall in the crisis years of 2014-2016, or because they directly affect where immigrants settle.

We have seen that the differences in distribution keys and the availability of suitable housing has created exogenous variation in how many refugees arrive in a district during the crisis - we will exploit this as the required variation across regions. Thus, we can introduce time fixed effects to flexibly accommodate changing employment trends in Germany. Our first stage and second stage estimation equations become

$$\delta L_{i,t}^{DE} = \beta\, I_{it} + \Gamma X_{it} + FE_i + FE_{t,r} + \epsilon_{i,t} \tag{2.3}$$

$$I_{i,t} = \alpha\, (T_t \times key_i) + \Gamma X_{it} + FE_i + FE_{t,r} + \epsilon_{i,t} \tag{2.4}$$

where key_i is the number of refugees allocated to a region above or below what a pure population key would prescribe, as defined in 2.1. T_t is a treatment dummy taking the value of 1 during the crisis years. The instrument $T_t \times key_i$ varies by year and region. Using T_t, the (exogenous) refugee immigration to Germany (in millions, so that is is roughly one in 2014-2016) provides the time variation in the explanatory variable. The region fixed effects FE_i control for important characteristics such a region's total population and pre-existing population from the refugee origin countries. X_{it} is a set of time varying controls at the level of the labor market region, such as GDP growth, population growth and (potentially) political variables.

The time fixed effects $FE_{t,r}$ are allowed to vary by Federal State. This allows us to account for the various institutional differences between states discussed before, which affect both immigration and labour market outcomes.

As an alternative, we can use a continuous variable I_t^{DE} which is total immigration into Germany, rather than the indicator taking the value of 1 during the crisis years and zero during other years. As we will see (in section 6, robustness), this does not have an impact on our estimates, which is to be expected since the number of arrivals is so much higher in 2014-2016 than before that using a dummy is just as good as this continuous variable.

This specification is similar to a difference-in-difference strategy in that it compares regions which receive more or fewer refugees due to the specific distribution key of their federal state as 'treated' areas, during the time of the refugee crisis relative to the time before. Immigration is affected by the distribution key and housing variables only during the crisis; if they are jointly determined with labor market trends, the region fixed effects address this issue. This assumes that while the instrument may be correlated with time-invariant individual characteristics (and possibly individual linear time trends), it is independent from time-varying shocks.

In an additional specification, we will not focus on the total employment change by region, but at specific sectors and how they are impacted by immigration. This is motivated by the fact that some sectors benefit from the demand generated by the immigrants. The specification therefore becomes

$$\delta L^{DE}_{i,t,s} = \beta_1\, I_{it} + \beta_2\, I_{it} \times D^D_s + \Gamma X_{it} + FE_i + FE_t + FE_s + \epsilon_{i,t,s} \tag{2.5}$$

$$I_{i,t} = \alpha\, (T_t \times key_i) + \Gamma X_{it} + FE_i + FE_t + FE_s + \epsilon_{i,t,s} \tag{2.6}$$

where $L_{i,t,s}$ is sectoral employment, rather than total employment in a region. D^D_s is a dummy taking the value one if sector s is likely to benefit from the demand generated by the specific need of the refugees - I simply used those sectors mentioned by the head of the federal employment agency's research institute, quoted in the introduction - construction, teaching, security and public administration. The coefficient β_2 thus measures the different impact that immigration to a region has on those sectors, compared to all other sectors. Immigration to the region is still instrumented as before, namely by the first stage equation 2.6.

2.4 Results

2.4.1 Demand effect of immigration

I will first present the results of the OLS regressions specified in equations 2.3 and 2.5. The variable of interest, immigration, is in these first regression the change in the population from the three most important refugee origin countries, Syria, Iraq and Afghanistan, relative to the regions's population. The outcome variable is the percentage change in employed persons of German nationality in the labor market region. It is allowed to be different for German individuals without a university degree.

The results are presented in table 2.4.1. Surprisingly, immigration is not correlated with this outcome variable in the OLS specification (1), when region fixed effects and region time trends are taken into account. This is possibly due to the fact that some of the factors attracting refugees, such as a pre-existing community, where less present in such regions that experienced above trend employment growth in the crisis years. However, we can not make any causal claims or rule out an effect, because of the endogeneity of immigration.

The results of the first stage regression of immigration on our instruments - the interaction of a time dummy for the years of the refugee crisis with the distribution keys and with available housing, are presented in table 2.4.2.

Table 2.4.1: Immigration and employment change in percent, OLS

	Employment change	
Immigrants (per hundred inhabitants)	-1.140	-2.641
	(21.367)	(19.763)
Region FE	Yes	Yes
Type × year FE	Yes	Yes
State × year FE	Yes	Yes
r2	0.1	0.17
N	2290	1782

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***.

The coefficient of roughly 0.05 means that a region with a distribution key of +1 would receive .05 more refugees per person during the crisis years than one with a value of zero (where the distribution key is exactly the regions share of the state's population). As discussed before, the mean absolute value of the distribution key is close to 0.1 – a region with a value of 0.1 would receive 0.005 more refugees per person. Since the baseline mean is one refugee per 100 inhabitants (0.01 per person), the distribution key leads to an increase of 50% over the counterfactual scenario in which the region has a distribution key value of zero, like the regions in the majority of states where only population is taken into account. This suggests that the states and regions over-comply significantly with the differences dictated by distribution key differences, where they exist.[7]

The coefficient of 0.04 on the second instrument means that during the refugee crisis, one additional empty apartment per 100 inhabitants meant roughly 0.04 additional immigrants per person. The average number of empty apartments per 100 persons in a labor market is about 2.5, with a standard deviation of 1.4, and the number of immigrants per 100 persons is 1. A region with a number of empty apartments which is one standard deviation above the mean would therefore receive 1.14 immigrants per hundred inhabitants, or about 14% more than average. Taken together, this shows that the instrument explains quite a bit of the variation in immigration.

Note that the region fixed effects and the state-by-year fixed effects together explain

[7] beyond overcompliance, there could also be path dependencies in where the refugees settle, magnifying initially small differences over time.

much of the remaining variation, as the R-squared is quite high.

Table 2.4.2: First stage: distribution keys, empty apartments and immigration

	Immigrants per person			
Crisis × distribution key	0.0551**		0.0459*	0.0323
	(0.025)		(0.027)	(0.025)
Crisis × empty apmnts.		0.0377**	0.0421**	0.0336**
		(0.015)	(0.017)	(0.017)
Region FE	Yes	Yes	Yes	Yes
Region × time trends				Yes
State × year FE	Yes	Yes	Yes	Yes
r2	0.756	0.749	0.761	0.785
N	2301	2303	2301	2301

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***.

Also of interest is the simple reduced form regression of our outcome variable on the instrument. The results of these regressions are presented in table 2.4.3. The coefficients are positive and, for the apartment instrument, weakly significant.

Finally we come to the estimates from the full IV regressions as specified in Equations 2.3 (second stage) and 2.4 (first stage) in table 2.4.4. I use both instruments alone in specification (1) and (2), and together in my preferred specification (3). This last specification allows us to exploit both instruments together, and the estimates are more precise as a result. The coefficient of immigrants per person on total percentage employment indicates that one immigrant more per 100 inhabitants would raise employment growth of German nationals by around 6% in the following year. Since annual employment growth during the crisis years was about 2%, this amounts to .12 percentage points of additional growth per year due to the refugee crisis.

A similar picture emerges when only employment of the less educated is considered (i.e. the non-university educated). Employment growth for this group was 4.3% higher in regions that receive one refugee more per 100 inhabitants (table 2.4.5).

Table 2.4.3: Instruments and employment change in percent, OLS

Outcome variable	Employment change, German nationals		
Model	(1) Key difference IV	(2) Appartments IV	(3) both IV
Distribution keys	13.08		5.986
	(16.242)		(16.794)
Empty apmnts.		35.81**	36.08**
		(15.152)	(15.344)
Region FE	Yes	Yes	Yes
Region × time trends	Yes	Yes	Yes
State × year FE	Yes	Yes	Yes
r2	0.0968	0.0977	0.0967
N	2290	2299	2290

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***.

Table 2.4.4: Refugee immigration and employment change in percent, IV

Outcome variable	Employment change, German nationals		
Model	(1) Key difference IV	(2) Appartments IV	(3) both IV
Immigrants (per hundred inhabitants)	2.3	7.2*	6.3**
	(3.0)	(3.9)	(3.1)
Region FE	Yes	Yes	Yes
Region × time trends	Yes	Yes	Yes
State × year FE	Yes	Yes	Yes
F statistic	14.8	61.2	35.9
N	2290	2290	2290

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***.

Table 2.4.5: Refugee immigration and employment change in percent, IV (less educated)

Outcome variable	Employment change, less educated German nationals		
Model	(1) Key difference IV	(2) Appartments IV	(3) both IV
Immigrants (per hundred inhabitants)	4.3	4.3*	4.3**
	(3.0)	(2.5)	(2.1)
Region FE	Yes	Yes	Yes
Region × time trends	Yes	Yes	Yes
State × year FE	Yes	Yes	Yes
F statistic	13.1	52.9	31.2
N	2290	2290	2290

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***.

2.4.2 Employment in specific sectors

Immigration could impact different industries differently - for example, we expect a positive impact in those sectors that expanded to manage the immigration, such as construction, teaching, security and administration. Using the specification in equations 2.5 and 2.6, we allow the coefficient on our variable of interest (immigrants to the region per person) to be different for those sectors. The results from this regression are presented in table 2.4.6.

As discussed earlier, I only have the breakdown for employment by sector and nationality for four years - 2007, 2010, 2013 and 2016, while in the main analysis presented above, I use yearly data and the last year of employment data is 2017. The panel is therefore shorter, and I aggregate refugee immigration as my main explanatory variable across these three-year periods.

Using sectoral data from this shorter time period, we find an insignificant main effect of immigration on employment (even if the coefficient is of a similar magnitude as in the main analysis). However, in the case of sectors benefiting from local demand, the coefficient is larger and significant. This suggests that indeed, these are the sectors especially benefiting from refugee immigration.

We take this as evidence for the hypothesis that it is the spending associated with refugees which lead to the increase in employment.

Table 2.4.6: Immigration and employment change in percent, IV - local demand

	(1) log emp. growth (both IV)
Immigrants (per hundred inhabitants)	4.31
	(16.54)
Immigrants × local demand	6.79**
	(2.85)
Region FE	Yes
Region × time trends	Yes
State × year FE	Yes
Sector FE	
F	
N	11643

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***.

2.4.3 Government spending and the GDP multiplier

Given tour hypothesis that it is the additional government spending of roughly 1 Thousand Euros per refugee per year that led to the positive employment effect, we also expect GDP growth to be positively affected. However, when we simply use GDP growth as an outcome variable in our main regression, we do not see a significant impact (see table 2.4.7).

Another possible way of analyzing the effect of additional government is to run a multiplier regression, assuming that additional government spending on refugees is around 1 Thousand Euros per person (the expenses required by the asylum seekers act, see the above section on Government spending). The empirical specification for my multiplier regression is

$$\frac{Y_{it} - Y_{it-1}}{Y_{it-1}} = \beta \frac{G_{it} - G_{it-1}}{Y_{it-1}} + FE_i + FE_{t,r} + \epsilon_{i,t} \tag{2.7}$$

Table 2.4.7: Refugee immigration and GDP growth, IV

Outcome variable	GDP growth		
Model	(1) Key difference IV	(2) Appartments IV	(3) both IV
Immigrants (per hundred inhabitants)	0.2	0.03	0.05
	(0.1)	(0.1)	(0.1)
Region FE	Yes	Yes	Yes
Region × time trends	Yes	Yes	Yes
State × year FE	Yes	Yes	Yes
F statistic	4.1	18.6	10.7
N	2290	2290	2290

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***.

The increase in government spending $G_{it} - G_{it-1}$ is approximated by the increase required by law given the number of asylum seekers during a year; then, I instrument it as before with our two instruments.

The coefficient β is interpreted as the government spending multiplier; if it is above one, the spending triggers other economic activity, while a multiplier below one indicates that the spending crowds out such activity.

But in the specification using both instrumental variables, the β coefficient is insignificant as before. Given the large standard error, we cannot conclude anything from its value. This may be due to the fact that, at the time of writing, GDP data is only available until 2016, and not as in the case of employment until 2017. It is also possible that the increase in spending is badly approximated when we simply assume that it follows the asylum seeker immigrant numbers, even though it is legally required to do so.

2.5 2017 Federal Elections in Germany

Immigration can also affect attitudes of the native population towards immigrants and towards parties opposing immigration, something that may be reflected in voting outcomes. The empirical strategy developed in this paper can in principle also be applied to this question.

Table 2.4.8: Immigration and employment change in percent, IV - local demand

	(1) GDP growth, both IV
Government spending increase	1.3
	(2.1)
Region FE	Yes
Region × time trends	Yes
State × year FE	Yes
Sector FE	
F	17.5
N	2275

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***.

To demonstrate this, I use voting outcomes for all German administrative districts in the last three federal elections - 2009, 2013 and 2017. In the last of these, the refugee crisis was one of the most important issues. A post-election survey of voters of the anti-immigration AfD party revealed that 92% of respondents thought that "the (AfD) party is mainly there to change the refugee policies with its initiatives", while 95% said that they feared a loss of German culture.[8] These responses show that far-right voters were indeed worried about immigration, but it is an open question which effect direct exposure to immigrants has on these attitudes.

Two recent papers studying this link are Otto and Steinhardt (2014) and Harmon (2017); they document that local immigration increases the vote share of rightwing parties. Steinmayr (2016) on the other hand shows that during the same refugee crisis studied here, proximity to refugee shelters lead to a decrease in the right-wing vote. These papers discuss the contact hypothesis, which proposes that contact with immigrants could reduce prejudice and thereby right-wing voting. I use larger geographic units than either of these papers, which means that my treatment is slightly different: from the perspective of the German voter, it is not so much direct contact with the immigrants as the experience of immigration to the larger region.

Studies of immigration and voting face the same problems of reverse causality as discussed in the context of labor market outcomes, since places with populations who are

[8] Infratest/dimap survey for the ARD, 24 Sept. 2017; link: https://wahl.tagesschau.de/wahlen/2017-09-24-BT-DE/umfrage-afd.shtml

less welcoming to immigrants may attract fewer of them in the first place. And since we have seen that even during the crisis, such pull-factors still played a role in the destination of refugees, my preferred specifications are again those that make use of the differences in distribution keys as an instrumental variable.

Firstly, I use a similar empirical specification as before, namely

$$rightwing_{i,t} = \beta\, I_{it} + FE_i + FE_{t,r} + \epsilon_{i,t} \tag{2.8}$$

$$I_{i,t} = \alpha\, (T_t \times instrument_i) + FE_i + FE_{t,r} + \epsilon_{i,t} \tag{2.9}$$

where the dependent variable $rightwing_{i,t}$ is the combined share of rightwing parties in electoral district i at time t (or its logarithm in an alternative specification). I_{it} is refugee immigration per 100 inhabitants as before, T_t is a time dummy taking the value 1 for the period after the refugee crisis (2017) and $instrument_i$ is the distribution key difference as defined before.

There are reasons to think that the linear and additive fixed effects model may not be appropriate in the German context. In particular, the AfD was founded only a few months before the 2013 federal elections, and it started out with a different profile - less extreme than its 2017 anti-immigrant incarnation. Other, older right-wing parties also never quite had the platform of the 2017 AfD (the most prominent previous such party, the NPD, is often described as neo-Nazi). Therefore, estimating fixed effects for the regions based on previous AfD or NPD voteshares is difficult, because these parties did not make the same political offering.

It may be more appropriate to use the logarithm of the right-wing voteshare as an outcome variable. This assumes that fixed region- and year characteristics, as well as refugee immigration to the region, impact the voteshare of right-wing parties in a multiplicative rather than additive way. For example, a region giving 2.5% and 5% of the vote to right-wing parties in 2009 and 2013 respective would give them 20% in 2017; in another region, the respective voteshares would only be 1%, 2% and 8%.

In simple regressions of (1) the right-wing voteshare or (2) its logarithm on year and region fixed effects, the second specification has the higher R-squared (.96 rather than .8 - see table 2.5.1). This is an indication that, indeed, the logarithmic specification is more appropriate in the German context.

Table 2.5.1: Additive vs. multiplicative fixed effects

Outcome	(1) Right-wing voteshare	(2) Log(Right-wing voteshare)
FE 2013	0.044***	1.40***
	(0.0)	(0.1)
FE 2017	0.12***	2.21***
	(0.0)	(0.0)
Region FE	Yes	Yes
R2	0.80	0.96
N	767	767

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***.

2.5.1 Results

The results from the panel regressions are reported in table 2.5.2. I find weak evidence for a positive effect of refugee immigration on the vote share of right-wing parties. My preferred IV specification (4) does give a positive point estimate that approaches significance; it is also larger than the OLS estimate (3), which we indeed expect to be downward biased due to the aforementioned reverse causality concerns. That the OLS estimate is nevertheless positive may be partially due to the inclusion of several sets of fixed effects - the raw correlation between refugee immigration and the right-wing voteshare (or its logarithm) is negative (-.35).

2.5.2 Alternative specifications

The limited comparability of the election "menus" facing German voters before and after the refugee crisis remains an important caveat for the interpretations of these results. To relieve us of this problem, we can use a simple cross-sectional model rather than the fixed effects model, and regress the vote share of the AfD in 2017 on immigration and suitable demographic and socio-economic control variables.

$$afd_i = \beta\, I_i + \Gamma X_i + FE_r + \epsilon_i \tag{2.10}$$

$$I_{i,t} = \alpha\, keydiff_i + \Gamma X_i + FE_r + \epsilon_{i,t} \tag{2.11}$$

Table 2.5.2: Immigration and voting, fixed effects

Outcome	Right-wing voteshare		Log(Right-wing voteshare)	
Model	(1) OLS	(2) IV	(3) OLS	(4) IV
Refugees p. hundred inhabitants	0.010	0.07	0.09***	0.3*
	(0.0)	(0.0)	(0.0)	(0.2)
Region FE	Yes	Yes	Yes	Yes
State × year FE	Yes	Yes	Yes	Yes
Type × year FE	Yes	Yes	Yes	Yes
N	765	765	765	765

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***.

The outcome variable is now afd_i, the vote share of the AfD party in 2017 (alternatively the increase of this share from 2013 to 2017 - this does not change the results). The vector of control variables X_i includes GDP per person, GDP growth, unemployment, population and foreign population share, as well as a classification into urban, high-density rural and low-density rural regions.

The results in these specification are insignificant, even if the magnitudes are similar to those in the panel estimations (see table 2.5.3).

Table 2.5.3: Immigration and voting, demographic and economic controls

Model	(1) OLS	(2) IV
Refugees p. hundred inhabitants	0.01	0.008
	(0.0)	(0.3)
Share of foreigners	-1.6	-1.6*
	(1.0)	(1.0)
Unemployment rate	10.0***	10.0**
	(1.8)	(4.0)
Other control variables	Yes	Yes
State FE	Yes	Yes
N	255	255

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***.

2.5.3 Interactions

One question of interest concerns the heterogeneous effects of immigration on the popularity of right-wing parties. In particular, immigration could have a different impact in areas which already have many immigrants, or in areas which have high unemployment. I test these hypotheses by interacting immigration, instrumented as before by the distribution key differences, with the share of foreigners in the population and with the unemployment rate.

The results are not significant (table 2.5.4). It is notable however that, in the IV specifications, the main coefficient is similar to before, while the coefficients on the interactions have the sign we expect from the literature - a large pre-existing population of foreigners diminishes the impact of immigration, while a high unemployment rate increases it.

Table 2.5.4: Immigration and voting, fixed effects

Outcome	Log(Right-wing voteshare)			
Model	(1) OLS	(2) IV	(3) OLS	(4) IV
Refugees p. hundred inhabitants	0.04	0.2	0.1	0.2
	(0.1)	(0.2)	(0.1)	(0.3)
Refugees × Share of foreigners	1.3	-6.0		
	(1.3)	(5.8)		
Refugees × Unemployment rate			-0.8	5.7
			(1.6)	(12.9)
Region FE	Yes	Yes	Yes	Yes
State × year FE	Yes	Yes	Yes	Yes
Type × year FE	Yes	Yes	Yes	Yes
N	765	765	765	765

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***.

The same is also true of the interactions of refugee immigration with the unemployment rate and with the share of foreigners in the population (table 2.5.5).

In conclusion, using the same strategy employed in the main analysis of the paper, I find weak evidence that the refugee immigration has contributed to the rise of right-wing parties in Germany. My relatively large units of analysis do not allow me to pick up the more local effects of direct exposure. It would be worthwhile to return to the question with more granular and local observations.

Table 2.5.5: Immigration and voting, demographic and economic controls, interactions

Model	(1) OLS	(2) IV
Refugees p. hundred inhabitants	-0.02	0.0006
	(0.0)	(0.1)
Refugees × Unemployment rate	0.5	1.9
	(0.4)	(4.4)
Refugees × Share of foreigners	0.09	-0.5
	(0.3)	(2.2)
Share of foreigners	-0.2	0.1
	(0.2)	(1.3)
Unemployment rate	0.8**	-0.8
	(0.4)	(3.2)
Other control variables	Yes	Yes
State FE	Yes	Yes
N	255	255

Significant at: p < 0.1: *; p < 0.05: **; p < 0.01: ***.

2.6 Robustness

2.6.1 Common Trends

The difference-in-differences instrumental variable (DDIV) strategy I use, which interacts an exogenous time-varying variable with a possibly endogenous cross-sectional variable, requires that the potential trends of both treatment and outcomes are independent from actual the value of the instrument. If the 'treatment' group (i.e. regions with high value on the cross-sectional variable) shows time trends which are different from the non-treatment group during years when the time-treatment dummy is zero, then different outcomes during years when this time-value is one cannot be interpreted as casual.[9] In my case, I argue that areas with non-zero values on the distribution key difference variable, and areas with more or fewer vacant apartments would have had similar developments to other areas in the absence of the immigration wave.

To address this concern, I will demonstrate that my cross-sectional instrumental vari-

[9] Jaeger et al. (2018) discusses instruments commonly used in the immigration literature along these lines; Christian et al. (2017) similarly critiques a paper on food aid and conflict (Nunn and Qian, 2014).

ables, both the distribution key differences and the number of empty apartments per person, do not determine either immigration or outcome variables – unemployment trends, GDP growth or GDP growth in sectors benefiting from refugee demand – in the period prior to the refugee crisis. This shows that, in fact, the 'treatment' and 'control' areas did follow common trends before the immigration wave. The regression specifications are

$$I_{i,t} = \alpha_t\,(instrument_i \times \mathbf{YD}_t) + \Gamma X_{it} + FE_i + FE_{t,r} + \epsilon_{i,t} \tag{2.12}$$

and

$$\delta unemp_{i,t} = \alpha_t\,(instrument_i \times \mathbf{YD}_t) + \Gamma X_{it} + FE_i + FE_{t,r} + \epsilon_{i,t} \tag{2.13}$$

$$GDPgrowth_{i,t} = \alpha_t\,(instrument_i \times \mathbf{YD}_t) + \Gamma X_{it} + FE_i + FE_{t,r} + \epsilon_{i,t} \tag{2.14}$$

$$GDPgrowth^{demand}_{i,t} = \alpha_t\,(instrument_i \times \mathbf{YD}_t) + \Gamma X_{it} + FE_i + FE_{t,r} + \epsilon_{i,t} \tag{2.15}$$

where $\mathbf{YD}_t$ is a set of year dummies. Thus, we investigate if the instrumental variables have an impact on immigration before the immigration wave with specification 2.12, and if it has an effect on employment growth (2.13) or GDP growth, either across all sectors in 2.14 or only in those sectors likely to benefit from refugee demand in 2.15. The tables of results from these regressions are available upon request; the figures 2.6.1, 2.6.2 ,2.6.3, 2.6.4, 2.6.5 and 2.6.6 show these same coefficients in graphical form.

It becomes clear that prior to the refugee immigration wave of 2014-16, the distribution keys and empty apartments per person did not influence either immigration or the economic development of a labor market region. Almost all of the coefficients before 2015 are insignificant.

When I do not normalize the empty apartment variable and immigration by the size of the LMR, the picture for immigration looks more dramatic - figure 2.6.7 provides the coefficients of a regression of the total number of immigrants on the total number of empty apartments, interacted with a set of year dummies.

2.6.2 Discretization

Another way of testing the robustness of our results is to discretize the instrumental variables. Since we are using the time-federal state fixed effects in all our regressions (for

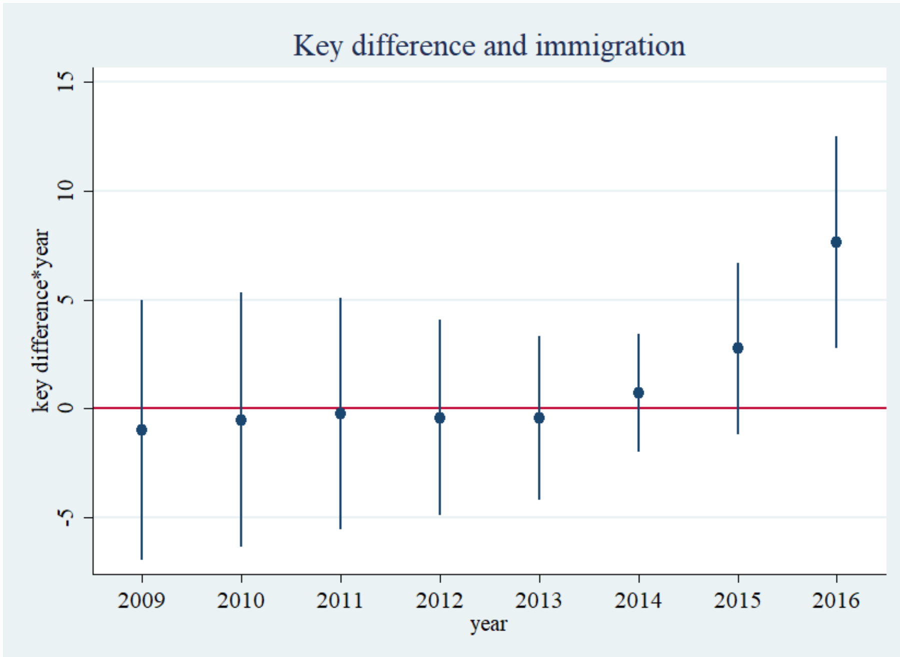


Figure 2.6.1:

Figure 2.6.2:

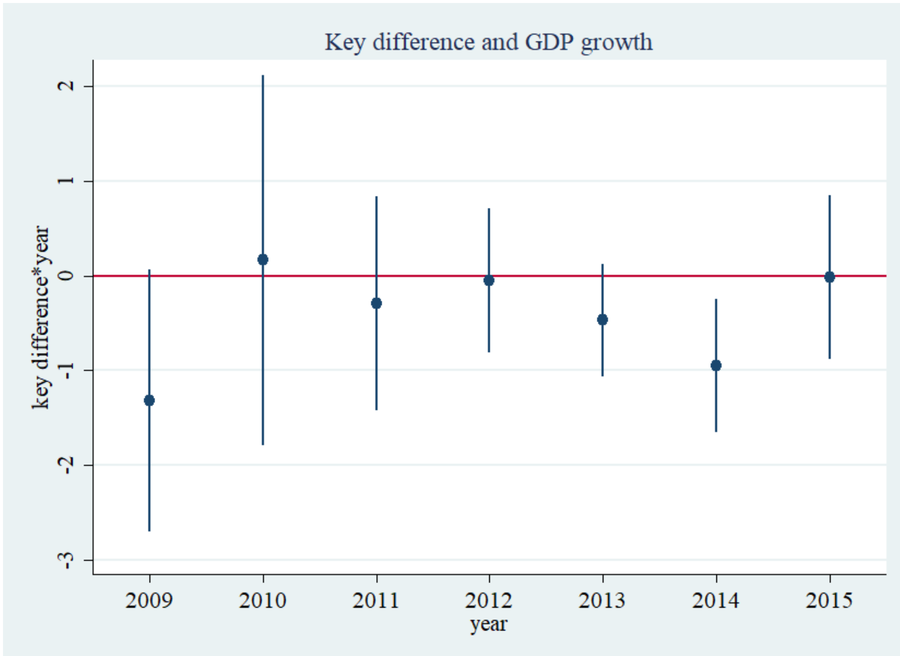


Figure 2.6.3:

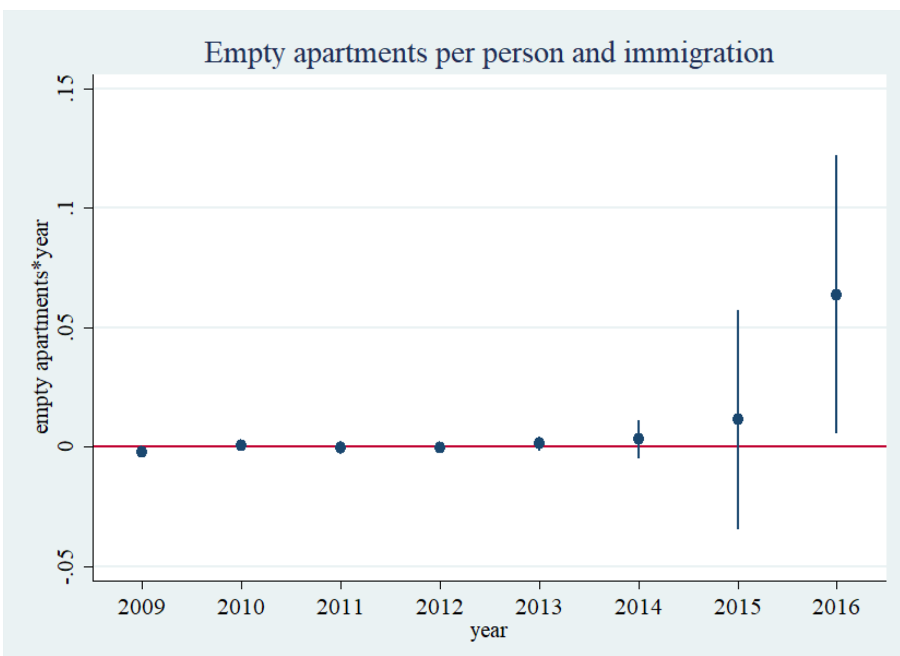


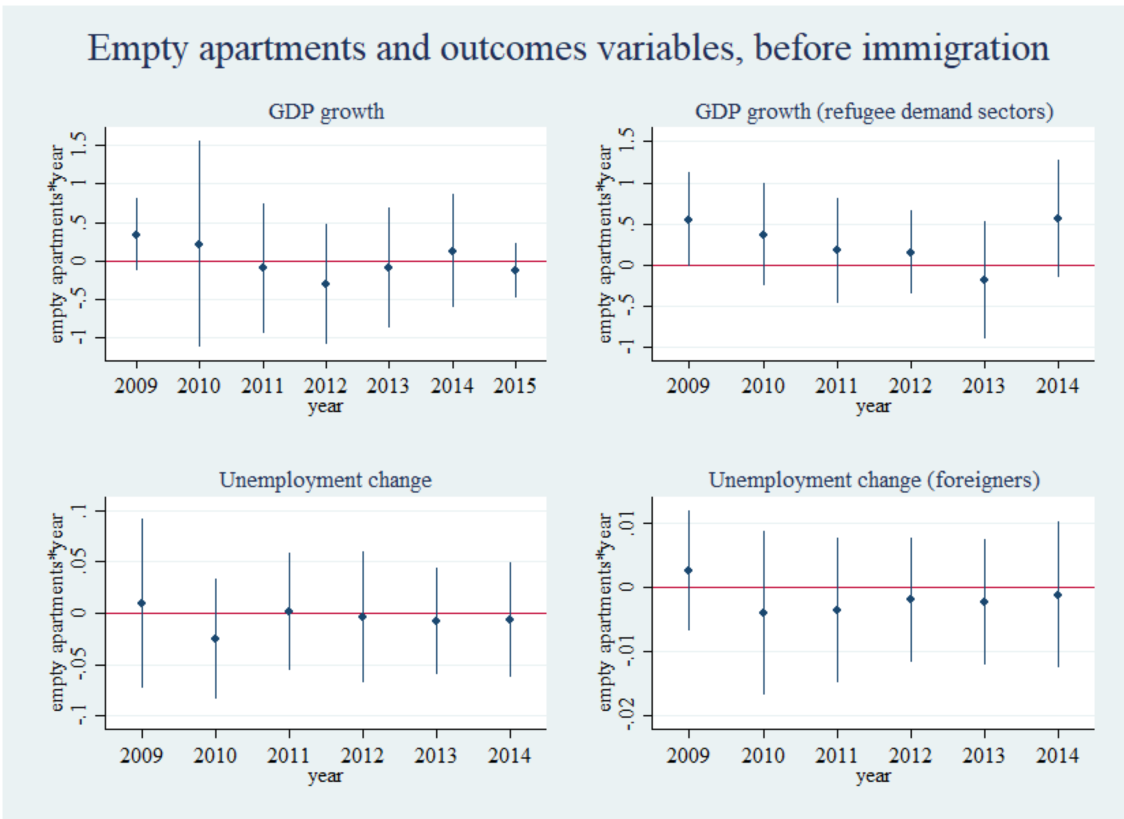
Figure 2.6.4:

Figure 2.6.5:

reasons detailed in the background section), when we assign labor market regions to treatment groups, we have to do so using their value of key_i and $eapp_i$ relative to other regions in the same state. This allows us to construct e.g. treatment and control groups which have higher or lower values on the instrumental variable, relative to other labor market regions in their state, or also to look at more quantiles of the distribution.

For the key difference instrument, I define all regions as treated which do not have a negative key difference. A regression of immigration per person on this dummy (which is one only in the treated regions and during the crisis years), shows a strongly positive and significant impact of treatment on refugee immigration. For the empty apartments variable, those regions are defined as treated which are above the within-state median of empty apartments. Here again, the effect on refugee immigration is strongly significant. Both regressions are shown in table 2.6.1, as well as a regression with both dummies. They can serve as an alternative first stage specification.

In the second column, $eapp_i$ is transformed into a categorical variable indicating the quintiles. It is interacted with the same time dummy. The results do not quite show the expected monotonicity; in fact, the first and second quintiles have indistinguishable coefficients (the coefficient of the first quintile in dropped in the regression), and the



Empty apartments and outcomes variables, before immigration
GDP growth
GDP growth (refugee demand sectors)
Unemployment change
Unemployment change (foreigners)
empty apartments*year
year

Figure 2.6.6:

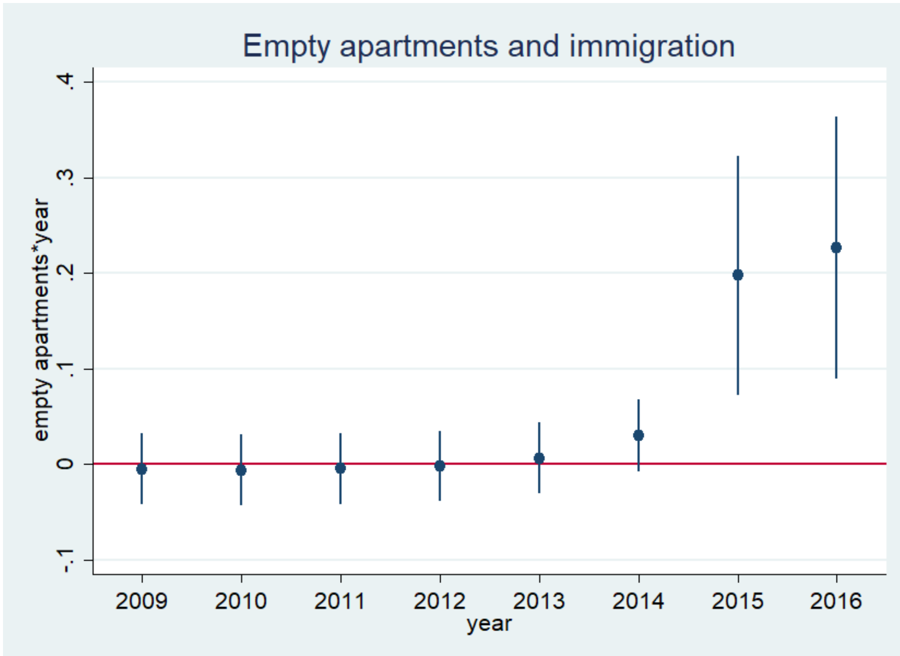


Figure 2.6.7:

Table 2.6.1: First stage: distribution keys, empty apartments and immigration - discretization

	Immigrants per person			
Crisis × distribution key dummy	0.16***		0.14**	0.11**
	(0.0)		(0.1)	(0.0)
Crisis × empty apmnts. dummy		0.062**	0.057**	0.046*
		(0.0)	(0.0)	(0.0)
Region FE	Yes	Yes	Yes	Yes
Region × time trends				Yes
State × year FE	Yes	Yes	Yes	Yes
r2	0.75	0.75	0.75	0.78
N	2303	2303	2303	2303

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***.

coefficients on the fourth and fifth quintiles are equally very similar. They are however positive and (weakly) significant for the upper quintiles.

Table 2.6.2 shows the results of the full IV regression where the sample is split into treatment and control groups. The first stage regression are the ones 2.6.1, otherwise the regressions are the same as the ones specified in equation 2.4, that is, our main regressions showing the effect of immigration on German employment. As you can see, the results are very similar to the ones obtained from our main specifications, but less significant.

Table 2.6.2: Refugee immigration and employment change in percent, IV - discretized treatment

	(1) emp. change	(2) emp. change	(3) emp. change
Immigrants (per hundred inhabitants)	1.4	10.6	7.8*
	(2.8)	(6.6)	(4.3)
Region FE	Yes	Yes	Yes
Region $\times$ time trends	Yes	Yes	Yes
State $\times$ year FE	Yes	Yes	Yes
N	2290	2290	2290

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***.

2.7 Conclusion and discussion

This paper shows that the refugee immigration wave of 2014-16 has so far had a significant positive impact on overall native employment. Especially employment in certain sectors benefited from the demand generated by the refugees and by the associated government spending.

While this is a positive tentative conclusion with respect to the welfare of German nationals, it is probably too early to make any final judgment about the eventual effects. Integration of the new immigrants will require that many more of them do find employment. A recent working paper (Evans and Fitzgerald, 2017) suggests that in the United States, refugees integrated relatively well – young refugees show little difference in achievement with their native peers, while older ones found employment rather slowly but ultimately at a high percentage. If this comes to pass, the patterns of substitution or complementarity between native and immigrant labor may turn out to be very different. Better data will also allow us to study the impact in more detail, and to pick up weaker effects with better precision. The IAB will eventually make updated individual level panel data available, including nationality, employer IDs, wages and employment status.

The consequences of the recent immigration wave in Germany can only be fully assessed in a few years time. In the meantime, this paper highlights the short-term impact of this important and controversial immigration.

Bibliography

Arbeitsmarkt (19. Dezember 2016). Arbeitsmarkt: Flüchtlinge finden kaum arbeit. ZEIT ONLINE, http://www.zeit.de/gesellschaft/2016-12/wenige-fluechtlinge-arbeitsmarkt-jobs-andrea-nahles.

Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *The quarterly journal of economics*, 118(4):1335–1374.

Borjas, G. J. and Monras, J. (2017). The labour market consequences of refugee supply shocks. *Economic Policy*, 32(91):361–413.

Card, D. (1990). The impact of the mariel boatlift on the miami labor market. *Industrial & Labor Relations Review*, 43(2):245–257.

Card, D. (2001). Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics*, 19(1):22–64.

Cengiz, D. and Tekguc, H. (2017). Is it merely a labor supply shock? impacts of syrian migrants on local economies in turkey. In *ASSA 2018 Annual Meeting*, pages 1–68.

Ceritoglu, E., Yunculer, H. B. G., Torun, H., and Tumen, S. (2017). The impact of syrian refugees on natives' labor market outcomes in turkey: evidence from a quasi-experimental design. *IZA Journal of Labor Policy*, 6(1):5.

Christian, P., Barrett, C. B., et al. (2017). Revisiting the effect of food aid on conflict: a methodological caution. Technical report, The World Bank.

Clemens, M. A. and Hunt, J. (2017). The labor market effects of refugee waves: Reconciling conflicting results. Technical report, National Bureau of Economic Research.

Dustmann, C., Schönberg, U., and Stuhler, J. (2017). Labor supply shocks, native wages, and the adjustment of local employment. *The Quarterly Journal of Economics*, 132(1):435–483.

Evans, W. N. and Fitzgerald, D. (2017). The economic and social outcomes of refugees in the united states: Evidence from the acs. Technical report, National Bureau of Economic Research.

Foged, M. and Peri, G. (2016). Immigrants' effect on native workers: New analysis on longitudinal data. *American Economic Journal: Applied Economics*, 8(2):1–34.

Gehrsitz, M. and Ungerer, M. (2017). Jobs, crime, and votes: a short-run evaluation of the refugee crisis in germany.

Geis, W. and Orth, A. K. (2016). Flüchtlinge regional besser verteilen. *Ausgangslage und Ansatzpunkte für einen neuen Verteilungsmechanismus. Gutachten für die Robert Bosch Stiftung. Hg. v. Robert Bosch Stiftung. Köln. Online verfügbar unter http://www. boschstiftung. de/content/language1/html/publikationen. asp*.

Glitz, A. (2012). The labor market impact of immigration: A quasi-experiment exploiting immigrant location rules in germany. *Journal of Labor Economics*, 30(1):175–213.

Harmon, N. A. (2017). Immigration, ethnic diversity, and political outcomes: Evidence from denmark. *The Scandinavian Journal of Economics*.

Hong, G. and McLaren, J. (2015). Are immigrants a shot in the arm for the local economy? Technical report, National Bureau of Economic Research.

Jaeger, D. A., Ruist, J., and Stuhler, J. (2018). Shift-share instruments and the impact of immigration. Technical report, National Bureau of Economic Research.

Nunn, N. and Qian, N. (2014). Us food aid and civil conflict. *The American Economic Review*, 104(6):1630–1666.

Ottaviano, G. I. and Peri, G. (2012). Rethinking the effect of immigration on wages. *Journal of the European economic association*, 10(1):152–197.

Otto, A. H. and Steinhardt, M. F. (2014). Immigration and election outcomes—evidence from city districts in hamburg. *Regional Science and Urban Economics*, 45:67–79.

Peri, G. and Sparber, C. (2009). Task specialization, immigration, and wages. *American Economic Journal: Applied Economics*, 1(3):135–169.

Raumabgrenzungen (2015). Laufende raumbeobachtung - raumabgrenzungen. Bundesamt fü Bauwesen und Raumordnung, http://www.bbsr.bund.de/BBSR/DE/Raumbeobachtung/Raumabgrenzungen/AMR/Arbeitsmarktregionen.html?nn=443270.

Steinmayr, A. (2016). Exposure to refugees and voting for the far-right:(unexpected) results from austria. *IZA DP No. 9790*.

Appendix

2.A Network instrument

2.A.1 Size of pre-existing community (network), immigration, and other economic trends

An alternative first-stage regression could have the following specification:

$$I_{i,t} = \alpha\, N_{i,t-1} + \Gamma X_{it} + FE_i + FE_{t,r} + \epsilon_{i,t} \tag{2.16}$$

where $N_{i,t-1}$ is the network size, he number of inhabitants of the same nationality as the immigrants who already lived in region i at a previous point in time. As discussed, a variation of this instrument has been used extensively in the literature. The network size does indeed influence migration in our current application, before the refugee crisis as well as during (see table 2.A.1).

However, we have to investigate if this instrument is correlated with other economic outcomes before the immigration wave. Regressions of the form

$$Y_{i,t} = \alpha_t\, (N_{i,t-1} \times \mathbf{YD}_t) + \Gamma X_{it} + FE_i + FE_{t,r} + \epsilon_{i,t} \tag{2.17}$$

where $Y_{i,t}$ are outcome variables such as GDP growth or the change in the unemployment rate, and $\mathbf{YD}_t$ is a set of year dummies, reveal that regions with a large network cannot be said to follow the same trends in economic development before the refugee immigration, as too many of the interaction coefficients are significantly different from zero (see table 2.A.2).

Table 2.A.1: Pre-existing networks and immigration

	Immigration
Network size (3 origin countries)	0.644***
	(0.200)
Network size × crisis years (2014-16)	0.222*
	(0.116)
Region FE	Yes
State × year FE	Yes
r2	0.809
N	2313

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***.

Table 2.A.2: Correlations of immigrant networks with economic trends, pre-refugee crisis

Outcome variable	Unemployment change	GDP growth
Network size × 2009	0.161	-5.052*
	(0.198)	(2.823)
Network size × 2010	0.289	-8.809***
	(0.180)	(2.359)
Network size × 2011	0.242	-1.549
	(0.157)	(2.842)
Network size × 2012	0.00898	-2.196
	(0.124)	(2.201)
Network size × 2013	0.297**	-2.480
	(0.116)	(1.568)
Network size × 2014	0.189**	-1.318
	(0.081)	(1.268)
Region FE	Yes	Yes
State × year FE	Yes	Yes
r2	0.622	0.546
N	2031	2016

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***.

2.A.2 Size of pre-existing community (network) and the composition of the new immigrant population

Since I am not using individual-level data, I can not investigate selection effects in detail. However, I do have data on the total population by gender and by type of residence permit. This allows me to investigate if the network instrument or the instrument I use in this paper - vacant apartments per person multiplied with the total number of immigrants to Germany - is correlated with the share of immigrants who are female or with the share who receive a particular residence permit.

There is no effect of either instrument on the share who are female, but the share who receive the type of residence permit which is most secure and most tightly linked to the asylum policy (permit for reasons of international law, humanitarian or political considerations) is slightly higher where the network is stronger. Other possible residence permits are permits for educational, work or family reasons as well as temporary suspensions of deportation and others (which are rarely issued). This breakdown is difficult to use because different regions have different policies with respect to the permits they issue, even when faced with the same cases of asylum seekers. [10]

These effects can be seen from table 2.A.3. The effect is not strong and it does not prove that selection is an important issue that would invalidate the network instrument, given that the breakdown into different permits is probably quite unreliable. It could also be that the network helps immigrants obtain the more secure type of permit, even when the composition of immigrants in a region with a strong network would be exactly the same as in other regions. This would however point to another problem: the network would also help immigrants in other ways, such as finding employment, as this paper and others argue. The native population would then be impacted by the network size during an immigration wave, even if the primary effect of the network size on immigration were absent. For all of these reasons I am using a different instrument in this paper, which the partially managed nature of a refugee wave permits.

[10] I use the increase in total population rather than in population with a specific permit in my main analysis, because of these problems and because I am interested in the effect of overall immigration.

Table 2.A.3: Network size and the composition of the immigrant population

Outcome variable	Female	Female	Ref. permit	Ref. permit
Network size	0.001		0.04*	
	(0.001)		(0.02)	
Empty apmnts.		0.001		-0.002
		(0.001)		(0.003)
Region FE	Yes	Yes	Yes	Yes
State $\times$ year FE	Yes	Yes	Yes	Yes
r2	0.176	0.176	0.166	0.165
N	2243	2243	1990	1990

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***.

Chapter 3

Reputational Incentives under Heterogeneous Demand Fluctuations

with Milena Petrova

Using data from an online home services marketplace, we study reputational incentives in an environment with heterogeneous seller demand fluctuations. Our strategy exploits the fact that some professionals work in highly seasonal occupations, while others do not. We show that sellers facing an upcoming demand downturn become less likely to receive a positive review, and we argue that this is because they value their reputations less and exert lower effort. Robustness checks rule out several other explanations for this phenomenon, e.g. reverse causality, selection of sellers into jobs or differences in the cost of seller effort. Neither the economic literature nor the platforms relying on reputation systems have sufficiently considered seller heterogeneity and the optimal way to account for it in the information given to the buyer and in seller incentives.

3.1 Introduction

Online platforms have been remarkably successful in generating activity in many diverse markets such as the sale and re-sale of various new and used goods (eBay, Amazon) and services such as short-term rentals of cars and property (Uber, AirBnB). The apparent issues of anonymity and lack of trust had made many doubtful of their viability in the beginning. The sellers and buyers are typically separated geographically, and the interaction is one-time only.

To a large extent, the success of online platforms can be credited to their employment of *reputation systems* (or *reputation mechanisms*). These are a set of rules by which buyer-seller interactions are reviewed publicly. Even if individual buyers interact only once with the seller, the mechanism establishes a dependence between current performance and future demand. By subscribing to a marketplace with a public reputation system, the seller agrees that the outcome of all interactions is made public, which in turn informs the decision of future buyers. Knowing this, current buyers are less hesitant to hire them.

Despite the ubiquity of reputation systems, the approach to their design is often "one size fits all", with limited consideration of seller heterogeneity and the economic environment. The predominant way in which individual reviews are aggregated and displayed is by taking a simple average, thus making each review equally improtant. The biggest review systems for restaurants - Tripadvisor, Yelp and GoogleMaps - use this simple approach. Amazon is one of the few online retailers that recognize the need for a more fine-tuned reputation system: it gives higher visibility of reviews received in the last 12 months, to prevent sellers from lowering quality when they have already accumulated a large amount of positive reviews. In their working paper on the economics and the tech industry, Athey and Luca (2018) note that platforms have to consider more carefully the way in which review information is made available to buyers, and how this affects seller incentives. Economists have only recently started to addressed optimality of reputation systems in specific economic environments, and how this can be achieved by review aggregation and display (see Jin et al. (2018)).

In this paper, we provide evidence that when the economic environment and actors are not homogeneous, a simple reputation system fails to incentivize a consistent effort level. We consider an environment in which sellers are differently exposed to demand fluctuations. Because reputation works by linking current and future demand, the value placed on a good reputation by the seller is low when future transaction opportunities are few. This lowers the incentive to exert effort, which in turn leads to a lower probability that the buyer is satisfied by a project's outcome. Consistent with the hypothesis, our empirical results indicate that sellers who can expect a more significant demand downturn in the upcoming months are less likely to receive a positive review.

We work with proprietary data from MaistorPlus, a Bulgarian online marketplace for home services. MaistorPlus is similar to other well-known platforms connecting clients, or *buyers*, and professionals, or *sellers*, such as oDesk (IT and business professionals), TaskRabbit (for home owners and low skilled labor), and Thumbtack (for a large variety of local services). The seller's account has information on his scope of expertise, and it contains a public record of his activity on the platform: the number of times he was hired and the corresponding average of the received reviews.[1]

[1]We refer to the individual interaction feedback as a *review*, while the cumulative of the reviews is the seller's *reputation*. In this sense, the review is an incremental change in reputation.

This information is an important determinant of their chances to be contacted and eventually hired by buyers. Our empirical analysis demonstrates the positive returns to reputation, especially when demand is high.

The source of variation in our data comes from the seasonality of demand for certain jobs, which is due to weather conditions permitting outdoor work. The majority of jobs are posted between July and November, which we call the high demand season. The sellers themselves specialize in different services, meaning that they are heterogeneously exposed to the demand seasonality. For example, someone specializing in plumbing services would see a more steady flow of demand compared to someone specializing in roof repairs, which mostly take place in the summer and autumn. As a result, the seasonal fluctuation of demand, which translates to a reputational incentive to exert effort, varies at the level of the individual seller. The analysis focuses on the seller' behavior at the end of the high demand season, when the discounted sum of work opportunities in the immediate future is the lowest.

Our measure of demand heterogeneity is seller demand *Smoothness*: the average monthly demand in the low period relative to the average monthly demand in the high period. A smoothness measure closer to one implies lower exposure to seasonality (most sellers are affected at least to some extent). We construct this variable using overall demand in the categories in which the seller is active, rather than the sellers' individual activity. *Smoothness* therefore measures exogenous demand conditions rather than the sellers' endogenous choices, so that our results are not confounded by reverse causality (after a positive review, the sellers may increase their activity).

The end of the high demand season is associated with a lower likelihood of a positive review across all sellers. This may be due to reasons other than lower effort, as the end of the high demand season is potentially associated with disruption in material supplies, bad weather or last-minute projects that are harder to perform well. By interacting seller-specific demand *Smoothness* and the time variable *High season end*, we are sure to find an effect which is driven by heterogeneous demand conditions at the level of the individual seller. We find that the sellers whose demand drops substantially in the upcoming low demand season are less likely to receive a positive review. The average seller in our sample has a 4 percent lower probability of a positive review in the last months of the high season. For sellers with low demand smoothness at the 10th percentile of the distribution, we see a drop of 16 percent, while sellers at the 90th percentile see an increase of 18 percent in the probability of a positive review. [2]

Our hypothesis assumes that the channel through which future demand affects the outcome is seller effort, via the value of reputation. Since we use fixed effects on the level of the individual seller, we can rule out that our results are driven by selection of professionals. We perform a number of robustness checks to rule out alternative explanations. In the first such test, we include time-varying dummies for categories of work that are highly seasonal. This addresses the concern that these categories of work may be less likely to receive positive reviews at the end of the season, and be at the same more sought out by low *smoothness* professionals. We also rule out the possibility that sellers who drop out of the platform at the end of the high season drive our results. Additionally, we propose two specifications that test indirectly whether the cost of effort, as a function of demand or of the size of the seller, confounds the results.

[2] The results in the reverse direction are weekly significant in some specifications: the end of the low demand season is associated with a higher likelihood of a positive review, but less so for sellers with more seasonal demand.

Given the current reputation mechanism, our results can be used to inform buyers better about seller incentives. Consider roof repairs, a highly seasonal job due to weather conditions. Towards the end of the high demand season, a homeowner would be better off with a seller specializing in plumbing and fixing roofs on the side (high *Smoothness*), rather than with the roof specialist (low *Smoothness*), all else equal. Equally, a homeowner looking for a roof specialist early in the season should not be as concerned by a negative review that was received in November as much as by one received in June. The results suggest that design the system itself may be optimized in a way that provides more consistent incentives for seller effort. For example, the reputation system can aggregate seller reviews through a weighted average that puts higher weights on reviews received under circumstances similar to the ones facing the particular seller.

Our work is related to a recent literature which uses data from online marketplaces to test formal reputation models, to study reputation dynamics, and to evaluate the effectiveness of reputation systems. The surveys of Bajari and Hortacsu (2004), Bar-Isaac and Tadelis (2008), Cabral (2012) and most recently Tadelis (2016) document this body of work. A substantial number of papers in the literature attempt to test whether the source of informational asymmetry in the market is due to moral hazard or adverse selection. For example, Cabral and Hortacsu (2010) examine the dynamics of seller reputation on eBay and find strong evidence for moral hazard: the first negative feedback received lowers the returns to reputation once and for all, and sellers react by lowering their effort and eventually exiting. However, the authors cannot rule out an underlying moral hazard plus adverse selection model. While the objective of our paper is not to rule out adverse selection, it is seller effort, rather than seller type, which is affected by the demand fluctuation. In the robustness checks, we rule out two alternative explanations for the effect we find that are based on seller type (seller size and exiting sellers).

The effectiveness of reputation systems has raised considerable research interest, with the majority of work focusing on the ability of reputation mechanisms to elicit truthful reviews. For example, Fradkin et al. (2017) investigate the informativeness of reviews on a hospitality marketplace, AirBnB, where review reciprocity causes significant bias. Horton and Golden (2015) also document reputation inflation, on oDesk, a service marketplace. Only recently have economists considered more explicitly how seller reviews should be aggregated, with the objective of providing future buyers with the most up to date information and sellers with the right incentives. The work of Jin et al. (2018) considers the optimal design of a reputation system in an environment with changing service quality. Because quality is exogenous, the effect of the reputation mechanism design on seller incentives is not a part of their analysis. We are not aware of other empirical work studying the heterogeneity of reputational incentives across sellers.

Although their work concerns the credit rating industry, Bar-Isaac and Shapiro (2013) are closer to the topic of our work as they investigate how the exogenous variation of the business cycle, a common demand trend, drives reputational incentives for Credit Rating Agencies (CRAs). The value of reputation is shown to depend on the economic fundamentals varying over the business cycle. Ratings accuracy is counter-cyclical as effort is lower during periods of high economic activity, which is consistent with our findings that the probability of a positive review is influenced by demand fluctuations.

The paper proceeds as follows. Section 2 provides background on our data source, the MaistorPlus marketplace, including information on the categories of jobs available and their seasonal variation, the sellers and the reputation mechanism. Section 3 presents the empirical implementation, as well as the results and robustness checks. Section 4 concludes.

3.2 MaistorPlus and data generation

We work with company data form the MaistorPlus marketplace, founded in Sofia, Bulgaria in 2012, and our sample covers the period between January 2013 and July 2015.[3] The marketplace connects buyers to subscribing home service sellers and is financed by sellers' subscription fees and by advertisements.

3.2.1 Job activity and reviews

Demand on the marketplace is generated by buyers who sign up freely and post what we call *jobs*: home repair projects for which they want to hire a service provider. We observe 4,167 jobs, as well as data at the job-seller level of interaction. The total value of all jobs in our sample is 12.6 million Euros.

The buyer provides a description of the job, specifies the job category (such as carpentry, roof repairs, construction, etc), an estimated budget, and a start date. All sellers active in that job category are potentially available and notified of the job; those who are actually available can message the buyer. The buyer is free to inspect the seller profiles and their reputation when he decides who to contact or hire. In the data, the buyers contact

[3]The website of the market place is: http://maistorplus.com/

1.3 sellers on average, and hire someone for 27% of posted jobs (see Table 3.2.1).

Table 3.2.1: Job activity on MaistorPlus

Data	Observations	Per job
Distinct jobs	4,167	-
- potentially available sellers	333,639	80.0
- available sellers	22,379	5.37
- contacted sellers	5,354	1.28
- hired sellers	1,126	0.27
- positive reviews (total and per hire)	749	0.66
- total value (Euro)	12,650,600	2,785

Only buyers who have hired a seller through the marketplace are allowed to post a review, which can be positive, negative, or neutral, or missing (1/0/-1/.). As it is customary in the empirical reputations literature, we separate the reviews into positive and non-positive reviews.[4] Out of 1,126 jobs where someone was hired in our sample, we observe that the buyer left a positive review in 749 cases.[5]

For buyer reviews to carry information, it is necessary that the good or service in question can be competently evaluated by the buyer. Services performed by expert professionals, such as home repairs, can have aspects of credence goods, as the buyer only observes whether the immediate problem is fixed, but not the quality of the work or whether problems were over-diagnosed. Seller professionalism, as well as explanations of how the problem arose and how to present future problems, are therefore highly

[4]Consistent with the previous literature (e.g. Nosko and Tadelis (2015) and Dellarocas and Wood (2008)), we view no feedback and neutral feedback as a non-positive review.

[5]The buyers can also post a textual review of the interaction. However, we do not work with this metric because it is highly collinear with the point review.

valued by the buyers and the predominant themes of the textual part of the reviews. Providing this observable aspect of the service is still costly and affected by fluctuations in the value of reputation.

3.2.2 Service sellers

There is a total of 864 active sellers in our sample, of whom 251 were hired at least once. They subscribe to one of three fixed-term plans, ordered by their annual cost: Start (100 Euro), Pro (150 Euro) and Pro+ (250 Euro). The more costly the plan, the more options the sellers have for the number of categories in which they are active (3/4/6 categories). The sellers can change these at any time and at no cost, potentially adjusting their activity to seasonal demand fluctuations. While the buyers can infer, to some extent, the seller size and scope of activity by observing the current active categories, they do not observe history of category changes, and therefore seller's demand smoothness.

Table 3.2.2 presents summary statistics on the experience of the sellers present in the main analysis: the sellers who were hired at least once in the duration of our sample. The average seller is hired 4.5 times and has 3 positive reviews, both with high standard deviation. The average time between the last two messages of interest is about 9 days, but some sellers appear to have not been available much longer. The average activity span (time between the first and last message observed in the data) is 19 months out of 38 months in the data. Periods of high inactivity are likely for sellers who work in seasonal job categories, and among those who use the marketplace relatively rarely because they have sufficient job referrals from outside of the marketplace.

Table 3.2.2: The experience of the MaistorPlus sellers on the marketplace.

At latest observed time period:	Mean	St. dev	Min	Max
- activity span on marketplace (months)	19.19	8.56	0	30
- time between last two messages (days)	8.65	39.5	0	362
- total times hired	4.48	7.02	1	49
- total positive reviews	2.98	5.17	0	38

3.2.3 Seasonality and reputation

Demand in the home services sector is highly seasonal because some types of services, for example outdoor work, can only be done during the months when the weather permits. There are also services that are demanded right before the start of fall and winter, such as services related to heating and insulation. Examples of demand seasonality in the different job categories can be found in Figure 1.

To establish the duration of the high demand season, we look at the number of jobs posted on the platform each month during the period 2013-2015, *N. jobs*. Table 3.2.3 has the results from two regressions that help us do that. In the first regression $S1$, we use indicators for the individual months. Months July through November experience a significantly higher level of demand compared to other months. We define a seasonal variable *High season* equal to 1 for months months July-November and use that in the second regression $S2$. The explanatory power of $S2$ is still very high, and in fact the adjusted R^2 is higher compared to $S1$.

Implicit in our hypothesis that an upcoming, temporary demand downturn lowers effort are the assumptions that reputation is more valuable when demand is high, and that professionals discount demand in the more distant future strongly enough. In the appendix 3.A, we present a small

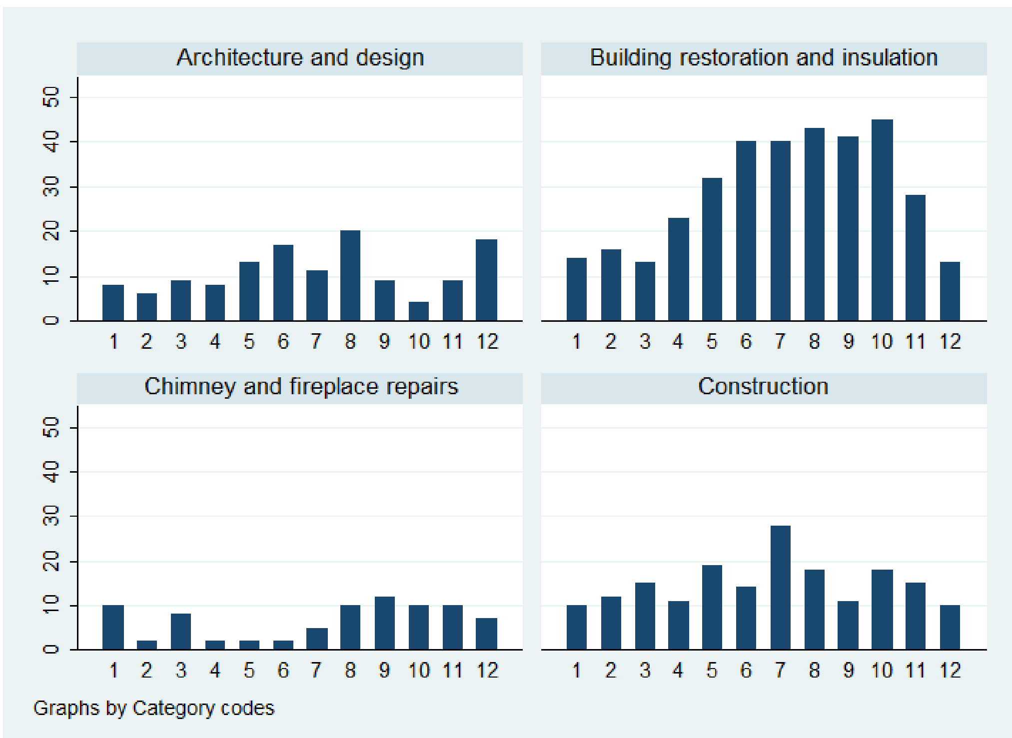


Figure 3.2.1: Number of jobs posted by both in different categories, year 2014

Table 3.2.3: Monthly demand on the market-place and the *High season* dummy.

Dependent variable	N. jobs	N. jobs
Specification	(S1)	(S2)
Independent variables		
February	-0.061	
	(0.311)	
March	0.236	
	(0.316)	
April	0.277	
	(0.297)	
May	0.319	
	(0.403)	
June	0.410	
	(0.426)	
July	0.988***	
	(0.314)	
August	1.016***	
	(0.338)	
September	0.945***	
	(0.317)	
October	0.828**	
	(0.305)	
November	0.801**	
	(0.323)	
December	0.250	
	(0.322)	
High season		0.711***
		(0.090)
Constant	4.410***	4.615***
	(0.299)	(0.094)
Year fixed effects	Yes	Yes
R^2	0.88	0.85
R^2 Adjusted	0.81	0.83
N. observations	36	36

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***.
Robust standard errors.

theoretical model making these assumptions in a more formal way, and we argue for their validity here.

Ex-ante, it is not obvious whether reputation is more valuable during high or low demand periods. On the one hand, the value of reputation could be higher in the high demand period because there are more buyers with a high willingness to pay for a seller with a high reputation. On the other hand, reputation could be more valuable in the low demand season when there are fewer jobs and more intense competition between the sellers.

We resolve this question by going to the data. We construct a panel at the professional-date level and a variable measuring the unconditional returns to reputation as the number of times a seller was contacted by buyers in a given month, *Monthly contacts*.[6] The variables of interest in specification $C1$ are our measure of reputation, *Percent positive reviews*, lagged one month, and the *High season* indicator. Their interaction is included in specification $C2$. Other control variables include the seller's activity on the platform during that time (how many jobs he was invited to, how many times he was available), and date and seller fixed effects. The estimates are presented in Table 3.2.4. The benefits of a good reputation are indeed positive: the number of monthly contacts increases with *Percent positive reviews*. The results of the second regression support the claim that reputation is more valuable when demand is high as the coefficient on the interaction between *Percent positive reviews* and *High season* is positive and significant. If *Percent positive reviews* increases by 20 percentage points, *Monthly contacts* in the *High demand* period increase by 1.3 percent.

Table 3.2.2, presenting summary statistics on the usage and experience

[6]The hiring decision is joint and depends on other potential projects that the seller may be considering. Therefore, we prefer to use as dependent variable the number of times the buyers have expressed preference for hiring the seller, i.e. *Monthly contacts*.

Table 3.2.4: Returns to reputation and the *High season*.

Dependent variable	Monthly contacts	Monthly contacts
Specification	(C1)	(C2)
Independent variables		
Percent positive reviews	0.038*	0.016
	(0.020)	(0.021)
High season	0.115***	0.107***
	(0.018)	(0.018)
High season*Percent positive reviews		0.065***
		(0.014)
Messages of availability	0.422***	0.422***
	(0.007)	(0.007)
Constant	-0.071	-0.069
	(0.051)	(0.051)
Fixed effects		
Date	Yes	Yes
Seller	Yes	Yes
R^2	0.77	0.77
N. observations	16,912	16,912

Significant at: p < 0.1: *; p < 0.05: **; p < 0.01: ***. Robust standard errors. The continuous dependent variables and regressors (except *Percent positive reviews*) are transformed by taking their natural logarithm, and their coefficients can be interpreted as elasticities. The coefficient on *Percent positive reviews* is a semi-elasticity.

of professionals on MaistorPlus, shows why it is credible that the temporary seasonality can affect a professional's valuation of their reputation to a measurable degree. The platform is young, and has low overall demand. The average professional is active on the platform for a year and eight months, and has been hired 4.5 times. This shows that, clearly, MaistorPlus is not the main source of jobs for the professionals, and they may be quite uncertain about platform activity – both their own and demand – even a few months into the future. It is this uncertainty about platform use, rather than time discounting of future income, that makes a strong discount factor plausible.

3.3 Econometric analysis

Our econometric analysis is based on the difference in differences methodology. The results indicate that indeed sellers who experience a drop in demand are less likely to receive a positive review towards the end of the high season. We also see a weekly significant result in the reverse direction: sellers who experience an increase in demand are more likely to receive a positive review towards the end of the low demand season. We perform a number of robustness checks that corroborate this conclusion.

3.3.1 Demand smoothness

Central to empirical analysis is the variable measuring how the seller's heterogeneous demand fluctuates, $Smoothness_i$. Let k denote the individual job categories and i denote the sellers. The data allows us to observe the set $\{K_i\}$ of categories that seller i is available in. For example, *Construction* is one such category for seller i if the seller has expressed availability to a job in that category. Let $Demand_{k,Season}$ denote the number of unique jobs in category k during the respective season.[7] As the platform is still growing during the period of observation, we group two *Low* and two *High* seasons to improve the representativeness of the demand pattern. Assuming that demand grows multiplicatively, our definition of $Smoothness_i$ is not distorted by this. For the demand in the *High season*, we use the jobs posted on the marketplace during July 2014-Nov 2014 and July 2015-Nov 2015. For demand during the *Low season*, we use jobs during Dec 2013-June 2014 and Dec 2014 - June 2015.

For the seller i, the monthly demand during any given season, $Demand_{i,Season}$, is the sum of demand over his active categories $\{K_i\}$ in that season, $Demand_{k,Season}$, divided by the duration of the season.[8] The seller's demand $Smoothness_i$ is the ratio of of demand in the low season to demand in the high season. Constructed this way, the demand measure is exogenous to the seller's actual activity and reputation. In other words, $Smoothness_i$ is defined as:

$$Smoothness_i = \frac{Demand_{i,Low}}{Demand_{i,High}} = \frac{\frac{1}{5}\sum_{k \in K_i} Demand_{k,Low}}{\frac{1}{7}\sum_{k \in K_i} Demand_{k,High}}$$

A demand $Smoothness_i$ value of 1 would indicate no fluctuation: seller i faces the same demand conditions during the $Lowseason_t$ and the $Highseason_t$. The lower is the value of $Smoothness_i$, the more the seller is exposed to the seasonality of demand. Most sellers experience a drop in demand during the low season, and the average value of smoothness is 0.59. A few sellers

[7] Our data rangers from January 2013 to December 2015 but there was a significant change in the marketplace rules in July 2015. The change affects the activity of the sellers, hence we cannot use data after it for the main analysis. However, it does not affect demand incidence.

[8] The low and high demand seasons have different duration, 5 and 7 months respectively. Dividing the total demand by the season length is necessary to avoid a mechanical difference due to the different length of the seasons.

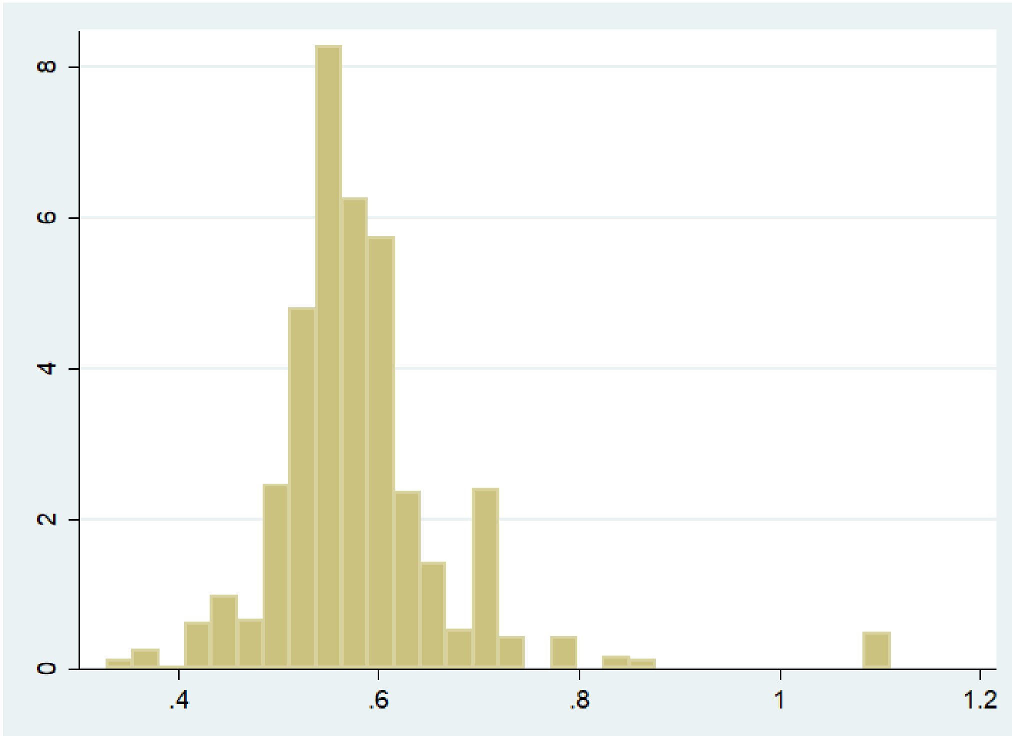


Figure 3.3.1: Histogram of $Smoothness_i$ for all sellers in the sample.

experience reverse seasonality with $Smoothness_i$ maximum of 1.59. Figure 2 contains the histogram of $Smoothness_i$. The distribution is can be described as relatively bell-shaped with a longer tail on the right.

3.3.2 Main specification and results

We test the hypothesis that seller demand heterogeneity affects seller incentives to exert effort, which in turn affects the probability that a job is done well and the seller receiving a positive review. This effect should be especially strong at the end of the high demand season, when the discounted sum of future demand is the lowest.

To this end, we work with a sample of completed jobs, indexed by the hired seller i, the date defined as month-year t, and the job characteristics k.

The outcome variable of interest is the probability that the seller i receives a positive review for job k in period t, $Pr(Positive\ review)_{ikt}$. The variable $High\ season\ end_t$ which indicates if the job was posted in the two months of the high demand season (October and November) and it is the common trend among all sellers. We interact it with seller $Smoothness_i$ to test whether sellers heterogeneous demand fluctuation affect the probability of $Pr(Positive\ review)_{ikt}$.

Specification (1) is:

$$\begin{aligned} Pr(Positive\ review)_{ikt} = \quad & \beta_1 High\ season\ end_t + \beta_2 Smoothness_i * High\ season\ end_t \\ & + \gamma X_{it} + FE_i + FE_k + FE_t + \epsilon_{ikt} \end{aligned}$$

The coefficient of interest is β_2. If it is positive, sellers with lower demand fluctuations (a higher $Smoothness_i$) are more likely to receive a positive review at the end of the high season. Or, the more the seller is exposed to seasonal demand fluctuations, the less likely he is to receive a positive review at the end of the high season.

As controls, we include covariates measuring seller experience and a number of fixed effects. The variable X_{it} is the *Percent positive reviews*. We also include fixed effects for each seller i, for the job category, budget and start date of the job k, and for the time period t. Since $Smoothness_i$ is constant over the sellers, it is absorbed by the seller fixed effects.

In specification (2), we allow for the complementary effect by including $Low\ season\ end_t$ dummy equal to 1 for the months May and June and its interaction with $Smoothness_i$. We expect coefficient β_4 to be negative, suggesting sellers who do not experience significant demand fluctuations are is less likely to make exceptional effort to improve their reputation before the beginning of the high season.

Specification (2) is:

$$
\begin{aligned}
Pr(Positive\ review)_{ikt} = \quad & \beta_1 High\ season\ end_t + \beta_2 Smoothness_i * High\ season\ end_t \\
+ \quad & \beta_3 Low\ season\ end_t + \beta_4 Smoothness_i * Low\ season\ end_t \\
& + \gamma X_{it} + FE_i + FE_k + FE_t + \epsilon_{ikt}
\end{aligned}
$$

We employ a linear probability model to make the interpretation of the estimated effects easier. Our estimates of (1) and (2) can be found in Table 3.3.1.[9] We find that the *High season* end_t is associated with a general drop in the probability of a positive review. This effect is countered by demand $Smoothness_i$, as the estimated coefficient on the interaction is positive. For the median seller with $Smoothness_i = 0.57$, the probability of receiving a positive review at the end of the high season is 4 percent lower: $-0.99 + 0.57 \times 1.66 = -0.04$. sellers at the 10th percentile see a drop of 16 percent, while those at the 90th percentile actually see an increase of 18 percent in the probability of positive review at the end of the high season.

The results of specification (2) are in line with specification (1), but the additional variables are not significant. It is possible that this is because of the small sample size. The *Low season* end_t is associated with a higher likelihood of a positive review, and sellers with higher demand $Smoothness_i$ are less likely to receive a positive review at the end of the low demand season.

Conditional on the seller being hired, we find that the variable *Percent positive reviews* is associated negatively with the probability of receiving another positive review. A potential explanation is that sellers are strategic in building and using up their reputation. Strategic reputation building is

[9] The results of probit and logit specifications are consistent and available upon request.

Table 3.3.1: Linear probability model for receiving a positive review.

Dependent variable	Positive review	Positive review
Specification	(1)	(2)
Independent variables		
High season end	-0.986**	-0.919**
	(0.439)	(0.446)
Smoothness*High season end	1.659**	1.532**
	(0.653)	(0.666)
Low season end		1.057
		(0.860)
Smoothness*Low season end		-2.319
		(1.448)
Percent positive reviews	-0.272***	-0.274***
	(0.065)	(0.064)
Constant	1.381***	1.396***
	(0.265)	(0.264)
Fixed effects		
Job category, budget, start	Yes, Yes, Yes	Yes, Yes, Yes
Date	Yes	Yes
Seller	Yes	Yes
R^2	0.41	0.41
N. observations	1,126	1,126

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***. Robust standard errors. The continuous regressors (except *Percent positive reviews*) are transformed by taking their natural logarithm, and their coefficients can be interpreted as elasticities. The coefficient on *Percent positive reviews* is a semi-elasticity.

indirect evidence of moral hazard, and moral hazard as predominant source of asymmetric information has been recorded in online environments such as eBay by Klein et al. (2016).

3.3.3 Robustness checks

We consider the robustness of our results to alternative channels of causality: selection between sellers and jobs, a pure end-game effect, and cost of effort.

Selection between sellers and jobs

As already mentioned in the **MaistorPlus and data generaton** section, the buyers do not observe the heterogeneous sellerl demand fluctuations, hence the opportunities for selection on the part of the buyers is limited. However, it is possible that sellers of different $Smoothness_i$ self-select into different jobs k, at different times t, and this interferes with our results.

Consider the indicator $Seasonal\ category_k$ equal to 1 for jobs with high demand incidence in June to November.[10] It may be that jobs in seasonal categories are more difficult to complete successfully. It is also possible that their successful completion is affected by the time period, for example: seasonal jobs may suffer from worsening weather conditions at the end of the high demand season. In addition, sellers in more seasonal lines of work (and therefore lower demand $Smoothness_i$) may be more likely to take on seasonal jobs, especially at the end of the high demand season. These effects present an alternative relationship between the probability of a positive review and the explanatory variables of interest: job

[10] To classify the jobs, we regress monthly job incidence on the $Highseason_t$ dummy. These regression are available upon request.

seasonality affects the effort level needed for a positive review, and low demand $Smoothness_i$ sellers select into seasonal jobs towards the end of the high demand season. To test this, we augment (1) by interacting the $High\ season\ end_t$ and $Smoothness_i$ variables individually and jointly with the $Seasonal\ category_i$ dummy.[11] Our first robustness specification ($R1$) is:

$$\begin{aligned} Pr(Positive\ review)_{ikt} = \quad & \beta_1 High\ season\ end_t + \beta_2 Seasonal\ category_k \\ & + \beta_3 Smoothness_i * High\ season\ end_t \\ & + \beta_4 Seasonal\ category_k * Smoothness_i \\ & + \beta_5 Seasonal\ category_k * Smoothness_i * High\ season\ end_t \\ & + \gamma X_{it} + FE_i + FE_k + FE_t + \epsilon_{ikt} \end{aligned}$$

The results of ($R1$) can be found in Table 3.3.2. Jobs in seasonal categories are not more or less likely to get a positive review, neither in general nor only at the end of the season. The probability of a positive review for a seasonal job is also not affected by the demand $Smoothness_i$ of the hired seller, neither overall nor at the end of the high demand season. The main effects of interest - the $Highseasonend_t$ and its interaction with $Smoothness_i$ - remain significant and of similar higher magnitude.

Leaving the platform

In the period for which we have data, the platform is relatively young and some sellers in our sample become inactive after a while. For these sellers, the incentive to exert effort in the last months of their subscription is even

[11] Individual category fixed effects are no longer included because of collinearity with the *Seasonal category* dummy.

Table 3.3.2: Robustness check with seasonal category fixed effect and interactions.

Dependent variable	Positive review	Positive review
Specification	(1)	(R1)
Independent variables		
High season end	-0.992**	-1.507**
	(0.413)	(0.480)
Seasonal category		-0.037
		(1.111)
Smoothness*High season end	1.626**	2.333***
	(0.683)	(0.421)
Seasonal category*High season end		0.859
		(1.033)
Seasonal category*Smoothness		-0.058
		(1.885)
Seasonal category*Smoothness*High season end		-1.310
		(1.646)
Percent positive reviews	-0.220***	-0.211***
	(0.063)	(0.061)
Total times hired	-0.030	-0.030
	(0.037)	(0.036)
Constant	1.360***	0.942
	(0.249)	(4.857)
Fixed effects		
Job category, budget, start	Yes, Yes, Yes	No, Yes, Yes
Date	Yes	Yes
Seller	Yes	Yes
R^2	0.41	0.39
N. observations	1,126	1,126

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***. Robust standard errors. The continuous regressors (except *Percent positive reviews*) are transformed by taking their natural logarithm, and their coefficients can be interpreted as elasticities. The coefficient on *Percent positive reviews* is a semi-elasticity.

lower, as they may be "cashing in" their reputation. If the end of their use of the platform coincides with the end of the high demand period, it may be that this is partially driving our results.

We restrict our sample to only those sellers who are still active on the platform in 2015 and 2016, which brings down our observations to 867 from 1,126. The results from estimating our model on the restricted sample, ($R2$), are presented in Table 3.3.3; all coefficients on the variables of interest are very similar to those in our main specification. This is not surprising, since the mean of demand $Smoothness_i$ of the two groups, the sellers who continue to be active and those who drop out, is not statistically different.

Cost of effort

A potential confounding factor in our results is the cost of effort, which may also change when demand is high, and differentially so for for sellers with different demand smoothness. With no information on seller costs, we are unable investigate this directly. Instead, we propose two regressions providing indirect evidence that the cost of effort is the driving factor of our results.

Firstly, we conjecture that the cost of effort fluctuates with current demand, and not with the discounted sum of demand in future periods. In ($R3$), we use the full set of bi-monthly indicators on their own and interacted with the *Smoothness* variable.

Table 3.3.3: Robustness check with sample restricted to sellers still active in 2015 and 2016.

Dependent variable	Positive review	Positive review
Specification	(1)	(R2)
Independent variables		
High season end	-0.992**	-1.035***
	(0.413)	(0.396)
Smoothness*High season end	1.626**	1.685***
	(0.681)	(0.649)
Percent positive reviews	-0.220***	-0.230***
	(0.062)	(0.067)
Total times hired	-0.029	-0.024
	(0.037)	(0.039)
Constant	1.360***	1.265***
	(0.249)	(0.245)
Fixed effects		
Job category, budget, start	Yes, Yes, Yes	Yes, Yes, Yes
Time (year)	Yes	Yes
seller	Yes	Yes
R^2	0.37	0.31
N. observations	1,126	867

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***. Robust standard errors. The continuous regressors (except *Percent positive reviews*) are transformed by taking their natural logarithm, and their coefficients can be interpreted as elasticities. The coefficient on *Percent positive reviews* is a semi-elasticity.

$$
\begin{aligned}
Pr(Positive\ review)_{ikt} = & \sum_{t=2}^{T=6} \beta_j Bimonthly\ dummy_t \\
& + \sum_{t=2}^{T=6} \alpha_j Smoothness_i * Bimonthly\ dummy_t \\
& + \gamma X_{it} + FE_i + FE_k + FE_t + \epsilon_{ikt}
\end{aligned}
$$

The estimation results of ($R3$) indicate no evidence to support a cost of effort based on the level of current demand. Indeed, the probability of positive review is significantly lower only in October and November. Furthermore, we see no differential effect for sellers of different demand $Smoothness_i$.

Another way to look at the issue of fluctuating costs is to consider the relationship between the size of the seller's operations and $Smoothness_i$. If sellers with lower $Smoothness_i$ are also smaller, they may be relatively more overstretched during the high demand season. In this sense, our estimations would be suffering from omitted variable bias, where $Smoothness_i$ picks up the effect of the omitted variable measuring seller cost of effort. To remedy this, we use the mean budget size for all jobs for which the seller was available, $Average\ budget_i$, as a metric for seller size, and interact it with the $High\ season\ end_t$ dummy.

$$
\begin{aligned}
Pr(Positive\ review)_{ikt} = & \ \beta_1 High\ season\ end_t + \beta_2 Seasonal\ category_k \\
& + \beta_3 Smoothness_i * High\ season\ end_t \\
& + \beta_4 Average\ budget_i * High\ season\ end_t \\
& + \gamma X_{it} + FE_i + FE_k + FE_t + \epsilon_{ikt}
\end{aligned}
$$

Table 3.3.4: Robustness check with bimonthly fixed effects.

Dependent variable	Positive review	Positive review
Specification	(1)	(R3)
Independent variables		
High season end	-0.992**	
	(0.413)	
Smoothness*High season end	1.162**	
	(0.486)	
Feb& March		1.000
		(1.230)
Apr&May		0.888
		(1.222)
Jul& Jun		-1.369
		(1.062)
Aug& Sep		-1.750
		(1.192)
Oct& Nov		-2.035*
		(1.050)
Feb& March*Smoothness		-1.701
		(2.103)
Apr& May*Smoothness		-1.531
		(2.096)
Jun& Jul*Smoothness		2.403
		(1.805)
Aug& Sep*Smoothness		2.782
		(2.026)
Oct& Nov*Smoothness		3.316*
		(1.763)
Fixed effects		
Job category, budget, start	Yes, Yes, Yes	Yes, Yes, Yes
Date	Yes	Yes
Seller	Yes	Yes
R^2	0.37	0.39
N. observations	1,126	1,126

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***. Robust standard errors. Constant, *Percent positive reviews* and *Total times hired* omitted from table for brevity.

Table 3.3.5: Robustness check with measures of seller size.

Dependent variable	Positive review	Positive review
Specification	(1)	(R4)
Independent variables		
High season end	-0.992**	-1.363
	(0.415)	(1.078)
Smoothness*High season end	1.626**	1.814**
	(0.683)	(0.705)
Percent positive reviews	-0.220***	-0.222***
	(0.063)	(0.063)
Total times hired	-0.029	-0.030
	(0.037)	(0.037)
Average budget*High season end		0.033
		(0.090)
Constant	1.360***	0.349
	(0.249)	(0.820)
Fixed effects		
Job category, budget, start	Yes, Yes, Yes	Yes, Yes, Yes
Date	Yes	Yes
Seller	Yes	Yes
R^2	0.37	0.37
N. observations	1,126	1,126

Significant at: $p < 0.1$: *; $p < 0.05$: **; $p < 0.01$: ***. Robust standard errors. The continuous regressors (except *Percent positive reviews*, total offers and average budget) are transformed by taking their natural logarithm, and their coefficients can be interpreted as elasticities. The coefficient on *Percent positive reviews* is a semi-elasticity.

The results of ($R4$) are presented in Table 3.3.5. *Average budget*$_i$ interacted with *High season end*$_t$ is not statistically significant, which suggests that our results are not suffering from bias due to omitting seller cost of effort that is also correlated with seller demand heterogeneity.

3.4 Conclusion and outlook

Reputation systems are essential for online economic activity because they provide buyers with information and sellers with incentives. However, the approach to designing such system is uniform and without much consideration for the economic environment and seller incentive heterogeneity. The economic literature has only recently started to consider the question of optimal aggregation and display of review information, and the effect that would have on seller behavior. Our work aims to inform this question by documenting how heterogeneous seller demand fluctuations lead to inconsistent effort provision on an online services marketplace.

Apart from the general contribution of our results, there are specific implications for reputation mechanism design in environments with heterogeneous demand fluctuations. For example, firms in the tourism industry (restaurants, hotels, etc) are similarly susceptible to heterogeneous and fluctuating demand and therefore incentives. The majority of buyers in these markets, who are tourists, rely exclusively on online reputation when making choices. Buyers choosing between hotels in tourist and residential areas may not be fully aware of the different demand conditions, and therefore incentives, that these establishments face. Making certain reviews more prominent, either by increasing their weight or visibility, can provide buyers with more useful information and sellers with more consistent incentives.

Bibliography

Athey, S. and Luca, M. (2018). Economists (and economics) in tech companies. Technical report, National Bureau of Economic Research.

Bajari, P. and Hortacsu, A. (2004). Economic insights from internet auctions. *Journal of Economic Literature*, 42(2):457–486.

Bar-Isaac, H. and Shapiro, J. (2013). Ratings quality over the business cycle. *Journal of Financial Economics*, 108(1):62–78.

Bar-Isaac, H. and Tadelis, S. (2008). *Seller reputation*. Now Publishers Inc.

Cabral, L. (2012). Reputation on the internet. *The Oxford Handbook of the Digital Economy*, pages 343–354.

Cabral, L. and Hortacsu, A. (2010). The dynamics of seller reputation: Evidence from ebay. *The Journal of Industrial Economics*, 58(1):54–78.

Dellarocas, C. and Wood, C. A. (2008). The sound of silence in online feedback: Estimating trading risks in the presence of reporting bias. *Management science*, 54(3):460–476.

Fradkin, A., Grewal, E., Holtz, D., and Pearson, M. (2017). The determinants of online review informativeness: Evidence from field experiments on airbnb.

Holmström, B. (1999). Managerial incentive problems: A dynamic perspective. *The review of Economic studies*, 66(1):169–182.

Horton, J. and Golden, J. (2015). Reputation inflation: Evidence from an online labor market. *Work. Pap., NYU*, 1.

Jin, G., Lee, J., Luca, M., et al. (2018). Aggregation of consumer ratings: an application to yelp. com. *Quantitative Marketing and Economics*, 16(3):289–339.

Klein, T. J., Lambertz, C., and Stahl, K. O. (2016). Market transparency, adverse selection, and moral hazard. *Journal of Political Economy*, 124(6):1677–1713.

Nosko, C. and Tadelis, S. (2015). The limits of reputation in platform markets: An empirical analysis and field experiment. Technical report, National Bureau of Economic Research.

Tadelis, S. (2016). Reputation and feedback systems in online platform markets. *Annual Review of Economics*, 8(1).

Appendix

3.A Theoretical framework

In this section, we present a general principal-agent model of moral hazard with imperfectly observed, costly effort and fluctuating demand. Like in Holmström (1999), payment is not contingent on performance, and the agent is incentivized to exert effort by career concerns. Agents do not have different types, however; the project's success, which is what the principal observes, is an increasing, stochastic function of the agent's costly effort. The principal makes the project's success public information by posting a review, thus influencing the agent's reputation and the demand for his services in future periods.

The new element is the seasonal demand fluctuation, which affects the agent's incentive to exert effort. When the seller faces high demand in the next period, the return to a good reputation is also high. The model is in partial equilibrium in the sense that we only consider the actions of an individual seller in the face of given demand, rather than the supply and demand of services in the market, taking into account the optimal behavior of all market players.

3.A.1 Demand smoothness and reputation

Let the subscript i identify the seller and the subscript t the period of observation. Let d_{it} be the monthly and D_{it} be the discounted sum of monthly demand of seller i at time t, and δ be the common discount factor. Then, D_{it} is defined as:

$$D_{it} \equiv \sum_{m=1}^{\infty} \delta^m d_{i,t+m},$$

The value to seller i at time t of subscribing to the marketplace is denoted as V_{it}. It is a function f of the total discounted demand D_{it}, the seller's reputation r_{it}, and unobserved factors ξ_i:

$$V_{it} = f\left(D_{it}, r_{it}, \xi_i\right),$$

We assume that the function f is increasing in the first two arguments:

$$\frac{\partial V_{it}}{\partial D_{it}} = \frac{\partial f\left(D_{it}, r_{it}, \xi_i\right)}{\partial D_{it}} \geq 0$$

$$\frac{\partial V_{it}}{\partial r_{it}} = \frac{\partial f\left(D_{it}, r_{it}, \xi_i\right)}{\partial r_{it}} \geq 0$$

and that their cross-derivative is also positive:

$$\frac{\partial^2 V_{it}}{\partial D_{it} \partial r_{it}} = \frac{\partial^2 f\left(D_{it}, r_{it}, \xi_i\right)}{\partial D_{it} \partial r_{it}} \geq 0.$$

Our first two assumptions are straightforward: the value of subscribing to the marketplace increases with demand, since demand increases the number of jobs available, and in reputation, since a higher reputation increases the probability to be hired for a given job.

The cross-derivative tells us that a decrease in reputation is less harmful when the discounted sum of future demand is low. Alternatively, an increase in reputation would be more valuable when future demand is high. This condition would be satisfied e.g. if in the function f, all future discounted demand flows would simply be linearly multiplied by a probability of being hired $p(r_{it})$, where $\partial p(r_{it})/\partial r_{it} \geq 0$. The assumption of a positive cross-derivative, indicating higher returns to a high reputation when demand is high, is supported by our empirical analysis in Table 3.2.4.

For a seller who does not face any seasonality, the value of D_{it} is stable. However, for a seller who faces seasonally fluctuating demand, this value reaches a minimum during the last month of the high demand season. More specifically, assume demand during the *Low season* is smaller by a factor of $S_i \in (0, 1)$, the demand *Smoothness* variable. This means that d_{it} is simply d in the *High season* and $S_i \times d$ in the *Low season* . Then D_{it}^{Nov} in November is smaller than $\overline{D}_{it}$, the value for future demand averaged over all other months, by

$$\overline{D}_{it} - D_{it}^{Nov} = \sum_{s=0}^{\infty} \sum_{m=1}^{6} d \left[\delta^{12s+m} \left(\frac{1 - S_i}{2} \right) + \delta^{12s+m+6} \left(\frac{S_i - 1}{2} \right) \right]$$

This difference is positive for the typical $S_i \in (0, 1)$, and it decreases in S_i. It relies on a strong enough month-to-month discount factor δ to be meaningful. Similarly, let $\overline{V}_{it}$ be the average value of the marketplace to the seller, and V_{it}^{Nov} be the value at November. Thus, the value of subscribing is least sensitive to a change in reputation during the month of November:

$$\frac{\partial \overline{V}_{it}}{\partial r_{it}} > \frac{\partial V_{it}^{Nov}}{\partial r_{it}} > 0$$

Define the marginal change in reputation due to a negative (non-positive) review by r^-. The associated loss in value, x_{it}, suffered by the seller is:

$$x_{it} \equiv \frac{\partial V_{it}}{\partial r_{it}} \times r^-.$$

x_{it} captures the penalty that a seller suffers from a negative review, and serves to discipline moral hazard.[12] Defining $\bar{x}_{it}$ as the average disutility of a negative review, and x_{it}^{Nov} as the disutility suffered in November, we know:

$$\bar{x}_{it} > x_{it}^{Nov}.$$

This difference decreases in S_i, the seller-specific demand *Smoothesss*:

$$\frac{\partial(\bar{x}_{it} - x_{it}^{Nov})}{\partial S_i} < 0.$$

Thus, the model as specified predicts that the penalty of a negative (non-positive) review is lowest at the end of the high demand season, and decreases in the smoothness of seller demand.

3.A.2 Effort and reputation

Consider buyer j who hires seller i to complete a project, whose outcome can be either a "success" or not. The seller's effort e_i, with cost $c(e_i)$, determines the probability of success $\pi(e_i)$. The buyer observes whether it is a success but does not observe the effort level. The buyer pays the seller p_j irrespective of the project outcome. If the project is not a success, the

[12] For simplicity we assume that the reputation does not change when a project is completed successfully.

buyer leaves a negative (non-positive) review, which damages the reputation of the seller and lowers his value of subscribing to the platform by x_{it}.

Let v_{ij} be the value of project j to seller i, defined as:

$$v_{ij} = p_j - c(e_i) - (1 - \pi(e_i))x_{it}.$$

The seller maximizes v_{ij} by setting the optimal effort level e_i. The first order condition is:

$$c'(e_i) = \pi'(e_i)x_{it}.$$

Both the cost of effort and the probability of success increase in the level of effort: $\frac{\partial c(e_i)}{\partial e_i} > 0$ and $\frac{\partial \pi(e_i)}{\partial e_i} > 0$. This establishes the following:

$$\frac{\partial e_i}{\partial x_{it}} > 0.$$

In the data, we are able to observe the determinants of x_{it}: the demand smoothness of individual sellers $Smoothness_i$ and the timing of the job post t in the demand cycle. We do not observe effort e_i directly, but we observe whether the buyer has left a positive review when the job was successfully completed. The model predicts that at the end of the high demand season, sellers who experience a larger demand drop (low $Smoothness_i$) are less likely to receive a positive review because they exert less effort.