The London School of Economics and Political Science

Essays in Applied Microeconomics

Javad Shamsi

London, May 19, 2025

A thesis submitted to the Department of Economics at the London School of Economics,
for the degree of Doctor of Philosophy

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of approximately 51,700 words (excluding bibliography and appendices).

Statement of conjoint work

I confirm that I am the sole author of all chapters of this thesis.

Statement of inclusion of previous work

I acknowledge that Chapter 3 is a revised version of a paper I submitted at the conclusion of my MRes in 2020.

Abstract

This thesis investigates how significant external shocks-namely digital disruption, immigration, and international sanctions-affect labour markets, shape firm dynamics, and ultimately alter political landscapes. Employing different empirical strategies, including quasi-experimental methods, machine learning, and text-based computational analyses, it sheds light on the multifaceted ways in which these shocks reverberate throughout economies and societies.

The first chapter, *Digital Disruption and Entrepreneurial Opportunities*, is based on my job market paper. It explores how online food delivery platforms such as UberEats and Deliveroo have transformed the UK restaurant industry. By compiling a unique dataset and leveraging a staggered rollout of platform entry, this study uses a dynamic difference-in-differences approach to identify causal effects. The findings demonstrate that digital platforms reduce traditional barriers to entry, facilitating a 35% growth in restaurant numbers-predominantly independent and minority-owned enterprises-and broadening the diversity of cuisines offered. This chapter underscores how digital technologies, rather than always favoring large incumbents, can also create inclusive pathways for smaller, diverse entrepreneurs to thrive.

The second chapter, *Understanding Multi-Layered Sanctions: A Firm-Level Analysis*, investigates how Iranian firms adapt to complex sanctions regimes. By applying computational linguistics to corporate transcripts, the study constructs a firm-level measure of sanctions exposure, revealing that sanctions not only harm politically connected firms, but also more extensively burden non-connected firms given their larger market presence. The analysis finds that heightened sanction exposure depresses firm valuations, sales, and investment, driven predominantly by lost export opportunities and higher import costs. These results challenge the notion of "smart" sanctions, highlighting the broad and often unintended consequences for the wider economy.

The third chapter, *Immigration and Political Realignment*, examines the influx of migrants following the EU's 2004 enlargement and its implications for the UK's political fabric. Through a shift-share instrumental variable design that exploits industry-level migration flows and regional employment structures, the chapter shows that heightened immigration exposure contributes to greater support for right-wing, anti-immigration parties and the 2016 Brexit Leave vote, at the expense of traditional Labour support. Although immigration bolsters local economies by increasing activity and employment, cultural and identity-based considerations overshadow economic benefits, propelling a shift in voter alignment and political discourse.

Acknowledgements

My PhD years, though challenging, have been among the most rewarding of my life. This journey would not have been possible without the support and encouragement of many remarkable individuals whom I was fortunate to meet along the way.

I am deeply indebted to my supervisors, Tim Besley, Guy Michaels, and Xavier Jaravel, for their guidance and support. Tim has been involved in my research from the very beginning. He instilled in me a passion for meaningful research and fundamentally shaped how I think as an economist. His mentorship has had a lasting impact on my professional development. Guy has supported me throughout the highs and lows of the PhD, offering invaluable advice and encouragement at every stage. Xavier has inspired me through his insightful feedback and dedication to innovative research. I am sincerely grateful to all three of my supervisors, not only for teaching me research, but also for helping me grow into a better person.

One of the greatest privileges of being at LSE was the opportunity to engage with a vibrant academic community. I benefited enormously from conversations and feedback from faculty members including Gharad Bryan, Michael Callen, Daniel Sturm, John Van Reenen, Catherine Thomas, Kate Smith, Swati Dhingra, Stephen Machin, Alan Manning, Steve Pischke, Mark Schankerman, Alessandro Gavazza, Maitreesh Ghatak, Thomas Sampson, Eduard Boehm, and Torsten Persson.

I also wish to thank my friends at LSE, who made this journey not only academically more rewarding but also more enjoyable; Azhar Hussain, Ali Chobdaran, Amirabas Salarkia, Bijan Aghdasi, Andres Fajardo Ramirez, Peter John Lambert, Tiernan Evans, Negar Mohammadi, Bernardo Mottironi, Michelle Rao, Soroush Sabet, Yasaman Saeidi, Shadi Farahzadi, Yannick Schindler, Lachi Singh, Hamza Syed, Abhijit Tagade, Linchuan Xu, Arnaud Dyevre, Hugo Reichardt, Ola Aboukhsaiwan, and many others whose names space does not permit me to include, but whose friendship I deeply value.

To my parents, who taught me love; my sister, who stood by me every step; and Reyhan, who believed in me when I did not.

Contents

List of Figures

LIST OF TABLES

Chapter 1

A New Order? Digital Disruption and Entrepreneurial Opportunities

Digital marketplaces like Amazon, Alibaba, UberEats, and DoorDash have become significant sources of income for many entrepreneurs. Yet, it remains unclear how this transformation affects the entrepreneurial landscape and the future of work. Do these marketplaces democratize commerce by lowering barriers to entry and empowering small businesses? Do they expand access to entrepreneurship for marginalized groups? Or do they mainly benefit large, established firms that can better leverage economies of scale and algorithms? Understanding these questions is essential for assessing how digital technology will shape future opportunities and economic equality.

The answers to these questions remain unclear both theoretically and empirically. Theoretically, online marketplaces lower barriers to entry, encouraging entrepreneurship. If these lowered barriers are more equitable than traditional systems, underrepresented groups-who often face higher barriers-could benefit the most. However, they might also favor large firms due to reasons like enhanced search capabilities that help consumers find the best firms more easily, network effects that amplify the reach of established businesses, and algorithmic sorting that prioritizes popular firms. This could result in the rise of "superstar" firms and deter smaller entrepreneurs. Empirically, it is hard to study this question because digital platforms typically impact markets all at once, leaving few clear control groups, and in cases where comparisons might be possible, the necessary data is often unavailable.

In this paper, I assemble a novel dataset from multiple sources to examine this trend toward digital marketplaces through the lens of Food Delivery Applications (Food Apps) in the UK. Food delivery services like UberEats and Deliveroo are prime examples of digital marketplaces. They are a new fixture in the restaurant sector, which is a key contributor to both economic value and employment. Unlike many digital platforms that launch nationwide, food delivery applications expand in stages due to regional logistics, providing a quasi-experimental setting. I have compiled a novel dataset that tracks the staggered rollout of two major food delivery services in the UK, UberEats and Deliveroo, from 2013 to 2023, enabling me to analyze the impact using a dynamic difference-in-difference framework.

The two mentioned forces-lowering entry barriers and fostering "superstar effects"-are particularly relevant in this context. On the one hand, food apps may lower entry barriers. They allow businesses to operate with less physical space, eliminate the need for personal delivery fleets, and provide infrastructure for payments and marketing. I analyzed restaurateurs' posts on Reddit, a popular online discussion forum, using a large language model and found support for this. Many cite lower fixed costs as the main benefit of these platforms. Data also shows that app-partnered restaurants often have less expensive, smaller locations, and allocate less space to dining areas. These cost savings can especially benefit marginalized groups who might struggle more to cover the barriers in the traditional setting, promoting equality in entrepreneurship.

However, these apps may also create a "superstar" effect favoring larger or established restaurants. By broadening customer reach-a point emphasized in Reddit discussions-they intensify competition as more firms vie in an expanded market. The reduction of information asymmetry through ratings and reviews makes it easier for top-performing restaurants to stand out. This dynamic could lead to the rise of superstar firms, discouraging smaller entrepreneurs from starting a business.

To understand these competing forces, I developed a theoretical model that includes both lower entry costs and the superstar effect. The model clarifies the mechanisms at play and demonstrates their ambiguous net impact, prompting the need for empirical analysis to determine how Food Apps shape market structure and entrepreneurial opportunities. Accordingly, I conduct an empirical analysis leveraging the staggered spatial rollout of Food Apps in the UK. I trace their impact across three interconnected layers-firms, entrepreneurs, and the product market-and organize my empirical findings accordingly.

First, at the firm level, Food Apps significantly expand the overall size of the market, with the number of restaurants growing by 35% after four years of rollout. This growth leads to increased employment in the sector and is primarily driven by the entry of small and independent businesses. These findings align closely with the reduced entry barriers channel described in the model, where small entrepreneurs, often constrained by limited access to finance and lack of economies of scale, benefit the most from these platforms. Consumer data corroborates the market expansion, indicating that users choose food apps in addition to dining in rather than as a substitute. Nonetheless, there is also a higher rate of restaurant closures, aligning with the notion that intensified competition from market integration forces out less productive firms.

Second, at the entrepreneurial level, ethnic minority entrepreneurs gain more from these platforms. By inferring entrepreneurs' backgrounds based on their names sourced from Companies House, I find that all ethnic groups except White British entrepreneurs experience significant positive impacts from the expansion of food delivery apps.

Third, at the product market level, opening up food entrepreneurship to different ethnic groups results in greater diversity in the products offered, benefiting consumers. I show entrepreneurs often create dishes that reflect their backgrounds. As more diverse entrepreneurs enter the market, the variety of cuisines grows. This is evidenced by a 4% increase in the number of cuisines available through these platforms and a 6% decrease in the Herfindahl-Hirschman Index (HHI) based on cuisine types. This increased diversity counters concerns that platforms might lead to standardization or homogenization of culinary offerings. Instead, the platforms promote culinary diversity, enriching consumer choices.

I also explore mechanisms that are consistent with the disproportionate benefits for minority entrepreneurs. One hypothesis is that these groups are less productive and only enter the market when barriers are lowered. However, the data does not support this explanation: migrant-run app-partnered restaurants exhibit comparable productivity-measured by Google Maps ratings-to both migrant-run restaurants that are not on Food Apps and app-partnered restaurants that are not run by minorities.

A more plausible explanation is that these groups face greater barriers to entry in the traditional setting, such as limited capital, networks, and discrimination in face-to-face interactions. Food delivery apps reduce and level these barriers, creating more equal opportunities. Supporting this, I find that food apps enable minority entrepreneurs, who are more likely to face capital constraints, to open businesses in more affordable areas-a pattern not observed among non-minority entrepreneurs. This aligns with descriptive evidence from Reddit, where restaurateurs cite the reduced need for prime locations as a key reason for joining these platforms. Moreover, inferring customers' backgrounds from Google Maps reviews, minority-run restaurants on platforms do not appear to attract a different racial clientele than offline establishments, suggesting that changes in customer demographics are not driving the benefits to minority entrepreneurs.

My empirical strategy, using the staggered rollout of two major food delivery applications, helps us control for multiple potential confounding factors. First, it accounts for location-specific differences that remain constant over time, such as the baseline rate of entrepreneurship or purchasing habits in different economic areas. Second, it adjusts for time-related effects that influence everyone equally, such as the rise in remote work increasing demand for food delivery. Third, it accounts for trends in outcome variables that differ across locations but follow a consistent pattern, like rich and urban locations exhibiting different trends than others. This last issue is managed through a specification that includes the interaction of local economy indicators with time-fixed effects.

Despite these controls, unobserved trends might have influenced where platforms chose to expand first. Anecdotal evidence and discussions with industry experts suggest that platforms decide where to roll out based on whether a region has enough customers to justify the overhead cost of entry. To test this, I conducted a machine learning exercise using over 30 spatial variables, including level indicators and trends, to predict rollout dates. The results show that variables like urbanization and income levels are key predictors. This suggests that rollout decisions are based on level variables, which are accounted for by location-fixed effects, rather than underlying trends.

To address this potential endogeneity issue more rigorously, I take three additional steps. First, I control for other local economic indicator variables interacting with time to account for the possibility that rich and poor regions might be on different trends. Second, I conduct an event study, which does not reveal any pre-existing trends, providing reassurance about the validity of the rollout assumption. Third, I use other industries as placebo controls, serving as proxies for local businesses, and find no significant impacts on them, further supporting the robustness of my findings. To also address recent econometric critiques of staggered difference-in-differences research designs, I confirm the robustness of the results by employing various alternative estimators.

This paper relates to several strands of literature. First, this paper engages with the literature on how digital technologies, often characterized as high fixed costs, benefit large firms and increase industry concentration (Hsieh and Rossi-Hansberg, 2023; Lashkari *et al.*, 2024; De Ridder, 2024; Aghion *et al.*, 2023). For example, De Ridder (2024) explains that technologies like IT reduce marginal costs but raise fixed costs. This shift leads to slower productivity growth and more market power for big firms. In contrast, I demonstrate a case where the digital economy helps small and independent businesses enter the market, particularly benefiting minority entrepreneurs. While the platform itself might be characterized as high fixed costs and low marginal costs, it enables operation within it with a low fixed cost. This reveals that IT technology is not necessarily limited to benefiting top firms but can also level the playing field.

Second, the paper contributes to the literature on digital marketplaces' impact on entrepreneurship. Much of the existing work focuses on gig economy workers-such as drivers and couriers-who provide standardized, low-barrier services with limited brand differentiation (Hall and Krueger, 2018; Koustas, 2018; Chen *et al.*, 2019; Cook *et al.*, 2021; Jackson, 2022). However, few examine how these platforms impact entrepreneurs who produce and sell differentiated goods, incurring fixed costs and making strategic decisions about product offerings and locations. Existing research often looks at niche platforms. For instance, Carballo *et al.* (2022) analyzes Peruvian firms and shows that a purely informational online platform reduces search costs in trade, benefiting smaller firms engaged in exporting. Other studies highlight how digital technologies can "level the playing field" for women entrepreneurs by mitigating challenges in face-to-face interactions (Poole and Volpe, 2023; Cong *et al.*, 2022; Sicat *et al.*, 2020; Pergelova *et al.*, 2019). My study advances this literature by examining a widely used platform, employing its staggered rollout as a research design, and showing how these applications reduce barriers for small businesses, particularly benefiting ethnic minority entrepreneurs.

Third, this paper builds on research about the economic impact of food delivery applications. I examine their effects on market structure, employment, and cuisine diversity. Previous studies, such as Raj and Eggers (2023); Raj *et al.* (2023); Raj and Choe (2023), show that platform penetration increases competition and exit rates among less efficient businesses while benefiting young and independent establishments by reducing search costs and enhancing digital capabilities.

Fourth, this paper connects to the literature on talent misallocation and the resulting loss of potential. Hsieh *et al.* (2019) highlight how race- and gender-based barriers result in talent misallocation across occupations. Similarly, Bell *et al.* (2019) and Aghion *et al.* (2017) show that children from disadvantaged backgrounds face higher obstacles to becoming innovators, leading to "lost Einsteins." Akcigit *et al.* (2017) provide further evidence, showing that this correlation between parental income and inventor success holds historically. My paper expands this literature by addressing how these barriers extend to less high-status sectors, like the restaurant industry, and how digital technology can mitigate them. Reducing barriers in such industries is still very important, as entrepreneurship and firm ownership have been shown to be key in reducing the racial wealth gap (Lipton, 2022; Fairlie and Robb, 2007).

Finally, I contribute to the literature on how digital platforms influence the spatial distribution of economic activities. Fan *et al.* (2018) find that e-commerce reduces the fixed cost of market entry and the impact of distance on trade, and boosts production in smaller cities, while (Couture *et al.*, 2021) report limited economic benefits in rural areas. In urban contexts, most studies focused on Airbnb (Almagro and Domínguez-Iino, 2024; Calder-Wang, 2021; Garcia-López *et al.*, 2020; Schaefer and Tran, 2020). Specifically, Almagro and Domínguez-Iino (2024) finds that Airbnb expansion leads to an increase in tourism-focused amenities (e.g., restaurants) at the expense of local amenities. Looking at Uber, Gorback (2020) shows that ridesharing services enhance amenities and housing prices in areas with driving accessibility but poor transit options. Building on this literature, my study reveals how food apps, by reducing the necessity for prime locations, enable restaurants to relocate to more affordable areas within neighborhoods, thus redistributing economic activity spatially and potentially mitigating location-based barriers for small businesses.

The subsequent sections of this paper are structured as follows. The following section details the study context and data sources. In Section 1.2, I introduce a model to guide the analysis and provide evidence supporting the assumptions about how food apps influence entrepreneurship. Section 1.3 outlines the research design, focusing on the staggered rollout of food delivery applications and the methodological approach. The empirical findings in Section 1.4 show that technology boosts market entry, especially for small businesses, and ethnic minorities, who traditionally faced entry barriers, benefit disproportionately. Finally, I demonstrate how the rise in entrepreneurship among ethnic minorities spills over into the product market, leading to greater product diversity. Section 1.5 concludes.

1.1 Context and Datasets

Food delivery applications have grown fast worldwide. In 2024, global revenue is expected to hit $1.2 trillion (Statista, 2024), with the UK market projected at $50 billion.

These applications broadly provide two types of services to businesses. In the first model, consumers order through the app, and the platform coordinates and handles the delivery on behalf of the restaurant. In the second model, consumers order through the app, but the establishment handles the delivery, with the platform just facilitating matching and information exchange. This paper focuses on the first model, examining Uber Eats and Deliveroo, the two biggest platforms based on this business model, which launched in the UK in 2013 and 2016, respectively.[1]

I now introduce the datasets used in my analysis, highlighting their role in the empirical exercise, and providing descriptive evidence where applicable.

1.1.1 Consumers

I use two datasets to track consumer use of food delivery apps.

Fable Spending Data. I use the Fable dataset to track consumer spending on food delivery apps. It covers 3820,000 monthly UK users and includes over a billion bank transactions from January 2016 onward. The data captures the first part of each consumer's postcode and full merchant postcodes, allowing me to see spending on Deliveroo and

[1]Just Eat is another major player in the UK market, but it is not considered in this study for two reasons. First, Just Eat predominantly operates under the second business model, particularly in its early stages, which is not the focus of this analysis. Second, because Just Eat relies on restaurants to manage their own deliveries, its expansion across the UK was more rapid and lacked the staggered rollout seen with Deliveroo and Uber Eats (Keeble *et al.*, 2021), making it less suitable for the empirical strategy employed here.

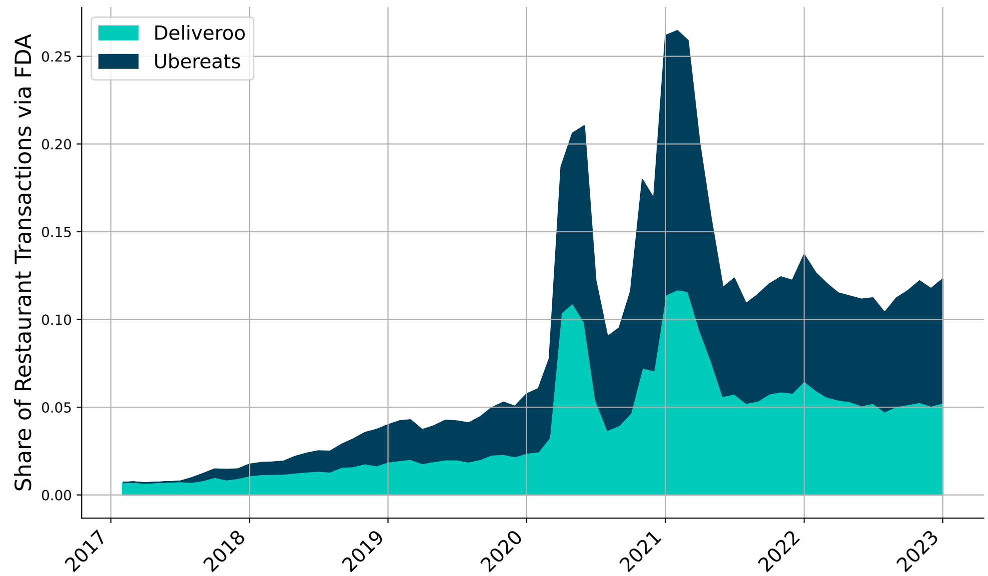
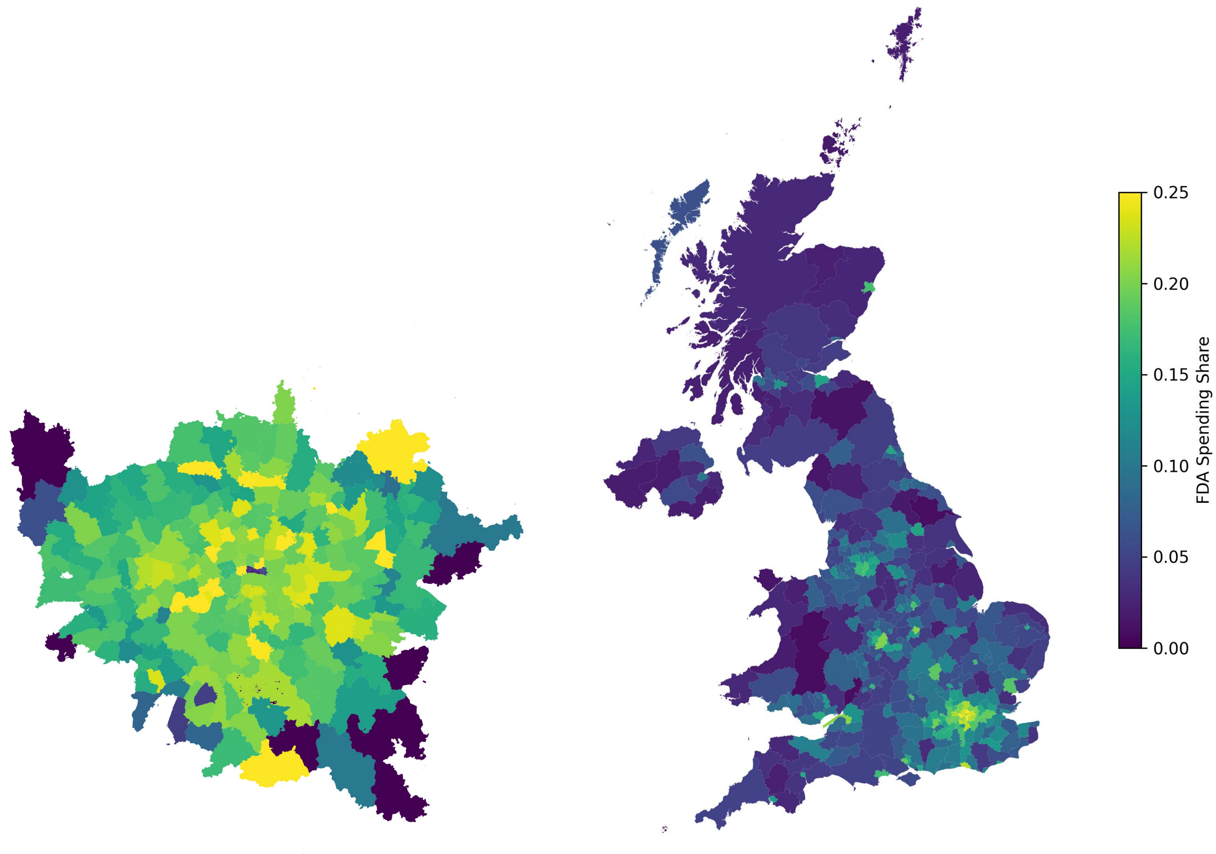


Figure 1.1. *Notes:* This figure illustrates the trend in the share of transaction values made through various food delivery applications (Food Apps) from 2017 to 2023. Data is sourced from Fable is limited to transactions in GBP and for MCC codes pertaining to eating establishments. The stacked areas represent the proportion of transaction values corresponding to Deliveroo and Uber Eats.

UberEats across different regions and over time. Table A1 columns (1) and (3) compare Fable users to the UK population. The sample includes a relatively higher proportion of younger and wealthier individuals, which should be considered when interpreting the results.

Using Fable, Figure 1.1 shows the market share of Food Apps in the restaurant sector over time, focusing on UberEats and Deliveroo. Usage has steadily increased, with two major peaks during the UK's nation-wide COVID-19 lockdowns. Although the pandemic accelerated Food App adoption, the trend has since stabilized without a notable decline.

National trends give an overall view but hide key regional heterogeneity. Figure 1.2 shows how food delivery application use varied across local authority districts in 2022. In some areas, more than 30% of restaurant spending went through UberEats and Deliveroo, while other areas saw little use. The box-and-whisker plot in Figure A1 complements this by showing the distribution of food apps usage over time, highlighting persistent and possibly widening geographic disparities.



(a) Greater London **(b)** UK

Figure 1.2. *Notes:* The figure represents a geographical visualization of the Food Delivery Application penetration (Deliveroo and UberEats) across postal districts in greater London (panel a) and various local authority districts in the UK (panel b) for the year 2022. The data, derived from Fable, includes transactions associated with the restaurant industry. Each district's Food App penetration is calculated as the proportion of outgoing spending tagged with references to food delivery services Uber Eats and Deliveroo relative to total outgoing spending in the restaurant industry within the spatial unit. This map was created by correcting mislabeled or outdated district names using a bespoke mapping function, ensuring alignment with the most current administrative boundaries as defined in the 2022 local authority district dataset.

Kantar's Worldpanel.[2] I use Kantar's Worldpanel data, both the Take-Home Purchase Panel and the Out-of-Home Purchase Panel, for this analysis. Though smaller in sample size, Kantar's Worldpanel offers richer details compared to Fable, allowing me to

[2]The analysis and findings have been undertaken independently based on data supplied by Kantar's Worldpanel Take Home Purchase Panel and Out of Home Panel, and Kantar's Worldpanel does not endorse the efficacy or accuracy of this analysis, interpretations, and findings. All errors and omissions remain my responsibility as the author of this publication.

examine substitution behavior, such as whether increased Food App usage leads to a decrease in other restaurant spending methods. Kantar's Worldpanel tracks household purchases through its fast-moving consumer goods (FMCG) panel, which covers about 30,000 British households. The Take-Home data focuses on food and drinks bought intended for "take-home" consumption, including items from supermarkets, convenience stores, and smaller vendors. The panel records both in-store and online purchases by scanning barcodes, and capturing product details like price, size, and nutrition.

The Out-of-Home (OOH) panel is a smaller subset of 7,500 individuals. It tracks food for consumption outside the home. This includes "on-the-go" food -which may have been purchased from the same sources as at-home food-as well as all meals from restaurants and takeaways, even those eaten at home. Participants record their purchases using a mobile phone app. Although multiple members from one household can participate, over 85% of households are represented by just one individual. To ensure consistency, I aggregate purchases from multiple household members into a single household-level record as in O'Connell *et al.* (2022).

Table A1 provides summary statistics for demographics and spending variables in the Kantar Worldpanel dataset, comparing them with the Fable dataset and official statistics from the ONS. Relative to ONS, Kantar's income distribution shows better alignment than Fable's, suggesting a more representative economic profile. In terms of age, Kantar occupies a middle ground, being younger than ONS's distribution but older than Fable's. The Kantar panel also reports a 60% female share.

Fable shows higher spending on Food Apps compared to Kantar's Worldpanel. Several factors explain this difference. First, since Kantar's data is based on self-reporting, there may be variations in how participants record their purchases (for example, a food app order might be classified under general delivery). Second, Fable's sample is slightly younger, over-representing those more likely to use delivery services. Additionally, the share of food delivery app spending within total restaurant spending is lower in Kantar's data because it encompasses a broader range of purchases, such as snacks and non-restaurant items, which dilutes the proportion attributed to delivery platforms. Lastly, Fable captures only bank card transactions, potentially increasing the reported share of delivery platforms by excluding cash transactions more common in non-Food App purchases.

1.1.2 Restaurants

I collected data on restaurants from various sources, including scraped listings, official records, and market research companies.

Full records of all restaurants on Deliveroo, and UberEats. I construct a comprehensive dataset of all restaurants listed on UberEats and Deliveroo by systematically scraping both platforms each quarter, from the first quarter of 2021 through the first quarter of 2024. For each restaurant, I recorded the name, type of cuisine, and location[3]. An example of the scraped data for each platform is presented in Figure A3. I will use these data, combined with additional sources for the pre-2021 period, to track platform expansion into different locations, as detailed in subsequent sections.

While it is possible that some registered outlets were not captured in our searches, the number of identified outlets in the last batch, 2024 Q1, aligns with reported figures of approximately 63,000 for UberEats in 2023 and 50,000 for Deliveroo in 2024 (John Lewis Partnership, 2023; Deliveroo, 2024), bolstering confidence in the dataset's completeness.

Restaurants on Google Maps. I compiled a dataset of over 180,000 restaurant listings from Google Maps, likely covering most, if not all, restaurants in the UK, including those on delivery platforms and those that are not. The dataset includes key details such as cuisine type, average ratings, price indicators, and reviewer information. I construct the data by leveraging both the official Google Maps API and web scraping techniques. Figure A4 panel (a) displays a sample restaurant listing on Google Maps along with the extracted data points.

One key advantage of this dataset is the inclusion of reviews. I extracted over 6 million reviews from all listed restaurants in the UK. This data helps me to infer the ethnicity of each reviewer based on their names. Figure A4 panel (b) showcases a sample of restaurant reviews on Google Maps along with the extracted reviewers' names.

I matched about 60% of app-partnered restaurants with their corresponding entries on Google Maps using names, coordinates, and the Google Places API service, as detailed in Section A1.3. Table A2 compares app-partnered restaurants with non-partnered ones, showing that app-partnered restaurants tend to have higher prices, more reviews, lower average ratings, and are generally newer.

Local Data Company (LDC). LDC is a commercial research consultancy specializing in retail locations throughout Britain. LDC's data includes detailed information such as business types, exact locations, names, opening and closing dates, and cuisine type. They collect this data by physically surveying premises every 6 or 12 months. In addition, LDC continuously updates its information by monitoring news sources to capture interim changes and keep the database up to date between surveys.

With data on both cuisine types and exact locations, LDC allows for cuisine type analysis at different spatial levels. Since the dataset includes restaurant names, we can match

[3] If a restaurant stays on the platform, it will appear in multiple waves of our data. However, tracking it across these waves is challenging due to naming variations and incomplete names in earlier datasets.

establishments directly with other sources at the restaurant level using these names, as explained in Section A1.3. Its focus on physical inspections, rather than just registration records, also helps capture the real operations of businesses, including cases where a single registered entity operates under multiple brands.

UK Business Structure Database (BSD). This database is an annual extract of the Inter-Departmental Business Register (IDBR). It includes almost all UK businesses registered for Value Added Tax (VAT) or with at least one employee under the PAYE system. The dataset contains key details such as the first half of the postcode, employment numbers, turnover, industry classification (SIC), legal status (e.g., sole proprietor, partnership), foreign ownership, company start date, and termination date.

Due to the anonymized nature of the data, I cannot directly match individual restaurants with other sources. Instead, for local authority-level analysis, I rely on aggregated data at the unit level (individual sites or enterprises).

I utilize the BSD as an official registry of legally recognized businesses, offering a baseline to cross-verify findings from LDC, which relies on market listings and advertised brands rather than formal legal entities. This comparison distinguishes expansions in registered firms from the creation of multiple virtual brands by a single registered firm, a practice known as "multi-branding". Further, because BSD encompasses all industries, not just restaurants, it allows me to conduct placebo tests on non-restaurant sectors to ensure that observed effects are indeed driven by Food Apps' influence on the restaurant sector.

Company House. The Company House dataset provides information on business directors, including names, ages, and nationalities. While it does not directly include ethnicity, I infer these attributes using name-based analysis, a common method in economic research. The basic premise is that names can provide clues about race, reflecting cultural traditions or established naming conventions. Typically, this method involves training a model on a large dataset of names annotated with race or ethnicity labels. Once trained, the model can infer race or ethnicity for names in an untagged dataset. I detail the procedures for this inference in Appendix A1.2.

A limitation of the Company House dataset is that it covers only incorporated firms at the enterprise level. However, as demonstrated in Table A3, which shows both local units (establishments) and enterprises (firms) from IDBR, this limitation does not significantly constrain the analysis. Incorporated firms constitute 79,000 of the 101,000 total enterprises (approximately 78%) and 95,200 of the 118,000 local units (about 81%). This indicates that the dataset captures the majority of businesses, with unincorporated firms representing a smaller portion. The second limitation-having data only at the enterprise level-is also less consequential, as the difference between enterprises and local units among incorporated firms is less than 16%, and even smaller when focusing on smaller restaurants. Additionally, since non-incorporated businesses are probably more likely to be immigrant-owned firms, the findings on ethnic minority entrepreneurs may actually underestimate the true effects.

1.1.3 Other Datasets

The Business Register and Employment Survey (BRES). To look at employment, I use BRES. BRES is a vital source of official employment statistics, providing detailed information on the number of employees and employment across different industries and regions in the UK. [4]

Price Paid Data. This dataset is a comprehensive dataset published by HM Land Registry, detailing property transactions in England and Wales since 1995. It includes key information such as the transaction date, price paid, property type, and full address details including postal code, local authority, district, and county. I use this data to construct an index to assess whether the Food App allows restaurants to relocate to more affordable areas.

To construct the index, I calculate the median property price for all transactions in each postcode over the past 10 years, adjusting for inflation. The median number of transactions for each postcode was six transactions. For postcodes with no transactions, which are rare, I impute the missing price data using nearby postcodes with known values. I use haversine distance to identify the 10 closest postcodes and take the median of their transaction prices to fill in the missing values.

Valuation Office Agency (VOA) Data. I use data from the Valuation Office Agency to obtain detailed information on commercial properties at the postcode level. The VOA, an executive agency of HM Revenue and Customs, assesses properties for council tax and business rates in England and Wales. The dataset includes estimated property valuations, the number and types of rooms, and the floor space of each room. While the Price Paid Data provides actual transaction prices at the postcode level, offering direct insights into property costs, the VOA data offers government-estimated valuations of commercial properties. I match this information to restaurants using their postcodes. Since most postcodes contain only one restaurant, matching is straightforward. For postcodes with multiple restaurants, I average the property characteristics. This helps me analyze how property features relate to restaurant operations and their participation in food delivery applications.

[4]When using this dataset, I use data from 2015 onwards. This is necessary because the figures from 2015 to 2022 include businesses registered for PAYE but not for VAT, which makes them inconsistent with pre-2015 data.

Reddit Data. I also use data from Reddit, an online platform where users discuss topics in communities called subreddits. I collect posts from two subreddits: *r/Restaurateur* and *r/RestaurantOwners*. These communities consist of restaurant owners and industry professionals sharing experiences and advice. I identify posts that indicate an intention to use food delivery apps. Analyzing these posts using a large language model, as will be discussed shortly, helps me understand the motivations and concerns of restaurant owners regarding the adoption of these platforms.

1.2 Conceptual Framework

In this section, I provide an intuitive framework that captures the two mechanisms through which food delivery applications might impact restaurants and entrepreneurship. The formal model and detailed derivations are presented in the Appendix A1.4.

I build upon the firm heterogeneity model in monopolistic competition introduced by Melitz (2003). Consumers have preferences represented by a standard Constant Elasticity of Substitution (CES) utility function with elasticity $\sigma > 1$ over a continuum of differentiated products, where each variety is denoted by ω and collectively represented by the set Ω. This specification leads to a demand function for each variety as follows:

$$q(\omega) = YP(\omega)^{-\sigma}P^{\sigma-1}$$

Where P is the aggregate price index, defined as:

$$P = \left(\int_{\Omega} P(\omega)^{1-\sigma}d\omega\right)^{\frac{1}{1-\sigma}}$$

On the production side, firms face uncertainty regarding their productivity levels. To enter the market, a firm must pay an entry cost. Upon entry, each firm draws its productivity φ from a Pareto distribution with shape parameter θ. After observing their productivity, firms decide whether to produce or exit the market. Operating firms incur a fixed production cost f_d and have constant marginal costs inversely proportional to their productivity. Labor is the only input in production.

There are two key equilibrium conditions in this framework. First, the Zero Cut-off Profit condition determines the productivity threshold φ^* below which firms cannot profitably operate in the market. Firms with productivity $\varphi < \varphi^*$ choose to exit. Second, the Free Entry condition ensures that the expected profits from entering the market equal the entry cost. It determines the market equilibrium price index.

I explore two mechanisms through which food delivery applications might affect this equilibrium:

1.2.1 Reduction of Fixed Costs

Food delivery applications can reduce the fixed production cost f_d by providing essential services such as logistics, payment systems, marketing, and customer service infrastructure. Also, restaurants no longer need to invest heavily in physical space, delivery fleets, or administrative overhead, lowering the barriers to operation and allowing more firms, especially smaller and less capitalized ones, to enter the market.

Empirical evidence supports this assumption. Analyzing restaurateurs' experiences shared on Reddit, detailed in the Appendix A1.5, I found that restaurant owners frequently cite cost reductions as a key motivation for adopting food delivery applications. Using a large language model to classify the content of these posts, I observed that key reasons include reductions in marketing expenses, leveraging platform-provided infrastructure, lowering on-premise delivery costs, and decreasing premises costs.

These insights align with additional data on the physical characteristics of restaurants partnering with food delivery applications. As shown in Figure A6, these restaurants are located in cheaper areas, are smaller in size, and allocate less space to dining compared to non-partnered establishments. This pattern persists even after accounting for postal districts, suggesting even within a postal distinct food app restaurants tend to sort into cheaper and smaller areas. By relying on app-driven visibility and logistical support, these restaurants can avoid the fixed costs associated with large, high-exposure locations.

1.2.2 Superstar Effects

There are several reasons to believe that food delivery platforms might lead to a superstar effect and a winner-takes-all dynamic in the restaurant industry. First, these platforms expand consumer reach, allowing restaurants to serve customers beyond their immediate locality while they reduce information asymmetries through features like ratings, and reviews. This will help the top firms to be known and thus help the top player benefit the most. Similarly, algorithmic sorting can disproportionately benefit top-performing restaurants. Third, network effects can amplify the success of popular establishments, as increased orders and reviews further enhance their visibility on the platform.

In my model, I capture this potential for a superstar effect by assuming a decrease in the shape parameter θ of the Pareto distribution of firm productivities, making the distribution more fat-tailed. A lower θ implies greater heterogeneity among firms and increases the likelihood that highly productive firms will dominate the market. While I do not model the specific mechanisms leading to this decrease in θ, this approach allows the model to represent various underlying reasons that might contribute to a winner-takes-all outcome.

Empirical evidence further substantiates this modeling assumption. Analysis of posts from Reddit shows that "Expanding Customer Reach" is the most cited reason for adopting food delivery applications. This expansion allows firms with higher productivity or better offerings to access a larger market, amplifying their competitive advantages. As these firms attract more customers from a broader area, they can capture a significant share of the market, potentially at the expense of less productive competitors.

1.2.3 Propositions

Based on the developed framework, I derive the following propositions. Detailed mathematical derivations and proofs for each proposition are provided in Appendix A1.4.

Proposition I. *If a new technology disproportionally benefits superstar firms, i.e., leading to a more fat-tailed distribution of firm productivities—meaning it decreases the shape parameter θ of the Pareto distribution—it will decrease the equilibrium number of firms in the market.*

This proposition reflects the superstar effect, where increased market integration, algorithmic sorting and etc favor highly productive firms, leading to market concentration and a reduction in the total number of operating firms.

Proposition II. *If a technological improvement reduces the fixed cost of production f_d, it will increase the equilibrium number of firms in the market.*

By lowering f_d, food delivery platforms make it feasible for more firms to operate profitably. Next, I consider what happens when a technology does both: lowers fixed costs and benefits superstar firms.

Proposition III. *If a new technology both disproportionately benefits superstar firms (decreasing the shape parameterθ) and reduces the fixed production cost f_d, then the equilibrium number of firms in the market will increase if and only if the proportional decrease in the fixed cost is sufficiently large relative to the decrease in θ. Specifically, the increase in the number of firms will occur when:*

$$\frac{\Delta f_d}{f_d} > \left(\frac{\sigma - 1}{1 + \theta - \sigma}\right) \frac{\Delta\theta}{\theta}$$

where Δf_d and $\Delta\theta$ represent the absolute decreases in f_d and θ, respectively.

Thus, the net effect on the number of firms is ambiguous. It depends on which force is stronger: the reduction in fixed costs or the superstar effect. If the fixed cost reduction is larger, more firms enter, promoting entrepreneurship and diversity. If the superstar effect dominates, the market becomes more concentrated, and fewer firms operate.

Also, we can see that higher σ amplifies the impact of the superstar effect, as consumers become more sensitive to price and quality differences among products. This heightened

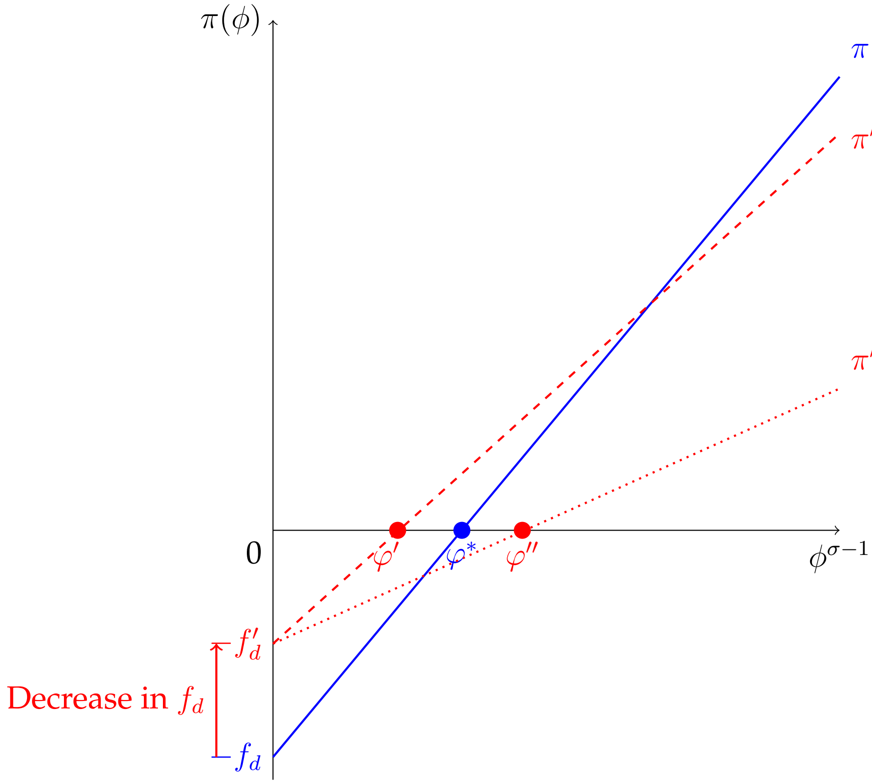


Figure 1.3. *Notes:* This figure shows a schematic representation of the effect of a technological change that simultaneously reduces fixed costs (f_d) and alters the productivity distribution to reflect superstar effects on the profit as a function of productivity. The reduction in fixed costs is represented by an upward shift in the y-intercept, while the change in slope captures the impact of superstar effects as well as the reduction in the fixed cost. Depending on the magnitude of these two opposing forces, the productivity cutoff (φ^*), below which firms exit the market, may shift to the right (φ') or to the left (φ'').

sensitivity leads consumers to substitute more readily towards highly productive "superstar" firms, allowing these firms to capture a larger market share and further dominate the market.

Figure 1.3 illustrates how such a technology impacts the productivity cutoff φ^* in the model. When fixed costs decrease due to the introduction of food delivery platforms, the profit function shifts upward (reflecting lower entry barriers), but the slope also decreases due to the reduction in the price index (see Appendix A1.4). In one scenario (dotted line), the slope becomes sufficiently flatter, and the cutoff shifts right to φ^*, allowing more firms to enter as reduced fixed costs make the operation profitable for less productive firms. In another scenario (dashed line), the cutoff shifts left to φ^*, indicating that the dominance of "superstar" firms prevents less productive firms from entering despite the lower fixed costs.

The theory demonstrates that the net impact on the number of restaurants and the entry of new entrepreneurs depends on the relative strength of these opposing forces. Because the outcome is ambiguous, we need empirical investigation to understand how food delivery platforms affect the restaurant industry. This is the focus of the next section.

1.3 Research Design and Rollout of Platforms

To examine the impact of food delivery applications on the restaurant industry, I utilize an event study design based on the staggered rollout of Deliveroo and UberEats across regions. This staggered rollout provides a quasi-experimental setting, allowing us to isolate the causal effects of these platforms by leveraging variations in their rollout dates. The baseline specification I will estimate is:

$$y_{st} = \alpha + \Sigma_j \beta_j \mathbb{1}[t = E_s + j] + \mu_s + \lambda_t + X_s \times \lambda_t + \epsilon_{st} \quad (1)$$

In this equation T represents the year, S is the spatial unit (either a Postal District or a Local Authority, as discussed further later), and E_s is the year when spatial unit S gained access to the food delivery apps. This approach compares the pre- and post differences in outcomes between regions (or individuals residing in regions) where a food delivery application was introduced and those in regions where the food delivery application has not yet been introduced or will not be introduced. The specification includes both region and year fixed effects, as well as controls for region population and GDP (X_s) interacted with time. Assuming that, in the absence of the platform rollout, outcomes would have followed similar trends, and that the treatment effects are uniform across locations and time, the coefficient β represents the average treatment effect on the treated (ATT) due to the introduction of Food Apps.

With these assumptions, the two-way fixed effects (TWFE) model allows us to address several potential concerns that could otherwise impede causal interpretation. Firstly, it rules out the possibility that time-invariant fixed differences in individual spending behavior or regional restaurant market features are driving the results. For instance, one might suspect that richer, more urban areas and their residents have different baseline outcomes. By incorporating location fixed effects, I can mitigate these concerns.

Secondly, the results are unlikely to be influenced by the outcomes that evolve uniformly across individuals or restaurants in different locations. For instance, global trends such as the increased reliance on takeaway foods due to the rise of remote working arrangements may affect all individuals and restaurants across different locations in a similar manner. Time-fixed effects help account for this.

However, the rollout of platforms might still be correlated with trends in the outcome variable. This is relevant because the rollout decisions were strategic rather than random. For example, denser urban areas gained access to both major food delivery applications earlier, suggesting that the expansion decision is not random. Nonetheless, the absence of pre-trends, which I will discuss, makes it unlikely that platforms timed their rollout to coincide with sudden shifts in local demand. Instead, the evidence indicates that level variables, rather than trends, influenced where the platforms decided to expand.

Specifically, several factors that do not concern trends in outcome variables guided the rollout sequence. First, establishing an office or operational infrastructure incurs fixed costs, so platforms prioritized markets that were sufficiently dense to justify these initial investments-leading them to focus on larger urban areas first. Second, scale constraints due to limited platform capacity, especially in the early stages, influenced the sequence of expansion. Third, to capitalize on network effects, platforms aimed to simultaneously attract a critical mass of users and restaurants, which was more feasible in tech-savvy areas with a higher concentration of restaurants.

A simple machine learning-based feature selection process supports this claim that level variables, rather than underlying trends, are deriving the rollout dates. In this exercise, as detailed in section A1.6, I employ the best selector method and consider a list of variables comprising both fixed location characteristics and level and trends of economic indicators as potential predictors. The analysis indicates that variables such as population, GDP, and urbanization are the best predictors of rollout dates rather than trends.

I take extra measures to ensure the validity of the parallel trends assumption. First, I apply the estimator from Borusyak *et al.* (2024) to detect any pre-existing trends. In the robustness checks, I (1) re-estimate the main analysis using alternative methods from De Chaisemartin and d'Haultfoeuille (2020), Callaway and Sant'Anna (2021), and Sun and Abraham (2021); (2) include local economic indicators to account for different trends based on the economic situations of regions; (3) add rollout-group-specific linear time trends; (4) test placebo industries and find no significant results; and (5) repeat the analysis excluding the COVID-19 period and excluding London.

1.3.1 Defining Market

Defining geographic markets in the context of food delivery applications is complex due to the fluid nature of delivery boundaries. This complexity arises primarily for two reasons. First, regardless of how spatial units are defined, a restaurant located near the borders of one unit may serve customers in neighboring units, complicating the assignment of restaurants and consumers to specific markets. Second, platforms dynamically adjust delivery zones based on factors such as demand, traffic, and courier availability.

To address this issue, I use two geographic units for analysis: local authorities and postal districts. Local authorities are administrative regions in the UK, with around 400 in total. A postal district, on the other hand, is defined by the first half of a postcode (e.g., "NW1" for "NW1 0QA"), known as the outward code. I employed Geographic Information Systems (GIS) and the National Statistics Postcode Lookup to map restaurant geolocations and postcodes to their respective postal districts. Although sizes vary, 2011 census data shows the median population of a postal district was 22,574, with an average of 24,714 (Office for National Statistics, 2015).

Each spatial unit offers distinct advantages for analysis. The local authority level provides a broader market definition, which helps mitigate concerns about spillover effects. Additionally, some outcome variables are only available at the local authority level, or using postal districts might give rise to issues related to small sizes in each cell. Moreover, policies and economic decisions are often made at this broader administrative level, making findings particularly relevant for local policymakers.

However, the local authority level approach comes with drawbacks, such as reducing spatial variation and potentially masking important differences within the area. Local authorities can be quite diverse in their economic structure, demographics, and geographic size, which can obscure local heterogeneity. On the other hand, postal districts offer more spatial variation and better align with how food delivery applications define their delivery areas, making them a useful unit of analysis. Given these strengths and weaknesses, where data permits, analyses will be conducted at both the local authority and postal district levels to fully leverage the benefits of each approach.

1.3.2 Pinpointing the Rollout Date

I use the earliest rollout of Deliveroo or Uber Eats as a benchmark for food delivery application presence. Capturing this rollout is challenging. First, platforms rarely announce expansions through press releases or media coverage in a systematic way. A more important conceptual challenge arises from platforms' non-standard rollout strategies. In some cities, a platform might launch services comprehensively, while in others, it might adopt a gradual, neighborhood-by-neighborhood approach. Different platforms can also adopt distinct rollout strategies. For example, Deliveroo might launch city-wide, while UberEats might opt for a phased neighborhood approach. This inconsistency makes it difficult to determine the appropriate spatial unit for recording rollout-whether at the city, borough, or postal district level.

To overcome these challenges and precisely pinpoint platform rollout dates into each spatial unit, I adopt a data-driven approach, utilizing distinct methodologies tailored to each platform's unique characteristics and rollout patterns.

Deliveroo: Deliveroo does not, at least systematically, disclose its rollout dates. To determine the rollout dates of this platform across different regions, I utilized two primary sources.

First, I systematically extracted location data for all restaurants listed on UberEats from the first quarter of 2021 through the second quarter of 2024 on a quarterly basis. This approach provided snapshot pictures of all restaurants on the platform at different points in time. Aggregating this information for all restaurants within a spatial unit, I estimate each platform's rollout date into that region based on the earliest restaurant among all restaurants. Table A4 shows the details of these scraped restaurants.

Second, for regions where Deliveroo began operations before 2021, I used a commercial dataset that arguably recorded the entire universe of restaurants on Deliveroo since its inception in the UK. This dataset, compiled through scraping exercises by data providers, starts from 2013 to 2021. The few instances where media coverage has provided rollout dates for Deliveroo (Daily Mail, 2019) align with the timing of restaurant appearances in this dataset, validating its accuracy.

UberEats: First, like for Deliveroo, I systematically extracted location data for all restaurants listed on UberEats from the first quarter of 2021 through the second quarter of 2024 on a quarterly basis. The presence of a restaurant in a particular postal district or local authority is used as an indicator of UberEats' operation in that region, even though the coverage may not be exhaustive. Table A4 summarizes these scraped restaurants.

To identify regions penetrated by UberEats prior to 2021, unlike Deliveroo, where I had access to an external dataset tracking restaurants on the platform, no comparable resource exists for UberEats. Instead, I relied on three additional sources and selected the earliest rollout date from these.

- I reviewed Uber's official announcements, which listed rollout dates for several regions on the Uber Newsroom website until August 2017. This provided 23 rollout dates, as listed in Table A5. Since city or region names do not always align with postal districts, I assign a postal district to the announced region if the majority of its spatial area falls within the mentioned city or region.
- UberEats maintains a coverage page listing regions it serves in the UK.[5] By leveraging the Internet Archive, I tracked historical versions of this page, identifying regions listed at various points in time. I retrieved coverage information for the following dates: 8 May 2020, 24 May 2020, 3 June 2020, and 30 September 2020. Figure A7 shows this.

[5]https://www.ubereats.com/gb/location

- For each of the regions listed in the coverage page of ubereats at 2021Q1, I used indexed pages in Google to find the indexed dates for that specific region. That is, I utilised Google search queries to determine the indexing dates of these city pages by searching for the presence of the city-specific URLs within Google's indexed pages. While Google reindexes periodically, the index date provides a point in time at which we can be certain that a link for that particular region existed, thus offering a conservative estimate for the earliest possible rollout date.

Finally, among these sources-systematic restaurant data extraction, official announcements, historical coverage analysis, and Google indexed dates-I selected the earliest indicated rollout date for each region as the rollout date of UberEats. This approach ensures that the earliest possible evidence of UberEats' operation in a region is used to determine the rollout timeline.

Using the identified rollout dates for each platform, panels (a) and (b) of Figure 1.4 show Deliveroo's and UberEatsr rollout across postal districts, highlighting considerable spatial variation. Appendix Figure A8 presents similar patterns at the local authority level. As expected, there is slightly less variation at this broader level, since the presence of a single restaurant in an area qualifies it as treated, leading to earlier treatment assignment across all units. Using alternative definitions, such as the second or third restaurant, yields similar results since many restaurants join around the same time. The comparison between panel (a) and panel (b) shows that Deliveroo, as expected, was the first to penetrate most areas. Appendix Figure A9 further explores this by directly comparing the rollout dynamics of both platforms, highlighting areas served exclusively by one or neither platform.

The steady expansion of food delivery applications has significantly increased their geographic and population reach over time. Figure A10 illustrates the coverage of postal districts and local authorities by UberEats or Deliveroo over time. Panel (a) shows a steady increase in postal district coverage starting from 2014, with a notable rise around 2016–2018, reaching about 78% by 2024. Panel (b) indicates a faster increase at the local authority level, nearing 100% coverage by 2022. This earlier coverage at the local authority level reflects its broader geographic scope, as discussed. To account for population differences across regions, Figure A11 combines platform rollout dates with postal district population data, illustrating that by 2023, nearly 90% of the population had access to at least one app-partnered restaurant.

Addressing Concerns of Market Boundaries and Accessibility A key concern in this context is the potential misclassification of areas as treated based on the presence of a single restaurant partnering with a food app. This may not reflect the broader accessibility of the platform for other entrepreneurs in the same area. For instance, other restaurants might lack the ability to join due to their specific location within the area, limitations in the platform's capacity, or if the initial partnership was an isolated experiment. Such factors might limit the ability of other businesses to join the platform, affecting the general applicability of the treatment effect.

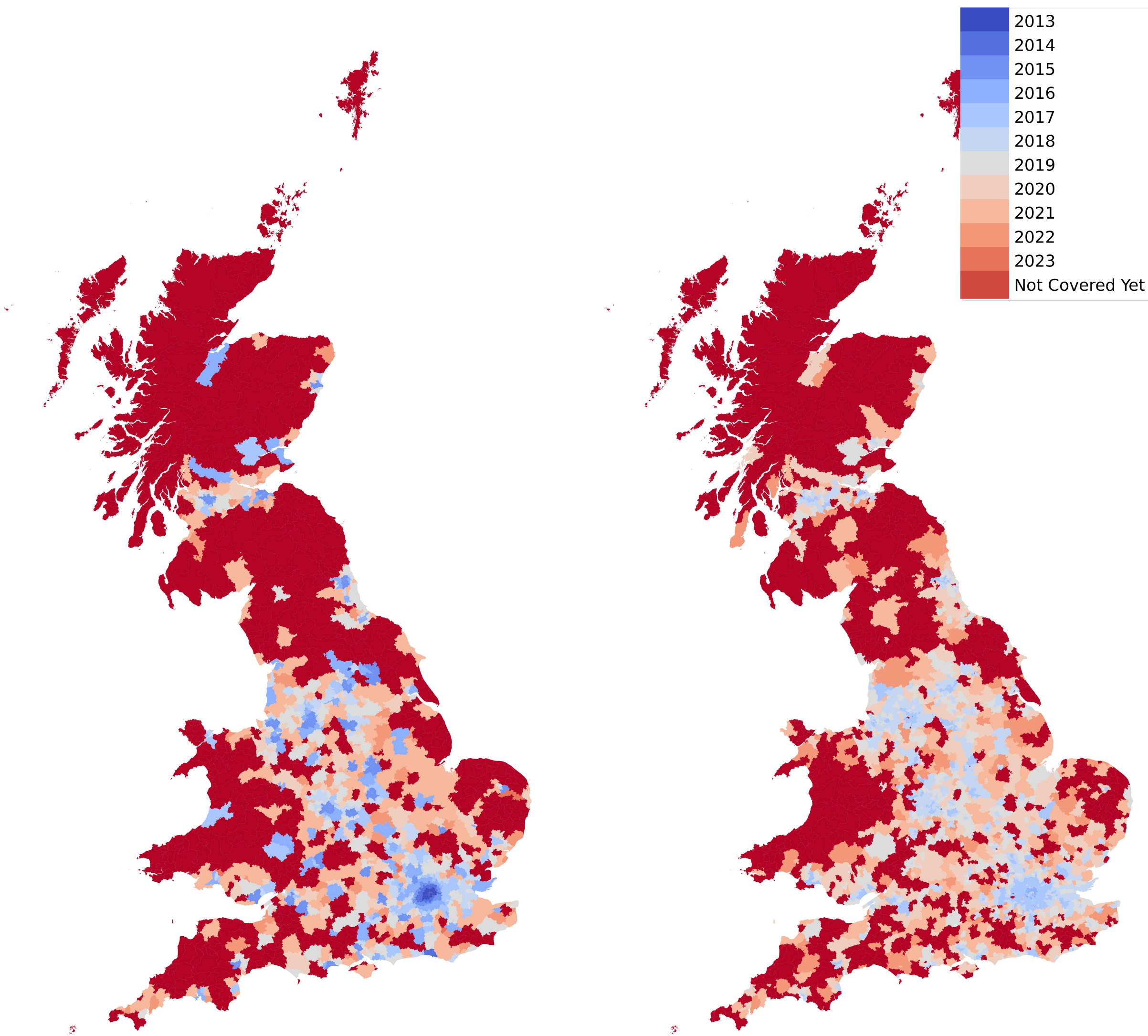
However, this issue, if present, likely understates the actual impact of treatment, as the measure does not fully capture the full scope of accessibility. This means any significant effects we find are conservative estimates. That said, to address this concern, I implemented three approaches to demonstrate that, on average, entrepreneurs in treated areas had access to partner with the platforms:

First, focusing on Deliveroo-for which I have detailed restaurant-level data before and after 2021-I analyzed the pattern of restaurant sign-ups following the platform's rollout. As shown in Figure A12, there is a significant spike in the number of restaurants joining Deliveroo during the initial rollout phase, followed by steady growth over time. This pattern indicates that the platform was onboarding multiple restaurants simultaneously, not just a single establishment. Therefore, it is unlikely that capacity constraints or isolated experiments prevented other entrepreneurs from partnering with the platform in treated areas.

Second, I redefined the rollout timing within postal districts by identifying the earliest time any restaurant within a five-mile radius of the district's centroid joined Deliveroo. The centroid was calculated using the weighted average of restaurant coordinates, with weights assigned based on the number of Google Maps reviews, giving more influence to highly-reviewed restaurants. The results remained consistent.[6]

Third, I assessed restaurant accessibility on both platforms across all UK postcodes, comparing postcodes in treated regions with those in control. Figure A13 confirms that treated postcodes have significantly higher accessibility, particularly at greater distances. When defined at the local authority level in panel (b), results remain similar, though the number of accessible restaurants decreases slightly in treated areas, which is expected due to the broader spatial classification. To further explore if the average number of accessible restaurants might mask variations in access across postcodes, I examined the extensive margin of access, specifically the share of postcodes with access to at least one Food App restaurant. Figure A14 shows that over 70% of postcodes in treated areas have platform access within 2km, compared to less than 10% in control regions.

[6] This analysis could not be conducted for UberEats, as regions treated before 2021 were identified at the region level rather than based on the presence of specific restaurants at the postal district level.



(a) Deliveroo Rollout

(b) UberEats Rollout

Figure 1.4. *Notes:* This map displays UK postal districts that have a minimum of one restaurant featured on Google Maps. Panel (a) depicts the introduction of the Deliveroo application, and panel (b) indicates the introduction of the UberEats application. The UK postal districts boundary file is sourced from: `https://longair.net/blog/2021/08/23/open-data-gb-postcode-unit-boundaries/`. A small number of postal districts could not be directly mapped due to updates in postal district definitions. These unmatched districts were associated with the closest matching district from the boundary file.

1.3.3 Validation of Rollout Date

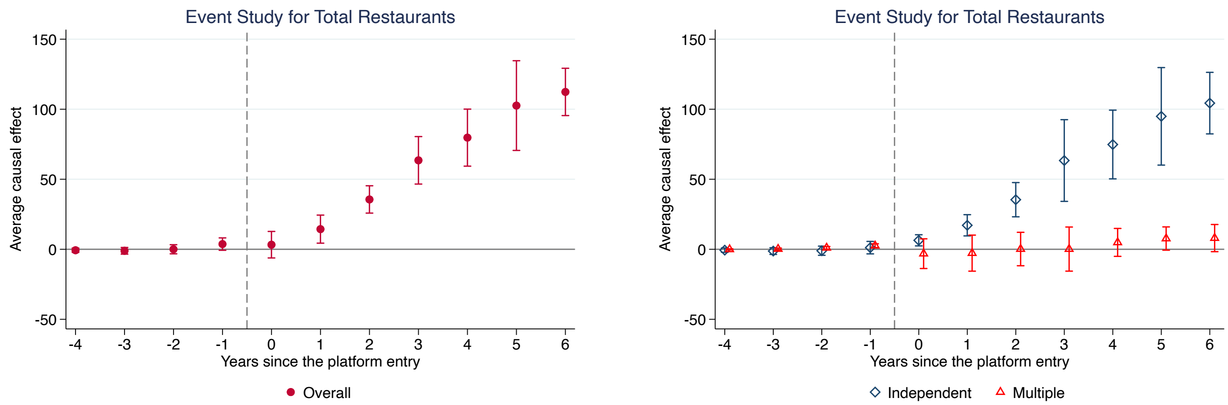
To assess whether the platform rollout measure captures not only access but also actual consumer usage, I conduct an event study regressing food delivery spending-measured using data from Fable and Kantar Worldpanel- on the staggered rollout measure, as specified in equation 1. The results, shown in Figures A15 and A16, indicate that spending on food delivery applications increases following their rollout in both datasets. This consistency provides robust evidence that the platform rollout measure accurately reflects consumer access and usage.

The graph shows near-zero pre-trend coefficients, ruling out two possibilities: mismeasurement of platform rollout, where consumers had access earlier than recorded, and meaningful use of platforms from non-residential locations, such as workplaces. This suggests that most orders are likely placed from home, or that both residential and workplace addresses gain access to food delivery applications around the same time.

In this chapter, I outlined the methodology for defining the market and pinpointing the rollout dates of the platforms. I showed that individuals in regions identified as having access to the platforms indeed have access to multiple restaurants. Furthermore, I used two spending datasets and showed my identified rollout dates align with the increase in spending on these platforms. In the following sections, I leverage this staggered rollout of platforms as a source of variation to explore the causal impacts of these e-marketplaces into three areas: first, on firms; second, on entrepreneurs; and third, on the product market.

1.4 Quasi-Experimental Effects of Platforms

This section explores the empirical evidence on how food delivery applications have transformed the restaurant industry. I structure the results into three parts: firms, entrepreneurs, and the product market. First, I analyze the firm-level effects, focusing on changes in the number of restaurants, including both new openings and closures, and examine how these trends differ across various types of establishments. Next, I turn to the entrepreneurs, examining the demographic and background characteristics of those who have benefited most from these platforms. If the costs associated with accessing digital platforms are more evenly distributed across demographic groups than traditional costs, these platforms could play a key role in making entrepreneurship more equitable. Then, I investigate the extent to which these impacts trickle down to the product market, influencing the variety of cuisines available to consumers. Finally, I explore underlying mechanisms and investigate the barriers in traditional brick-and-mortar settings that are potentially mitigated in an e-marketplace environment.



(a) Overall

(b) Independent vs Multiple

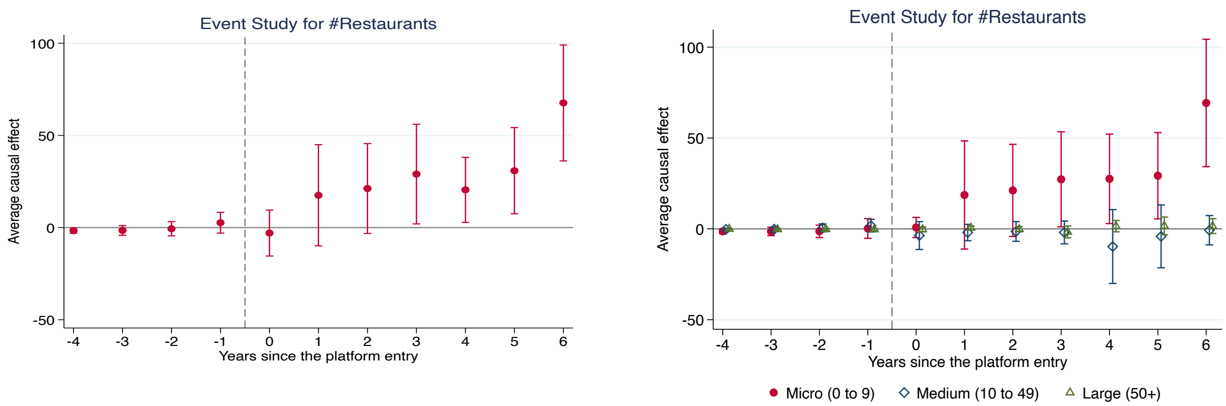
Figure 1.5. *Notes:* Panel (a) presents the average causal effect of food delivery application rollout on the total number of restaurants over time. Panel (b) shows the average causal effect on the number of independent versus multiple establishment restaurants. The x-axis represents the years since the platform rollout, and the y-axis shows the average causal effect. Data is sourced from the Local Data Company (LDC).

1.4.1 Impact on Firms: Restaurant Market Dynamics and Industry Expansion

I start by showing the impact on the number of restaurants. Figure 1.5 panel (a) shows a clear increase in the number of restaurants following platform rollouts in local authorities. On average, five years after the introduction of the platform, the number of restaurants rises by 100 units, which represents a 35% growth for the average local authority. These estimates are based on data from the Local Data Company (LDC) and control for local authority and year fixed effects, as well as interactions between local GDP and population by year.

Panel (b) highlights that this growth is driven by independent restaurants, with no significant change observed for chain restaurants, likely due to platforms' ability to provide essential infrastructure, such as delivery logistics and payment processing, which smaller establishments would otherwise struggle to afford. In contrast, establishments already well-known to consumers or firms that already enjoy high levels of brand recognition are less likely to benefit from the broader customer base provision that these technologies provide.

To accommodate zero values in the outcome variable, Equation 1 is specified and estimated in levels (although log transformations are still presented in Figure A17 for reference). However, to enhance interpretability, estimated level effects are converted into percentage changes. This transformation is done by calculating $P_j \equiv \hat{\beta}_j / E[\hat{y}_{st} | t = E_s + j]$, where $\hat{y}_{st}$ is the predicted outcome when omitting the contribution of the event dummies, i.e.,



(a) Overall

(b) By Size

Figure 1.6. *Notes:* Panel (a) presents the average causal effect of food delivery application rollout on the total number of restaurants over time. Panel (b) shows the average causal effect on the number of restaurants by size, categorized as micro (0 to 9 employees), medium (10 to 49 employees), and large (50+ employees). The x-axis represents the years since the platform rollout, and the y-axis shows the average causal effect. The data is sourced from an extract compiled from the Inter-Departmental Business Register (IDBR), accessed through NOMIS..

$$\hat{y}_{st} \equiv \sum_l \hat{\mu}_l \cdot I[l = s] + \sum_k \hat{\lambda}_k \cdot I[k = t].$$

Hence, P_j represents the period-j effect of platform rollout, expressed as a percentage of the outcome that would have occurred without platform presence. This approach follows the methodology used by Kleven *et al.* (2019). As shown in Figure A18, this model reveals a 35% increase in the number of firms.

To assess whether Food Apps disproportionately benefit small restaurants, Figure 1.6 explores the impact on the number of restaurants by business size, using data from the Inter-Departmental Business Register (IDBR). Unlike LDC data, which relies on field research, the IDBR compiles comprehensive business information from administrative sources like VAT and PAYE records, enhancing the validity of our findings. Panel (a) replicates and validates the previous analysis, while panel (b) indicates that smaller businesses, particularly those with fewer employees, experience the most substantial growth following the rollout of food delivery applications. This finding aligns with the notion that these platforms are particularly advantageous for small, independent restaurants, which can leverage the platforms to reach a wider audience without substantial capital investment.

Food Apps account for only a bit more than 10% of total restaurant sales (Figure 1.1), yet they have caused the number of restaurants to grow by 35%. To reconcile this we have to remember that new restaurants enabled by Food Apps do not necessarily rely exclusively on these platforms. Even a modest additional revenue from Food Apps can make opening a restaurant profitable. Second, and more importantly, Food Apps have led to the proliferation of small restaurants that, while accounting for a small share of sales, greatly contribute to the total number of establishments.

Consumer Spending Pattern Next, I examine how food delivery apps affect consumer spending patterns. A key question is whether consumers simply redistribute their existing spending across more restaurants, or whether food delivery applications stimulate additional spending, thereby expanding the market. The introduction of food delivery applications can create two main substitution effects: customers can either transition from dine-in to delivery, i.e., cross-channel cannibalization, or shift from home cooking to delivery, which expands the market.

To investigate this, I use Kantar's Worldpanel data and apply the same specification as in Equation 1, incorporating individual spending as the outcome variable with individual fixed effects. The results in Figure A22 show that Food App spending increases without reducing other types of restaurant spending. This suggests that the second substitution effect-shifting from home cooking to delivery-is driving the increase.

In other words, food delivery apps expand the overall market. For example, consumers may order delivery during bad weather, a situation in which they might otherwise avoid restaurant spending. This market growth is evident in higher spending in the restaurant industry (and correspondingly higher industry revenues) and an increase in the number of restaurants. It suggests that the effect of reduced barriers to entry dominates any potential superstar impact these platforms might induce.

Entry and Exit of Restaurants The previous analysis shows a net increase in the number of restaurants, but this result may be driven by different patterns of entry and exit. To fully understand how food delivery applications are affecting the market, it is important to look at these two factors separately. An increase in restaurants could come from high rates of new openings and few closures, or from a churn where many restaurants open but also close.

Figure 1.7 shows that both openings and closures have risen, likely due to heightened competition as food delivery applications expand. Less efficient restaurants may be forced out as consumers have more options and better ways to compare them. However, the number of new openings continues to exceed closures, leading to a net increase in restaurants, particularly among small, independent businesses (Figure A23). This suggests that while some restaurants exit, the reduced barriers to entry provided by food delivery applications-through delivery logistics, marketing, and payment systems-help new entrepreneurs enter the market, more than offsetting the rise in closures.

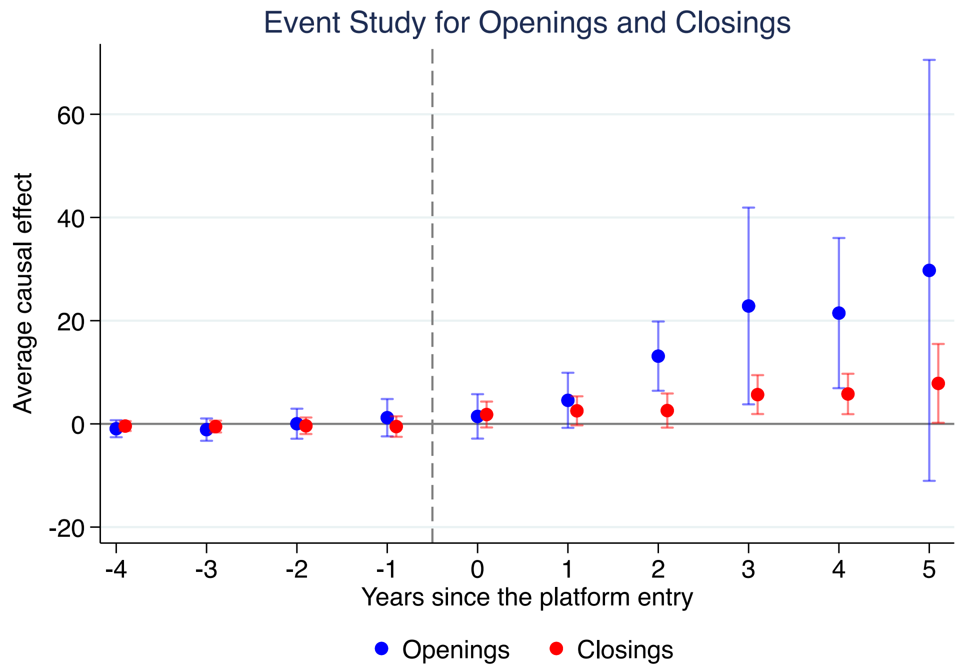


Figure 1.7. *Notes:* This figure presents the impact of food delivery applications on the number of restaurant closings and opening per year across Local Authority Districts (LADs). The analysis controls for postal district and year-fixed effects, as well as local economic indicators and population interacted by time. Data is sourced from the Local Data Company and covers the period from 2010 to 2023.

Employment I examine employment data to see if the increase in restaurants leads to more jobs. For example, employees of large chain restaurants may have left to start their own businesses, "stealing" customers from other establishments, leaving their previous positions unfilled. Higher competition might also push restaurants to cut costs, resulting in more establishments without additional jobs.

Figure A24 shows that employment in the restaurant industry has increased. Panel (a) shows an overall increase in employment, indicating that the rise in the number of restaurants is not driven by the aforementioned scenario but reflects genuine growth in the workforce.

Panel (b) suggests that this growth is even stronger for part-time positions, which aligns with the food delivery industry's need for flexibility. Delivery platforms face sharp demand peaks during mealtimes, while non-platform restaurants experience a steadier customer flow throughout the day as they cater to a broader range of needs. Non-platform restaurants also tend to have more predictable demand due to physical space cap and the deterrent effect of overcrowding, making sudden peaks less likely. These features make part-time workers a better fit for platform restaurants, which offer the flexibility to handle fluctuating order volumes without the long-term commitments of full-time staff.

Robustness I evaluate the stability of the findings through various robustness checks.

COVID-19. There is also a concern that the COVID-19 pandemic might distort the relationship between platform adoption and outcomes. While the pandemic certainly accelerated platform take-up, this is not necessarily an issue. On the contrary, it can be viewed as a useful force that induced variation in my treatment-food delivery applications. The key question is whether my results are driven by regions that adopted platforms due to COVID-19, and whether these regions are representative of the broader population. To address this, I re-estimate the main analysis without the COVID-19 period. As shown in Figure A19 panel (a), although the estimates become less precise, the magnitude of the effects remains consistent, suggesting that the pandemic does not drive the broader patterns in the data.

Excluding London. London is a unique region with higher restaurant densities and potentially different patterns of platform adoption compared to other areas. To ensure that the results are not disproportionately influenced by dynamics specific to London, I exclude local authorities within London from the analysis. Figure A19, panel (a), shows that the findings remain robust even without London.

Quantile Regressions. A limitation of this level-specification model is that results could be disproportionately influenced by larger values, especially in count outcomes like restaurant numbers, where distributions are often skewed and the mean may not accurately represent the central tendency. If there is impact heterogeneity-meaning the effects vary across different levels of the outcome variable-focusing on mean impacts could mask substantial differences at lower quantiles. To address this, I present quantile regressions of Equation 1, offering insights into the effect of platform rollout across the entire outcome distribution. While these regressions are based on a subsample, leading to somewhat wider confidence bands, the Figure in Appendix A1.1 demonstrates that median impacts are broadly similar to mean impacts. This finding mitigates the concern that our results are overly reflective of the upper tail of the distribution.

Winsorizing. Outliers in the data can have an outsized influence on regression estimates, particularly in skewed distributions. To address this, I Winsorize the data at the 5th and 95th percentiles, reducing the impact of extreme values. Figure A19, panel (a), demonstrates that the results remain consistent after Winsorization.

Alternative Estimators. To test the sensitivity of the results to different estimation approaches, I re-estimate the model using alternative difference-in-differences estimators, specifically De Chaisemartin and d'Haultfoeuille (2020), Callaway and Sant'Anna (2021), and Sun and Abraham (2021). These estimators account for potential variations in treatment effect assumptions. The results, consistent with the main specification, indicate that the platform rollout impacts are robust to changes in the estimation technique.

Placebo Tests with Similar Industries. To validate the exogeneity of the staggered rollout design, Figure A20 presents a placebo test, analyzing the platform's impact on other urban-related industries. No significant effects are found across sectors such as retail, cleaning, and hotels, suggesting the timing of platform rollout is not correlated with other local trends.

Expanded Placebo Tests Across All Industries. To address concerns about cherry-picking placebo industries, Figure A21 expands this test, showing the distribution of t-statistics for all three-digit SIC 2007 industries, with less than 5% of industries showing t-statistics higher than the restaurant sector. These outliers could potentially represent industries benefiting from externalities associated with a growing restaurant presence.

1.4.2 Impact on Entrepreneurs: Uneven Entrepreneurial Success Across Demographics

This section examines which demographic groups benefit the most from food delivery applications. To understand the uneven impact of these platforms, we must first explore the racial dynamics within the restaurant industry. Minorities are heavily represented in the restaurant workforce but their over-representation in the workforce does not seem to translate to entrepreneurship or managerial roles. As shown in Figure A25, although non-White British individuals make up less than 20% of the broader population, they account for nearly 40% of those employed in the restaurant industry. However, their share drops to below 20% in top managerial or ownership positions.[7]

Several factors can explain this disparity, highlighting the unique barriers that minority entrepreneurs face in setting up and owning restaurants. Limited access to finance and capital, often due to a lack of credit history, discrimination by financial institutions, or restricted access to networks, is a major challenge (Fairlie *et al.*, 2022; Bartlett *et al.*, 2022). Additionally, minorities may face discrimination in leasing commercial spaces, with landlords being less willing to rent to them or offering less favorable terms (Edelman *et al.*, 2017). This discrimination extends to regulatory hurdles and biased interactions with suppliers and customers (Combes *et al.*, 2016; Doleac and Stein, 2013; Leonard *et al.*, 2010). Cultural and language barriers further complicate the business environment,

[7]While our occupation classification does not directly identify ownership, it includes high managerial roles, which also encompass owner-managers.

particularly for migrant entrepreneurs (Azmat, 2013; Drori and Lerner, 2002). Navigating the regulatory landscape can be especially challenging for those unfamiliar with local laws and regulations. Furthermore, minority entrepreneurs are often confined to specific industries or market niches, such as ethnic food, where "niche entrapment" might limit their ability to expand into broader markets (Munshi, 2003; Patel and Vella, 2013).

Can Food Apps mitigate these disparities? Potentially, yes. In the theoretical framework, setting up a physical establishment involved uniform fixed cost barriers, while food delivery applications offered lower fixed costs, encouraging more entrepreneurship and leading to the creation of more firms, as observed in previous empirical analyses. However, food delivery applications not only lower fixed costs but also standardize them across demographics, effectively leveling the playing field and disproportionately benefiting those who face higher barriers in the physical setting. By reducing face-to-face interactions, Food Apps can limit discrimination and ease language barriers. They lower fixed costs, alleviating challenges related to raising capital and securing leases. Food Apps also offer broader customer access without the need for extensive marketing, helping minority entrepreneurs overcome traditional network and capital limitations. Integrated payment and logistics services further simplify regulatory navigation.

Minority Representation in Food Apps I begin by presenting descriptive evidence on the representation of restaurateurs from minority backgrounds on food delivery platforms. To do this, I matched data from the Company House directory with scraped data from Deliveroo and UberEats. The matching process is not straightforward due to discrepancies between trading and registered names or addresses. Despite this, I achieved a match quality of around 20%, focusing on matches where I am confident in their accuracy, minimizing false positives. The matching process is detailed in Section A1.3. Once matched, I inferred the backgrounds of restaurant directors based on their first and last names.

Figure A26 panel (a) shows that minority groups are more represented among restaurants partnered with food delivery applications. While British directors make up 50% of all restaurant directors, their representation drops to 22% in restaurants partnered with Deliveroo and UberEats, reducing to less than. In contrast, minority groups such as Middle Eastern, South Asian, East Asian, and African directors are more prominently represented. This finding is confirmed when analyzing nationality, as shown in panel (b). [8]

[8] There may be concerns that the matched sample of Deliveroo and UberEats restaurants is not representative of all restaurants on these platforms. However, this is unlikely to make minorities overrepresented on the platform; if anything, the opposite may be true. Minority-owned restaurants often use names that reflect their cultural heritage, which may complicate accurate matching using fuzzy algorithms. These names may include uncommon symbols, accented letters, or varying English spellings and transliterations.

While these descriptive patterns indicate that minorities are overrepresented on food delivery applications, this stylized fact does not imply a causal relationship. It is unclear whether food delivery applications actively increase opportunities for entrepreneurs from underrepresented backgrounds or whether other factors are driving this pattern. For instance, platforms may concentrate on areas with higher minority populations. The next section addresses this question through causal inference methods.

Food Apps' Impact on Minority Entrepreneurs I analyze the causal impact of food delivery applications across demographics of entrepreneurs. To do so, I employ the same dynamic event study framework. However, as I want to report and compare the net effect for entrepreneurs from specific backgrounds, I estimate the average treatment effect over time for different demographic groups. The impact for each demographic group (denoted as g) is estimated using the following equation:

$$y_{g,s,t} = \alpha + \sum_{g} \beta_g D_{st} \times \mathbb{I}(g) + \mu_s + \lambda_t + X_{s,t} \times \lambda_t + \epsilon_{g,s,t} \tag{2}$$

Here, $y_{g,s,t}$ is the number of ethnic group g directors in local authority s at time t, and D_{st} equals 1 if location s is treated at time t (i.e.,$t \geq E_s$). Unlike the previous specification, which estimates separate coefficients for multiple leads and lags, this approach focuses on a single coefficient, capturing the average effect over time for each demographic group. This choice allows me to report and compare net effects across groups more directly.

Figure 1.8 reveals significant variation in the platform's impact on different backgrounds, with entrepreneurs with African and Middle-Eastern-sounding names benefiting the most. This suggests that food delivery applications are effective in democratizing market access for immigrant entrepreneurs, providing them with a viable pathway to business ownership and success. I also examine the impact based on nationality. Figure A27 shows that immigrant entrepreneurs from the Middle East and Africa benefit the most, while European and British entrepreneurs benefit the least.

I further analyze the impact on entrepreneurs by gender and age. Appendix Figure A28 panel (a) shows no significant difference between female and male entrepreneurs. Panel (b) reveals that while the platforms offer opportunities across all ages, the impact diminishes for older age groups. This suggests that younger people, who may face more barriers in traditional restaurant operations while also having higher digital literacy, benefit from the platforms. Interestingly, there is also a significant but noisy positive impact on

Furthermore, minority-run businesses, particularly those owned by immigrants, are more likely to operate under trading names that differ significantly from their registered names or to undergo name changes or rebranding as they adjust their business models and update their information. This discrepancy between the names listed on UberEats and those registered with Company House could lead to a lower match rate.

Figure 1.8. *Notes:* The figure shows the impact of the platform on different background groups, reported as the percentage changes by computing $\Delta\hat{y}_m = \hat{\beta}_m / E(\hat{y_m}|D_{it} = 1)$, where $E(\hat{y_m}|D_{it} = 1)$ is the average predicted number of entrepreneurs from background m after the rollout of the platform when omitting the contribution of the treatment variable for the presence of the platform. The analysis controls for location and year-fixed effects, as well as local economic indicators and population interacted by time. Backgrounds are determined by inferring ethnicities from the first and last names of individuals using data from Company House.

the 60+ age group, hinting at a U-shaped relationship where both younger entrepreneurs with digital skills and older entrepreneurs with experience or capital gain the most.

1.4.3 Impact on the Product Market: Enhancing Cuisine Diversity

In this section, I examine whether the observed democratization of restaurant entrepreneurship extends to the product market, specifically affecting cuisine diversity and consumer options. To this end, I examine which cuisine types benefit most from Food Apps and how this affects cuisine diversity.

First, I provide descriptive evidence on the cuisine types of app-partnered restaurants. I compare restaurants on Food Apps to those not listed to examine differences in cuisine types. Using LDC data, I categorize restaurants by cuisine type based on specific keywords and match them with UberEats and Deliveroo listings. I use only trusted matches to ensure conservative estimates with a low false-positive rate. Figure A26 panel

(c) shows that non-UK cuisine types are more prevalent among app-partnered establishments. This observation is not necessarily causal; it may reflect a greater inclination of non-UK cuisine restaurants to join these platforms rather than Food Apps directly causing the creation of minority cuisine establishments.

To measure the causal impact of food apps on each cuisine type, I use the same research design based on the staggered rollout specified in Equation 2, but now the outcome variable is the number of restaurants offering a specific cuisine. I also follow the same technique and normalize the results by the predicted value of the outcome variable in the absence of treatment to estimate percentage changes in the number of restaurants specializing in each cuisine.

Figure 1.9 demonstrates substantial heterogeneity in the impact of food apps across different cuisine types, indicating that these platforms do not benefit all cuisines equally. (The corresponding results in levels are provided in Appendix Figure A35.) Interestingly, the variation in cuisine benefits partially corresponds to the patterns seen in entrepreneur demographics in Figure 1.8. For instance, cuisines from African and Middle Eastern regions benefit significantly, much like the entrepreneurs from these areas, while European cuisines and entrepreneurs show comparatively smaller effects.

To explore the reasons behind this pattern, I consider the concept of homophily. In the context of entrepreneurship, this implies that entrepreneurs may choose to offer cuisines that match their cultural backgrounds. This might be because individuals from a particular region have a comparative advantage in establishing restaurants that serve their native cuisine, due to specialized knowledge, skills, and cultural capital. Such comparative advantage arises from possessing specific human capital, including cooking techniques, traditional recipes, and cultural understanding. If Food Apps disproportionately assist underrepresented groups, and if homophily holds, we would expect a wider variety of cuisines available to consumers.

Homophily I start by showing the distribution of different cuisine types, inferred from Google Maps, based on the entrepreneurs' backgrounds inferred from their names on the Company House. Figure A34 shows that restaurant directors are more likely to offer cuisines that align with their own backgrounds. For example, more than 50% of Middle Eastern restaurants are run by people with Middle Eastern-sounding names while only close to 10% of them have white-sounding names.

To quantify the extent of this homophily, I conduct regression analyses. The specification is as follows:

$$y_{i,g} = \alpha_g + \beta_g \times \mathbb{I}\left[i \in g\right] + \epsilon_{i,g}, \quad \forall g$$

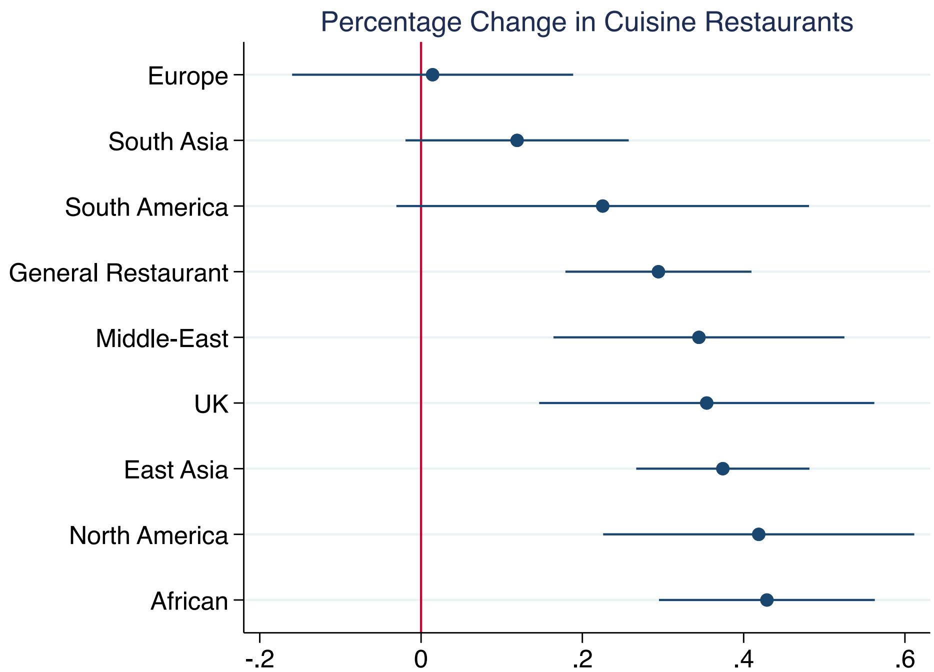


Figure 1.9. *Notes:* The figure shows the impact of the platform on different cuisine types, reported as the percentage changes by computing $\Delta\hat{y}_m = \hat{\beta_m}/E(\hat{y_m}|D_{it} = 1)$, where $E(\hat{y_m}|D_{it} = 1)$ is the average predicted number of cuisine m restaurants after the rollout of the platform when omitting the contribution of the treatment variable for the presence of the platform. The analysis controls for location and year-fixed effects, as well as local economic indicators and population interacted by time. Cuisine types are categorized as outlined in Table A8. Data is sourced from Companies House.

In this equation, $y_{i,g}$ is a dummy variable indicating whether the restaurant director i belongs to the group g, and $\mathbb{I}[i \in g]$ is a binary variable indicating whether the restaurant's cuisine corresponds to the group g. I run this regression for each background group. The directors' backgrounds are inferred from Companies House data, while the restaurant's cuisine type is based on Google Maps listings. I also perform this separately for the subset of directors of app-partnered restaurants.

The results, shown in Figure 1.10, reveal positive coefficients across all groups, confirming homophily among all demographics. Some groups exhibit stronger correlations; for instance, South Asian restaurants are 30% more likely to have a director with a South Asian-sounding name. Interestingly, the degree of homophily seems to be stronger for overall restaurants compared to those on food delivery applications [9].

[9] There might be several reasons for this. First, in traditional restaurants, face-to-face interactions make having a background that aligns with the cuisine type crucial for creating authenticity and signaling expertise, something missing in the online framework. Second, traditional brick-and-mortar restaurants typically

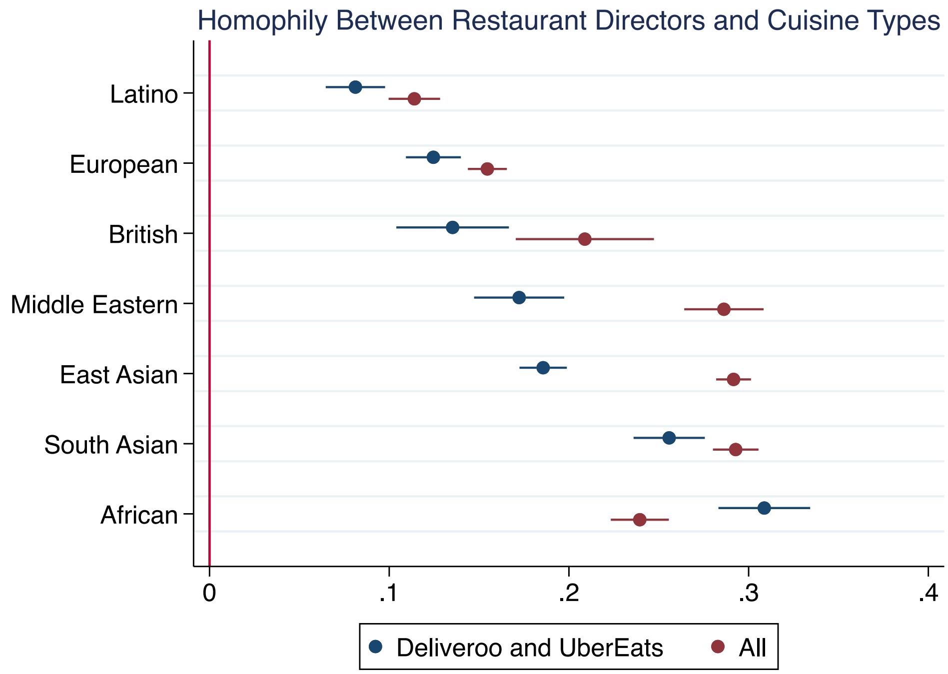


Figure 1.10. *Notes:* This figure illustrates the degree of homophily between restaurant directors and the cuisine type of their restaurant. Directors' backgrounds are inferred from their first and last names as described in the text, using data from Companies House. Restaurants listed in Companies House are matched to Google Maps based on name and postcode, with cuisine type inferred from Google Maps information. The probability of a director having a Muslim background is linked to Middle Eastern cuisines, and the probability of having a Hispanic background is linked to South American cuisines. Probabilities of having European and East Asian probabilities are the maximum values derived from subcategories (e.g., various European nationalities, East Asian, and Japanese), based on name analysis.

Cuisine Diversity To determine whether the heterogeneous impact across cuisine types leads to greater overall diversity, I employ different measures to quantify the diversity of restaurant offerings within postal districts over time. These measures are designed to capture both the concentration and the variety of cuisine types available to consumers.

First, I calculate the Herfindahl-Hirschman Index (HHI) to assess the market concentration of different cuisine types within each spatial unit and time period. In the absence of revenue or sales data that has information on the cuisine type, I use the proportion of restaurants belonging to each cuisine type as a proxy for market share. The HHI is computed using the formula:

attract customers from their local neighborhoods, with the surrounding area's ethnic composition influencing the types of restaurants that succeed. In contrast, food delivery applications weaken the tie between location and clientele, allowing them to serve diverse audiences beyond their local ethnic communities.

$$\text{HHI}_{st} = \sum_{i=1}^{K_{st}} \left(\frac{n_{ist}}{N_{st}} \right)^2$$

where n_{ist} is the number of restaurants of cuisine type i in postal district s at time t, $N_{st} = \sum_{i=1}^{K_{st}} n_{ist}$ is the total number of restaurants in postal district s at time t, and i indexes the different cuisine types, ranging from 1 to K.

A higher HHI indicates greater concentration (less diversity), while a lower HHI suggests a more diverse culinary landscape within the postal district. I define cuisine types at two levels of granularity-for example, one broad category like "Indian" and another more specific, such as "South Asian.".

Second, I calculate the number of distinct cuisine types present in each spatial unit and time period. This measure is defined as:

$$D_{st} = \left| \left\{ i \;\middle|\; n_{ist} > 0 \right\} \right|$$

where the notation $|\cdot|$ denotes the size of the set, and the set $\left\{ i \;\middle|\; n_{ist} > 0 \right\}$ includes all cuisine types i for which there is at least one restaurant in spatial unit s at time t. This measure provides a straightforward count of the variety of cuisine types, offering insight into the breadth of options available to consumers. I calculate this measure using both granular and broader classification.

Using the same research design as before, I now consider each of the four measures of cuisine diversity as the outcome variable. Table 1.1 presents the results for both levels of categorization. The findings indicate that the rollout of food delivery applications leads to an increase in cuisine diversity across all metrics: the HHI decreases, and the number of distinct cuisine types increases.

Robustness A potential concern with the disproportionate benefit of food delivery applications for ethnic minority cuisines is that the observed relationship may be driven by demographic shifts, with platforms expanding in already diverse regions that are attracting more migrants. This could suggest a spurious correlation rather than a causal effect. To address this, I perform two analyses. First, I restrict the sample to British nationals, who are either born in the UK or have lived there for many years, to ensure that the increase in minority cuisines is not simply due to new migration. The results remain consistent, indicating that the rise in minority cuisines is not tied to recent demographic changes. Second, I conduct a placebo analysis by examining spending on items from grocery stores indicative of specific cuisine types, such as falafel for Middle Eastern cuisine or curry for South Asian cuisine. If the demographic shifts were driving the observed increase in restaurant cuisine types, we would expect to see a similar effect. However, as

Table 1.1. Platform and Diversity, Postal District Analysis

	Broad Cuisine Categories		Detailed Cuisine Categories	
	HHI (Cuisine)	#Cuisine Types	HHI (Cuisine)	#Cuisine Types
FDP	-0.022	0.174	-0.021	0.422
	(0.004)	(0.033)	(0.004)	(0.066)
Mean of dep. variable	.339	4.89	.215	11.9
Location FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Spatial units	2050	2052	2050	2052
Observations	26745	28728	26745	28728

Notes: This table presents the impact of food apps on cuisine diversity at the postal district level, measured using different diversity metrics. The columns show results for both broad and detailed cuisine categories, using the Herfindahl-Hirschman Index (HHI) for cuisine concentration and the number of cuisine types. The analysis controls for location and year fixed effects, as well as local economy indicators interacted by time. Broad cuisine types are categorized as outlined in Table A8. Data is sourced from the Local Data Company (LDC).

Figure A36 shows, there is no significant correlation between platform rollout and spending on these items.

1.4.4 *Mechanisms Behind Differential Entrepreneurial Impact*

I explore why food delivery applications disproportionately benefit minority entrepreneurs. While I cannot empirically test each mechanism individually, I provide evidence on key factors that may drive these outcomes.

There are two main reasons why minorities might benefit more. The first mechanism aligns with the theoretical framework presented earlier, where setting up a physical establishment involves uniform fixed costs for all entrepreneurs. In this framework, food delivery applications reduce these fixed costs equally for everyone. Under this mechanism, minorities might have been less productive or less able to cover the high fixed costs in the past, and the lowering of these costs now allows them to enter the market.

The second mechanism involves the idea that different demographics face different levels of fixed costs in traditional settings. Minority entrepreneurs may have faced higher barriers than others, but they were not less productive. Food delivery applications not only lower fixed costs but also harmonize them across demographics, effectively leveling the playing field. That is, if the costs associated with digital platforms are more evenly distributed across backgrounds compared to traditional restaurant costs, then delivery platforms promote a more equitable business landscape across diverse demographics.

Each explanation has different policy implications. The first suggests a need for policies focused on improving productivity among minority entrepreneurs through training and resources. The second implies that policies should focus on removing structural barriers and reducing discrimination that lead to higher barriers for minorities.

To investigate these hypotheses, I focus on three key areas of evidence. First, I analyze restaurant productivity to assess whether minority-run restaurants are less productive, which assesses the first mechanism. Second, I examine the racial composition of customers to explore whether changes in customer demographics contribute to the benefits minorities receive, shedding light on potential reductions in discrimination and face-to-face biases. Third, I assess the impact in less expensive versus more expensive parts of the same neighborhood to see if operating in more affordable locations explains the disproportionate benefits to minorities. This considers whether capital constraints and difficulties in securing prime locations hindered minority entrepreneurs before. Both the second and third pieces of evidence relate to the second mechanism, where food delivery applications reduce and harmonize barriers that previously disadvantaged minorities. Lastly, I discuss a potential alternative mechanism involving the reduction of search frictions in the product market.

Productivity Measuring productivity in this context is challenging. Official datasets, such as business structure databases that report revenue and employment, allow the construction of various productivity measures but are anonymized and thus cannot be linked to other datasets. Acknowledging this limitation, I use average Google review scores as a broad proxy for restaurant productivity. I infer minority ownership by matching restaurant listings to Companies House data and identify platform presence through matches to platform listings. While Google reviews are an imperfect measure of productivity, they offer the best available alternative, as they capture consumer satisfaction and can be linked to entrepreneurs' backgrounds and platform use.

Table A6 shows that minority-run restaurants on food delivery apps are not less productive than others. Specifically, the results from Column (3) indicate that minority-run restaurants on food delivery apps do not exhibit significantly lower productivity, compared to either minority-run restaurants not on the app or non-minority-run restaurants on the app. The interaction term (Minority $\times$ Food App) of 0.11 offsets much of the review penalty associated with minority ownership (-0.16) and app presence (-0.13). This suggests that minority-run restaurants on food delivery apps perform similarly to their counterparts.

Racial Composition of Consumers To determine whether food delivery applications help minority-run restaurants overcome barriers or biases in face-to-face interactions, I analyze the racial composition of customers. If food apps mitigate such biases, we might expect minority-run or ethnic cuisine restaurants to attract a different racial composition of customers, particularly reaching more white customers.

For this analysis, I utilize a dataset of over 6 million Google Reviews left on UK restaurants. Google Reviews provide a rich source of customer feedback and include reviewer names, which I use to infer racial backgrounds. Additionally, by identifying platform-specific keywords in the reviews, I can determine whether an order was placed through a food delivery application. The result of the racial inference of customers depends on the demographic composition of reviewers and their propensity to leave reviews. While the tendency to leave reviews might vary among different ethnic groups, it is unlikely to differ systematically across different cuisine types within each ethnic group.

First, I conduct a non-parametric analysis to compare the racial distribution of customers across app-partnered and non-partnered restaurants. I find that the racial makeup of customers remains consistent across app-partnered and non-partnered restaurants, as well as between Food App and offline customers of app-partnered establishments, regardless of the ethnic cuisine offered. Figure A29 illustrates this non-parametrically, with the bars showing the racial profile of customers for non-partnered restaurants, app-partnered restaurants, and a subset of app-partnered restaurant customers confirmed to have placed orders through Food Apps.

To further investigate this issue, I perform a parametric analysis focusing specifically on white customers. To ensure a clear distinction between the two cases, I include reviews from restaurants not on the platform and reviews from app-partnered restaurants where the order is explicitly linked to a food delivery application. I regress an indicator variable for whether the reviewer is white on a set of dummy variables for different cuisine types and interaction terms between these cuisine types and an indicator for reviews associated with platform orders. The regression specification is as follows:

$$W_{ir} = \alpha + \sum_k \beta_k \text{CuisineType}_k + \sum_k \gamma_k (\text{CuisineType}_k \times \text{PlatformOrder}_i) + \epsilon_{ir} \quad (3)$$

where W_{ir} is an indicator variable equal to 1 if reviewer i of restaurant r is inferred to be white, CuisineType_k are dummy variables for each cuisine type k, PlatformOrder_i is an indicator variable equal to 1 if the review includes keywords suggesting it is about an order placed through the platform. Coefficients γ_k in this specification allow me to test whether the likelihood of a reviewer being white differs for platform orders compared to orders for restaurants that are not on the platform, across various cuisine types. A positive and significant coefficient on the interaction terms would indicate that platform orders are associated with a higher proportion of white customers for ethnic cuisines.

The results, depicted in Figure A31, indicate that the interaction terms γ_k are generally insignificant or, in some cases, negatively significant. This indicates that there is no evidence to suggest that food delivery applications help ethnic cuisine restaurants attract more white customers. For instance, the argument that Food Apps allow minorities to overcome face-to-face racial biases-by attracting customers who might otherwise avoid visiting their establishments in person-is not supported by this evidence. Also, the notion that Food Apps enable minority-run restaurants to enter predominantly white neighborhoods, which they might otherwise avoid, seems unlikely given the consistent racial makeup of users across platforms.

Therefore, the benefit to minority entrepreneurs is not from reaching different racial groups. Minority-run restaurants probably reach a wider geographic customer base using food apps, but not a different demographic one.

Differential Effects in Low-Cost Areas

I examine whether Food Apps allow entrepreneurs to open restaurants in cheaper, more deprived areas while still reaching customers. This option may have been less feasible in traditional settings.

To compare the net effect across different levels of physical space price within the same postal district, I divide each postal district into four units based on the quartile of the property price index and estimate the following equation:

$$y_{s,j(s),t} = \alpha + \sum_{j=1}^{4} \beta_j D_{st} \times \mathbb{I}\left[j(s) = j\right] + \mu_s + \lambda_t + X_s \times \lambda_t + \epsilon_{s,j(s),t} \tag{4}$$

Here, $y_{s,j(s),t}$ represents the number of restaurants in postal district s, quartile $j(s)$ of postal district s, at time t. D_{st} is an indicator variable equal to 1 if the area had access to the food delivery application at time t.

Figure 1.11 highlights the differential impact of food delivery application rollouts on restaurant numbers across various price level quartiles within the same postal district for White British vs non-White British. The largest increase-over 100%-occurs in the least expensive areas for non-White British entrepreneurs. For White British entrepreneurs, the impact is negligible, except for a positive effect in the most expensive areas.

A likely explanation is that minority entrepreneurs often face credit constraints and higher barriers to entry in high-rent areas. Food Apps reduce the need for prime locations, allowing restaurants to operate in lower-cost areas without losing access to customers. This enables minority entrepreneurs, who may have less capital, to open businesses with

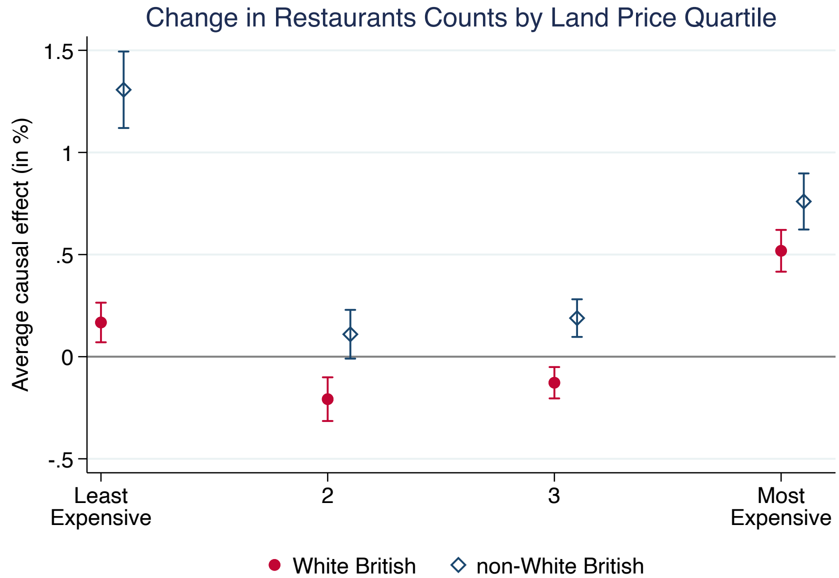


Figure 1.11. *Notes:* This figure illustrates the average causal effect of food delivery application rollout on the demographics of entrepreneurs categorized as White British and non-White British, across different land price quartiles within postal districts. The x-axis represents land price quartiles, while the y-axis shows the average causal effect in percentage terms. The estimation is based on the specification in 4, controlling for postal district and year fixed effects, as well as local economic indicators and population trends interacted with time. The data is sourced from Companies House, covering the period from 2010 to 2023.

lower fixed costs. This explains why non-White British entrepreneurs see a larger positive impact in less expensive areas.[10]

While the within-district analysis highlights the tendency for businesses to cluster in cheaper areas within a postal district, I also examine the broader effects across entire postal districts, treating each district as a single unit. Figure A33 confirms that the impact is stronger in cheaper districts and those with lower IMD levels. Combined, these findings indicate that Food Apps facilitate restaurant openings in both cheaper districts and the least expensive areas within districts.

Alternative Mechanism: Reduction in Search Frictions

An alternative explanation for the increase in cuisine diversity is that food delivery applications reduce search costs for consumers, directly impacting the product market. In traditional settings, consumers face high search costs when seeking products that match

[10]To address the limitations of Company House data, which excludes unincorporated businesses and uses registered addresses, I supplement it with LDC data, which includes both incorporated and unincorporated businesses and focuses on trading addresses. However, LDC lacks information on director ethnicity and cannot distinguish results for minority versus non-minority groups. Figure A32 shows stronger effects in cheaper neighborhoods and areas with lower IMD levels within postal districts.

specific preferences, especially niche or minority-preferred products like ethnic cuisines. Digital food delivery applications lower these search frictions by providing a centralized marketplace where consumers can easily find and evaluate a wide array of culinary options. Thus, platforms increase the visibility and accessibility of niche cuisines. This aligns with the "long tail" effect in the literature (Brynjolfsson *et al.*, 2006, 2011), which suggests that reducing search frictions disproportionately benefits niche products by connecting them with consumers who have specific tastes.

This mechanism suggests that the platforms first enhance the product market by making niche cuisines more accessible. As a result, entrepreneurs who specialize in these cuisines-often ethnic minorities-benefit from increased demand and choose to enter the market. Thus, the impact on entrepreneurs is a consequence of the initial effect on the product market.

Disentangling this mechanism from the reduction in the entry barrier effect discussed in this paper is challenging due to homophily-the tendency of individuals to produce goods aligned with their cultural backgrounds, which I documented. Because entrepreneurs often offer cuisines that reflect their heritage, any change in the product market naturally correlates with changes in the entrepreneurial landscape. This interdependence makes it difficult to determine whether the primary driver is reduced search frictions for consumers or lowered entry barriers for entrepreneurs.

Despite this complexity, several factors suggest that reduced search frictions are a less likely primary driver in this context. First, many cuisines that benefited from food apps, such as East Asian and British cuisines (Figure 1.9), are not niche in the UK, as shown by their prevalence (Figure A26, Panel (c)). Second, looking at where the product market and entrepreneurial impacts diverge uncovers an important insight. For instance, while British cuisine benefits modestly from platform expansion, White British entrepreneurs do not experience similar gains. If reduced search costs were the primary driver, the growth in British cuisine should lead to more opportunities for White British entrepreneurs. Instead, the benefits accrue disproportionately to ethnic minority entrepreneurs, even within British cuisine. This divergence implies that the platforms are specifically helping minority entrepreneurs, rather than changes in consumer search behavior alone.

Therefore, although I cannot fully rule out the alternative mechanism and acknowledge that both effects may coexist, the evidence points toward the reduction of entry barriers for minority entrepreneurs as the more significant factor. The platforms enable these entrepreneurs to overcome traditional obstacles, leading to increased representation in the market and a corresponding rise in the supply of ethnic cuisines.

1.5 Conclusion

Food delivery applications exemplify a broader shift in business and employment structures offered by digital marketplaces. This shift is characterized by lowered entry barriers that democratize market access, but it also introduces mechanisms that could lead to superstar effects. This paper delves into how these digital marketplaces influence the restaurant industry, particularly in democratizing market access. At the firm level, I find that these platforms reduce entry barriers, leading to an increase in restaurant numbers, driven by the entry of small, independent businesses. At the entrepreneur level, the most significant benefits are seen among minority and migrant entrepreneurs. For the product market, this democratization results in greater cuisine diversity, offering consumers a wider range of choices.

Why do minority entrepreneurs benefit more? It is not because minorities are less productive and only enter when barriers are low; the evidence indicates that minority-run restaurants on these platforms are just as productive as their counterparts. Instead, the key factor is that digital marketplaces not only reduce but also standardize and harmonize entry barriers. Thus, groups that face more challenges in traditional settings-like credit constraints-gain the most.

To understand the mechanisms behind this phenomenon, I examine whether reaching new customer demographics contributes to the benefits received by minority entrepreneurs. Analysis of the racial composition of customers shows no significant changes, suggesting that platforms do not primarily help minorities by expanding their customer base to different racial groups. Instead, the evidence indicates that platforms enable minority entrepreneurs to overcome the high costs of securing prime property locations. Food delivery apps make it viable for them to operate in lower-cost areas within the same neighborhoods, consistent with the notion that they face challenges in the brick-and-mortar context, such as credit constraints.

The results presented in this paper should be interpreted cautiously for several reasons. Firstly, my estimates are relatively short-term. On the consumer side, currently, younger, wealthier users drive food delivery application growth, but as other demographic groups adopt these services, their preferences could shift. For example, they might substitute food delivery for dine-in options, potentially reducing its overall impact. On the platform side, platforms may gradually achieve monopolistic positions, potentially altering their interactions with both restaurant owners and users.

Furthermore, this study does not capture the overall welfare effects of food delivery applications. A comprehensive assessment would need to consider various dimensions, including potential benefits such as reducing time spent on food preparation, supporting new work arrangements like working from home and creating employment opportunities for couriers, as well as potential downsides like health impacts.

How generalizable are these findings to other digital marketplaces? Restaurants, with their short life cycles and unregulated spatial patterns, are highly sensitive to urban changes, making them ideal for studying the economic impact of digital platforms. However, limitations exist. First, the two forces examined-lowered entry barriers and superstar effects-may behave differently elsewhere. While reduced entry barriers are common across digital marketplaces, the level of market integration varies. In food delivery, restaurants compete locally due to the perishable nature of the product. This explains why, despite some exits from the market, the reduction in entry barriers outweighs these exits, leading to a net increase in restaurants. In contrast, digital marketplaces like Amazon or Google Play involve national or global competition, potentially allowing a few large firms to dominate and limiting the benefits of lower entry barriers for smaller businesses. Second, the link between entrepreneurship and product diversity may be less direct in other contexts. In the restaurant industry, homophily ensures that increased minority entrepreneurship leads to more diverse offerings. This connection may be weaker on platforms like Amazon or Google Play, where products are less likely to reflect the backgrounds of entrepreneurs.

How can policymakers promote entrepreneurship, especially among underrepresented groups? This research demonstrates that the digital marketplace plays a pivotal role in transforming the food service industry by reducing entry barriers and fostering entrepreneurship. To harness these benefits, policymakers should focus on increasing digital literacy and providing the necessary infrastructure to ensure access to these platforms. Investing in secure and uninterrupted internet connectivity is particularly relevant for developing countries. Finally, addressing the specific challenges of traditional settings remains essential to creating a more equitable entrepreneurial landscape.

Chapter 2

Understanding multi-layered sanctions: A firm-level analysis

Sanctions are not just rising in popularity as a foreign policy instrument, but they are also growing in complexity. Recent examples in Russia and Iran demonstrate this trend, encompassing a range of measures from asset freezes to trade restrictions and targeted actions against key sectors and individuals. These measures often face unpredictable application, unclear interpretation, and evolving nature. Targeted governments may also respond with their countermeasures, including tailored macroeconomic tactics, subsidies, governmental contracts, and loans, further increasing uncertainty and complexity. As a result, corporations and individuals in target countries encounter diverse and often unforeseen challenges due to these sanctions, with the extent of impact varying according to their distinctive attributes and the nature of their operations.

The growing complexities of sanctions introduce added challenges for researchers attempting to address key policy-driven questions, such as the effectiveness of sanctions, their impact on the incentives of targeted countries, and the extent of their collateral damage. In scenarios with defined sanctions on specific entities, establishing distinct treatment and control groups is straightforward. However, as sanctions evolve to be more multifaceted and intertwined, discerning which firms are impacted becomes ambiguous. Without clear knowledge upfront about which firms are subject to sanctions, drawing a direct link between their performance and the sanctions becomes a formidable task. This obscurity also hinders assessments of whether sanctions successfully target political leaders' interests and persuade them to adjust their actions.

To account for these factors, we need a flexible framework that can incorporate these numerous, potentially ex-ante unknown, channels, capturing their interactive influences at equilibrium. This effort is further hampered by the scarcity of reliable data from sanctioned countries, often attributable to their lack of transparency. Compounded by political constraints, conducting surveys in such environments might be impracticable. Intriguingly, this opacity might itself be endogenous to the imposition of sanctions, as they may not be willing to disclose the extent to which sanctions have impacted them.

In this paper, I overcome these challenges by utilizing a text-based methodology to quantitatively measure the impact of sanctions on individual Iranian firms-a nation heavily under sanctions-using stakeholder perceptions. To this end, I first use a training library of sanction-related articles and a training library of non-sanction text to find two-word combinations (bigrams) that are frequently used in sanction-related texts. I also assemble a unique dataset composed of transcripts and reports from board meetings of publicly traded Iranian firms. I then use a natural language processing method to quantify sanctions exposure by counting instances of sanctions-related bigrams in discussions between firm management and financial analysts, with each bigram assigned a weight that reflects its relative importance to sanctions.

This approach is inspired by studies that aim to measure a firm's exposure to specific shocks, such as political risk, COVID-19, Brexit, and climate change (Hassan *et al.*, 2019, 2024, 2023; Sautner *et al.*, 2023).[11] The premise here is that company meetings serve as a forum for management to discuss current issues and for analysts to probe the company's challenges and thus offer a wealth of valuable information. Significant sanctions exposure, due to any reason like reliance on international supply chains or competition with imported substitutes, is likely to emerge in these dialogues.

This method offers a subjective risk metric, allowing the measurement of a firm's sanctions exposure without resorting to executive surveys, which are often impractical in the context of sanctioned countries. Given the intricate and multi-layered nature of comprehensive sanctions and the associated challenge of categorizing clear treatment and control groups from sanction documents, the flexibility of this approach stands particularly useful.

Using these new measures, I present a series of novel empirical findings. First, the average Iranian firm reports significant challenges due to sanctions. My main measure of sanction exposure, averaged across firms, intuitively evolves over time, reaching its apex in 2018 following the announcement from the Trump administration regarding its departure from the Joint Comprehensive Plan of Action (JCPOA) and imminent re-imposition of sanctions. It also intuitively fluctuates across industries, which further attests to its validity. Industries with deep ties to international supply chains, partnerships, and markets-such as architectural and engineering activities, technical testing and analysis, and computer electronics manufacturing-score high on this scale. In contrast, sectors like sports,

[11] Hassan *et al.* (2019) uses computational linguistics to measure U.S. firms' political risk via earnings conference calls, revealing heightened discussions during peak political risk periods. Hassan *et al.* (2024) employs a text-based method to capture the global impacts of Brexit uncertainty, highlighting anticipated regulatory and trade challenges. Hassan *et al.* (2023) determines firms' primary concerns about COVID-19, illustrating simultaneous demand and supply shocks. Sautner *et al.* (2023) leverages machine learning to gauge firms' attention to climate change exposures, predicting green innovation outcomes.

amusement, and creative arts and entertainment, which are less involved in global trade, register much lower values. To provide perspective on the severity of this shock, I compare it to the concern surrounding COVID-19. At its peak, the sanctions concern was 20% more severe than that of the COVID-19 shock, underscoring the substantial risk that sanctions represent to Iranian firms.

Most sanction proponents justify their use on the basis of providing incentives for policy reform for political decision-makers in the sanctioned country, as sanctions can be lifted in exchange for policy changes. According to this idea, modern sanctions should ideally target the economic interests of elite decision-makers while sparing non-decision-makers. However, my second finding suggests that with increasingly more complex sanctions, the idea of "targeted sanctions" appears to be a misnomer. Instead, sanctions impact politically-connected and non-connected firms alike, implying that sanctions may operate as 'blunt instruments', affecting the broader economy. I find for every $\$1$ loss inflicted on connected firms, an externality of $\$5$ is imposed on non-connected firms, primarily because non-connected firms represent a more substantial segment of the market.

Third, I examine the extent to which sanctions adversely affect Iranian firms. Initially, I study stock market reactions to unexpected sanction-related events. To do so, I utilize search intensity data for the topic "Sanctions against Iran" on Google Trends, and identify eight major events related to sanctions on Iran. For each identified event, I conduct an event study to assess the abnormal return of firms with higher exposure to sanctions. The results show a robust and quantitatively large impact of unfavourable news about sanctions on the returns of firms exposed to sanctions. Furthermore, I assess firm-level performance, showing that sanctions reduce firms' sales, and investments. Interestingly, the impact on hiring was relatively muted, a finding consistent with the notion that employment costs are often sticky in the short term, and aligning with results in Salehi-Isfahani (2023).

I next turn to investigating the potential mechanisms through which sanctions might operate. I undertake a systematic human audit with the help of two trained experts. These experts were recruited from PhD students specializing in Economics at Sharif University of Technology. These human auditors scrutinize the text fragments that underlie my sanction scores to pinpoint the specific channel through which sanctions impact the firm's associated decisions. The findings suggest that the most potent channels are the limitation of exports and the escalation of import costs.

I address three main concerns that could challenge the causal interpretation of my results. One potential challenge is that corporate executives could use the threat of sanctions as an excuse for underperformance. This challenge is addressed by turning to the stock market. If mentions of sanctions were merely a form of deception or cheap talk, then the stock market should not price sanction exposure during the advent of unanticipated sanction news. Reassuringly, the observed results suggest otherwise.

The second concern extends the first, focusing on how politically connected firms may also refrain from openly discussing sanctions due to their already familiarity with associated risks or political considerations, resulting in fewer references to sanctions in their case. To explore whether these firms do systematically underreport their sanction exposure, I again analyze stock market reactions. The analysis reveals that in the wake of unexpected sanction news, the market valuations of politically connected firms are adjusted in a way that aligns with their actual exposure to sanctions, similar to other firms. This suggests that the market acknowledges and factors in the vulnerability of even politically connected firms to sanctions, as indicated by my measurements.

The third challenge is that companies subject to sanctions may have inherent differences from other businesses, such as being vulnerable to various types of risks or having a trade-focused business model. It is possible that these other factors, rather than the sanctions themselves, could be responsible for the observed results. However, I argue that this challenge is mitigated by the inclusion of industry-fixed effects and the robustness of the results to a set of controls for firm-specific characteristics.

Taking together, my findings indicate that sanctions present a substantial challenge for Iranian firms, as evident in stakeholder discussions. These sanctions are growing in their complexity and impact firms through various mechanisms, leading to diminished stock market returns and declines in sales, investments, and hiring activities. While there is a noticeable variation in exposure to sanctions among Iranian firms, this variation does not align with any indicators of political connectedness. This highlights the indiscriminate nature of sanctions imposed on Iran.

This paper aligns with and contributes to several branches of literature. The first contribution is to the economics of sanctions literature, which primarily employs cross-country analyses to estimate the cost of sanctions on an entire economy (Yang *et al.*, 2004; Felbermayr *et al.*, 2019; Afesorgbor, 2019; Crozet *et al.*, 2021). A subsection of this literature leverages microdata to study sanction effects on individual firms. Crozet *et al.* (2016), Stone (2016), Ahn and Ludema (2020), and Nigmatulina *et al.* (2022) are notable examples in the case of sanctions on Russia. Two works examining sanctions on Iran are Haidar (2017) and Draca *et al.* (2023). Specifically, Haidar (2017) employs export customs data to investigate the impact of the imposition of United Nations export sanctions in 2008 on Iranian non-oil exports, revealing that despite sanctions, two-thirds of these exports were deflected to non-sanctioning countries. Meanwhile, Draca *et al.* (2023) adopts an event study approach to analyze the evolution of nuclear-related sanctions relief for Iran during the P5+1 negotiations in Geneva, culminating in the JCPOA agreement. Their research reveals that firms associated with sanctioned entities experienced notable positive returns during the diplomatic breakthrough.

Most of these paper's studies using firm or individual data compare sanctioned entities to non-sanctioned ones. Nevertheless, the evolution towards more intricate and multi-layered sanctions has blurred the demarcation between treatment (those exposed to sanctions) and control groups (those unexposed), as businesses may experience impacts through various channels that are not immediately apparent. This complexity necessitates a versatile analytical framework capable of capturing the diverse degrees of a firm's exposure to sanctions. To this end, I employ a text-based approach that analyses the perceptions of firms' stakeholders. This approach can account for the intricate nature of global trade relations, interconnectedness, and spillovers. Furthermore, this approach allows for the identification and decomposition of channels through which sanctions impact firms, providing more insight into the way sanctions operate.

The second significant contribution of this study is to the literature examining the political success of sanctions. A line of empirical research investigates how sanctions might bring political change and the conditions under which sanctions are more likely to fulfil the objectives set forth by the sender, largely building upon the cross-country analysis and dataset of Hufbauer (1990). Recently, Draca *et al.* (2023) analyzed the success of sanctions in targeting the economic interests of political elites in Iran. This paper's findings are in line with Draca *et al.* (2023)'s assertion that sanctions act bluntly, but the methodology differs. Here, I separately identify politically connected firms and those exposed to sanctions and directly test to evaluate the correlation between these two groups. I show when sanctions reach a high level of complexity, the concept of being 'smart' or 'targeted' loses its relevance. I also examine the channels through which sanctions operate, exploring if these differ between politically connected and non-connected firms.

Lastly, this work contributes to the growing field of economics literature that leverages text as data (Gentzkow *et al.*, 2019), specifically within the subset that utilizes text to gauge firms' susceptibility to particular shocks (Hassan *et al.*, 2019, 2024, 2023; Sautner *et al.*, 2023). I showcase the adaptability of text-based measurements in assessing firm-level shocks in a new context. I demonstrate this approach can be applied to a developing country undergoing sanctions, thus extending the utility of text-as-data methodology to broader contexts.

The remaining sections of the paper are organized as follows. Section 2.1 provides the historical context of sanctions on Iran, discussing the key events and developments that have shaped the imposition and impact of sanctions. In Section 2.2, I introduce the datasets used in the analysis. Section 2.3 presents the methodology and operationalization of the measure of sanction exposure, detailing the text-based approach. The section further demonstrates the validation and usefulness of the measure. Section 2.4 provides evidence of how precise sanctions are in hitting the interest of political decision-makers. Section 2.5 presents the empirical results on the economic impact of sanctions, including the analysis of stock market reactions, an assessment of the investment, sales, and employment patterns of firms exposed to sanctions, and the decomposition of sanction mechanisms. Section 2.6 concludes.

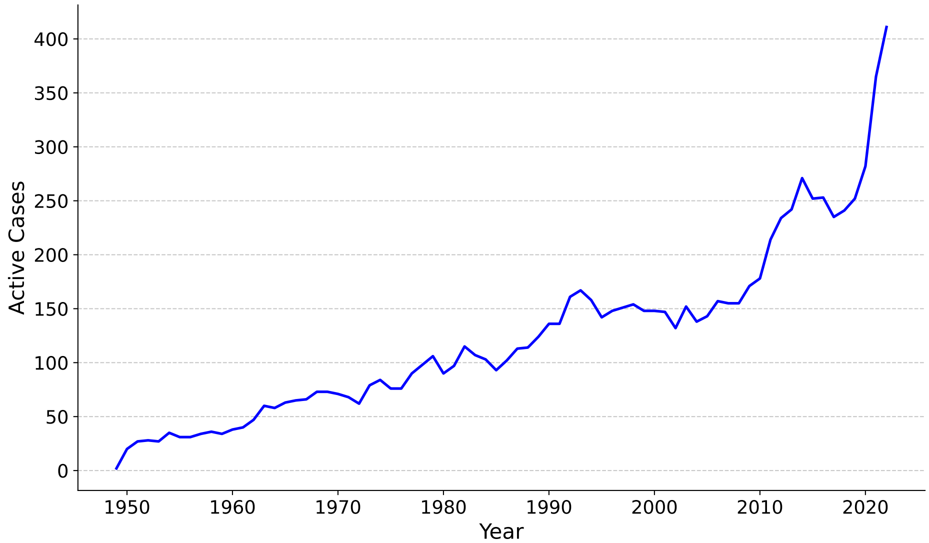
2.1 Historical Context

Sanctions on Iran have been imposed by various countries and international organizations since the Iranian Revolution of 1979. These sanctions have been put in place for various stated reasons, including support of terrorist groups, human rights records, nuclear programs and other perceived threats to international security. Over the years, the scope and severity of these sanctions have evolved, targeting different aspects of Iran's economy, political structure, and military capabilities. The strongest sanctions on Iran are imposed by the US, and the strongest sanctions the US has imposed are on Iran. Figure 2.1 Panel A shows sanctions have emerged as an increasingly prominent foreign policy tool in recent years, and Panel B indicates that Iran is by far the most targeted country for US sanctions.

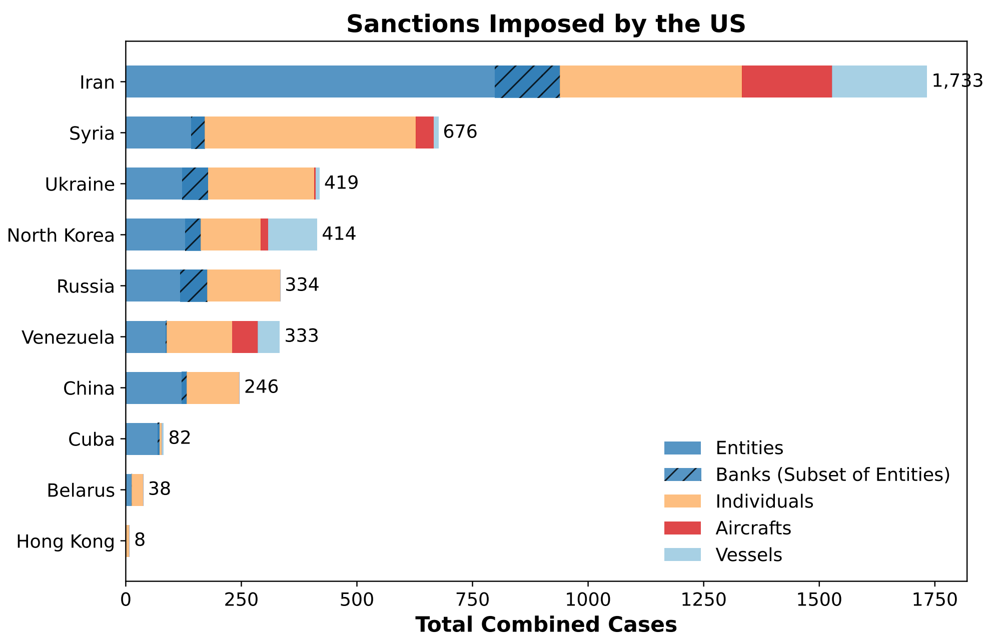
The history of sanctions against Iran can be traced back to 1979, when the United States imposed economic sanctions following the Iranian Revolution and the US embassy hostage crisis. These sanctions, based on the International Emergency Economic Powers Act (IEEPA), included freezing approximately 12 billion worth of Iranian assets held in US banks and a comprehensive ban on US exports to Iran, except for essential goods such as food and medicine.

In 1995, President Clinton issued Executive Orders 12957 and 12959, which expanded sanctions on Iran on the grounds of its support for terrorist groups, human rights abuses, and pursuit of weapons of mass destruction (WMD). The Iran and Libya Sanctions Act (ILSA) of 1996 further expanded these sanctions by penalizing foreign companies that invested in Iran's energy sector, thus extending the reach of US sanctions extraterritorially.

The UN Security Council adopted Resolution 1737 in 2006, imposing sanctions on Iran for its non-compliance with the International Atomic Energy Agency (IAEA) and its refusal to suspend uranium enrichment activities. These measures included asset freezes and travel bans for individuals involved in Iran's nuclear and missile programs, as well



(a) Number of active sanction cases over time



(b) Sanctions imposed by the US as of 2022

Figure 2.1. *Notes:* Panel A displays the number of active sanction cases from various sanctioning bodies over time. Each 'case' denotes a distinct imposition of sanctions, which might target an individual, firm, distinct entity, or even an entire sector within a nation. The data is from the third release of the Global Sanctions Data Base (Felbermayr *et al.*, 2020; Kirikakha *et al.*, 2021). Panel B illustrates selected countries on the Specially Designated Nationals and Blocked Persons list. The data is from Peterson Institute for International Economic; Office of Foreign Assets Control.

as restrictions on the trade of sensitive nuclear-related materials and technologies. In subsequent years, the UN Security Council passed additional resolutions, further targeting Iran's financial, transportation, and energy sectors.

In 2010, the United States and the European Union intensified the pressure on Iran by adopting the Comprehensive Iran Sanctions, Accountability, and Divestment Act (CISADA) and the EU Regulation 961/2010, respectively. These measures targeted Iran's energy and financial sectors, aiming to reduce its oil exports and access to the international banking system. The EU imposed a full oil embargo on Iran in 2012, while the United States tightened restrictions on the Iranian financial sector, including the Central Bank of Iran (CBI). These sanctions also severely limited Iran's international financial access; for example, in early 2012, the Belgium-based Society for Worldwide Interbank Financial Telecommunication (SWIFT) removed several Iranian banks from its system. It is important for my identification to highlight that sanctions placed on Iran were a mix of both targeted "smart sanctions" and broader comprehensive measures. These sanctions were arguably intentionally ambiguous, making it difficult for businesses and traders to understand the risks of conducting transactions with Iran. Diplomatic efforts to reach an agreement were shrouded in uncertainty.

In April 2012, the P5+1 nations (five permanent members of the UN Security Council and Germany) resumed negotiations on Iran's nuclear program with a meeting in Istanbul, which was deemed successful by both sides. Through a series of meetings in the following months, the first pivotal diplomatic milestone was reached in November 2013, when the parties reached a framework agreement in Geneva. Subsequently, extensive negotiations took place to finalize an agreement in which sanctions would be lifted in exchange for concessions on Iran's nuclear program.

Finally, in 2015, the US, EU and UN lifted many of their sanctions on Iran as part of the Joint Comprehensive Plan of Action (JCPOA), commonly known as the Iran Nuclear Deal, which aimed to limit Iran's nuclear program in exchange for lifting of sanctions. However, despite certifying Iran's compliance to Congress twice since taking office, President Trump announced in May 2018 that the US would be withdrawing from the JCPOA. This decision led to the re-imposition of US sanctions on Iran, including the "snapback" of secondary sanctions targeting non-US companies conducting business with Iran.

Several decades of sanctions have negatively impacted Iran's economy. The annual GDP growth rate of Iran, along with big events regarding sanctions on Iran in the last two decades are depicted in Figure 2.2. This figure suggests that sanctions are taking a toll on the Iranian economy, as depicted by the lower growth rate during epochs of sanctions.

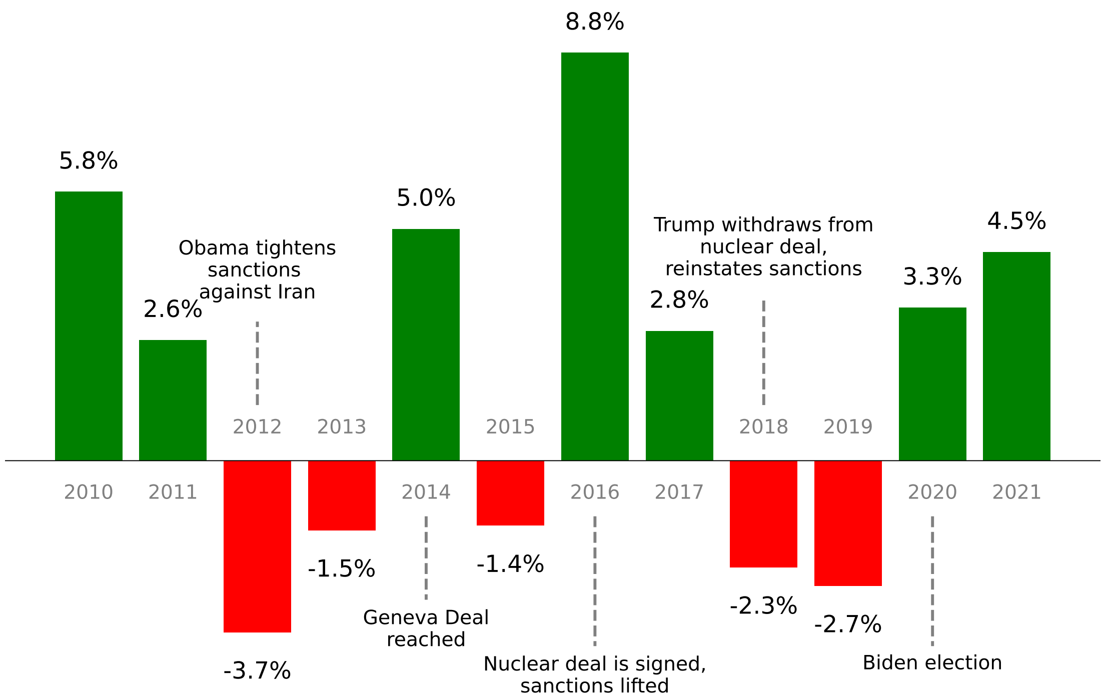


Figure 2.2. *Notes:* The figure displays the GDP growth rate of Iran, presented as a percentage over time. Key events related to sanctions imposed on Iran are highlighted to provide context for observed economic trends. The data has been sourced from the World Bank.

Sanctions on Iran are complex and have a multifaceted nature, as they are imposed by various entities and have varying levels of reach. I have conducted several interviews with business representatives inside and outside of Iran, and they both have reported that sanctions can impact businesses in various ways and that complying with sanctions requires navigating a complex web of regulations and guidelines, often with varying levels of enforcement and differing interpretations of the rules. Their insights revealed these difficulties exist for both Iranian firms and non-Iranian firms considering business endeavours in Iran.

For non-Iranian firms, entering into business ventures with Iranian counterparts necessitates meticulous due diligence to ensure compliance with sanctions regulations. This is partly due to fear generated due to the records of hundreds of millions of dollars of fines that have been levied against institutions like Credit Suisse AG in 2009, ING Bank N. V. in 2012, and BNP Paribas SA in 2014 by OFAC for allegedly violating US Sanctions. Even for transactions that are stated to be exempt from sanctions, such as those involving humanitarian aid, medical supplies, and certain agricultural products, businesses need to stay vigilant and seek expert advice to ensure compliance with the applicable regulations as these exemptions can be complex and subject to change.

On the other hand, sanctions also affect Iranian firms directly. Their experience, however, varies significantly based on factors such as industry sector, operational scale, and integration level with the global economy. Some Iranian firms may find workarounds or alternative sources of financing and trade, while others may struggle to survive under the weight of international sanctions. This feature of sanctions necessitates a more flexible approach to measuring a firm's exposure to sanctions, a topic explored further in subsequent chapters.

2.2 Data

I assemble a novel dataset to analyze the conversations between financial analysts and other market participants with firms' managers. The dataset comes from three sources. First, I use the text from information conferences, known as "Konferans-e Ettela'at-e Rasani", that firms hold periodically. When a firm plans to hold one of these conferences, it is announced in advance, and market participants, such as financial analysts, investors, and other stakeholders, have a few days to post their questions and concerns on a platform provided by the firm. The firm then answers all questions. I have obtained, through data scraping, all of these answers and questions in information conferences from the official outlet for stock market-related documents, the Codal website.

The second and third sources are related to annual meetings. Unfortunately, the full transcripts of these meetings are not accessible. Instead, I utilize summaries of the annual conferences from two major market analyst firms. These summaries are usually compiled by specialized market reporters who attend the conference and summarize the key takeaways and highlights of the meeting. This data is useful as it allows me to analyze the sentiments and concerns of analysts and market participants towards a firm and its performance, even if the full transcript of the conference is not available.

The period under consideration in my study spans from September 2016 to 2022, and my data consists of 5,500 meeting reports from 700 firms listed on the Tehran Stock Exchange. The average number of reports for each firm is 8.9 and the standard deviation is 4.6. Unfortunately, records for meetings held prior to September 2016 were unavailable, preventing their inclusion in this analysis. Since the number of meetings a firm holds depends on several factors, there is variation in the number of meetings per firm in my sample.

To collect data on the stock returns of the full universe of domestic publicly traded companies, I scrape daily information from the website of the Tehran Stock Exchange (TSE). The TSE is recognized for its financial access, depth, and efficiency among developing countries, according to the World Bank's Global Financial Development database (Čihák *et al.*, 2012). The TSE operates from 9:00 AM to 12:00 PM for three hours daily, five days a week (Saturday to Wednesday).

I collected firm-level data, such as employment, investment, and sales, by scraping income statements, cash flow statements and balance sheets, which were released on the official outlet Codal, and extracting the relevant information. I exclude non-annual financial documents as my firm-level analysis will be conducted annually. To ensure the reliability and accuracy of the collected data, a manual auditing process was conducted by human reviewers. Non-annual financial documents were excluded from the dataset to maintain consistency. As a result, I obtained an unbalanced panel dataset comprising data from 600 companies, covering fiscal years that concluded between June 2010 and July 2020.

I use GDELT (Global Database of Events, Language, and Tone), a large, real-time database of news, social media, and other publicly available data sources, to create a measure of sanction intensity over time. GDELT captures a wide range of events and information from all over the world, allowing me to analyze the volume and tone of news and social media mentions related to sanctions. I can then use this information to create a measure of sanction intensity, which can be used to track the evolution of the importance of sanctions over time.

The data on connected firms for this study is sourced from Draca *et al.* (2023), who focused on two principal actors targeted by sanctions due to their significant roles in Iran's nuclear program decision-making: the Islamic Revolutionary Guard Corps (IRGC) and Iran's Supreme Leader. Both are reported to control sizable conglomerates. The target group of firms is defined using sanction documents from the UN, EU, or US that state entities are owned or controlled by the IRGC or Setad. Specific identifiers from the Department of the Treasury's Specially Designated Nationals and Blocked Persons list (SDN list) help identify entities linked with the IRGC. Conversely, entities sanctioned due to links with Setad are identified through a detailed US Treasury press release. All TSE-listed assets of the IRGC and Setad entities defined in this process are then identified, resulting in a target portfolio of 50 firms, representing about 16% of the TSE's total market capitalization. These firms include ones that are fully owned by IRGC or Setad, as well as ones where these entities hold stakes.

2.3 Sanction Exposure

As argued previously, when sanctions evolve to become intricate and multi-faceted, the true nature of the exposure of firms to them is far more complicated than can be understood from accounting statements or sanction documents alone. This might partly be due to the deliberate policy of ambiguity from the sanctioning countries. In order to more accurately assess a firm's exposure to sanctions, I follow Hassan *et al.* (2019) and measure exposure to sanctions based on transcripts of firms' meetings. This can flexibly capture the exposure to sanctions through channels that are not measured using conventional methods and can best think of capturing the concerns of firms and investors by directly asking them. In particular, I measure the share of conversation between the meeting participants and firm managers that centres around sanctions.

These conversations are conducted in Persian. The decision at hand is whether to translate the dialogues into English for analysis or to analyze them in native Persian. Both approaches have their merits and limitations. Utilizing English text analysis tools allows for access to a larger pool of resources, tools, and libraries that have been extensively tested and optimized for text analysis. Conversely, the tools and resources available for processing Persian text are not as extensive as those available for English. Nevertheless, I decided to conduct the analysis in Persian. This decision was primarily due to the fact that translating Persian text to English can result in the loss of information, meaning and nuances in the original text, potentially impacting the accuracy and reliability of the analysis. As such, utilizing Persian text analysis enables a more accurate and reliable analysis of the data, even though it requires more extensive adaptation and utilization of existing resources.

I create a measure of overall sanction exposure by looking at announcement conference texts as well as reports on the annual meetings and measuring how much of it is related to sanction. Initially, to validate that sanctions-related discussions mirror real-world sanction shocks, I examined the frequency of sanction mentions, adjusted by the total word count in these documents, over a timeline. More precisely, I decompose each meeting document into a list of words and then count the number of occurrences of "sanction" or "JCOPA" and divide it by the total word count for that quarter's documents. I investigate if the evolution of mentions of "sanction" in firms' meetings over time aligns with the timeline of sanctions.

Figure 2.3 presents the frequency of sanction mentions, adjusted by the total word count in these documents, across firms listed on the Tehran stock exchange market. The media sanction intensity measure is also displayed, calculated as the percentage of global online news coverage monitored by GDELT mentioning sanctions and Iran. The two series display a highly positive correlation. Consistent with the timeline of sanctions, discussions about sanctions remained relatively low before 2018. However, a sudden increase was observed after President Trump's announcement to withdraw from the Iran deal on May 8, 2018, with a second peak on June 24, 2019, when further sanctions were imposed, including a sanction on the supreme leader. These results align with our prior

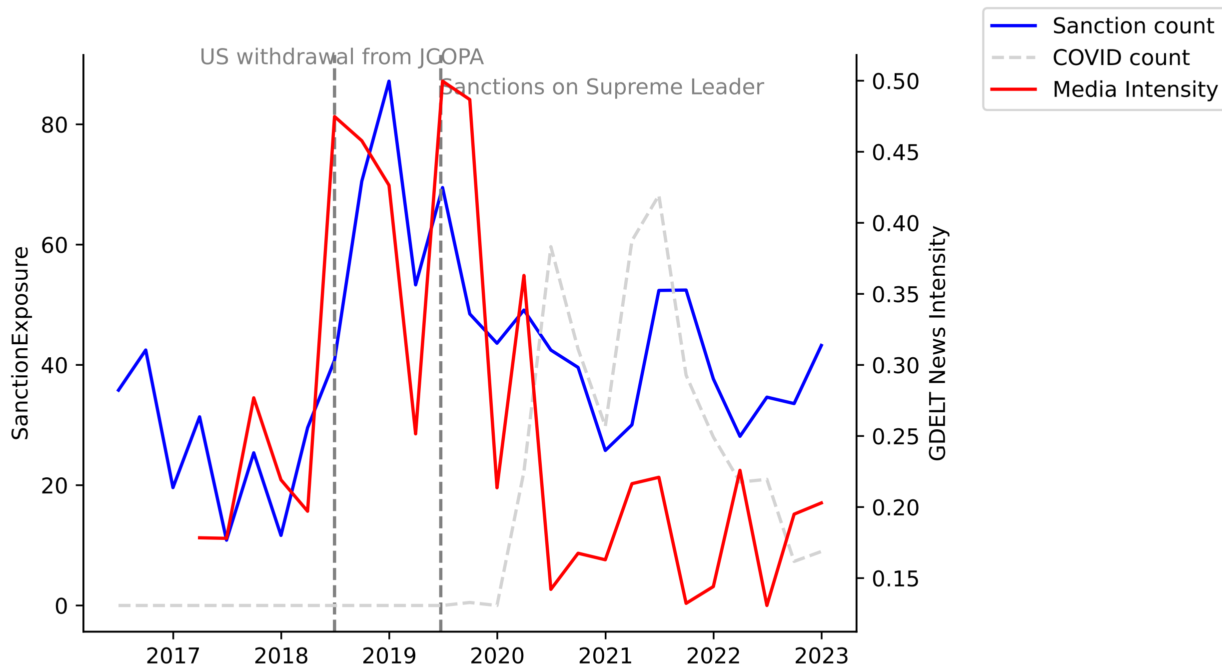


Figure 2.3. *Notes:* This figure illustrates the quarterly count of "sanction" (left axis) and the measure of news media sanction intensity (right axis). The measure of news media sanction intensity is built using GDELT and calculates the percentage of global online news monitored by GDELT that includes "Iran" and either "sanction" or JCPOA. The vertical line marks the quarter in which the Trump administration made the decision to withdraw from the JCPOA deal and reinstate sanctions.

expectations and lend support to the validity of our approach, indicating that discussions on sanctions in these corporate meetings offer a reliable reflection of the actual impact of sanctions on the firm. The occasional lag observed in this measure compared to the GDELT measure can likely be attributed to the time needed for recent news events to be addressed in subsequent meetings.

The fact that the timing of "sanction" mentions in meetings that are intended to address the pressing issues faced by firms lines up with the timeline of sanctions is significant not only because it validates the measure as an accurate indicator of sanctions exposure, but in its own right. These meetings are intended to concentrate discussions on the firm's actual risks and challenges, dedicating more time to the most consequential matters. This comovement suggests that sanctions do pose a real risk to the economy and are not just symbolic moves.

To contextualize the scale of these shocks, I draw a comparative benchmark using a similarly constructed shock for COVID-19[12], illustrated as a grey dashed line in Figure

[12]When determining exposure to COVID-19, I took into account all Persian spelling variations of 'COVID', 'Corona', and 'Coronavirus'.

2.3. The figure shows a swift surge in COVID-19 concerns immediately following 2020. Intriguingly, even at its zenith, the concern level regarding COVID-19 was 20 percent lower than the peak concern level about sanctions. This comparison underscores the considerable magnitude of Iranian firms' concerns about sanctions. The direct analysis of the impact of sanctions on firm valuation and performance will be discussed in the next section.

The method of solely focusing on mentions of "sanction" or "JCOPA" to understand the exposure of each firm to sanctions, echoing Hassan *et al.* (2024), is transparent, minimizes noise, and allows for comparative studies with other impactful events, such as the COVID-19 shock. However, this approach is arguably information-restrictive as it only looks at 'sanction' and may overlook other relevant terms associated with sanctions that don't explicitly mention 'sanction'. Thus, in my principal approach, instead of a pre-determined selection of words associated with sanctions, I use a computational linguistics-based sequence-classification method to assign to each bigram[13] a weight that indicates how strongly it is associated with discussions of the sanction. This is essentially utilizing tf-idf vectorization and follows Hassan *et al.* (2019). Using the alternative approach of only looking at mentions of "sanction" yields qualitatively comparable results.

The first step in constructing my measure is to identify those two-word combinations that are archetypes of discussions around sanction. To this end, I define two training libraries: S, composed of texts primarily focused on sanctions, and NS, containing typical non-sanction related text. While the process of constructing the measure is automatic, the library choice requires human discretion. I draw from the leading Iranian economic publications for my training libraries: Donay-e Eghtesad, Tejart Farda, Eghtesad Online, 90Eghtesadi, and Farsnews. This is partly because each source doesn't provide enough size and partly to minimise the role of human judgment by using training libraries from outlets with different political leaning. I selectively target articles tagged with 'sanction' or featuring the term in their title to constitute the sanction library. A randomized selection of non-sanction articles from these publications forms the non-sanction library. I then extract all adjacent two-word combinations from the texts of these two libraries, with all punctuation removed.[14] The resulting weighting term would be $1[b \in S/NS] \times \frac{f_{b,P}}{B_p}$.

[13]Bigrams are favoured in the literature and appear to be successful in their ability to strike a balance between effectively capturing relevant language patterns related to sanctions and maintaining analytical simplicity. Bigrams capture more context than unigrams, as unigrams can miss out on the context provided by adjacent words. For instance, the words 'New' and 'York' separately do not convey the same meaning as 'New York' together. Additionally, bigrams avoid the high dimensionality and sparsity issues associated with higher-order n-grams.

[14]I eliminate all words that contain pronouns, shortened pronouns, or two adverbs. I further eradicate all half-spaces, typically seen in two-part words in the Persian language. Despite experimenting with additional text preprocessing techniques, such as removing stop words and lemmatization, I did not find them

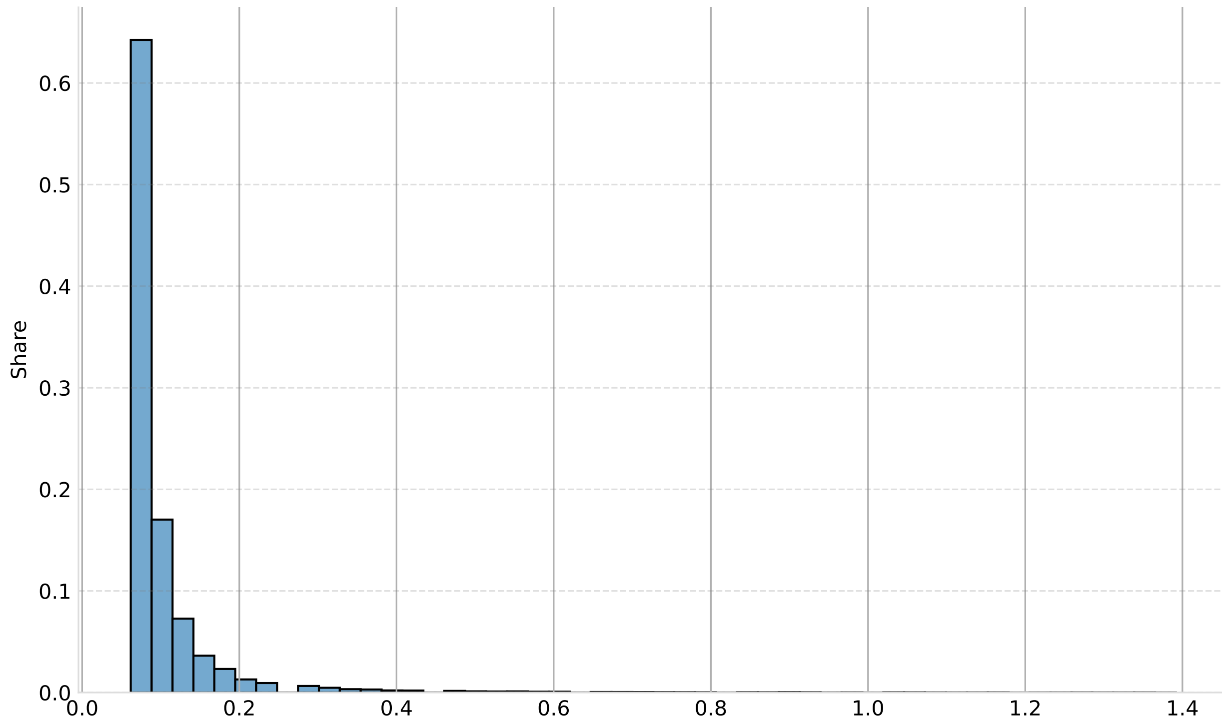


Figure 2.4. *Notes:* This figure displays a histogram of the weights, $f_{b,P}/B_P$, associated with bigrams $[b \in S/NS]$ derived from the textual analysis.

$f_{b,P}$ is the frequency of bigram b in the sanction training library, and B_p is the total number of bigrams in the sanction training library. When dealing with more than two training libraries, the first term can be reformulated into a more recognizable expression: log(# of training libraries / # of libraries in which the bigram occurs). The first component, known as "inverse document frequency" (idf), eliminates bigrams that also appear in non-sanction training libraries. The second component, known as "term frequency" (tf), gives more importance to bigrams that are commonly used in the training library for sanctions. These two terms combined are known as tf-idf, determining the weight of each bigram.

Table 2.1 displays the most frequent bigrams in S/NS, based on their term frequency $(f_{b,P}/B_P)$, that are strongly associated with discussions of sanctions and have the highest weight in our measure. These bigrams are exclusively related to sanctions, such as "from JCPOA", and "Lifting Sanctions". Figure 2.4 illustrates a histogram of the term frequency of these bigrams, which shows a highly skewed distribution with a median term frequency of 6.186×10^{-2} .

to have a meaningful impact on our results. Thus, I choose not to implement these methods in order to stay consistent with the methodology outlined by Hassan *et al.* (2019).

Table 2.1. Top 50 Bigrams and their Translation

Rank	Translated Bigram	Weight	Frequency	Rank	Translated Bigram	Weight	Frequency
0	From JCPOA	75.46	50	25	That sanctions	11.28	14
1	To JCPOA	46.90	5	26	Exit JCPOA	11.24	2
2	Revival of JCPOA	46.26	3	27	Revival negotiations	11.07	0
3	JCPOA is	40.77	2	28	Sanction and	10.77	62
4	JCPOA and	36.00	23	29	Economic sanctions	10.56	29
5	JCPOA in	30.80	5	30	Sanctions on Iran	10.35	6
6	JCPOA to	29.91	6	31	Action JCPOA	10.22	0
7	In JCPOA	28.01	1	32	Implementation of JCPOA	9.97	0
8	American sanctions	25.48	17	33	JCPOA from	9.63	2
9	Sanctions are	25.39	18	34	That JCPOA	9.63	1
10	Lifting sanctions	22.35	46	35	JCPOA commitments	9.46	0
11	These sanctions	21.93	4	36	Applying sanctions	9.17	22
12	Sanctions and	18.63	128	37	JCPOA agreement	8.70	6
13	Some sanctions	17.58	2	38	Preserve JCPOA	8.45	0
14	Nullify sanctions	17.58	0	39	Canceling the sanctions	8.45	0
15	Sanctions against	16.60	6	40	JCPOA negotiations	8.11	3
16	Sanctions in	16.56	34	41	With JCPOA	8.03	1
17	From the sanction	14.87	37	42	Comprehensive action	7.99	2
18	From sanctions	14.79	58	43	Sanction is	7.99	10
19	Sanctions to	14.15	26	44	Lifting the sanctions	7.90	13
20	New sanctions	13.39	14	45	That the sanctions	7.60	4
21	From the sanctions	13.01	37	46	And the sanctions	7.01	34
22	About JCPOA	12.97	0	47	And sanction	7.01	8
23	JCPOA is	12.63	0	48	That sanction	6.97	6
24	Sanction it	11.75	1	49	And JCPOA	6.93	1
25	That sanctions	11.28	14	50	Negotiations to lift	6.76	1

Notes: This table shows the translation of top 100 bigrams with the highest term frequency $f_{b,P}/B_P$ and receiving the highest weight in the construction of sanction exposure. The frequency column lists the frequency count of each bigram in all transcripts.

I create a measure of overall sanction exposure by looking at announcement conference text as well as reports on the annual meetings and measuring how much of it is related to sanction. I compose each document into a list of bigrams. Specifically, I decompose each meeting document for firm i in time t into a list of bigrams. I then count the number of occurrences of bigrams indicating discussion of sanctions, multiplied by the corresponding weight, and divide by the total number of bigrams in the transcript:

$$SanctionExposure_{i,t} = \frac{1}{B_{i,t}} \sum_{b=1}^{B_{i,t}} 1[b \in S/NS] \times \frac{f_{b,P}}{B_p} \tag{5}$$

In the above equation, b is a bigram in a document from firm i at time t, and $B_{i,t}$ is the total number of bigrams in that document. Ideally, this measure could be delineated for every firm and quarter. However, due to the limited sample size for individual firms, the majority of my analysis averages the data across all timeframes for each firm. Consequently, I omit the t subscript and predominantly work with $SanctionExposure_i$.

Hassan *et al.* (2019) suggests differentiating between a shock's first and second moment effects on a firm and introduces a method to do so. While I incorporate this distinction, my primary analysis relies on $SanctionExposure$. This decision is based on two main reasons. Firstly, my Persian sentiment and risk dictionary, compared to the one utilized in Hassan *et al.* (2019), might not be as comprehensive, possibly leading to information loss. Secondly, and most crucially, the study period witnessed various sanction-related events, some involving the "imposition" of sanctions, and others related to "lifting" sanctions. Hence, when firms discuss sanctions, they could be referring to either imposition, lifting, or a blend of both. If this variability is not considered, the results may become confounded. One potential solution involves examining each instance of the term "sanction", and applying a multiplier of -1 if the context is about "lifting" sanctions. Though this approach could alleviate the issue, it does not fully resolve it since it necessitates subjective judgment. Furthermore, it might not always be clear if the reference to sanctions pertains to their "lifting" or "imposing" or a discussion of both. For the sake of simplicity in notation, I have not explicitly detailed this adjustment, but it is applied in the following analysis.

With these caveats in mind, I employ the method outlined in Hassan *et al.* (2019) to differentiate between these first- and second-moment impacts. This method creates measures of sanctions risk and sentiment by analyzing word counts in relation to synonyms for risk or uncertainty and positive and negative tone words, respectively. More precisely, I count the number of bigrams indicative of sanction discussions within a 10-word window surrounding each occurrence of "risk" or "uncertainty" synonyms, and then divide this count by the total number of bigrams in the transcript.:

$$SanctionRisk_{i,t} = \frac{1}{B_{i,t}} \sum_{b=1}^{B_{i,t}} \{1[b \in S/NS] \times \frac{f_{b,P}}{B_p} \times 1[|b - r| < 10]\} \tag{6}$$

r here is the position of the closest synonym for risk or uncertainty.

To determine the firms that are winners and losers (reflecting first-moment impact) as opposed to those exposed to risks (indicating second-moment exposure), I use the same procedure, but this time I use the translation of Loughran and McDonald (2011)'s sentiment dictionary[15] to differentiate between positive and negative words.

[15]The English words were translated using Google Translate and then reviewed and edited by a certified English-to-Persian business translator. Some words were excluded from the translation as they did not have a one-to-one equivalent in Persian, while for some others, more than one Persian translation was considered. Despite these adjustments, the overall number of positive and negative words remained largely the same.

Table 2.2. Summary Statistics

	Firm-level outcomes			Firm-year outcomes		
	$\overline{SExposure_i}$	$\overline{SRisk_i}$	$\overline{SSentiment_i}$	$SExposure_{it}$	$SRisk_{it}$	$SSentiment_{it}$
	(1)	(2)	(3)	(4)	(5)	(6)
Mean	-100.85	40.88	16.41	-106.82	44.84	17.77
Median	-28.29	17.39	0.0	0.0	0.0	0.0
SD	171.15	63.36	32.57	286.89	109.45	58.84
N	678	678	678	3133	3133	3133

Notes: The table presents descriptive statistics for the variables included in the subsequent analysis. It provides information on the mean, median, standard deviation, and the number of observations for each variable. $\overline{SExposure_i}$, $\overline{SRisk_i}$, and $\overline{SSentiment_i}$ are averages for each firm in the sample.

$$SanctionSentiment_{i,t} = \frac{1}{B_{i,t}} \sum_{b=1}^{B_{i,t}} \{1[b \in S/NS] \times \frac{f_{b,P}}{B_p} \times \sum_{c=b-10}^{b+10} S(c)\} \tag{7}$$

In the above equation, S assigns a sentiment to each word c based on the following function, where S^+ is the set of positive-tone words and S^- is the set of negative-tone words.

$$S(c) = \begin{cases} +1, & \text{if } c \in S^+ \\ -1, & \text{if } c \in S^- \\ 0, & \text{otherwise} \end{cases}$$

Table 2.2 displays the average, median, and standard deviation of the variables in my analysis. The key variables of interest are my Sanction exposure, risk, and sentiment measures. For the purpose of this analysis, I also consider the firm-level averages (denoted by an overline) of the Sanction Exposure, Risk, and Sentiment variables. This set of variables is derived by taking the average of all available Sanction variable scores across all years for each firm.

The average $SanctionExposure_{it}$ by industry is presented in Figure 2.5. This metric was obtained by calculating the mean value of $SanctionExposure_{it}$ for all firms in each industry. The results reveal that the "Architectural and engineering activities", "Water transport", and "Mining of coal and lignite" industries have, on average, the highest proportion of time spent discussing political risk topics during conference calls. Conversely, the "Sport and amusement" and "Creative art and entertainment" industries exhibit the

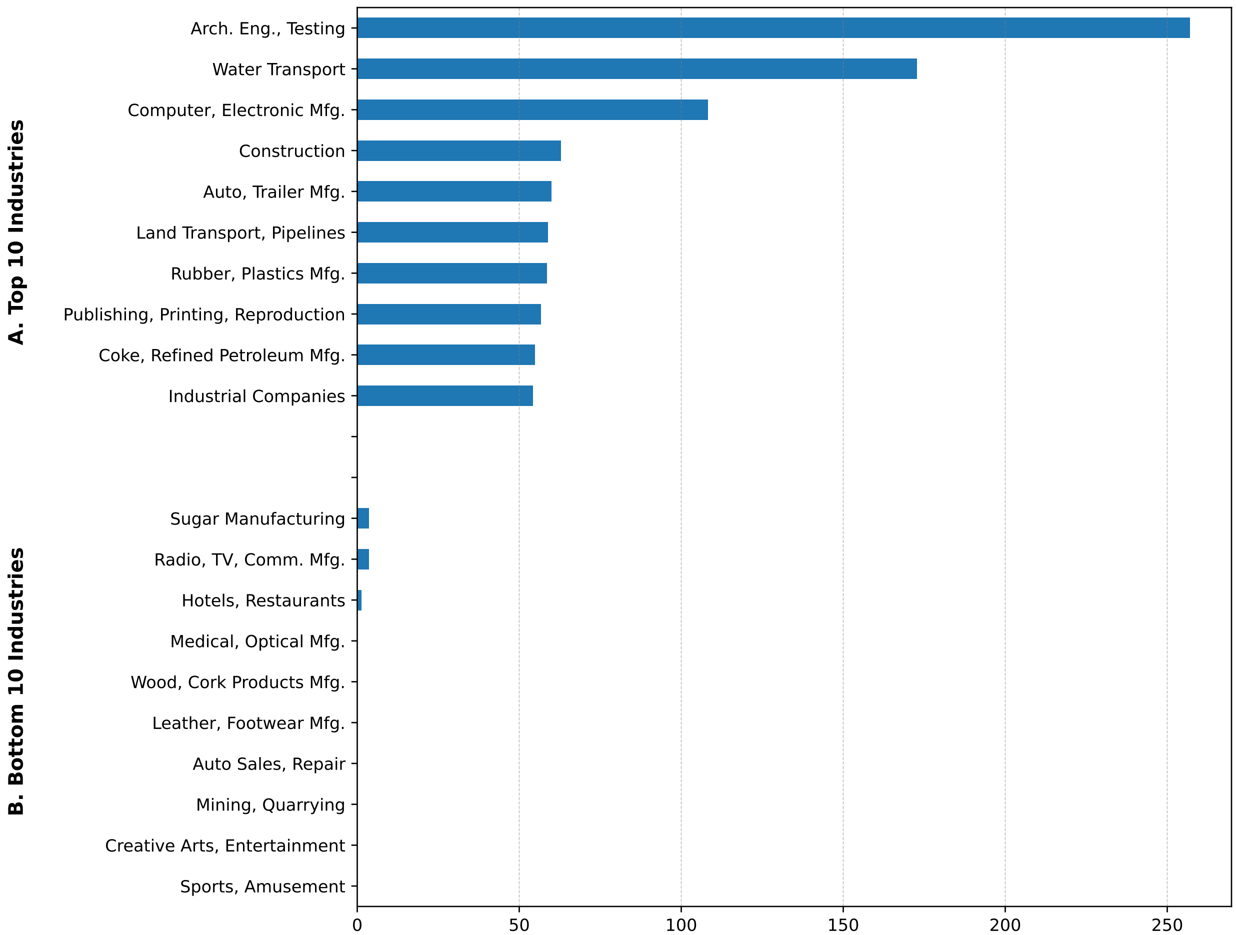


Figure 2.5. *Notes:* This figure displays the mean value of $SanctionExposure_{i,t}$ calculated across all firms in each industry. Higher mean values indicate that Iranian firms operating in that industry are potentially more exposed to the impact of sanctions, and may face greater challenges as a result.

lowest exposure to sanctions. These findings are consistent with the expectation that industries that are more dependent on international trade and connectedness to the outside world may be more sensitive to political risks.

As a final validation exercise, I scrutinized specific firms directly targeted by U.S. sanctions. After analyzing data from the United States Treasury documents, 15 publicly traded firms were identified as being explicitly mentioned in sanctions documents during the time frame of the analysis. To understand the impact of these events, I conducted an event study, focusing on changes in $SanctionExposure_{it}$, which measures the intensity of discussions about sanctions in company meetings, around the time the firms were mentioned in sanction documents. This was analyzed through the following model:

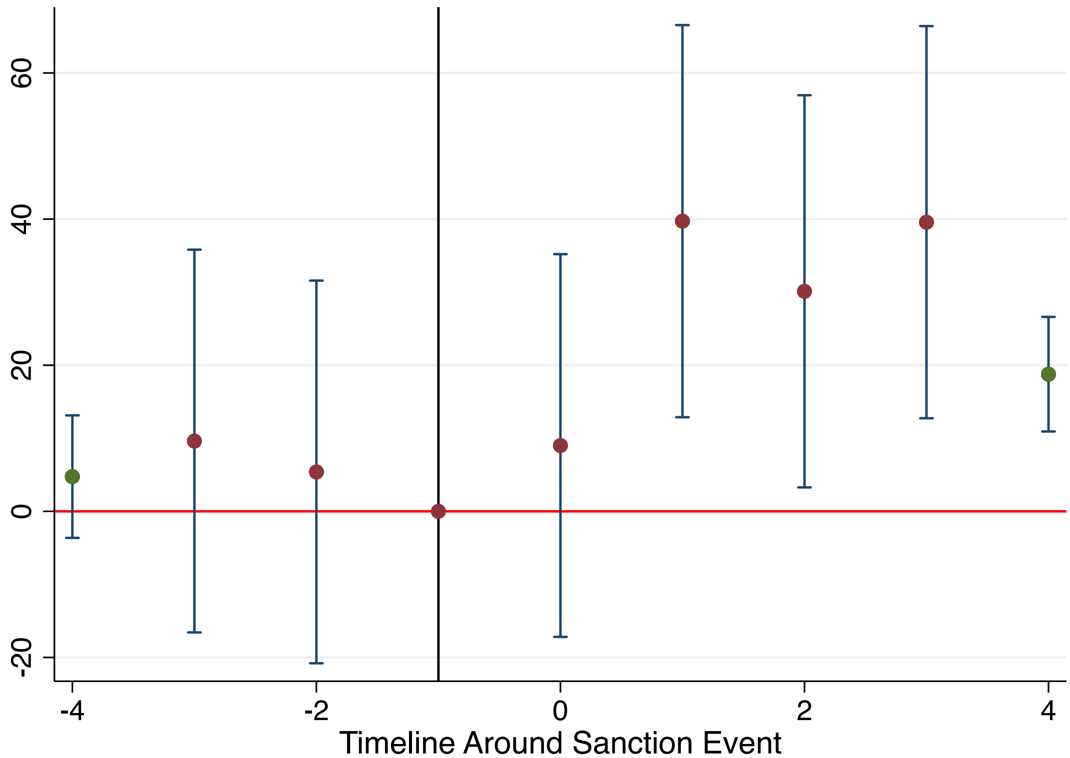


Figure 2.6. *Notes:* This figure presents the quarterly point estimates and 95 percent confidence intervals, depicting the impact of U.S. sanctions imposition on $SanctionExposure_{it}$. The X-axis denotes the quarters relative to the sanction event, illustrating the temporal dynamics surrounding the imposition. Notably, the first and final data points aggregate the effects for periods extending beyond the specified lead and lag quarters.

$$SanctionExposure_{it} = \alpha + \Sigma_j \beta_j \mathbb{1}[t = E_i + j] + \mu_i + \lambda_t + \epsilon_{it} \quad (8)$$

Here, t represents the quarter, and E_i is the specific quarter when the U.S. Treasury announced sanctions on an Iranian firm. Figure 2.6 showcases the results, demonstrating a noticeable increase in discussions about sanctions in company meetings following their sanctioning by the U.S. This finding validates the $SanctionExposure$ measure and boosts our confidence that discussions about sanctions in firms are relevant and reflective of the actual challenges these firms encounter.

2.4 Assessing Sanction Precision

Sanctions can exert pressure on a target government to modify its political behaviour in at least two ways. The first is through direct means by inflicting harm on the interests of political decision-makers, thereby persuading them to alter their behaviour. This is usually the rationale offered by those imposing sanctions. Alternatively, sanctions can work indirectly by inciting a popular revolt that overthrows the government, or by creating public frustration that places pressure on the government to change its behaviour.

In this section, I scrutinize these two ideas by investigating if the companies targeted by sanctions are indeed those that are more exposed to them. If we find that this is the case, it lends support to the first rationale for using sanctions, which involves directly inflicting harm on the interests of political decision-makers to persuade them to change their behaviour. To this end, I estimate the following specification:

$$SanctionExposure_i = \alpha_{j(i)} + \beta Target_i + u_i \tag{9}$$

In the above equation, $Target_i$ is a binary indicator that takes the value 1 if firm i is the target of the sanctions, and 0 otherwise. This equation thus assesses the correlation between being a primary target of sanctions and exposure to these sanctions. Target firms are identified as companies that are owned or controlled by political decision-makers and are the focus of sanctions with the aim of changing the behaviour of their owners. To identify these firms, I follow the definition from Draca *et al.* (2023). They identify target firms as those owned or controlled by Iran's IRGC or Setad, using official documents and Tehran Stock Exchange data, resulting in a list of 50 firms.

The findings are detailed in Table 2.3. These results emphasize the lack of correlation between different measures of political connection and exposure to sanctions in a significant way. This indicates that sanctions, in their current form, may not accurately pinpoint their intended targets. Instead, they seem to cast a wide net, affecting the economy of the targeted country in a more comprehensive way than initially intended or openly acknowledged. In particular, the study reveals that the economic interests of Iran's political elite were not exposed to sanctions any more than other publicly traded firms.

These findings suggest that complex, multi-layered sanctions may not be as effective as sometimes argued in exclusively impacting specific firms or individuals. Selectively 'activating or deactivating' their impact across various entities within the recipient nation appears not to be possible. As sanctions grow more complex, the notion of 'targeted' sanctions becomes less applicable. The subsequent chapter's findings reinforce this by illustrating that, in equilibrium, sanctions impact through various pathways that are somewhat indiscriminate. These channels, like financial limitations or restricted access to intermediary goods, affect a wide range of firms, not only the politically affiliated ones.

Table 2.3. Exposure and Target

	Sanction Exposure		
	(1)	(2)	(3)
Target Type Indicator:			
Broad-Spectrum Target Firms	8.954		
	(7.028)		
Direct Target Firms		21.505	
		(19.884)	
Indirect Target Firms		7.308	
		(7.442)	
IRGC Firms			31.886
			(22.922)
Setad Firms			6.770
			(7.329)
Observations	840	840	840

Notes: This table presents the relationship between firms' exposure to sanctions and various indicators of being the target of sanctions. Different definitions of the target at the firm level data, sourced from Draca *et al.* (2023) are merged with the Sanction Exposure measure, formulated as per equation 5. 'Direct' pertains to firms directly identified in smart sanctions documents or firms where targeted entities possess direct ownership. Indirect exposure ('Indirect') refers to firms not directly named but linked to directly targeted entities through ownership. The variables 'IRGC' and 'Setad' represent specific targeting types; 'IRGC' refers to firms linked to the Revolutionary Guard Corps of Iran, while 'Setad' denotes firms affiliated with Setad, a distinct Iranian entity. The initial column's 'Target' variable is a composite indicator that amalgamates all these categories. For a more in-depth understanding of the classification process, check Draca *et al.* (2023). Standard errors are shown in parentheses.

2.5 Economic Impact

I now turn to the real economic impact of sanctions and ask if firms that frequently report concerns about sanctions actually experience economic ramifications. Initially, I explore whether firms with high sanction exposure experience an excessive negative return following news about the imposition of sanctions. The underlying premise is that to the extent that news about sanctions is unanticipated, firms with greater exposure should display a more negative excess return, signalling a diminished future revenue stream. I then delve into firm-level performance and see if sanction leads into lower sales, investments, and hiring. Lastly, I probe the mechanisms underlying these effects.

Table 2.4. Comparison of Sanction Exposure Between Targeted and Non-Targeted Firms

	(1)	(2)	(3)	(4)	(5)
	Non-targeted firms		Targeted Firms		Difference
	N	$MeanExposure$	N	$MeanExposure$	(4)-(2)
Broad-Spectrum Target Firms	733	33.81	69	42.57	8.76
		(60.79)		(60.79)	(6.64)
Indirect Target Firms	741	34.03	61	41.09	7.06
		(60.71)		(60.71)	(7.03)
Direct Target Firms	794	34.37	8	53.88	19.51
		(60.17)		(60.17)	(18.57)
IRGC Firms	796	34.34	6	64.29	29.95
		(60.04)		(60.04)	(27.11)
Setad Firms	739	34.06	63	40.50	6.45
		(60.85)		(60.85)	(6.74)

Notes: This table presents the results of t-tests comparing firms' exposure to sanctions across various indicators of being the target of sanctions. Different definitions of the target at the firm level data, sourced from Draca *et al.* (2023) are merged with the Sanction Exposure measure, formulated as per equation 5. 'Direct' pertains to firms directly identified in smart sanctions documents or firms where targeted entities possess direct ownership. Indirect exposure ('Indirect') refers to firms not directly named but linked to directly targeted entities through ownership. The variables 'IRGC' and 'Setad' represent specific targeting types; 'IRGC' refers to firms linked to the Revolutionary Guard Corps of Iran, while 'Setad' denotes firms affiliated with Setad, a distinct Iranian entity. The initial column's 'Target' variable is a composite indicator that amalgamates all these categories. For a more in-depth understanding of the classification process, check Draca *et al.* (2023). Each entry in the table provides the number of firms, the mean and the standard deviation of sanction exposure for politically connected and non-connected firms, along with the difference in means and its associated standard error.

2.5.1 Stock Market Reaction

This section analyzes how stock markets responded to the events related to sanctions on Iran. The idea is that when investors were informed of this development, they recalibrate their expectations about the future of publicly-listed firms, leading to changes in stock prices during the event period. These stock price shifts mirror changes in investors' perceptions of both direct and indirect sanctions effects on Iranian firms, which can affirm that my measure transcends mere distraction or trivial rhetoric and contains substantive information.

There are numerous events related to sanctions and negotiations to lift sanctions between Iran and the West that could be examined through an event study approach. To avoid biases associated with arbitrary event selection, I adopt a systematic methodology as proposed by Amiti *et al.* (2020) to identify key events. Specifically, I pinpoint days with a peak in the number of Google searches for the term "Sanctions Against Iran," as depicted in Figure 2.7. This is based on search activities which includes searches in all available languages and countries where Google is accessible. By setting the time horizon to span the entire period under review, Google Trends provides a monthly intensity index of searches. To pinpoint the specific dates that correspond to the most significant surges in search volume, I refine the approach by conducting targeted Google Trends queries with narrower time frames surrounding each identified peak, obtaining a daily index of search interest for "Sanctions against Iran".

Subsequently, I then cross-reference these dates with media reports to identify significant sanction-related events around these periods. Two events are excluded from this analysis. Initially, the aftermath of the assassination of Iranian general Soleimani is omitted because it is not directly linked to the implementation or removal of sanctions. Furthermore, the event dated November 2018 is excluded due to its ambiguous nature regarding its positive or negative implications for sanctions. While the US ushered in the second wave of renewed sanctions in November 2018, the other signatories of the Iran nuclear agreement-France, Britain, Germany, Russia, and China-announced their plans to launch a "Special Purpose Vehicle" (SPV). This mechanism aimed to ease transactions with Iran, bypassing US sanctions, and was designed to "assist and reassure economic operators pursuing legitimate business with Iran." Given the ex-ante unknown nature of whether this event is positive or negative, it is not considered in the analysis.

Abusing notation and omitting time subscript t for each event, I run the following specification:

$$R_i = \alpha + \theta Sanction_i + \gamma X_i + u_i \tag{10}$$

Here, R_i refers to the four-trading-day return of firm i following the event, while X_i is a vector that includes industry fixed effects, firm-specific characteristics such as the size of the asset, and the firm's market betas, which is calculated by regressing monthly returns of the firms on the monthly Tehran Stock Market index (TEDPIX). The variable $Sanction_i$ represents either the firm-level averages of Sanction Exposure ($\overline{SExposure_i}$), Sanction Risk ($\overline{SRisk_i}$), or Sanction Sentiment ($\overline{SSentiment_i}$) for firm i. This strategy is valid if, absent the sanction events taking place during this window, no systematic differences would exist between the returns of the exposed versus non-exposed firms. In other words, we require the standard identification assumption $Cov(Sanction_i, u_i|X_i) = 0$.

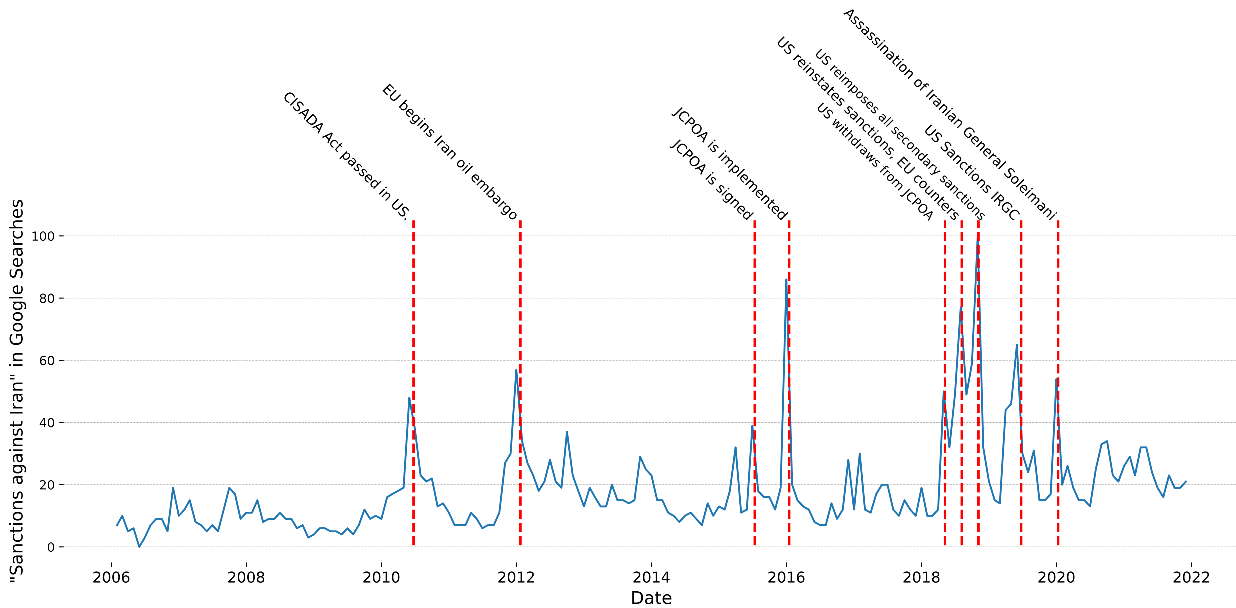


Figure 2.7. *Notes:* This figure displays the frequency of Google searches for "Sanctions against Iran," marked with significant related events: June 24, 2010: The Comprehensive Iran Sanctions, Accountability, and Divestment Act (CISADA) was passed under President Obama. January 23, 2012: The EU imposed an oil embargo on Iran and froze the assets of Iran's central bank. July 14, 2015: JCPOA Signed. January 16, 2016: Most UN sanctions on Iran were lifted following the JCPOA. May 8, 2018: The US withdrew from the JCPOA. August 7, 2018: The EU nullified US sanctions on countries trading with Iran, and the US reinstated non-oil sanctions. November 5, 2018: The US reimposed all secondary sanctions on Iran. June 24, 2019: The US sanctioned eight senior commanders of the IRGC. January 10, 2020: Sanctions were authorized on Iran's key sectors by President Trump. Each vertical line on the figure corresponds to these key events, providing a visual timeline of political events and public interest.

The result is tabulated in Table 2.5. While coefficients are noisy, the signs of coefficients generally align with expectations. Firms exposed to sanctions tend to exhibit a lower return when there is news about the imposition of sanctions and a higher return when news is about the lifting of sanctions. Specifically, columns 3 and 4, which indicate events leaning towards the lifting of sanctions, positively influenced the excess return of firms more exposed to sanctions. It's noteworthy that the evolution of the JCPOA was filled with uncertainties at every phase, so each major event conveyed fresh insights into the probability of sanctions being removed. Other columns demonstrate events associated with the imposition of sanctions led those sanction-exposed firms to experience a negative excess return.

Column 8 provides a parsimonious summary of former results by estimating the average impact of all sanction-related events on firms that are more susceptible to sanctions. To accomplish this, I introduce a variable, denoted as E_t, which assumes a value of zero

Table 2.5. Stock Market

	Stock Returns								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	24-06-2010 CISADA Act Passed	22-01-2012 EU Oil Embargo	14-07-2015 JCPOA Signed	16-01-2016 JCPOA Implement	09-05-2018 US Exits JCPOA	07-08-2018 US Part1 Sanctions	24-06-2019 US Sanctions IRGC	2010- 2020 Pooled Results	
Event Type	(Negative)	(Negative)	(Positive)	(Positive)	(Negative)	(Negative)	(Negative)	(Negative)	
$\overline{SExposure_i}$	-3.756	-1.705	8.069	4.056	-4.103	-7.993	-6.041		
	(4.425)	(3.658)	(4.329)	(4.037)	(2.912)	(4.664)	(4.741)		
$E_t \times \overline{SExposure_i}$								-2.849	-2.904
								(1.328)	(1.377)
$E_t \times \overline{SExposure_i} \times \overline{T_i}$									0.507
									(3.324)
Observations	219	263	398	436	501	510	542	868100	868100
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table shows the OLS estimation results for cross-sectional regressions in columns 1 to 7, and for time series analysis in columns 8 and 9, examining $StockReturns_{it}$ against $\overline{SanctionExposure_i}$ (abbreviated as $\overline{SanctionE_i}$). The stock returns are calculated as $\Sigma_{t=0}^{t=N} LogP_{i,t}/P_{i,t-1}$, where t is at a daily frequency, and $[0, N]$ represents the period of four trading days (including weekend days) following the event. Positive events are those in which an occurrence has enhanced the probability of lifting or easing sanctions on Iran. Conversely, negative events refer to incidents suggesting the imposition of stricter sanctions on Iran. These events are pinpointed through peaks in Google Trends analysis, tracking searches related to sanctions against Iran. The controls incorporate the firm's market betas, market cap and ISIC 2-digit dummies, and the standard errors are robust.

on days without any events. On event days, depending on whether the event conveyed a positive, neutral, or negative outlook regarding the possibility of sanctions being imposed, the variable takes on the values -1, 0, or +1, respectively. I then proceed to interact this variable with my measure of sanction exposure. The resulting negative coefficient confirms that firms with higher exposure to sanctions are likely to experience a decline in market value upon receiving news that hints at a potential escalation in the severity of sanctions.

The stock market's response to unanticipated sanction news confirms mentions of sanctions were not merely a form of deception or cheap talk, or an excuse for poor performance. To further verify that politically connected firms do not systematically underreport their sanction exposure, I analyze if the market valuations of politically connected firms are adjusted in a way that aligns with their measured exposure to sanctions, similar to other firms. This analysis, presented in the last column of the table, examines if sanction exposure impacts politically connected (target) and non-connected firms differently on event days. The zero coefficient for the interaction term, indicating no differential effect of sanction exposure on politically connected firms, implies that the market recognizes and accounts for their vulnerability to sanctions, consistent with my measurements, just as it does for other firms.

The computation of the full in-sample distribution of point estimates, as inferred from the last column and each firm's market cap, indicates that an average Iranian firm loses around 800B rial in response to unfavourable news about sanctions. Although this figure may not convey a direct interpretation on its own, comparing the total impact on politically connected versus non-connected firms is very insightful. Based on the estimation presented in the final column, the total effect on all politically connected firms amounts to 32000B Rial, whereas non-connected firms face a more substantial impact of 161000B Rial. This substantial difference underscores the externality associated with the enactment of sanctions aimed predominantly at connected firms. This externality signifies that for every $1 of intended damage on politically connected firms, an unintended loss of $5 is incurred by non-targeted firms. This considerable externality stems from the fact that sanctions impact politically connected and non-connected firms roughly in the same way, but there's a significantly larger number of non-connected firms. The externality multiplier remains the same when considering General Equilibrium effects and spillovers, provided these effects don't systematically differ between connected and non-connected firms.

The observed correlation between sanction-related events and the fluctuating stock market valuations of firms discussing sanctions in their meetings supports the notion that sanctions inflict damage on firms with a higher degree of exposure, subsequently affecting the entire economy. I discuss two other potential interpretations of these results, but the evidence at hand largely supports the original explanation.

Firstly, one could argue that this correlation merely illustrates how firms with distinct characteristics respond differently to macroeconomic shocks, rather than reflecting the specific impacts of sanctions. However, considering that my analysis accounts for various firm-level characteristics and industry-specific factors, this correlation is more likely to be indicative of sanctions' effects. Furthermore, the analysis takes into account a range of different events, making it highly improbable that certain macroeconomic shocks consistently coincide with sanction-related events.

An alternative interpretation suggests that sanctions, while impacting the values of firms exposed in the targeted country, primarily result in a resource reallocation among firms with different sanction exposures, without having a substantial effect on the overall economy. While some resource reallocation is plausible and perhaps even probable - for example, firms producing similar goods may gain an inadvertent advantage from sanctions due to decreased foreign competition, or the government may bolster support for certain firms to help it circumvent sanctions - it cannot completely counteract the effects of sanctions. The preceding chapter's evidence demonstrates that certain sanction-induced mechanisms -for example, no access to intermediary inputs- can more or less universally affect all firms, indicating that resource reallocation cannot comprehensively mitigate these impacts. Moreover, the aggregate effect of sanctions on the entire economy, as illustrated in Figure 2.2, suggests that resource reallocation cannot offset the net effect.

2.5.2 *Impact on Firm-level Investment and Sales*

Previous sections showed that sanctions pose a challenge for firms, at least to the extent that they are likely to be discussed by stakeholders and that equity markets may price these shocks accordingly. This section will delve into the specific impacts of sanctions exposure at the firm level, focusing on investment and sales, while also acknowledging several limitations and challenges that arise in this context.

The first limitation is that due to data availability, our $SanctionExposure$ measure does not cover a broad enough time period and does not extend far enough into the past. This means that for t prior to the re-imposition of sanctions in 2018, $SanctionExposure_{it}$ data is absent. This results in a constrained range in the imposition and lifting of sanctions over the timeframe for which $SanctionExposure_{it}$ data is accessible.

Secondly, the sanctions levied against Iran cover a period of more than four decades, during which the intensity of sanctions has fluctuated significantly. This extensive and variable period lacks clear "sanctions on" and "sanctions off" phases, complicating the application of a difference-in-difference approach.

Acknowledging these constraints, I adopt the following specification to estimate the effect of sanctions:

$$y_{i,t} = \delta_i + \delta_t + \beta SanctionExposure_i \times SanctionEpoch_t + \gamma X_{it} + u_{it} \tag{11}$$

This regression employs data from the decade spanning the Persian calendar years 1390-1400. $SanctionEpoch_t$ is an indicator variable assigned a value of one during the years 1393, 1394, 1397, and 1398, corresponding to periods of maximum sanction intensity. $SanctionExposure_i$ represents the average sanction exposure for each firm over time.

Table 2.6 displays the results of this analysis. The sales growth rate, represented as $\frac{\Delta Sales_{i,t}}{\Delta Sales_{i,t-1}}$, indicates the annual change in sales relative to the previous year's sales. The capital investment rate, denoted as $\frac{I_{i,t}}{K_{i,t-1}}$, is calculated annually using the perpetual inventory method, the details of which are provided in appendix A2.1.

Column 1 shows the base specification of the relationship between sales and $SanctionExposure$, and, as control, the year and firm fixed effects. As anticipated, we find a significant negative association between $SanctionExposure$ and the sales growth rate, implying that firms most exposed to sanctions tend to experience lower sales during periods of intensified sanctions. Column 3 highlights firms exposed to sanctions retrench investment when faced with sanctions. Columns 2 and 4 include $SanctionRisk$ and $SanctionSentiment$. Aside from $SanctionRisk$ in the final column, all other variables display anticipated signs. However, their correlations are notably weaker and lack statistical significance, which is in alignment with the discussions outlined in Section Three. The last two columns look at the effect of sanctions on employment. It shows a negative impact on employment, although the effect is small. The fact that headcount employment is less responsive to an external shock compared to other firm-level outcome variables is consistent with the idea that some firms may have been able to maintain employment levels by reducing hours or wages. Employment costs are costly and sticky in the short term and thus are typically viewed as short-term fixed costs, making adjustments like layoffs costly and disruptive. Additionally, firms may prioritize workforce continuity and skill retention, anticipating a recovery after the shock, whereas investment decisions can be more easily deferred or adjusted in response to changing conditions.

2.5.3 Decoding Sanction Channels

The current findings prompt an inquiry into the specific risks and impacts that firms attribute to sanctions. In this section, I try to identify major channels through which sanctions will affect firms in as systematic a manner as possible. I achieve this through a

Table 2.6. Firm-level Effects of Sanctions

	$\frac{\Delta Sales_{i,t}}{\Delta Sales_{i,t-1}} * 1000$		$\frac{I_{i,t}}{K_{i,t-1}} * 1000$		$\frac{\Delta Emp_{i,t}}{\Delta Emp_{i,t-1}} * 1000$	
	(1)	(2)	(3)	(4)	(5)	(6)
$\overline{SanctionExposure_i}$	-0.768		-1.625		-0.182	
	(0.236)		(1.382)		(0.170)	
$\overline{SanctionRisk_i}$		-0.213		1.351		0.262
		(0.452)		(3.382)		(0.493)
$\overline{SanctionSentiment_i}$		0.205		1.017		0.093
		(0.100)		(0.750)		(0.116)
Observations	4195	4195	3697	3697	1174	1174
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table displays the relationship between exposure to sanction and firm-level outcome variables. Sales, $\frac{\Delta Sales_{i,t}}{\Delta Sales_{i,t-1}} * 1000$ is the change in year-to-year sales over last year's value. Capital investment, $\frac{I_{i,t}}{K_{i,t-1}} * 1000$, is calculated recursively using a perpetual-inventory method. Details are in the appendix A2.1. All regressions include firm and year-fixed effects. Standard errors are clustered at the firm level.

structured human reading of texts utilized to construct SanctionSentiment and SanctionRisk. The method involves scrutinizing the paragraphs encapsulating each instance of the term "Sanction".

In accordance with the methodology established by Hassan *et al.* (2024) and Baker *et al.* (2016), a structured approach was developed to undertake the human reading of these text fragments. The first step involved conducting interviews with business professionals, both domestic and international, who engage in importing to and exporting from Iran, as well as executives and market experts in the Tehran Stock Exchange to identify the various channels through which sanctions could affect Iranian firms. The discussions uncovered that businesses could face impacts from sanctions on both the supply and demand fronts.

On the supply side, firms may struggle to acquire the necessary inputs and intermediaries that were previously supplied from abroad. Even if they manage to find a foreign supplier, they may encounter obstacles when trying to obtain the required foreign exchange or make payments to the supplier. Additionally, finding a shipment company willing to transport their inputs and an insurance company willing to insure the cargo could be a challenge.

On the demand side, firms may struggle to find customers for their products or services, as many export destinations and foreign customers may no longer be accessible to them. Even if they find foreign customers, issues with money transfer and shipment could persist. Domestic consumers, including both the government and private sector, may also be affected by the sanctions.

After considering all these channels, the potential categories of impact were streamlined to:

(1) Restrictions on Money Transfers and Access to Financial Markets: Economic sanctions can hinder a company's ability to transfer money to and from Iran and limit its access to global financial markets. This can affect both their ability to pay their suppliers and receive payments from their consumers, as Iranian banks may be unable to make or receive payments to foreign banks where suppliers or consumers are located.

(2) Increased Logistics and Transportation Costs: Economic sanctions can raise the cost of logistics and transportation, as shipping companies and airlines may be hesitant to do business with Iran and shipping to and from Iranian ports, and insurance companies may not provide coverage for transportation to and from the country, leading to higher costs and potential delays in delivery. This also affects both the supply and demand sides for business.

(3) Other Export Limitations and Restrictions: Sanctions may hinder Iranian firms' ability to export goods through other means, e.g., by discouraging potential business partners from conducting transactions with them. This can manifest in traditional buyers halting purchases from Iranian businesses either due to government instructions or their own assessment of risks and compliance costs.

(4) Import Costs, Supply Chain Disruption, and Lower Foreign Investment: Economic sanctions can elevate the expenses associated with importing both goods and capital into Iran, while simultaneously creating disruptions in supply chains. This makes it difficult for companies to obtain the essential capital and inputs they need for their operations.

(5) Foreign Exchange and Currency Volatility: Economic sanctions can increase the exchange rate and currency volatility, making it difficult for Iranian companies to conduct international business and manage financial risk. The foreign exchange market in Iran is multi-layered, and firms planning to buy intermediaries might face obstacles in acquiring the required foreign currency due to market disruptions or government-imposed restrictions. Similar challenges happen for exporter firms trying to exchange their foreign currency for domestic currency.

(6) Reduced Government Support: Economic sanctions can limit the resources available to the Iranian government to support businesses and invest in infrastructure and public projects. As one of the largest customers in the economy, a reduction

in government demand can limit a company's access to government services and resources, potentially hindering its ability to operate and grow. This can be particularly problematic for businesses that rely on government contracts or subsidies, as reduced government spending can lead to a contraction in these markets.

(7) Lowered Demand and Market Contraction: Sanctions can impact the entire Iranian economy, leading to reduced demand for goods and services, whether for consumer-focused companies or those selling to other businesses. This can lead to a decrease in a business's revenue and profitability.

(8) Other channels: There could be other specific ways in which sanctions can affect businesses operating in Iran, depending on the type of business, industry, and partners involved.

Subsequently, an instruction manual was composed for two independent human auditors, who were recruited from the Ph.D. program of an Iranian economics department (Sharif University of Technology). The manual comprised elaborate step-by-step directives for classifying fragments into each of the eight topic categories. In addition, the study requested the auditors to flag fragments in which the meeting participants mentioned that sanctions had limited or no impact on the firm or fragments that the auditors found challenging to classify. Each auditor was asked to classify all fragments. The study found that the auditors agreed on the classifications most of the time, and in cases of disagreement, a third auditor was invited to provide judgment.

The transcripts presented in Table 2.7 provide sample excerpts on each topic related to sanctions. Upon reading the text, it becomes clear that the discussions primarily focus on specific channels through which the firm in question could potentially be impacted by sanctions. To illustrate the distribution of these topics, Figure 2.8 shows the proportion of each pre-defined category in the discussions of sanction risks. The horizontal axis represents the topic categories, while the vertical axis displays the proportion of each topic relative to all other specific topics mentioned by the firm. The figure shows, in equilibrium, the most prevalent channels through which sanctions are hitting Iranian firms are Export limitation, followed by increased import costs and increased logistics costs.

Additionally, the plot reveals that concerns over sanctions extend beyond politically connected firms. This is illustrated by the hashed area within each bar, which represents the proportion of concerns over sanctions originating from politically connected firms (data from Draca *et al.* (2023)). It becomes evident that, in equilibrium, the majority of concerns over sanctions arise from non-politically connected firms to some extent because most of the firms are not politically connected. This result lines up with the findings of chapter 2.4 that politically connected firms do not exhibit higher levels of sanctions exposure,

Table 2.7. Firms' Meetings Excerpts by Category

Company	Time	Translations of Excerpts
Challenge: Restrictions on Money Transfers and Access to Financial Markets		
IASCO	2018-08	Although the sanctions make transferring currency from exports challenging, past experience with sanctions has led to the development of new channels for money transfer- alternative foreign currency transfer methods have mitigated the impact...
Challenge: Increased Logistics and Transportation Costs		
Farsnov Cement Co	2018-09	The company is in talks with the Government Shipping Company to continue exporting at a similar rate as last year despite the harsh US sanctions. Additionally, our export product buyers have suggested alternative transportation methods...
Challenge: Other Export Limitations and Restrictions		
Pars Oil Co	2018-10	Question: Has the company experienced any issues with export sales due to the upcoming sanctions? Answer: It's uncertain how the upcoming sanctions will affect the company's exports as it depends on the specific mechanism of the sanctions, making it impossible to make a specific prediction...
Challenge: Higher Import Costs and Supply Chain Disruption		
Iran Tire Co	2018-10	Question: Has acquiring raw materials from overseas become problematic for the company since the sanctions? While half of the intermediate goods are sourced from foreign suppliers, the company aims to secure its raw materials on time despite the obstacles...
Challenge: Foreign Exchange and Currency Volatility		
Zagros Pharmed Pars Co	2018-08	Question: If sanctions are imposed, what exchange rate does the company use to import raw materials, and have there been any obstacles in obtaining them at this rate? Moreover, if the company utilizes the discounted central bank rate, what is the likelihood of this rate being liberalized?
Challenge: Reduced Government Support		
Persian Railway Transportation	2018-09	How has the estimated decrease in government oil production and export due to the sanctions affected the company's operations? Is there any alteration in the rate received from the National Iranian Oil Products Distribution Company per kilometer/ton transported?
Challenge: Lowered Demand and Market Contraction		
Persian Railway Transportation	2018-09	Will the decrease in fuel oil exports due to the oil sanctions lead to a reduction in the demand for transporting these materials to export terminals?
Challenge: Other channels		
IKCO	2018-10	Question: Is there a possibility that the production of Peugeot and Suzuki products will cease due to the current and future sanctions and the departure of foreign companies from Iran?

Notes: This table presents selected excerpts from Iranian publicly traded firms' meeting for each channel through which sanctions can affect firms. The original discussions were conducted in Persian and the English translation is reported here.

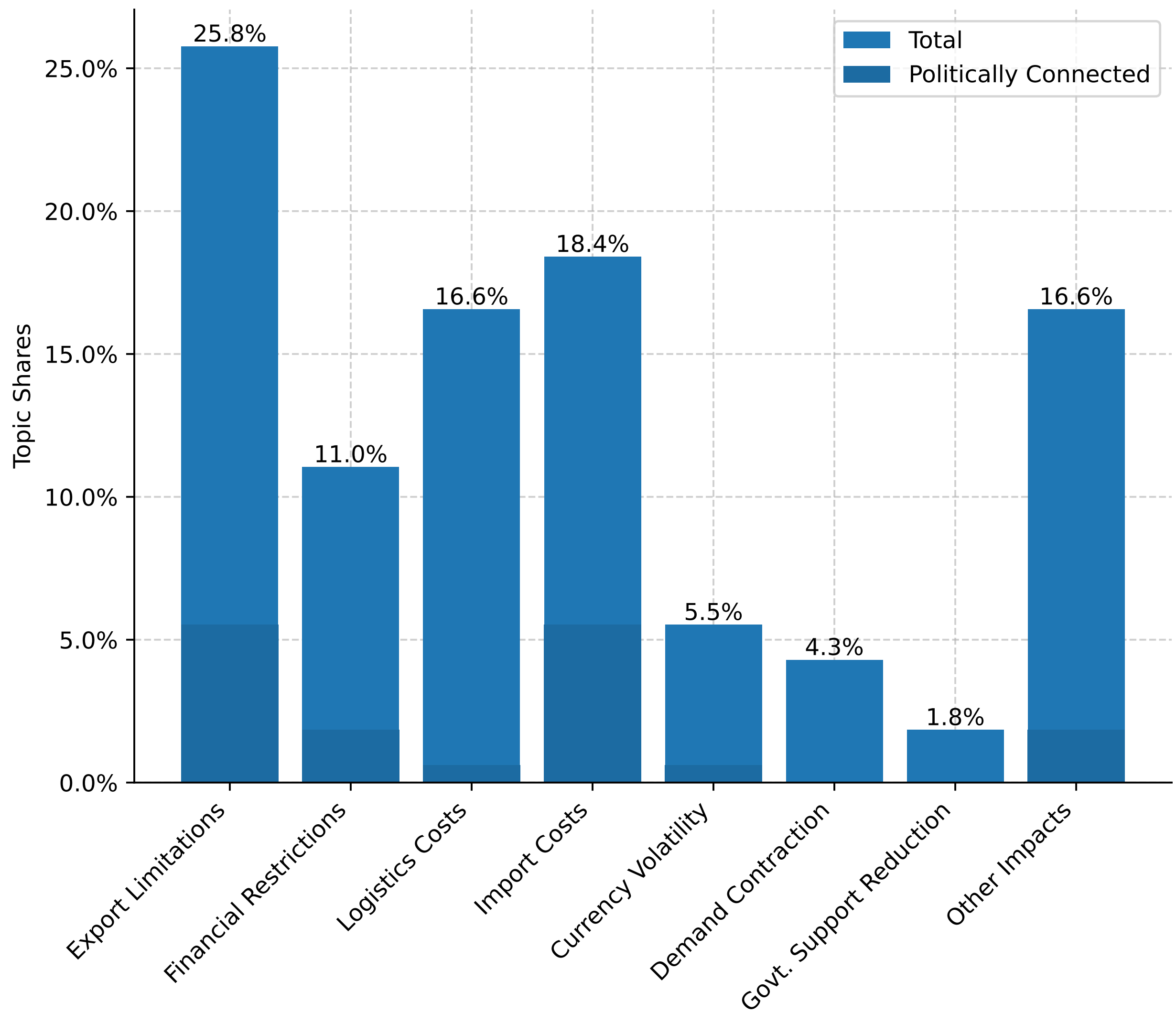


Figure 2.8. *Notes:* This figure illustrates the proportion of each topic category discussed during Iranian firms' meetings when the topic of "sanctions" was raised. The hashed area shows the share of mentions of "sanctions" that are from politically connected firms. The definition of connected firms is taken as the most extensive definition from Draca *et al.* (2023).

Figure 2.9 illustrates the progression of concerns related to sanctions over time. For each quarter, the figure displays the percentage of sanction-focused discourse dedicated to each channel. The graph suggests the relative share of each mechanism has remained remarkably consistent over time. The steadfastness of these thematic proportions suggests that businesses have settled into a rhythm of expectation and response regarding sanctions, possibly reflecting a market that has, to some extent, adapted to the persistent state of economic containment.

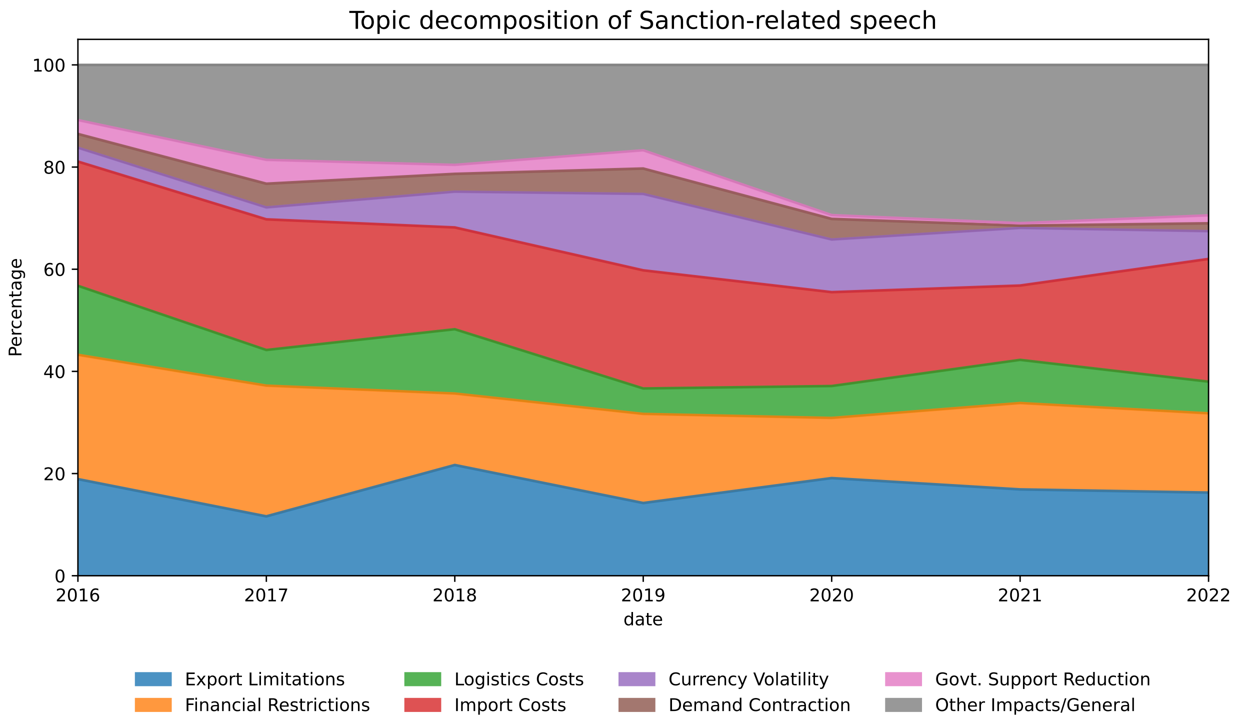


Figure 2.9. *Notes:* This figure displays the evolution over time of types of discussions in meetings of Iranian firms, focusing on sanctions. It breaks down the proportion of dialogue related to each of the eight identified channels through which sanctions exert their impact on firms, highlighting how the emphasis on different channels has shifted throughout the observed period..

2.6 Conclusion

This research, set within the context of sanctions imposed on Iran, looks into the challenges that firms face in an economy subject to different layers of sanctions. Utilizing a text-based methodology, I explore the economic effects of comprehensive sanctions at the firm level, offering unique insights into how these effects spread and the extent of their economic impact on the targeted firms. The study underscores that Iranian firms face considerable challenges due to sanctions, with the concern over sanctions at certain points surpassing even the anxiety induced by the COVID-19 pandemic.

A key finding of this research challenges the claim by some proponents of sanctions that they solely or predominantly target political decision-makers. The analysis reveals that politically-connected and non-connected firms are equally affected, indicating that sanctions often act as blunt instruments inflicting harm broadly. Such outcomes suggest that the multi-layered nature of sanctions hinders their precision, resulting in a widespread adverse effect. This study's evidence shows that for every $1 of loss inflicted on connected firms, an externality of $5 is imposed on non-connected firms.

Further, firms with higher sanctions exposure experienced a pronounced reduction in stock market value in response to unexpected sanction events. This investigation extends to explore the effects of sanctions on firms' investments and sales, establishing that sanctions can induce a considerable negative impact on the overall economic performance of firms. The sanctions have predominantly impacted firms by limiting access to export destinations and escalating import costs. This affirms that sanctions can create significant economic disruptions even when the political objectives may not be entirely realized.

The indiscriminate nature of multi-layered sanctions calls for a shift in focus: from the mere quantity, marked by successive additions of blanket sanctions, to the quality, which emphasizes strategic implementation. If the primary aim of sanctions is to target political decision-makers, it's important to recognize that the increasing complexity of these sanctions eventually leads to a loss of precision and thus diminishing returns. Beyond a certain point, they may even yield negative returns, as their impact on the non-political elite risks alienating the local population, and the added complexity of each new layer exacerbates the difficulty of their eventual reversal.

There are a few important points to consider regarding this analysis. Firstly, it's worth noting that the data only pertains to listed firms. This means that the overall impact on a national level may be more negative if unlisted firms, such as farmers, were also negatively affected on average. Alternatively, the impact may be less negative if the sanctions led to new businesses entering more exposed sectors.

Second, this paper employs a comparative analysis between firms more exposed to sanctions versus those less exposed within the Iranian economy. However, it's crucial to acknowledge that sanctions will likely exert substantial influence across the entire economy, not just on the firms directly exposed. Consequently, while the differential impact on more and less exposed firms can be discerned through this analysis, the total effect of sanctions on Iranian firms and the broader economy, taking into account these spillover and GE effects, is not quantified in this study.

Third, it is worth mentioning that this study does not directly delve into the effect of sanctions on households. While the focus is primarily on firms, the downstream effects on households may be substantial and warrant further exploration. The consequences of sanctions on household income, consumption patterns, and living standards constitute an important area for future research.

In conclusion, this paper contributes to the literature on the economics of sanctions, their political effectiveness, and the emerging field of text-as-data in economics. The findings emphasize that while multi-layered sanctions do impair firms in the recipient nation, they inadequately target politically connected entities. This research, hence, underscores the pressing need for more focused and careful considerations in the application of sanctions as a foreign policy tool, given the extensive, severe, but indiscriminate economic consequences they can impose on firms and, by extension, economies of the target countries.

Chapter 3

Immigration and Political Realignment

Immigration has become a contentious issue in many countries. What, then, are the electoral repercussions? Recent studies show that immigration can benefit the electoral prospects of right-wing, and sometimes far-right parties (Tabellini, 2020; Dustmann *et al.*, 2019; Halla *et al.*, 2017). However, the adoption of anti-immigration rhetoric by right-wing parties, rather than their left-wing rivals, presents a puzzle. Immigrant competition for jobs and potential wage suppression predominantly happens among the unskilled labour sector-a demographic traditionally inclined toward left-wing ideologies. Parties on the left, with their agendas centred on economic redistribution, might have been the more apparent recipients of support in the wake of this economic dislocation. Yet, intriguingly, it is the nativist and ethno-nationalist populists with low redistributive agendas who have seized this narrative.

In this paper, I delve into these dynamics within the UK context, focusing on the European Union (EU) immigrants from new member states. I show the sudden influx of immigrants changes attitudes toward anti-immigration stances and sways voters toward right-wing anti-immigrant parties. A similar anti-immigration rhetoric response is observed on the supply side of politics as political parties take more localist anti-immigrant rhetoric in response to immigration. The findings do not indicate significant negative effects on wages or undue strain on the welfare state attributable to immigration. Rather, I explain these results by showing the sudden influx of immigrants increases the salience of immigration in politics and makes this issue and broader cultural concerns the primary points of political contention. I empirically show that in response to immigration, the non-economic dimension becomes the driver of voting and group clustering. The results suggest that, in the wake of an immigration shock, working-class voters may pivot away from left-wing political parties, which would maximize their economic well-being, and lean toward nationalist parties that resonate with their national and cultural affiliations.

My research design uses the EU 2004 enlargement as a natural experiment. In 2004, ten Eastern European countries joined the EU. My identification strategy is based on the arrival of migrants from EU accession countries in different industries, proxying their comparative advantage, which UK locations are deferentially exposed to through predetermined industry specialization. This research design approximates an ideal experiment that would randomly assign a different number of migrants across different locations. To further address potential endogeneity, I instrument for the growth in migration from accession countries to the UK in each industry using migrants' growth in other pre-2004 EU members. This approach, which is inspired by the ideas presented in the paper by Autor *et al.* (2013), allows me to isolate the supply-driven variation in exposure to immigration and study its effects on voting.

The exclusion restriction underlying this approach assumes that the common within-industry component of rising immigration from accession countries in the UK and other European countries arises from the relative skills of accession countries' workers in different industries and occupations. This assumption posits that UK locations specializing in industries for which other pre-2004 European countries' industries attracted a high level of immigration are not unobservably different from other UK locations. To test this assumption, I perform several falsification tests using the lagged outcome variable. Across a range of specifications, results consistently support the assumption.

Using this measure, in the second part of the paper, I show people more exposed to immigration tend to vote more for nationalist and anti-immigration parties. I focus on the electoral outcome of the strongly eurosceptic UK Independence Party (UKIP), which directly reflects natives' demand for anti-immigration policies. Using both aggregate and individual-level data, I document that UKIP gained significant support in regions heavily impacted by immigration in general, European, and local elections. I then show immigration shifts people's attitudes to become more anti-immigration and socially conservative. I see a similar pattern concerning the 2016 Brexit referendum. Regions with heightened exposure to immigration demonstrated a marked inclination toward supporting Brexit, a trend robustly validated through both aggregate and individual-level data assessments. My counterfactual analysis suggests that if immigration had not been a factor, the outcome of the Brexit referendum would have been different.

The third part of the paper explores various potential mechanisms behind these trends, distinguishing between economic and cultural factors. The labor market analysis shows that immigration boosts economic activity and lowers unemployment without notably affecting wages, except for a modest impact at the lower end of the wage distribution. Additionally, I observe that immigrants reduce the pressure on the welfare system, contradicting claims that they burden the system. Despite these economic benefits, the evidence points to cultural dynamics as the primary drivers behind the observed shift in voting behavior. Specifically, in response to immigration, voters become more socially conservative and particularly more anti-immigrant in their attitudes and tend to vote based on factors other than the traditional left-right class dimension.

In the fourth part of the paper, I present evidence of an analogous shift on the political supply side. Using several techniques from natural language processing, I show that in their political speeches, UK parties have increasingly focused on immigration, often portraying it negatively. To shed light on the broader potential shift in cultural values, I use Enke (2020) measure and observe cultural polarization of political rhetoric. According to this metric, Conservative speeches have increasingly shed their universalistic undertones in recent years, while Labour speeches have adopted a more inclusive, universalistic rhetoric.

Finally, I make a case that all aforementioned results can be explained by a shift in voter alignment, transitioning from traditional economic considerations to cultural nuances, in reaction to immigration. I show while disagreements on redistribution policies show a downward trend in the UK, disagreements around cultural policies, particularly concerning immigration, have intensified. This shift in public discourse is further corroborated by clustering analysis, revealing a realignment of voter clusters from economic to cultural dimensions throughout the study. I see this clustering along cultural dimensions is stronger in regions hit hardest by immigration.

The shift from class-based to identity and culture-based politics risks sidelining the critical focus on redistribution and the welfare state in the face of rising economic inequalities. As identity-driven narratives gain predominance, policymakers may find it increasingly difficult to implement policies aimed at economic efficiency or equity if such policies are at odds with the dominant identity-driven political narratives. This transition can amplify polarization on matters such as immigration, globalization, and nationalism, fostering extreme policy stances. Such conditions are fertile grounds for the rise of populism, where political leaders might leverage identity concerns to rally support, potentially at the cost of overlooking detailed economic strategies.

This work builds on and integrates several literature strands. I contribute to the literature on electoral repercussions of immigration that predominantly finds increased immigration increases support for right-wing parties. For example, Tabellini (2020) found that although immigration in the interwar United States conferred economic gains on the host community, it concurrently amplified support for conservative politicians and anti-immigrant policies. Mayda *et al.* (2022) identified a similar trend with low-skilled immigration to the United States from 1990 to 2016. Analogous results emerge in European contexts, including Austria (Halla *et al.*, 2017; Steinmayr, 2021), Italy (Barone *et al.*, 2016), Spain (Mendez and Cutillas, 2014), and Germany (Otto and Steinhardt, 2014). The predominant methodology within this literature is the "shift-share" empirical design, which integrates historical settlement patterns across regions with the contemporaneous national migration influx. This approach aims to address reverse causality issues, specifically the tendency of potential immigrants to avoid regions perceived as unwelcoming, and omitted variable bias, the idea that intertwining factors can concurrently shape immigration patterns, economic dynamics, and political attitudes.

This work contributes to the literature on the political effects of immigration in several ways. Methodologically, I introduce a novel quasi-experimental shift-share design based on the industry composition of each region and the comparative advantage of immigrants across industries. I further instrument this measure with industry-specific immigration to other non-UK EU countries. This approach leverages a new variation in exposure to immigration, previously unexplored, and addresses some limitations of traditional shift-share instruments that rely on historical settlement patterns[16]. The traditional shift-share instrument based on prior settlements is in particular not suited to study Eastern European migration to the UK as evidence indicates that the historical distribution of Eastern European migrants in the UK is not a strong predictor for later inflows of later migrants. This study also broadens its scope to study the supply side of politics, analyzing political responses to immigration at a granular, sub-national level. Lastly, the paper provides evidence on the realignment of voters along the cultural dimension, offering insight into the mechanisms through which immigration intensifies anti-immigrant sentiment and subsequently bolsters support for right-wing factions. It also provides an answer to the puzzle of why anti-immigration sentiment predominantly translates into heightened support for right-wing parties, rather than left-leaning ones.

A subset of the previously mentioned literature focuses on the impact of immigration within the context of the UK, particularly in relation to the Brexit referendum (Becker *et al.*, 2016, 2017; Colantone and Stanig, 2018). Notably, Becker *et al.* (2017) and Colantone and Stanig (2018) find no positive correlation between EU immigration and the leave vote in the Brexit referendum. However, this study reveals that this is due to not accounting for the selection of immigrant locations. Upon isolating exogenous immigration shocks, it becomes evident that immigration impacts voting behavior. The findings align with those of Becker *et al.* (2016) and Viskanic (2017), who observed an increase in UKIP's vote share following the influx of Eastern European migrants. Compared to these studies,

[16]According to Borusyak *et al.* (2022), these widely used shift-share instruments based on historical settlement patterns ultimately resemble traditional difference-in-differences models, contrasting regions with and without historical settlements. Such an approach may not sufficiently control for unobserved time-varying confounding shocks. In contrast, this paper's research design pivots on the exogeneity of shocks and, as will be demonstrated, possesses a sample size of shocks substantial enough to mitigate the challenges commonly associated with traditional approaches.

this paper adopts a new measure of exposure to immigration based on a novel shift-share instrument and sheds some new light on the underlying mechanism behind these electoral dynamics. Moreover, this paper intersects with adjacent literature exemplified by Carreras *et al.* (2019), which explores the cultural and economic divisions underlying Brexit. Unlike related works that primarily use correlational analysis, this research employs causal inference to more accurately assess these dynamics.

This work is related to a recent new line of research focusing on identity in economics. This literature acknowledges individuals' multiple identities and explores how these are prioritized based on economic factors. Shayo (2009) models identity choice as a balance between societal status and group alignment costs, suggesting that social identity formation and economic conditions are interlinked. Grossman and Helpman (2021) apply this to trade policy, showing how economic changes can shift self-identification and influence protectionist tendencies. Bonomi *et al.* (2021) introduce multiple political dimensions (economic left versus economic right, culturally liberal versus conservative), indicating that the salience of the issue, shaped by economic shocks, can redirect social identities and influence political alignments, transitioning the traditional left–right divide to a liberal–conservative one. Besley and Persson (2019) explores how voters' beliefs and party affiliations evolve with economic shifts and the salience of non-economic factors, like immigration, highlighting a dynamic interplay between economic conditions, social identity, and political landscapes.

The empirical results of this paper align with and complement several theoretical papers within this literature. Notably, based on Gennaioli and Tabellini (2023), when socially conservative voters, often less skilled, are more exposed to immigration, or when the salience of immigration issues increases, voters are more likely to align with their cultural identity rather than their economic class. The central theme of this paper complements these theories by providing empirical evidence that immigration can impact party support and reshape political cleavages, underscoring the increasing importance of cultural factors in political decision-making by overshadowing the traditional emphasis on class-based politics. This finding aligns with one of the scant empirical investigations in this area, Danieli *et al.* (2022), which highlights how people's priorities have shifted from economics toward cultural issues over time.

Finally, this paper stands at the intersection of political economy and computational linguistics, contributing to a burgeoning literature that employs text data to parse complex socio-political phenomena (Wilkerson and Casas, 2017; Gentzkow *et al.*, 2019). While previous research has applied computational methods from natural language processing to trace the portrayal of immigration in parliamentary discussions (Nguyen *et al.*, 2015; Card *et al.*, 2022), these studies have not explored how political parties causally respond to local immigration shocks. This work is among the pioneering efforts, alongside a select few such as Bhatiya (2023), to apply text analysis for examining the degree of political responsiveness to constituency-level shocks. I employ a range of text metrics to assess legislators' engagement with immigration issues and also incorporate the approach devised by Enke (2020) to quantify the universal values in legislators' rhetoric.

The subsequent sections of this paper are structured as follows: Section 3.1 outlines the study's data and context. Section 3.2 introduces the immigration exposure measure and delineates the empirical approach. Section 3.3 examines the immigration impact on voting patterns at both aggregate and individual levels. Section 3.4 explores the underlying mechanisms of this impact, suggesting a cultural rather than economic influence, as immigration appears not to detrimentally affect the economy yet significantly alters cultural attitudes. Section 3.5 scrutinizes the political supply side, utilizing natural language processing to assess political reactions to immigration surges. In Section 3.6, evidence is presented to support the thesis that an immigration shock catalyzes a transition in voter alignment from conventional class-based distinctions to cultural identity considerations. Finally, Section 3.7 discusses the implications and offers concluding remarks.

3.1 Setting and Data

In this section, I provide context for the study by discussing the background and political context of the EU and immigration in the UK. Specifically, I examine the EU enlargements in 2004 and 2007 and the influx of migrants from accession countries to the UK, as well as the political manifestation of these events. I then describe the data sources and variables used in the analysis.

3.1.1 Background

The roots of the European Union (EU) can be traced back to the post-war 1950s. The Treaty of Rome in 1957, signed by Belgium, France, West Germany, Italy, Luxembourg, and the Netherlands, initiated the European Economic Community (EEC), a customs union that embedded free labor mobility into its framework. This set the stage for the EU as we know it.

The UK initially hesitated to join the ECC but later made two applications in 1963 and 1967 that were vetoed by France. The UK ultimately joined the EEC in 1973. A referendum followed in 1975 due to Labour's promise to reevaluate ECC membership and consult the public on these new terms. The public was asked if the UK should remain in the European Community. The affirmative response by a margin of 34.5 percent confirmed the UK's membership under the revised conditions.

Upon joining, the UK was instrumental in driving forward economic integration, particularly through its pivotal role in establishing the Single Market in 1986, advocating for the free movement of goods, services, capital, and labor. In 1992, the EEC transitioned into the European Union with the Maastricht Treaty. The UK was cautious about further deeper political integration, opting out of the Euro currency and the Schengen Area.

On May 1, 2004, ten countries including eight from Eastern Europe, alongside Malta and Cyprus, joined the EU, expanding the EU's population by nearly 75 million. This was the largest enlargement of the EU since the UK joined in 1973. Bulgaria and Romania also joined on January 1, 2007, adding an additional 30 million people to the EU. Upon their accession to the EU, the UK Tony Blair's government was one of the few member countries that did not impose temporary restrictions on the arrival of migrants from these new member states (hereinafter referred to as "NMS").

Evidence suggests the actual number of migrants from these countries coming to the UK was much higher than what the UK government had anticipated. Figure 3.1 shows the number of migrants based on the country of origin over time, using data from the Annual Population Survey (APS). As the figure shows, prior to 2004, the majority of EU-born migrants to the UK came from "EU-14" countries that had joined the EU before 2004. However, after 2004, there was a significant increase in the number of migrants from the NMS. The population of NMS-born residents in the UK increased by more than a factor of 10, from an estimated 160,000 in 2004 to 1,850,000 in 2017. This stands in contrast to the more gradual increase in migrants from EU-14 countries. Notably, after the Brexit referendum, the number of migrants from NMS began to decline. These features indicate that the expansion of the EU represents a sudden significant shock to the influx of migrants to the UK.

The other remarkable aspect of this new wave of migration is that the spatial distribution of these new migrants within the UK is also different from the spatial distribution of migrants from these countries who entered the UK before 2004. This is evident in Figure A1, which shows the share of NMS migrants as a share of each local authority population. These distinctive characteristics motivate my empirical analysis to use the 2004 and 2007 EU enlargements as a natural quasi-experiment.

As the European Union's influence expanded, so too did the opposition within the UK to further integration. This opposition is best reflected in the United Kingdom Independence Party (UKIP). Originally established as the Anti-Federalist League in 1991, this single-issue Eurosceptic party was rebranded as its current name in 1993, broadening its manifesto to encompass a wider right-wing agenda with the primary objective of withdrawing the UK from the EU. Although UKIP struggled to secure seats in the UK Parliament due to the first-past-the-post electoral system, they achieved greater success

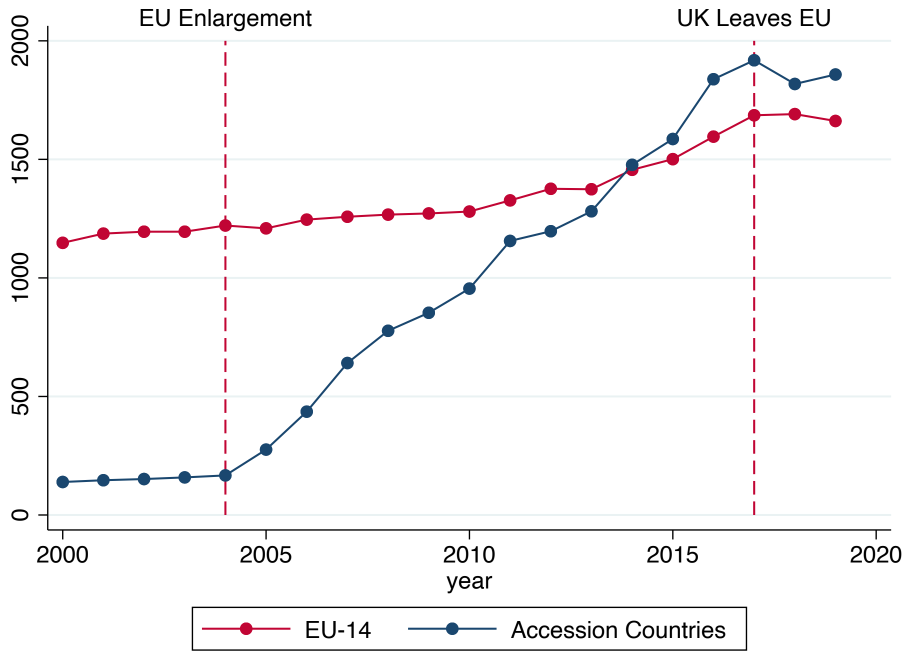


Figure 3.1. *Notes:* This graph shows EU-born migrants in the UK over time. Data is sourced from ONS, Population by Nationality, and Country of Birth. The estimated population of residents in the United Kingdom is categorized by country of birth, excluding those living in communal accommodations such as hostels or care homes. Estimates are based on the Annual Population Survey (APS), comprising wave 1 and wave 5 of the Labour Force Survey (LFS), plus annual sample boosts. The sample boosts are included primarily to improve geographical coverage. For statistics relating to accession countries before 2004, data is sourced from the quarterly Labour Force Survey, since the original dataset does not include this information. Accession countries refer to those that joined the UK in 2004.

in European Parliament (EP) elections. This success can be attributed to two main factors: the implementation of proportional representation in European elections, and the fact that European elections tend to focus voters' attention on EU-specific issues. In the 2014 EP elections, UKIP secured a victory with 26.2% of the vote.

UKIP's rise in the UK mirrors broader trends observed in several other Western countries, reflecting a growing wave of populist and Eurosceptic sentiment. This phenomenon is characterized by skepticism toward globalization signified by supranational institutions like the European Union, concerns over national sovereignty, and often a tough stance on immigration. Similar movements have gained traction in countries like France with the National Rally (formerly National Front), Germany with the Alternative für Deutschland (AfD), Italy with the League (Lega Nord), and the United States with the election of Donald Trump, who capitalized on themes of anti-immigration and anti-establishment rhetoric. These parties and leaders typically channel public frustration over economic dislocations, perceived loss of cultural identity, and dissatisfaction with the political status quo.

Before the 2015 general election, Prime Minister David Cameron, seeking to appease the Eurosceptic wing of his party and counter the UKIP threat, made a strategic pledge to renegotiate the UK's terms with the European Union and to hold an in-out EU membership referendum, should the Conservatives secure a majority. This move was largely seen as an attempt to reunite his party and retain votes that might have otherwise gone to UKIP. He made this promise in the light of predictions showing the most likely scenario would be a hung parliament. Contrary to widespread expectations, the Conservatives won an outright majority. This unexpected electoral result forced Cameron to uphold his referendum promise, ultimately leading to the 2016 Brexit referendum.

Throughout the Brexit referendum campaign, the issue of immigration emerged as a pivotal and divisive issue, particularly emphasized by the Leave campaign and UKIP. Many proponents of Brexit adeptly tapped into public concerns over rising immigration levels, framing the EU's free movement of people as a loss of British control over its borders. They argued that the UK should regain control over who enters the country and adopt an "Australian-style points system" that treats EU and non-EU migrants equally. As illustrated in Figure A4, in the lead-up to the election, the level of concern regarding immigration significantly increased, surpassing economic issues. Additionally, the disparity in concern between Labour and Conservative voters expanded considerably, a notable change from 2001, when immigration was only a minor issue. Polling data also revealed that a significant driver for the Leave vote was the desire to regain control over immigration and borders, with 33% of Leave voters indicating this as their primary motivator, based on an election day survey of 12,369 voters by Ashcroft (2016). Ultimately, the UK voted to leave the EU by 52% to 48% on June 23, 2016, after a contentious 10-week campaign.

While Brexit was a culmination of concerns over immigration, this pattern seems to begin to intensify with the rise in migrants from NMS. Captured in Figure 3.2, this escalation is evidenced by three interwoven indicators – public opinion, media representation, and parliamentary focus – all of which collectively illustrate how immigration became a pivotal issue in the UK, ultimately peaking during the lead-up to the Brexit referendum. Public opinion, as depicted in the top panel, reflects a growing perception among the populace that immigration was a top issue facing the country. The middle panel's portrayal of media representation echoes this sentiment, revealing a parallel increase in the frequency with which immigration was featured in the nation's most widely-read newspapers. The bottom panel, showing parliamentary focus, indicates that the issue was not only a matter of public and media concern but also a significant topic of legislative discussion, with mentions of immigration in the House of Commons spiking alongside the other indicators. The timeline of these indicators provides a suggestive narrative: as the number of NMS immigrants grew, so did the salience of immigration as a political and societal issue, a trend that reached a critical point with the Brexit decision.

Contrary to the focus on immigration by the Leave campaign and its salience throughout the campaign, data suggests that areas with a higher proportion of foreign-born residents were, paradoxically, more inclined to vote Remain in the EU (Becker *et al.*, 2017; Colantone and Stanig, 2018). A plausible rationale is that immigrants often gravitate toward regions with more inclusive cultures and robust economies, as exemplified by London, which absorbed a significant portion of net migration from NMS and voted predominantly for Remain. The subsequent chapter will delve into the causal relationship between immigration and the rise of anti-immigration sentiment.

3.1.2 Data Sources

The data on the composition of employment at local authorities or constituency comes from the Office for National Statistics' (ONS) Business Register and Employment Survey (BRES). The BRES is an annual business survey that provides employee and employment estimates at detailed geographical and industrial levels. It is the official source of employee and employment estimates by detailed geography and industry in the UK.

To categorize workers according to the type of firm for which they work, I use the two-digit Standard Industrial Classification (SIC). Data on the number of migrants from NMS at the national level is obtained from the Labour Force Survey (LFS), a quarterly survey conducted by the ONS that provides information on the employment status and characteristics of the UK population. The LFS is a large, nationally representative sample survey that is widely used to produce official statistics on the UK labor market. [17]

I estimate the annual bilateral gross migration flows from NMS to other European countries in each industry for the period 2004-2016 using data from the European Union Labour Force Survey (EU-LFS). The EU-LFS is a large household sample survey conducted by Eurostat that aims to provide quarterly results on the labor participation of people aged 15 and over, as well as those outside the labor force, in 35 participating countries. It is the largest European survey of its kind and is widely used to generate official statistics across the labor markets of European countries.

[17] Data prior to 2006 are reported using the SIC 1992 classification, while data from 2006 onwards use the SIC 2007 classification. I use Office for National Statistics proportional mapping between these two classifications. A proportional mapping provides the most accurate correspondence when the focus is on aggregate or mean measures, like in our case.

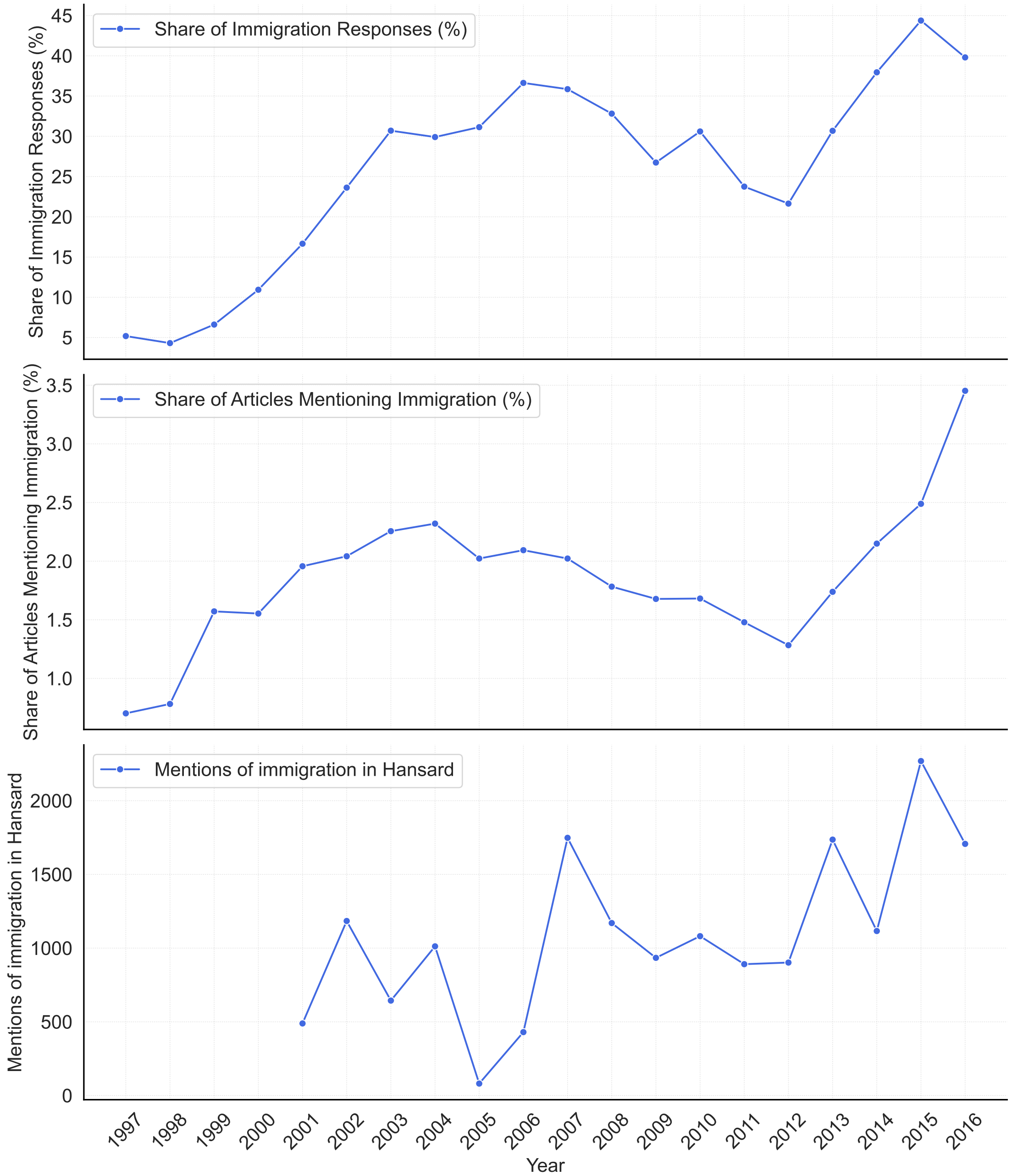


Figure 3.2. *Notes:* The top panel quantifies public opinion, displaying the annual percentage of respondents identifying immigration as one of the top three critical issues in the UK (Source: Ipsos Mori). The middle panel examines media representation, showing the weighted proportion of articles mentioning immigration in the three highest-circulation UK newspapers during this period: The Sun, Daily Mirror, and Daily Mail. The bottom panel offers a parliamentary viewpoint, illustrating the frequency of immigration mentions in the Hansard records by MPs in the House of Commons.

For the examination of anti-immigration sentiments, I utilize individual-level data from the British Election Study (BES), specifically focusing on Wave 8, conducted between May 6 and June 22, 2016, immediately preceding the Brexit referendum on June 23. This wave encompasses responses from a substantial sample size of 31,409 participants. It not only captures the vote intention in the referendum but also provides a comprehensive dataset that includes attitudes toward immigration, among other variables. To add a geographical dimension to the analysis, I categorize each respondent based on their place of residence, assigning them to their corresponding local authority.

Labor market analyses rely on the Annual Population Survey (APS) for regional unemployment and economic activity rates, alongside the Annual Survey of Hours and Earnings (ASHE) for detailed hourly wage data across wage distribution quantiles, broken down by local authority of residence.

For the individual-level analysis of voting, I use data from the UK Understanding Society panel survey. Understanding Society is a panel survey of households in the UK that collects data on a wide range of topics related to social, economic, and health issues. It is conducted by the Institute for Social and Economic Research (ISER) at the University of Essex and began in 2009, with ongoing waves of data collection every year. Understanding Society builds on the British Household Panel Survey (BHPS), a similar panel survey conducted from 1991 to 2009. The panel for Understanding Society consists of around 40,000 households, with approximately 80,000 individuals participating. The survey includes detailed information on demographics, employment, education, health, and other topics, as well as measures of attitudes and beliefs, including those related to immigration.

To examine parties and the evolution of their ideological positioning, I use data from The Chapel Hill Expert Survey (CHES). CHES is an ongoing initiative led by the Center for European Studies (CES) at the University of North Carolina at Chapel Hill. It systematically collects expert assessments on the ideological positions and policy stances of a wide range of political parties in Europe. I use CHES data to determine parties' stances on economic and cultural issues over time.

In addition, to further delve into MPs' positions on immigration and social values, I use collections of the UK Parliamentary Debates (Hansard) extracted through web-scraping their official website (https://hansard.parliament.uk). These records are released under the Open Parliament License, facilitating their use for research with appropriate attribution. By analyzing the language and tone used in these debates, I can better understand parties' cultural views, in particular on immigration, and how they may have changed in response to immigration.

3.2 Empirical Strategy

This section introduces the concept of immigration exposure, which measures the extent to which a region is affected by immigration. I will then explain the method I use to construct an instrument for immigration exposure, which helps to identify the causal effect of immigration on my outcome of interest. Finally, I will present the results of the first stage and test the validity of the instrument.

The 2004 EU enlargement introduced a large influx of immigrants to the UK. The immigration exposure measure, or immigration shock, is designed to capture the shock felt by the average worker within distinct local labor markets. This quantification is achieved by weighting the national industry-level migration growth from NMS by the local area's share of employment in that industry. Essentially, this index reflects how much each local labor market is exposed to migration from NMS, based on its industrial composition. If a location has a high share of employment in industries that are facing significant NMS migration, its import exposure index would be high.

More precisely, this measure comprises two elements: national-level shocks and predetermined local exposure shares. The shocks are calculated from the variation in the flow of migrants from NMS over time across different industries. Each shock represents the national-level change in the number of migrants from accession countries in each 2-digit industry, comparing year t to 2004. This approach hinges on the idea that NMS individuals possess a comparative advantage in certain sectors relative to UK workers. This comparative advantage, coupled with the sudden EU enlargement, naturally inclines them toward employment in certain sectors. The industry-level shocks are then combined with exposure shares, s_{ik}, which are calculated based on the specialization of industries in different locations. Consequently, the measure of the immigration shock at the regional level is calculated as follows:

$$\Delta IM_{it} = \sum_k s_{ik} \frac{\Delta IS_{k,t}}{L_k} = \sum_k \frac{L_{ik}}{L_i} \frac{\Delta IS_{k,t}}{L_k} \tag{12}$$

In the above formula i indexes regions, k indexes industries, and t indexes times. The national-level change in the number of migrants from NMS in each 2-digit industry k between periods t and 2004, represented by $\Delta IS_{k,t}$, is normalized by the total number of workers in the same industry in the UK, represented by L_k. The region-specific shock is then calculated as the weighted sum of these changes in immigration share across industries, with the weights reflecting the respective significance of each industry within that region. I look at the net change in immigration (i.e., the net flow) over this time frame since 2004 is the year at which 10 out of 12 NMS joined the EU. While I include immigrants from Bulgaria and Romania who joined the EU in 2007 in shock, their exclusion does not change results qualitatively.

To avoid simultaneity bias, I use start-of-period shares (i.e., shares in 2004) in the above formula. While lagging the shares by more periods could help to isolate cleaner time-varying shock variation, it might also reduce the predictiveness of the exposure measure and thus reduce the efficiency of the analysis. Notably, the shares from before 2004 are reported in a different industry classification version, and using a mapping to convert them to the current classification will introduce additional noise. However, as these shares do not vary significantly over time, the year in which they are calculated has minimal impact on the results.

To address the concern that changes in UK industry demand may affect the influx of migrants, I use a non-UK exposure variable, $IM^O_{i,t}$, as an instrument for the immigration exposure $IM_{i,t}$. This variable is constructed using data on contemporaneous industry-level growth of migrants from NMS to other existing European countries. The idea behind using this instrument is that the flow of migrants from NMS to the UK might be influenced by changes in both UK supply and demand conditions, which may have direct effects on our outcome variable in UK regions. However, the flow of migrants to other European countries is influenced only by the comparative advantage of migrants and some domestic supply and demand shocks. The instrument is calculated as follows:

$$\Delta IM^O_{it} = \sum_k \frac{L_{ik}}{L_i} \frac{\Delta IS^O_{k,t}}{L_k} \tag{13}$$

where $\Delta IS^O_{k,t}$ is the change in NMS migrants for other European countries for 2-digit industry k between periods t and 2004. This expression can be motivated by the fact that other European countries in the EU are similarly exposed to the influx of migrants from accession countries, which is driven by the comparative advantage of these workers in certain industries. This approach is based on the logic presented in Autor *et al.* (2013). Conceptually, this instrument leverages multiple sets of shocks. One can treat industry shocks from each individual country as an independent as-good-as-randomly assigned instrument. However, to align with the approach used in Autor *et al.* (2013), I use the average migration across ten EU members as my instrument.

Before the 2004 expansion, the EU comprised 15 countries, including the UK. In my analysis, I focus on 10 of these countries (EU10): the Republic of Ireland, Sweden, Greece, Spain, Finland, Italy, Portugal, the Netherlands, Luxembourg, and France. These nations imposed no or relatively mild restrictions on NMS migrants compared to Belgium, Denmark, Austria, and Germany, which are excluded.[18] Including all EU members in the instrument doesn't markedly affect the results. Although some included countries had transient restrictions on migrants from accession nations, these were comparatively lenient than those in the omitted nations and were phased out within a few years. Moreover, these restrictions were generally uniform across sectors, hinting that the migrant composition across industries remained unaffected. The key findings remain robust, even when the instrument is restricted to only Sweden and Ireland - two countries that, akin to the UK, avoid any entry restrictions from the outset.

The identification of shift-share instruments hinges on the exogeneity of either the shocks, the shares, or both. Conventional shift-share instruments in immigration literature, using pre-settlements patterns, are generally perceived as leveraging exogenous shares (Goldsmith-Pinkham *et al.*, 2020). However, in our context, shares are unlikely to be exogenous as they are equilibrium shares that could measure the location's exposure to any unobserved demand or supply shocks across industries (e.g. China import competition or automation). Instead, I rely on the exogeneity of the shocks to establish the validity of my identification approach, as formalized by Borusyak *et al.* (2022).

Following the framework established by Borusyak *et al.* (2022), the validity of this instrument is anchored in specific identification conditions. The first condition is the relevance condition, such that the instrument has power. More precisely, we should have $E[\Delta IM_{it} IM^O_{it} | X_{it}] \neq 0$. Figure 3.3 plots the relationship between actual and predicted immigration exposure in each local authority. This parallels the first-stage regression in the later analysis, conducted without any controls. The t-statistic and R-squared are 4.9 and .5, respectively, revealing the substantial predictive power of the other EU countries' instrument for changes in immigration exposure for the UK.

Building on the numerical equivalence in Borusyak *et al.* (2022), when a shift-share research design leverages exogenous variations in shocks, the exclusion restriction can be written as an orthogonality condition between the underlying shocks and shock-level unobservable. Omitting the time subscript for brevity, Borusyak *et al.* (2022) formalize this condition as follows:

$$(\frac{1}{I}\sum_i \Delta IM^O_i \epsilon_i \xrightarrow{p} 0) \iff (\frac{1}{K}\sum_k \hat{s_k} \frac{\Delta IS^O_k}{L_k} \bar{\epsilon_k} \xrightarrow{p} 0) \tag{14}$$

where $\hat{s_k} = \frac{1}{I}\Sigma_i s_{ik}$ and $\bar{\epsilon_k} = (\Sigma_i s_{ik}\epsilon_i)/\Sigma_i s_{ik}$. Casting the exogeneity assumption as a condition on shocks, we can see that the consistency of my estimates can be inferred from the law of large numbers as applied to the equivalent shock-level regression. This means that

[18] Moreover, Germany's data at the 2-digit industry level is unavailable in the EU-LFS for the study period.

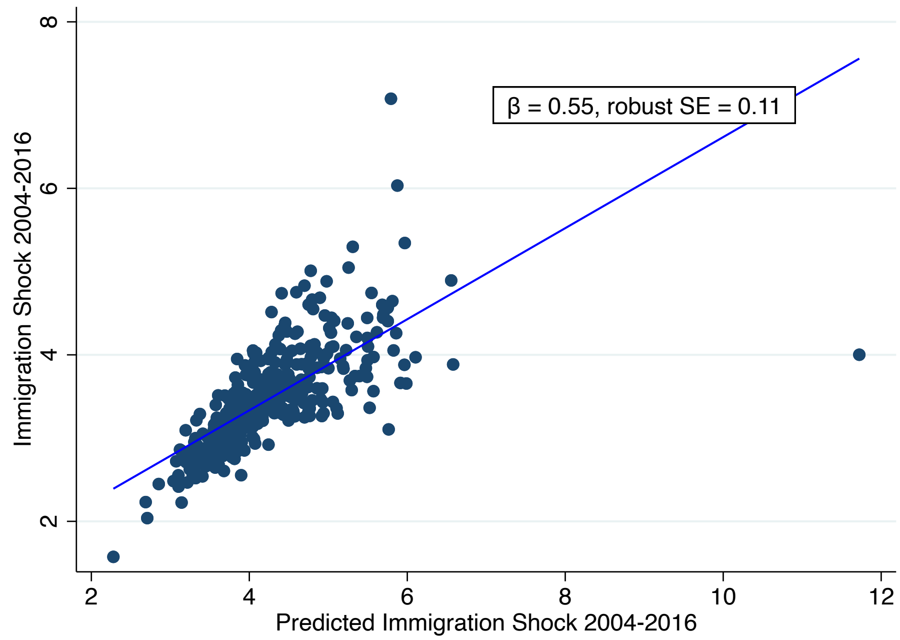


Figure 3.3. *Notes:* This graph depicts the first-stage relationship between actual and predicted immigration shocks in local authorities across the UK from 2004 to 2016. The actual immigration shock is derived from industry-specific changes in NMS immigration within the UK, weighted by the representation of each industry in the local authorities, as detailed in equation 12. In contrast, the predicted shock is computed using similar industry weights but combined with the change in immigration in each industry to other non-UK, pre-2004 EU countries. Each data point corresponds to one of the 390 local authorities.

my shares are allowed to be endogenous. Falsification tests using lagged outcome variables will confirm the as-good-as-random assignment of shocks, validating the shift-share instrument. According to Borusyak *et al.* (2022), the concentration of industry exposure as measured by the inverse of its Herfindahl index (HHI), $1/\Sigma_{n,t}s_{nt}^2$, corresponds to the effective sample size. As I will discuss, the inverse of the HHI of the weights $\hat{s}_n$ is 389, reassuring that my effective sample size is large enough.

Table A1 presents the distribution of the instrument, which is based on migration from NMS to other European countries. The distribution appears to be regular, with a significant amount of variation. The effective sample size is 389. The second column only includes shocks in 2016, as the Brexit specification only considers cross-sectional variation in immigration exposure in 2016. As expected, the effective sample size for this subset is smaller, which is an important factor to consider in the cross-sectional analyses focusing only on 2016.

It is important to note that in a shift-share design, the assumption of independent and identically-distributed (iid) observations is unlikely to hold. As a result, conventional standard errors may not be valid in the presence of exposure-based clustering, as pointed out by Adao *et al.* (2019) and Borusyak *et al.* (2022). In the table appendix Table A2, I follow Adao *et al.* (2019) to correct for standard errors for the main analysis. These potentially more conservative standard errors do not significantly differ from the baseline standard errors.

One potential threat to the identification is that the immigration from NMS countries to other European countries might not only reflect the comparative advantage of immigrants but also demand shocks that are common between the UK and other European countries. A related concern is that migration shocks might be confounded by other unobserved characteristics. For example, migrants from NMS might tend to work in industries that are concentrated by routine jobs, which are already on a different labor market trend. To address these concerns, I will control for a range of technological shocks and conduct a series of falsification tests to confirm my assumption that I have a quasi-random shock assignment with a large enough effective sample size.

Furthermore, the fact that the decision to expand the EU was made collectively by countries outside the UK supports my assumption that the influx of immigrants after 2004 was driven by supply rather than demand. As previously mentioned, both the composition and spatial distribution of immigration after 2004 differed significantly from the pattern of immigration prior to that year. Furthermore, if demand were a significant factor in determining immigration patterns, my estimates of the effect of immigration on anti-EU sentiments would likely be downward biased. This is because negative shocks to a particular industry would result in that industry receiving fewer immigrants, and regions specialized in that industry would be more likely to support anti-EU platforms. Therefore, my results can be considered conservative estimates.

There has been a recent discussion following the observations by Jaeger *et al.* (2018) on shift-share instruments, pointing out potential issues when there's a slow adjustment process and a high serial correlation in the immigrants' country-of-origin distribution. They suggest that this setup might blur the distinction between immediate reactions to new immigrant arrivals and delayed responses to previous inflows. For several reasons, these concerns do not significantly apply to this analysis.

First, I am exploiting an exogenous structural break in the pattern of immigration that dramatically changed the country-of-origin mix of immigrants, as evident in Figure 3.1. This means the serial correlation of immigrant flow with the flow before 2004 is very low. This argument is supported by the findings of Jaeger *et al.* (2018), which suggest that shift-share instruments are still consistent when there is a structural break in their aggregate components.

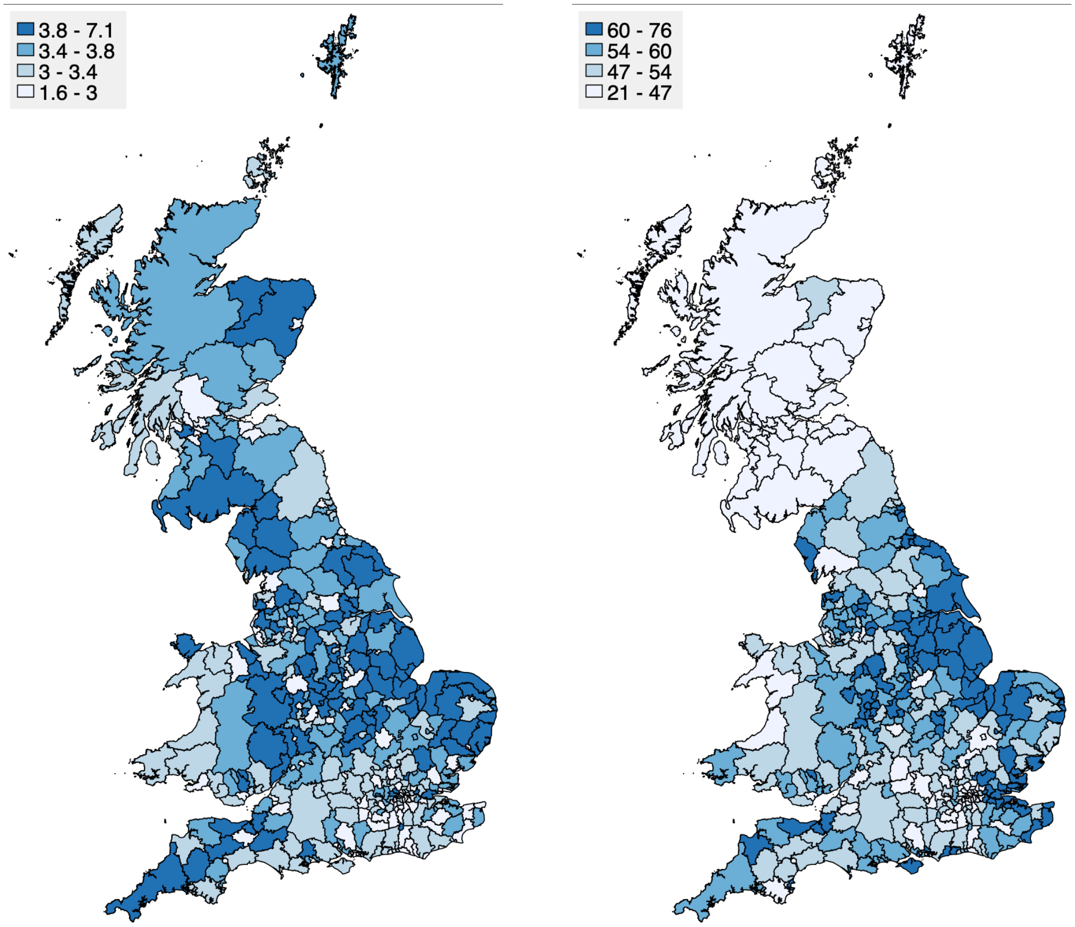
Second, the concerns raised by Jaeger *et al.* (2018) are unlikely to apply in my setting because I do not use past settlement patterns as shares but rather the employment structure. This further reduces the issue of serial correlation. Third, general equilibrium adjustments are much more relevant for wages, as the adjustment in the capital may gradually offset the initial negative effect of immigration on wages and lead to subsequent return and positive wage growth. Priorly, there is no reason to expect such dynamic adjustments in electoral outcomes in response to immigration.

Immigration exposure by location authority reveals a considerable amount of geographic variation in its strength. In Figure 3.4 panel A, I report immigration exposure in 2016, the year of the EU referendum. The results indicate that locations in the Midlands and Northern England are hit hardest by immigration, with some strong effects elsewhere. This shock spans from a low of 1.34 in the City of London to a peak of 6.5 in South Holland (East Midlands), averaging out at 2.74, accompanied by a standard deviation of 0.58, reflecting its dispersion.

As a point of comparison, Figure 3.4 panel B displays the geographic variation in the Brexit vote. Brexit vote tends to be high in locations in the Midlands and Northern England, where the immigration exposure is also high. Interestingly, these places are known as "The Red Wall", a term describing constituencies that historically supported the Labour Party (but "turned blue" in the 2019 general election). While not crucial to the identification strategy, these facts provide context and help to better understand the role of immigration in the Brexit vote. The histogram in Figure A2 plots the range of variation in the immigration shock measure for 2016, indicating a significant amount of variation in this measure.

3.3 Voter Behavior and Attitudes

Using several survey data and official election results, this section studies how immigration affects voting decisions. I establish that regions with higher exposure to immigration exhibit a significant tilt toward right-wing anti-immigration UKIP party and the Leave campaign in the 2016 Brexit referendum. I will also conduct placebo tests to ensure that my results are not being driven by some underlying, long-term factor that impacts both immigration and anti-EU sentiment. In the following chapter, I will delve into the mechanisms driving this political realignment.



(a) Immigration Shock

(b) Leave Vote Share

Figure 3.4. *Notes:* This map shows the spatial distribution of immigration shock and Leave vote across local authorities in the UK. Panel A illustrates the strength of the immigration shock in 2016 at the local authority level, with darker shades indicating a stronger shock. Panel B shows the Leave vote share in the 2016 Brexit referendum, with darker shades representing a higher percentage of votes for Leave.

3.3.1 Voting Patterns: Administrative Data

The initial investigation into the political ramifications of immigration begins with an analysis of its impact on voting decisions. The focus here is primarily on the electoral performance of UKIP. The party's vote share is often interpreted as a barometer for British Euroscepticism, a sentiment that culminated in the 2016 EU referendum's leave vote. To analyze the relationship between exposure to immigration and support for UKIP, I employ a pooled difference-in-difference approach. The core of this analysis is reported in the following equation:

$$y_{i,r,t} = \alpha_i + \eta_{r,t} + \beta \Delta IM_{it} + \epsilon_{i,r,t} \tag{15}$$

where $y_{i,r,t}$ represents the share of UKIP in location i, in region r, in the election held at time t. The immigration shock, ΔIM_{it}, is instrumented using the variable IM^O_{it}, as described previously. Throughout the paper, I look into three different types of elections. Except for general elections, which are reported at the constituency level, the spatial unit of my analysis would be the local authority.

Before 2004, the number of migrants from NMS was minimal (as illustrated in Figure 3.1). As a result, in the construction of immigration shock, ΔIM_{it}, it is practically equivalent to considering the level value of migrants or the change from 2004. Given the natural experiment of the EU enlargement occurred in 2004, shocks prior to this year are set to zero, aligning with the negligible NMS immigration to the UK before this period. The adjusted immigration shock formula, shown in equation 16, refines the definition by setting pre-2004 shock values to zero and maintaining the post-2004 immigration exposure as previously defined. This empirical specification exploits the national-level, time-varying shocks that impact different industries when immigrants enter the UK labor market, as well as the variation in employment composition across places. The instrument defined in equation 13 will also be refined accordingly and the same approach will be used throughout the paper.

$$\Delta IM_{it} = \begin{cases} 0 & t < 2005 \\ \sum_k \frac{L_{ik}}{L_i} \frac{\Delta IS_{k,t}}{L_k} & t \geq 2005 \end{cases} \tag{16}$$

I start by presenting the results of the OLS relationship between immigration exposure and vote for UKIP in Table 3.1, panel A. Column 1 shows the effect of immigration exposure on UKIP vote share in European elections held in 2004, 2009, and 2014. The results indicate that local authorities that experienced a significant influx of migration from NMS saw a significant increase in UKIP vote shares. Specifically, a one standard deviation increase in immigration shock would increase the UKIP vote share by 1.6 percent.

The analysis extends to local and general elections, as presented in Columns 2 and 3, respectively. Each electoral context offers distinct dynamics and complexities. For instance, local elections are more frequent compared to their European and general counterparts, ensuring that in any given year, certain local authorities are actively engaged in council elections. However, within the scope of this study, only three instances each of European and general elections were observed. Additionally, turnout in local and general elections is generally higher than in European elections. Conversely, local and general elections, unlike European ones, employ a system of First-Past-The-Post (FPTP), potentially incentivizing strategic voting. Moreover, the issue of immigration, intrinsically tied to EU dynamics, assumes greater prominence in European elections. Consequently, UKIP's performance in these elections might more accurately mirror the electorate's stance on immigration. Despite these electoral nuances, the results from Columns 2 and 3 consistently indicate that heightened immigration is correlated with increased support for UKIP.

Panel B provides estimates of the effect of immigration exposure on the vote for UKIP by instrumenting the immigration exposure by similarly constructed measures using immigration change from NMS to other pre-2004 European countries. The results are similar in magnitude to those obtained through OLS, suggesting that the source of bias may not be significant. By comparing the "Average effect in the last election" and "Mean of dependent variable" rows in columns 2 and 3, it is clear that a large portion of support for UKIP in local and general elections can be attributed to the immigration shock. It is important to note that the vote share for UKIP in European elections, as indicated in the "mean of dependent variable" row, is much higher than in other elections. This is due to the use of a proportional voting system in European elections, which benefits smaller parties like UKIP, as well as the greater salience of issues related to Europe in these elections, as mentioned before.

I perform a pre-trend falsification test by examining the relationship between the immigration shock and the performance of UKIP in previous elections to confirm the orthogonality of my shocks. Specifically, I regress the outcome variable at different points in time on the immigration exposure in the latest election year in my sample period (which is 2016, 2015, and 2014 for local, general, and European elections, respectively). This specification allows for the impact of immigration to be different at different times. The lagged dependent variable serves as a proxy for unobserved error terms ϵ_{it}, and the lack of a relationship supports my identification strategy. I estimate the following equation:

$$y_{i,r,t} = \alpha_i + \eta_{r,t} + \sum_{t\in[2000,2016]} \beta_t \times Year_t \times IM_{i,2016} + \epsilon_{i,r,t} \tag{17}$$

In Figure 3.5, I plot out the estimated coefficients $\hat{\beta}_t$, which are coefficients of the interaction of the immigration shock in the last election year and a set of year fixed effects, over time for local, European, and general elections. We would not expect the exposure in 2016, just before the Brexit referendum, to predict the support for UKIP in prior elections. As shown in the figure, the relationship is indeed absent in elections before the referendum. All three plots suggest that I cannot reject the hypothesis that there is no relationship between the lagged outcome variable and current shocks. Panel A suggests that immigration exposure in 2016 only had a significant effect on UKIP electoral outcome

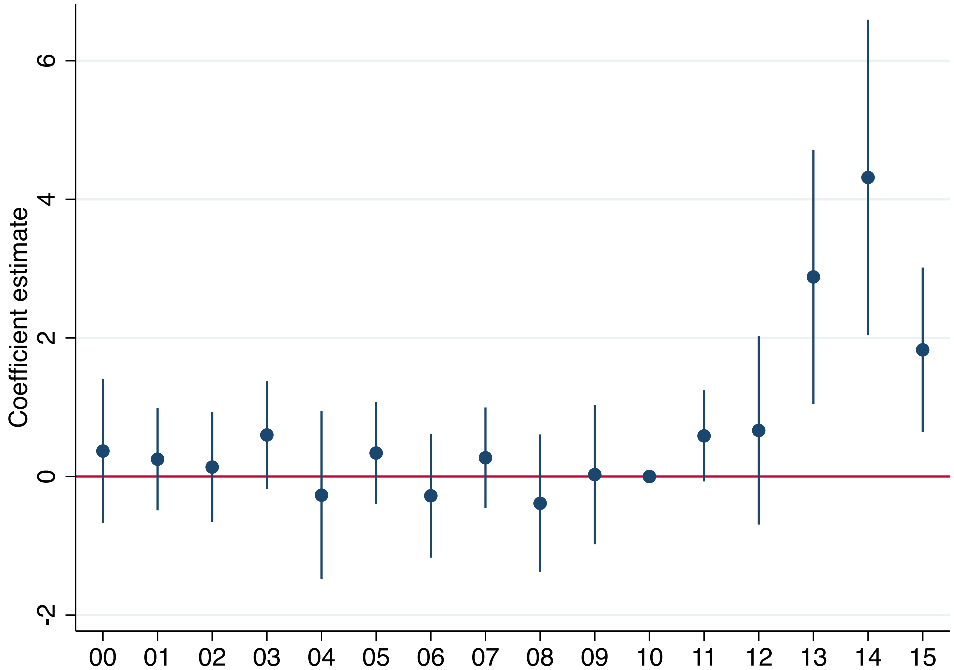
Table 3.1. Effects of Immigration on the Electoral Performance of UKIP

	(1) European elections	(2) Local Elections	(3) General Elections
Panel A. OLS			
Immigration Shock	1.636	1.279	2.181
	(0.464)	(0.520)	(0.297)
Avg effect in the last election	5.238	4.097	6.874
Standard deviation	.9922	.7760	1.349
Mean of dependent variable	22.3	4.49	6.03
Panel B. 2SLS			
Immigration Shock	1.407	0.992	2.293
	(0.555)	(0.779)	(0.291)
F-stat	196	254	406
Avg effect in the last election	4.505	3.178	7.226
Standard deviation	.8532	.6020	1.418
Mean of dependent variable	22.3	4.49	6.03
LA/Constituency FE	Yes	Yes	Yes
Region-Year FE	Yes	Yes	Yes
Spatial units	347	346	566
Observations	1041	3263	2047

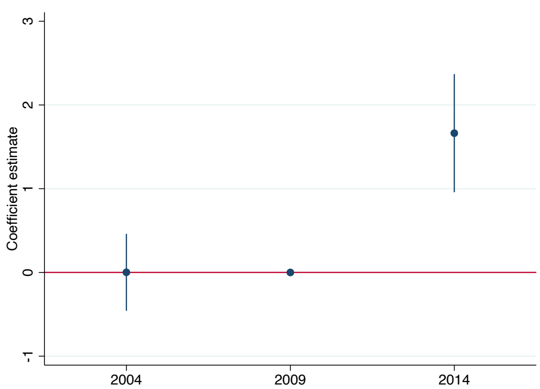
Notes: This table presents the estimated effects of immigration shocks on the electoral performance of the UK Independence Party (UKIP) across different types of elections: European, local, and general. The immigration shock variable is constructed using industry-specific changes in immigration, weighted by the industry composition of each region. The exact construction of the immigration shock and its instrument is explained in the text. F-stat refers to the Kleibergen-Paap rk Wald F-statistic for weak instruments. Robust standard errors, clustered at the local authority (for Local and European Elections) or constituency level (for General Elections), in parenthesis.

in the few years prior to 2016. Specifically, the constructed shock is not statistically associated with support for UKIP before 2013. Panels B and C, which look at European and General elections respectively, also show that the exposure measure in the last election year only explains the outcome in the last election year.

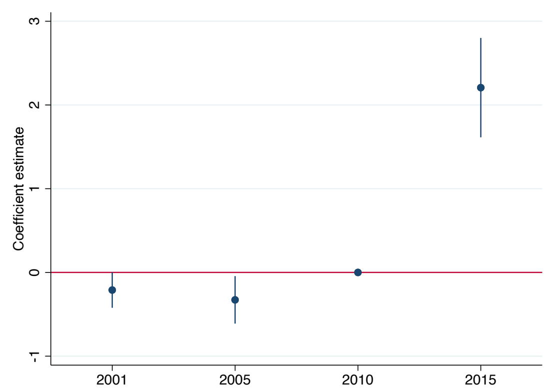
As an alternative specification, I estimate the model in first differences, separately for each period. This approach has the advantage that by focusing on a precise, fixed time frame between two consecutive elections for each regression, it ensures that the analysis



(a) Local Election



(b) European Election



(c) General Elections

Figure 3.5. *Notes:* Analysis of Pre-trends in Votes for UKIP. This figure presents the impact of immigration shocks in the last election year on the percentage of votes for UKIP in English and Welsh local, European, and general elections from 2000-2015 in panels a, b, and c, respectively. The graph shows point estimates of the interaction between the immigration shock and a set of year-fixed effects, while controlling for local authority district fixed effects and region-by-year fixed effects. Standard errors are clustered at the local authority (for Local and European Elections) or constituency level (for General Elections), and 90% confidence bands are shown.

captures the net average effect specific to that interval. This addresses the concern articulated by Jaeger *et al.* (2018) that my estimates may conflate short and long run responses. Furthermore, this refined approach facilitates the execution of a pre-trend test, adding another layer of robustness to the analysis. It is worth highlighting that as immigration exposure is zero before 2004, for any period post-2004, the level of immigration exposure and its change from 2004 would essentially be the same. The model estimated is as follows:

$$\Delta y_{i,t} = \alpha_{j(i)} + \beta \Delta IM_{i,t} + \epsilon_{it} \tag{18}$$

I will estimate this model using European, Local, and General elections. When analyzing local elections, it is important to consider the fact that these elections take place at least every 4 years, but not all local governments hold elections at the same time. Some local governments elect all of their local councilors every 4 years, while others elect half of their councilors every 2 years, and some elect one-third of their councilors every year. Instead of running different regressions for every combination of two elections, which would result in few observations and many coefficients, I consider four different periods: 2000-2003, 2004-2007, 2008-2011, and 2012-2015. Each local authority in each of these periods had at least one election. When there is more than one election, I take the average.

Table 3.2 displays the outcomes of the model estimated using first differences. The initial three columns demonstrate that both OLS and 2SLS estimates yield coefficients that are consistent in sign and magnitude across various election types, underscoring the robustness of the statistical associations. Specifically, the first column analyzes European elections between 2004 and 2014, the second focuses on general elections from 2005 to 2015, and the third examines local elections spanning from 2004-2007 to 2012-2015. Regardless of the election type and the estimation method (OLS or 2SLS), the findings consistently indicate that regions experiencing a substantial influx of immigrants are more likely to support UKIP, a party representing anti-immigration politics. This trend confirms that the impact of immigration shock extends beyond merely influencing attitudes, manifesting clearly in voting behaviors that favor anti-immigration parties. It indicates that exposure to immigration influences people's opinions, and these altered attitudes become significant considerations in their voting decisions.

The subsequent three columns in Table 3.2 explore the link between past changes in the electoral outcome of UKIP and future changes in immigration exposure. This analysis acts as a falsification test, aiming to verify that the observed results are not confounded by any long-term common factors that may be affecting both the success of UKIP and the increase in immigration exposure. The lack of significant findings in these columns lends

Table 3.2. First Difference Estimation

	Main analysis			Pre-trend analysis		
	(1)	(2)	(3)	(4)	(5)	(6)
Election:	European	General	Local	European	General	Local
	2014-2004	2015-2005	(2012-15)-(2000-3)	2004-1999	2005-2001	(2004-7)-(2000-3)
Panel A. OLS						
Current Imm. Shock	1.729	1.983	2.735			
	(0.442)	(0.345)	(0.635)			
Future Imm. Shock				-0.019	-0.149	-0.006
				(0.325)	(0.093)	(0.170)
Panel B. 2SLS						
Current Imm. Shock	2.045	2.919	3.032			
	(0.612)	(0.394)	(0.941)			
Future Imm. Shock				-0.274	-0.237	0.088
				(0.495)	(0.117)	(0.212)
F-stat	77.9	261	75.3	77.9	292	75.3
R-Squared	347	573	346	347	570	346

Notes: This table displays the outcomes of first-difference estimations examining the effects of immigration shocks on the electoral performance of the UKIP across various election types. The analysis is conducted separately for European, local, and general elections and captures the net effect specific to each time window. For local elections, the analysis is segmented into four distinct periods (2000-2003, 2004-2007, 2008-2011, and 2012-2015) to accommodate varying election cycles across local governments. When multiple elections occur within a period, the average outcome is considered. The last three columns serve as a falsification test, exploring the relationship between past changes in UKIP's electoral outcomes and future changes in immigration exposure. F-stat refers to the Kleibergen-Paap rk Wald F-statistic for weak instrument. Standard errors are clustered at the local authority (for Local and European Elections) or constituency level (for General Elections) and are presented in parentheses.

weight to the assertion that the identified effects are capturing the period-specific effects of immigration exposure.

Alternative Standard Errors: In light of discussions by Adao *et al.* (2019) and Borusyak *et al.* (2022), accounting for correlated errors in shift-share research designs is crucial. The findings remain significant across various inference methods designed to mitigate biases stemming from correlated unobservables among locations sharing similar characteristics. Table A2 presents standard errors as derived from the methodologies of Adao *et al.* (2019) which doesn't show significant differences from the conventional standard errors.

Other Parties' Support: Should the rise of UKIP be attributed to immigration, it's crucial to discern which parties are bearing the brunt of this shift. Such insights not only deepen our understanding of immigration's impact on the political landscape but also are crucial for traditional political parties to refine their electoral strategies, focusing on appealing to those who might be swayed by UKIP's messaging. Table A3 examines the effect of immigration shock on the support for the Conservative and Labour parties. The evidence suggests that immigration is helping UKIP gain support at the expense of the Labour Party. That is, panel A of the table indicates a significant loss of support for the Labour Party in areas with a higher level of immigration shock in European, local, and general elections. Estimates are similar in magnitude in OLS and 2SLS estimates.

On the other hand, panel B of Table A3 shows no evidence of an effect of immigration shock on the support for the Conservative Party. This is consistent with the idea that societal focus shifts from class-based distinctions to cultural-based distinctions will harm traditional left-wing parties the most. This is because these parties traditionally focus on class struggle, advocating for the working class against the capitalist elite. A shift toward cultural issues might dilute their traditional class-based message, especially among conservative voters, causing them to lose votes.

Brexit Referendum: Given the substantial impact of immigration on UKIP support and the critical position of UKIP in shaping the narrative around the Brexit referendum, it's crucial to investigate the direct impact of immigration on this defining political event. The Brexit referendum was not just a reflection of UKIP's political agenda, but also a crucial indicator of public sentiment toward immigration. To unravel the extent to which immigration shock contributed to the Leave campaign's success, the following baseline specification is employed:

$$y_i = \alpha_{j(i)} + \beta IM_{i,2016} + \epsilon_i \tag{19}$$

where y_i is the vote share for the leave option in local authority i. The results of this analysis are presented in Table 3.3. All regressions, except for column 1, include fixed effects $\alpha_{j(i)}$ for NUTS-1 region j in which local authority i is situated. I exclude local authorities in Scotland as I suspect the political landscape in Scotland can be very different from the rest of the UK. However, the results are mainly robust to the inclusion of Scotland. I also drop Northern Ireland and Gibraltar as the largest and smallest 'local authority' by order of magnitude.

The point estimates across all specification and estimation methods indicate a strong positive relationship between exposure to immigration shock and leave vote share. The effect is quite substantial; according to the second column, two regions situated within the same NUTS-1 region but differing in exposure to immigration shock by one standard deviation are expected to vary by 5% in support of the leave campaign. This suggests that a modest decrease in the magnitude of the immigration shock may have resulted in a different outcome in the referendum.

To further strengthen the validity of the results, the analysis progressively incorporates additional controls. In column 3, adjustments are made for demographic variables while column 4 also controls for other factors impacting the labor market throughout the study period. These include the volume of imports from China between 1990 and 2007, and changes in routine occupations, as proxied by their baseline employment shares. These controls reduce the magnitude of the coefficient on immigration shock but also make it more precisely estimated. These patterns strengthen the presumption that the pattern of migration from NMS across different industries is a supply-driven force that is largely unrelated to other industry shocks. Interestingly, this specification does not find a relationship between exposure to Chinese import competition and support for the leave campaign, which is in contrast to some previous studies that have suggested a link between these factors.

Finally, column 5 looks at the effect on turnout and finds a modest effect. This could indicate that the referendum held significant importance for locations hit by immigration, possibly due to concerns about the implications of Brexit on immigration policies, and rights to live, work, and move freely.

Counterfactual Analysis: The findings reveal a positive causal relationship between immigration and the Leave vote, necessitating further investigation to determine whether this impact extends to altering major political events. To undertake this counterfactual analysis, I rely on the results in the fifth column of Table 3.3. I evaluate the political consequence of these estimates by constructing a counterfactual leave vote share that would have occurred in the absence of increases in migration. The counterfactual leave vote share at the national level can be expressed as:

$$Leave\hat{S}hare = \sum_i E_i(L_i - \beta \widetilde{IS}_{i,2016}) \tag{20}$$

where β is the 2SLS coefficient estimate of the effect of immigration on the leave vote share, E_i and L_i are the electorate size and the observed leave share in local authority i, respectively. $\widetilde{IS}_{i,2016}$ is the estimated immigration shock that can be attributed to the supply-driven component of the increase in migration from accession countries in local authority i. The calculation of $\widetilde{ISi}, 2016$ involves multiplying the local authority i observed immigration shock by the partial R-squared from the first-stage 2SLS regression, valued at 0.51 in our base case (refer to Figure 3.3). This $\widetilde{IS}_{i,2016}$ variable is a consistent estimate of the contribution of the supply component of migration to changes in the actual increase in migration, assuming the instrument's validity and absence of measurement error.

The analysis does not account for the turnout effect, given the uncertainty regarding how immigration-induced new voters might vote compared with the existing voter base. In creating the counterfactual scenario, I also assume that other factors, including observed covariates and unobserved factors reflected in the error term, remain constant despite removing the supply-driven migration increase from new EU countries. The results suggest that the leave vote share in the counterfactual world, where there is no immigration from accession countries, would be 48.1%. This finding implies that a modest decrease in immigration shock could have been sufficient to tip the balance toward the remain camp in the Brexit referendum.

3.3.2 *Voting Patterns: Individual Survey Data*

Now, I use Understanding Society panel data to extend the analysis to the individual level and see whether the same pattern holds at the individual level. Using panel data at the individual level allows me to control for respondents' fixed characteristics, such as ethnicity, cohort, and education. Leveraging the longitudinal aspect of the data, in Figure A3, I use a Sankey diagram to visualize some descriptive information about where supporters of UKIP and the Leave campaign came from. Panel A shows a substantial flow from Conservative to Leave, and a smaller but significant flow from Labour to Leave. UKIP supporters exhibit an almost exclusive flow toward Leave, validating using UKIP as a proxy for anti-EU and anti-immigration policies. Panel B depicts the flow of people in terms of their party support. It maps people's party support in 2015 to their previous party support. It reveals two critical trends: a substantial share of UKIP's support base comprised individuals previously outside the traditional two-party preference, and there was a considerable flow from Labour to UKIP. These patterns suggest that UKIP's appeal transcended traditional party lines, possibly tapping into broader concerns among voters that are not strictly defined by the conventional left-right political spectrum.

Table 3.3. Effects of Immigration on Brexit Referendum

	Leave vote				Turnout
	(1)	(2)	(3)	(4)	(5)
Panel A. OLS					
Immigration Shock	7.074	5.126	2.645	1.881	0.447
	(1.969)	(1.217)	(0.908)	(0.805)	(0.250)
Panel B. 2SLS					
Immigration Shock	7.401	4.780	2.959	2.134	0.691
	(2.393)	(1.201)	(0.721)	(0.618)	(0.279)
R-Squared	.216	.428	.745	.783	.853
Observations	348	348	348	345	345
Region Fixed Effects	No	Yes	Yes	Yes	Yes
Demographics	No	No	Yes	Yes	Yes
Initial composition of immigrants	No	No	No	Yes	Yes
Routine Jobs	No	No	No	Yes	Yes
Import Competition Exposure	No	No	No	Yes	Yes

Notes: This table examines the direct impact of immigration on the Brexit referendum. All regressions control for NUTS-1 regions. Columns 2-4 add three sets of controls. First, they add demographics which include employment share of manufacturing, construction, and agriculture, and the share of people 20-44 years, 45-59 years, and people over 60 years old. Second, they control the share of employment in routine jobs at the baseline as well as the vote share of UKIP in the 2004 European election. Finally, the last set of covariates controls for the growth rate of migration from EU15 countries and non-EU countries (2001-2011) as well as the initial NMS resident share. Standard errors are clustered at NUTS-1 level, and presented in parentheses.

While Figure A3 provides insight into which party UKIP supporters and Leave campaigners previously supported, it does not directly explain how immigration impacts voting behavior. To probe this dynamic, I examine the relationship between individual voting patterns and the degree of immigration shock encountered in their local areas. This inquiry is formulated through the estimation of the following econometric model:

$$\Delta y_{j,t} = \alpha_j + \eta_t + \beta \Delta IM_{i(j),t} + \epsilon_{jit} \tag{21}$$

In Table 3.4, I report the results of the individual level analysis. The preferred specification is the last column, which includes individual-fixed effects as well as region-wave-year time fixed effects. By including individual fixed effects, the model capitalizes on within-individual variations in immigration exposure over time, while controlling for constant individual characteristics. Other included fixed effects account for time-varying demand and supply shocks at the governmental region and national level. The results show that individuals who experienced a significant influx of immigration in their local area are more inclined to support UKIP. Both OLS and 2SLS methods validate this finding, which also mirrors the aggregate-level analysis. Both individual and aggregate analyses indicate a remarkably consistent effect size; a one-standard deviation rise in immigration shock increases the likelihood of voting for UKIP or UKIP vote share by around 2%.

While the analysis indicates a causal relationship between the immigration shock and increased voting for UKIP, it does not specifically identify if these UKIP voters are the ones who have developed more anti-immigration attitudes, as the Understanding Society lacks direct queries on immigration attitudes or social policy preferences. Nevertheless, this finding, in conjunction with the patterns I have documented previously, aligns with the notion that an immigration shock elevates the salience of immigration in the political sphere and media discourse, potentially shaping individuals' beliefs toward anti-immigration stances, which then crystallize into a distinct voting pattern that diverges from the traditional left-right ideological spectrum.

Table 3.5 extends the analysis and looks at support for the leave campaign and turnout at the 2016 Brexit referendum. The leave campaign variable is constructed using a number of questions that ask individuals about their perception of the EU. The results on the effect of immigration on support for the leave campaign, represented in the first three columns, indicate that the immigration shock is driving people toward voting leave in the referendum. Results are consistent when estimated using both OLS and 2SLS methods. The last column of the table shows that immigration does not appear to have any significant effect on turnout in the referendum. It is worth noting that individual fixed effects could not be included in this analysis because the relevant data was only collected in one wave. Instead, a rich set of demographic variables was included.

In table A4, I run a couple of placebo tests to investigate whether the results found in the previous analyses hold up when considering different time periods. Specifically, I regress measures of anti-EU attitudes prior to 2016 on the 2016 immigration shock. It is expected that the 2016 immigration shock should not be correlated with pre-period attitudes. The results show that out of the four different variables tested, only one of them appears to have a significant relationship. This suggests that the previous findings on

Table 3.4. Individual-level Analysis (I)

	(1)	(2)	(3)	(4)	(5)
	Support for UKIP				
OLS Estimates:					
Immigration Shock	0.026	0.025	0.016	0.015	0.023
	(0.006)	(0.006)	(0.005)	(0.005)	(0.007)
2SLS Estimates:					
Immigration Shock	0.089	0.089	0.020	0.019	0.073
	(0.024)	(0.024)	(0.009)	(0.009)	(0.028)
Observations	236,312	236,310	220,202	220,196	220,196
Local Authority FE	Yes	Yes	No	No	Yes
region x wave x time FE	Yes	Yes	No	Yes	Yes
individual FE	No	No	Yes	Yes	Yes
region x year FE	No	No	Yes	No	No
Demographics	No	Yes	No	No	No

Notes: This table examines the relationship between individual-level voting behavior and local immigration shock, specifically focusing on support for UKIP. Demographic variables include age, income decile, highest qualification, current employment status, and occupation. Standard errors, adjusted for clustering at the local authority, are shown in parentheses.

the relationship between immigration shock and support for the leave campaign and for UKIP are robust and not simply due to some other common factor driving both variables.

The observed impact of exposure to immigration on the shift toward right-wing, anti-immigration parties and supporting the Leave vote in the referendum can be due to several reasons. The upcoming analysis in section 3.4 will first demonstrate that neither labor market dynamics nor pressure on the welfare system fully explains this shift. It will then examine how this trend may reflect a shift in voters' attitudes toward immigration. In Section 3.6, the discussion broadens to reveal a more comprehensive transformation: cultural alignment is identified as the dominant divide in response to immigration, overshadowing traditional economic factors in the electoral decision-making process. This observation underscores that economic incentives are no longer the predominant determinants of political leanings.

3.4 Unveiling the Underlying Mechanisms

This section explores the mechanisms behind the observed relationship between immigration and UKIP support. There are several potential drivers of this pattern. For

Table 3.5. Individual-level Analysis (II)

	Support for Leave Campaign			Turnout
	(1)	(2)	(3)	(4)
OLS Estimates:				
Immigration Shock	0.074	0.057	0.053	-0.009
	(0.011)	(0.009)	(0.009)	(0.011)
2SLS Estimates:				
Immigration Shock	0.095	0.069	0.065	0.001
	(0.014)	(0.013)	(0.012)	(0.014)
Observations	33,140	33,138	33,134	26,487
region x wave x time FE	Yes	Yes	Yes	Yes
qualification and age FE	No	Yes	Yes	No
economic activity status FE	No	Yes	Yes	No
income decile FE	No	No	Yes	No
employment sector FE	No	No	Yes	No
individual FE	No	No	No	Yes

Notes: This paper examines the effect of immigration on the individual-level support for the Leave campaign and voter turnout during the 2016 Brexit referendum. Support for the Leave campaign is measured using questions about opinions on leaving the EU. The outcome variable in the initial three columns is support for the Leave campaign and the last column outcome variable is Referendum turnout. Demographic variables include age, income decile, highest qualification, current employment status, and occupation. Standard errors, adjusted for clustering at the local authority level, are shown in parentheses.

instance, the observed notable rise in the salience of immigration issues leading up to the referendum (Figure 3.2) alone could account for the voter shift toward UKIP. Such an increase in the salience of a cultural issue can elicit a heterogeneous response among the electorate by amplifying the importance of cultural considerations and thereby motivating a segment of voters with socially conservative inclinations to prioritize the cultural issues in their identity and voting behavior. This channel is conceptualized in detail in Bonomi *et al.* (2021).

While the increase in immigration's salience serves as a plausible explanation for the shift, it is imperative to probe into other potential mechanisms that might also contribute to this trend. To this end, I differentiate between economic and cultural factors, while acknowledging their potential interplay. Economically, immigration is usually recognized for its overall positive contributions to the economy (Dustmann and Preston, 2019). However, beneath the surface of these aggregate benefits, specific concerns arise regarding job competition and wage pressures, particularly at the low end of the income distribution. On the cultural front, immigration introduces a broad spectrum of societal changes through the arrival of individuals from varied cultural, racial, religious, linguistic, and social backgrounds. This influx of diversity, while enriching in many respects, also poses challenges to societal cohesion and integration. Following the frameworks of Dustmann and Preston (2007) and Alesina and Tabellini (2024), this paper concentrates on three main areas: the labor market repercussions, the impact on the welfare system, and the hurdles to cultural integration.

The analysis reveals that neither labor market dynamics nor welfare system pressures fully account for this political shift. Instead, I show immigration shock shifts public attitudes toward immigration, characterized by an increase in anti-immigration sentiment and a heightened perception of immigration as a critical issue. Interestingly, immigration shock also seems to reduce the demand for redistribution and shifts voter values toward authoritarianism.

3.4.1 Labour Market Impact

A primary concern regarding immigration is its potential effect on the native labor market, particularly through job competition. This could lead to the displacement of native workers from the labor market (the extensive margin) or downward pressure on wages (the intensive margin). To assess the impact of immigration on the labor market, I first examine the extensive margin by analyzing its effects on economic activity and the unemployment rate. Local labor markets within the UK are interconnected due to internal migration, capital flows, and trade. Consequently, our estimates should be interpreted as indicators of relative regional improvement or decline. Using data from the Annual Population Survey spanning 2000 to 2016 at the local authority level, I estimate these impacts using equation 15. Results are reported in table 3.6. The findings in column 1 suggest a potential increase in the economic activity rate across both OLS and 2SLS estimations. Further analysis in columns 2 and 3 reveals that this increase is largely driven by men. Additionally, I observe a reduction in unemployment (column 4), which holds for both females and males, with a stronger effect for males (columns 5 and 6). Notably, when examining individuals aged 50 and older in column 7, a demographic that largely supported Brexit, I find no evidence of increased unemployment. The magnitudes of these effects are significant: immigration appears to decrease the average unemployment rate-initially observed at 5.51% across local authorities during the study period-by roughly 0.9%, as derived from 2SLS estimates in column 4. Additionally, immigration increases

Table 3.6. Effects of Immigration on the Employment

	Economic Activity			Unemployment Rate			
	(1) All	(2) Male	(3) Female	(4) All	(5) Male	(6) Female	(7) 50 and Older
Panel A. OLS							
Immigration Shock	0.363	0.324	0.175	-0.106	0.058	0.163	0.343
	(0.248)	(0.312)	(0.272)	(0.125)	(0.186)	(0.223)	(0.210)
Average effect	.443	.396	.213	-.12	.070	.199	.419
Panel B. 2SLS							
Immigration Shock	0.699	1.009	-0.005	-0.770	-0.691	-0.454	-0.185
	(0.387)	(0.479)	(0.400)	(0.223)	(0.307)	(0.290)	(0.325)
F-stat	219	216	215	204	241	274	212
Average effect	.854	1.23	-.00	-.94	-.84	-.55	-.22
Mean of DV	78.3	83.9	72.9	5.51	6.47	5.91	4.54
LA FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Spatial units	346	345	346	316	347	347	347
Observations	6592	6587	6582	5891	4888	4628	3272

Notes: This table presents the estimated impacts of immigration shocks on the economic activity rate and unemployment rate using Annual Population Survey data. Some data points are excluded due to the Office of National Statistics determining insufficient precision in the statistics. The term "F-stat" refers to the Kleibergen-Paap rk Wald F-statistic, which is used to test for weak instruments. The table presents robust standard errors, which are clustered by local authority, in parentheses.

the economic activity rate, which averaged 78% across local authorities, by approximately 0.8%, according to the 2SLS estimates presented in column 1.

To investigate the intensive margin, I analyze the impact of immigration on hourly wages using data from the Annual Surveys of Hours and Earnings for each local authority from 2000 to 2016. Results are reported in table 3.7. While based on column 1 the overall effect on wages appears negligible, there is some evidence that those at the lower end of the income distribution, particularly the 25th percentile reported in column 5, might experience a slight negative wage impact due to immigration. This finding broadly aligns with previous research by Dustmann *et al.* (2013) and Becker and Fetzer (2018), which indicates that immigration can depress wages at the lower end of the distribution while

Table 3.7. Effects of Immigration on the Wage Distribution

log(Hourly Pay):	(1) Avg	(2) 90th Pct	(3) 75th Pct	(4) Med	(5) 25th Pct	(7) 10th Pct
Panel A. OLS						
Immigration Shock	-0.006	0.008	-0.010	-0.007	-0.008	-0.003
	(0.003)	(0.009)	(0.003)	(0.003)	(0.003)	(0.002)
Average effect	-.62%	.847%	-1.0%	-.78%	-.89%	-.28%
Standard deviation	.710	.957	1.19	.886	1.00	.326
Panel B. 2SLS						
Immigration Shock	-0.000	0.017	-0.007	-0.007	-0.008	0.001
	(0.005)	(0.015)	(0.005)	(0.005)	(0.004)	(0.002)
F-stat	220	101	205	216	216	213
Average effect	-.03%	1.85%	-.74%	-.74%	-.89%	.152%
Standard deviation	.041	2.09	.844	.839	1.01	.172
Pre-log mean of DV	15.0	22.8	17.6	11.8	8.46	6.99
LA FE	Yes	Yes	Yes	Yes	Yes	Yes
Region-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Spatial units	348	327	344	348	347	346
Observations	7427	1615	7216	7428	7427	7411

Notes: This table presents the estimated impacts of immigration shocks on wage distribution, with the dependent variable being the log of hourly wages at the mean and also various percentiles within the earnings distribution of a local authority, as derived from the Annual Survey of Hours and Earnings. Some data points are excluded due to the Office of National Statistics determining insufficient precision in the statistics. The term "F-stat" refers to the Kleibergen-Paap rk Wald F-statistic, which is used to test for weak instruments. The table presents robust standard errors, which are clustered by local authority, in parentheses.

slightly increasing them at the upper end. However, the magnitude of this impact appears minimal, with the strongest effect observed at the 25th percentile of the wage distribution, where wages might decrease by an average of 0.9% due to immigration. Considering that immigration was shown to boost economic activity and reduce unemployment, it seems unlikely that this slight wage pressure can be the main driver behind increasing opposition to immigration. These findings are consistent when examining annual wages, as shown in Table A5.

A potential concern is that immigration can alter the demographic composition of a population, potentially leading to differing impacts across various groups. For example, the positive impact in the local labor market might accrue to the immigrant population themselves and might not extend equally to natives. Conversely, if immigrants largely complement native workers rather than directly competing with them, the positive effects of immigration on native unemployment could be greater, and the potential negative wage impacts could be significantly reduced. Unfortunately, the available data does not permit a precise distinction between the effects on native British individuals. However, the scale of immigration is likely not large enough to significantly alter the composition of unemployment or economic activity rates across different groups.

3.4.2 Pressure on the Welfare System

During the Brexit campaign, arguments that immigrants place undue strain on the welfare system were common. This sentiment is reflected in survey evidence, such as the 2014 European Social Survey, which indicated that 43% of British respondents believed immigrants take out more than they contribute to health, welfare, and taxation, compared to only 31% who believed the opposite. This perception stands in contrast to the findings of Dustmann and Frattini (2014), who demonstrated that immigrant groups arriving after 1999 have made positive fiscal contributions. Specifically, they calculated that recent immigrants from EU accession countries contributed nearly £5 billion between 2001 and 2011.

I explore how the immigration shock affected the number of claimants for major benefit types in each local authority. Using the log of the number of claimants as the dependent variable from the Work and Pensions Longitudinal Study (2000-2016), I estimate equation 15. Results, reported in Table 3.8, suggest that local authorities experiencing higher levels of immigration witnessed a decline in demand for most benefits. This aligns with the idea that immigration can stimulate economic growth, create jobs, and reduce long-term reliance on social benefits. Specifically, consistent with previous labor market findings, EU accession immigrants may alleviate labor shortages in critical sectors, boosting the local economy and generating employment opportunities. These immigrants often possess skills that complement the native workforce, increasing productivity and promoting economic growth, this will cause a lower dependence on welfare benefits. Additionally, as NMS immigrants tend to be younger with fewer dependents, their initial demand for social support services like income support and incapacity benefits is typically lower. An exception is observed with the Job Seeker's Allowance, where immigration appears to increase the number of claimants. This could be due to the initial employment hurdles immigrants face, such as language barriers, unrecognized qualifications, or a lack

Table 3.8. Effects of Immigration on the Welfare

log(Benefit Type):	(1) All	(2) Carers Allow.	(3) Disab. Living	(4) Incap. Benefit	(5) Income Support	(6) Job Seeker
Panel A. OLS						
Imm. Shock	0.002	-0.002	-0.024	-0.083	0.045	-0.008
	(0.009)	(0.013)	(0.013)	(0.024)	(0.011)	(0.022)
Average effect	.330%	-.23%	-3.2%	-11.%	6.11%	-1.0%
Standard deviation	.294	.206	2.93	10.1	5.45	.938
Panel B. 2SLS						
Imm. Shock	-0.033	-0.031	-0.034	-0.215	0.061	-0.078
	(0.028)	(0.017)	(0.017)	(0.044)	(0.018)	(0.033)
F-stat	38.4	221	61.2	38.4	221	56
Average effect	-4.5%	-4.1%	-4.6%	-29.%	8.34%	-10.%
Standard deviation	4.04	3.74	4.17	26.1	7.45	9.51
Pre-log mean of DV	1389	819.	939.	2600	1911	2467
LA FE	Yes	Yes	Yes	Yes	Yes	Yes
Region-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Spatial units	348	348	348	348	348	348
Observations	5916	5905	5213	5916	5901	5914

Notes: This table presents the estimated impacts of immigration shocks on various types of welfare benefits using data from the Work and Pensions Longitudinal Study (WPLS). The analysis examines the log of the annual count of benefit claims, as recorded by the Office for National Statistics (ONS). Some data points are excluded due to the Office of National Statistics determining insufficient precision in the statistics. The term "F-stat" refers to the Kleibergen-Paap rk Wald F-statistic, which is used to test for weak instruments. The table presents robust standard errors, which are clustered by local authority, in parentheses.

of local work experience, leading to a temporary higher dependency on the Job Seeker's Allowance. However, this effect is expected to be short-lived.

The decline in demand for welfare benefits suggests two potential dynamics. First, it indicates limited migration directly into the welfare system. Second, it implies that immigration may stimulate the local labor market, drawing the native population into employment and reducing reliance on benefits overall. While the data limitations prevent me from disentangling the precise effects on UK-born versus foreign-born workers, the negative effect on the net number of claimants makes it unlikely that there's a substantial increase in absolute claimant numbers from migrant populations. This suggests that immigration is not likely to be placing undue pressure on the benefits system.

If economic factors fall short in accounting for the political shift observed, what then is the driving force behind this shift? The following section of this paper posits that a transformation in voters' cultural attitudes provides a cogent explanation. However, alternative explanations exist, such as the heightened salience of immigration as an issue and the consequent shift in priorities among the electorate. Under this scenario, when immigration becomes more visible or is perceived as impacting local economies or social structures, political parties and candidates that emphasize immigration issues may gain traction not because individuals inherently change their ideologies, but because they prioritize the immediate challenges. This strategic voting can temporarily align voters with parties or candidates that promise to address these concerns, reflecting strategic voting based on current priorities rather than a deep-seated change in social attitudes or political identities. Nonetheless, subsequent analysis will demonstrate that support for UKIP signifies a fundamental change in attitudes, potentially heralding more durable consequences.

3.4.3 *Cultural Concerns*

Immigration may influence the voting choices of native populations through more than just economic factors, a perspective acknowledged across economics, political science, and sociology. Natives frequently worry, fueled by the anti-immigrant rhetoric of politicians, that immigrants from significantly different backgrounds fail to assimilate into new cultural norms, potentially challenging the societal fabric and integrity. Economic anxieties may interact with cultural fears, amplifying negative native perceptions. Some research has attempted to illuminate the role of cultural factors in anti-immigration sentiment by examining the cultural distance between immigrants and the host society. However, in my case, since all immigrants originate from the same origin, differing minimally in their distance to the culture of the host country, I cannot use cultural differences among immigrants to assess the impact of cultural factors. Instead, I can look at how the cultural attitude of voters will evolve in response to immigration. However, we should keep in mind that these cultural attitudes can themselves be influenced by economic factors.

I explore whether shifts in individual-level voting patterns may reflect changes in attitudes and social preferences toward immigration. Further exploration in Chapter 3.6 considers whether these shifts signify a broader transition in identity emphasis from class to culture. Figure A4 represents the evolution of concerns among Conservative and Labour party supporters toward immigration and the economy between 2001 and 2015.

In the earlier period, the economy overwhelmingly preoccupied supporters of both parties, while immigration concerns were relatively marginal. By 2015, a pronounced pivot is observed: immigration concern has markedly increased and has replaced the economy as the point of contention with the disparity between the parties' supporters regarding it reaching a significant 18%. The graph displays a clear shift in the political landscape with immigration becoming a prominent issue, especially for Conservative supporters, indicating a significant realignment of priorities over the 14-year span.

To see whether this increase in anxiety about immigration is caused by immigration, I use data from Wave 8 of the British Election Study (BES), the wave leading to the referendum, scrutinizing individual perceptions and attitudes toward immigration. Specifically, this research utilizes four variables: the belief in immigration's benefits to Britain's economy (Econ) and cultural life (Cultural), the perception of immigration trends (Change), and the stance on immigration policy (Policy). Higher values on Change indicate a stronger perception of increasing immigration, while higher values for the other three variables suggest more favorable views on immigration. Now, I estimate the following specification:

$$y_j = \alpha + X_j + \beta \Delta IM_{i(j),2016} + \epsilon_i \tag{22}$$

where y_j is one of the four aforementioned metrics reflecting immigration attitudes and perceptions of individual j in local authority i. The immigration shock, $\Delta IM_{i(j),2016}$, represents the shock in local authority i that individual j lives in 2016 and is instrumented using the variable $IM^O_{i(j),2016}$. All regressions have a rich set of individual demographics as a control.

Table 3.9 presents findings. The table displays two sets of estimates: Ordinary Least Squares (OLS) and Two-Stage Least Squares (2SLS), both considering an 'immigration shock' variable, which reflects a measure of local-level immigration exposure. The OLS estimates show a negative association between immigration shock and all four measures of cultural attitudes toward immigration, suggesting that areas experiencing higher immigration shock are associated with more negative views on these aspects. The 2SLS estimates, which account for potential endogeneity, reinforce these findings with slightly larger magnitudes of the coefficients. While these patterns might simply show immigration is shifting attitudes, the question is whether voters are prioritizing these new attitudes in their voting decisions and shaping their identities based on these cultural and social preferences. Further analysis in subsequent sections suggests these results align more with a shift in political cleavages, from class-based to cultural distinctions, prompting voters to decide which party to support based on their immigration preference.

Table 3.9. Public Attitudes

	Immigration Preference				RedistPref	AuthScale
	Econ	Cultural	Change	Policy		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. OLS						
Immigration Shock	-0.114	-0.142	0.036	-0.167	-0.033	0.171
	(0.030)	(0.034)	(0.013)	(0.054)	(0.035)	(0.047)
Panel B. 2SLS						
Immigration Shock	-0.120	-0.156	0.045	-0.181	-0.126	0.179
	(0.039)	(0.044)	(0.018)	(0.068)	(0.042)	(0.060)
Observations	17,284	17,443	17,572	16,996	16,817	16,541
Demographics	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table presents regression results using data from Wave 8 of the British Election Study, specifically examining the public's stance on immigration, redistribution, and cultural issues. Column (1) 'Econ' reflects responses to the survey question assessing the perceived economic impact of immigration. Column (2) 'Cultural' is based on the question evaluating immigration's influence on cultural life. The survey question regarding perceptions of whether immigration levels are rising or falling informs Column (3) 'Change'. Column (4) 'Policy' relates to views on the policy of allowing families of residents into Britain. The 'RedistPref' variable in Column (5) is scored on a 0-10 scale, formulated by combining and standardizing five variables to gauge attitudes toward redistribution, where 10 indicates the highest preference for redistribution. Likewise, the 'AuthScale' variable in the final column is based on a 0-10 scale, aggregating and normalizing five variables that explore individuals' liberal versus authoritarian values, with 0 representing libertarian views and 10 indicating authoritarian tendencies. The independent variable is the immigration shock experienced in 2016 at the local authority, controlling for individual demographics such as household income, age, educational attainment, and job zone, while incorporating fixed effects for various governmental regions. Standard errors are clustered at the local authority level.

Columns 5 and 6 of Table 3.9 reveal that immigration not only shifts the attitudes of voters regarding immigration but also influences broader economic and social attitudes, leading to a decreased demand for redistribution and an increase in authoritarian sentiments among voters. Specifically, individuals in areas with higher exposure to immigration are found to be more receptive to reductions in domestic public spending and position themselves more authoritatively on the authoritarian-liberal spectrum.

The observed shift in redistribution preferences, triggered by immigration-a factor ostensibly disconnected from fiscal redistribution-may initially appear counterintuitive. Nonetheless, the literature offers two compelling interpretations. First, Alesina *et al.* (2023) found that prompting individuals to think about immigrants can significantly diminish support for redistributive policies, a pattern that is particularly pronounced among less educated and right-wing respondents. The authors suggest it is rooted in a reluctance to redistribute wealth toward individuals perceived as outsiders or foreigners. Second, Bonomi *et al.* (2021) posits that significant immigration influxes can pivot societal identity from class-based to culture-based distinctions. As cultural aspects become more dominant, they play a greater role in shaping policy preferences. The emphasis on cultural identity blurs class distinctions and thereby dampens redistributive conflict.

The findings of this section align with the narrative proposed by Bonomi *et al.* (2021), illustrating how immigration acts as a catalyst for the transformation of societal identity from class-based to culture-based. This transition can be driven by two mechanisms. First, immigration shock increases the salience of immigration issues, serving as a stand-in for wider cultural issues (as illustrated in Figure 3.2). Second, individuals negatively impacted by immigration could be predominantly conservative, potentially as a result of their lower education. This shift from class to cultural identity leads to voting patterns that reflect cultural preferences, explaining the rise in the support of the UKIP.

This pivot toward cultural identity causes voters to move their beliefs in the direction of stereotypes, increasing polarisation and conflict about issues like immigration. Conversely, individual beliefs about redistribution become less polarised. This phenomenon can explain why voters exposed to immigration become anti-immigrant and demand less redistribution. If this transformation toward cultural identity is indeed occurring, it anticipates a corresponding shift in the political arena's supply side. The next section will therefore explore the adjustments made by political parties in response to immigration dynamics.

The compilation of evidence reviewed thus far underscores the significant role of cultural factors, as opposed to economic ones, in shaping the preferences of native populations and in the political realignment in reaction to immigration. This aligns with Tabellini (2020), which demonstrates a notable positive influence of immigration on the employment rates and occupational earnings of native individuals. Nevertheless, we cannot entirely overlook the role of economic factors. First, immigration appears to precipitate modest economic drawbacks, predominantly affecting the lower echelons of the native workforce in the short term. Second, economic insecurities can be voiced through cultural concerns, often exacerbated by political figures and media outlets, which can lead to natives harboring skewed perceptions about immigrants and their impact.

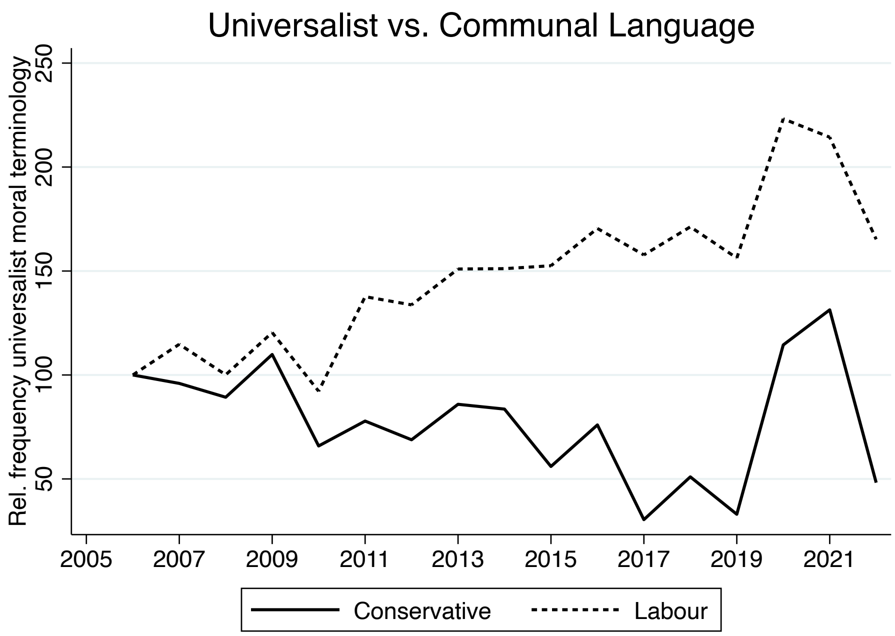
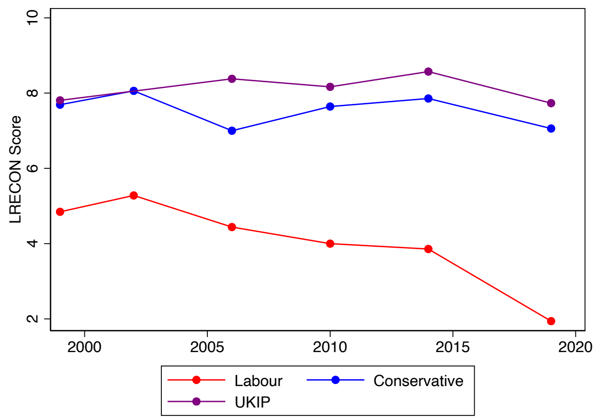


Figure 3.6. *Notes:* This graph illustrates the trend in the relative frequency of universalist versus communal moral rhetoric in speeches within the UK Parliament from 2006 to 2022. The solid line represents the relative frequency of universalist rhetoric in combined speeches delivered by Conservative MPs. In contrast, the dashed line indicates the relative frequency of universalist language in speeches by Labour MPs. The methodology for this computation is adapted from Enke (2020). For clarity and comparison, the frequencies for each party are normalized, setting the value to 100 in the initial year of the plot (2006).

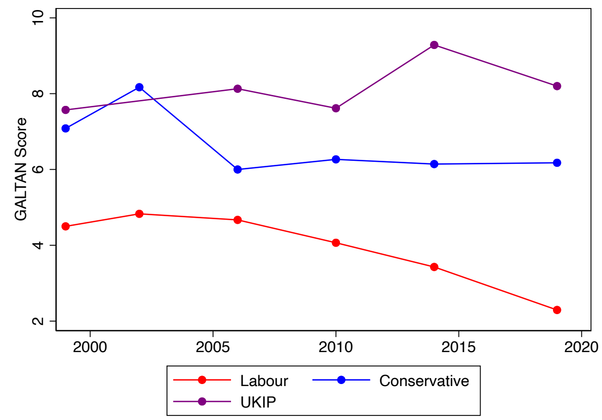
3.5 Party Responses to Immigration

In this section, I investigate if an analogous development has taken place on the political supply side. That is, whether the UK parties have shifted their focus in political activities toward prioritizing cultural issues over economic ones in response to immigration. Concurrent with the rise in NMS immigration, Figure 3.6 illustrates a cultural polarization in political rhetoric in the UK, as captured by the metric developed in Enke (2020). The figure indicates that parliamentary speeches have become less universalistic over recent years for Conservatives, with a notable increase in universalism for Labour.

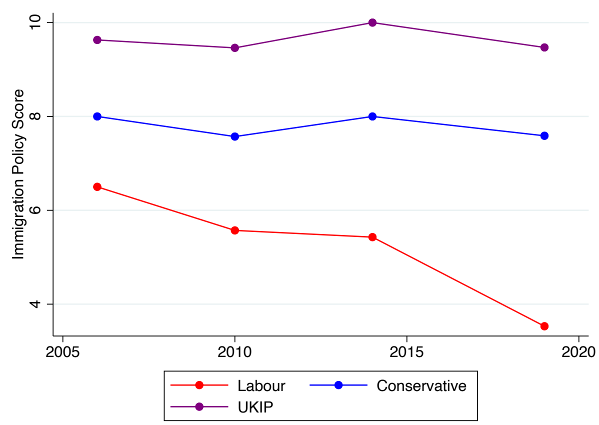
The measure used in Figure 3.6 measures universalism relative to the communal moral values of MPs. To get a more comprehensive view and measure the ideological positions of political parties on other margins, I use data from the CHES. In Figure 3.7 Panel A, I present the evolution of positions on economic issues for the parties UKIP, Labour, and Conservative over the same time window. It appears that Labour has become increasingly left-wing over time, while the other two parties do not show any clear trend. Interestingly, while UKIP and the Conservatives both align to the right of Labour on economic issues,



(a) Party's position on economic issues



(b) Party's position on social and cultural values.



(c) Party's position on immigration policy.

Figure 3.7. *Notes:* Parties' Scores over Time. Panel A measures the party's position on social and cultural values such as personal freedoms, abortion rights, same-sex marriage, tradition, and stability on a scale of zero to ten, with a higher score indicating a more traditional/authoritarian stance. Panel B measures the party's position on economic issues including privatization, taxes, regulation, government spending, and the welfare state on a scale of zero to ten, with a higher score indicating a belief in a reduced role for government. Panel C measures the party's position on immigration on a scale of zero to ten, with a higher score representing a more restrictive policy on immigration. .

there seems to be no substantial distinction between the stances of these two parties on this margin.

In Panel B, I present the trend for parties' positions on social and cultural values. As with the economic positions, it is only Labour that becomes increasingly more progressive over time. Panel C focuses specifically on the parties' positions on immigration. As expected, UKIP is almost as anti-immigrant as possible, while the Conservative party is positioned between Labour and UKIP. Like the other two panels, only Labour exhibits a change in its position over time, moving toward a more pro-migrant stance. In sum, similar to Enke (2020) measure reported in Figure 3.6, CHES scores along different dimensions also exhibit a divergent trend between main political parties in the UK.

While the timing of the rise in immigration and increase in the political salience of immigration (reported in Figure 3.2) suggest that this polarisation might have happened due to an immigration shock, a direct causal relationship has not yet been established. My next step is to explore how MPs may adapt their local political positions and rhetorics in response to immigration shocks within their constituencies.

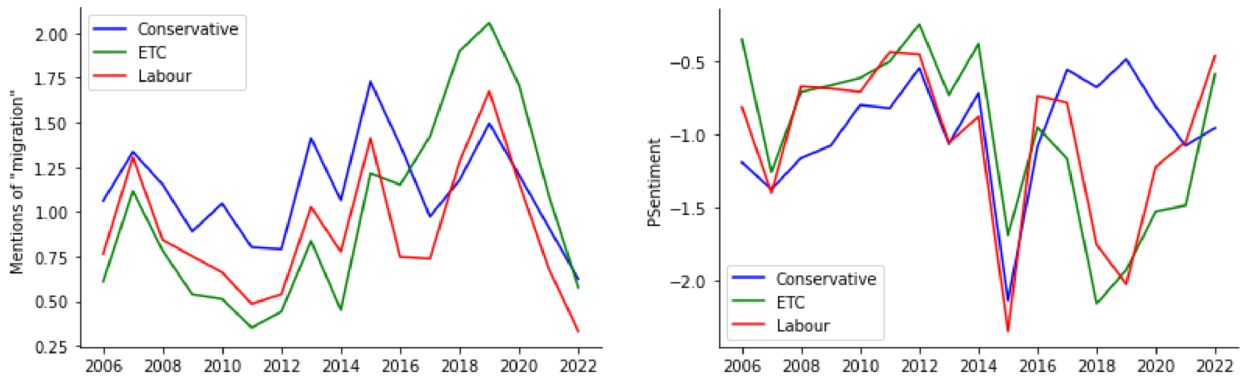
To investigate this possibility, I analyze the relationship between the exposure of a constituency to immigration and the engagement with immigration topics in Parliament by the MP of that region. I apply natural language processing techniques to Parliamentary speeches to construct three indicators for each constituency and year that illuminate various aspects of the political discourse surrounding immigration.

Frequency Measure: This metric measures the density of selected keywords indicative of discussion around migration and minority issues[19] within an MP's parliamentary discourse over a specified year. It is calculated by tokenizing speeches to extract words, filtering out non-alphabetic characters to focus solely on textual content, and then counting occurrences of relevant keywords. The aggregate frequency of these keywords is then normalized by the total word count of the MP's annual contributions, yielding a relative frequency measure. This metric, termed $MigrationTalk_{i,t}$, quantifies the extent to which MPs engage with the designated topics within their parliamentary language, offering an objective metric for thematic emphasis.

Sentiment Measure: The sentiment score captures the emotional resonance and evaluative tone of parliamentary discussions on immigration by identifying the presence of relevant keywords within MPs' tokenized contributions. For each keyword, a snippet - spanning 10 words before and 10 words after each keyword- is extracted to capture the surrounding sentiment. Leveraging the `NLTK` library's sentiment analysis tools, which assign sentiment values to words, a compound sentiment score is calculated for each contribution, ranging from -1 (highly negative) to +1 (highly positive)[20]. This process aggregates scores across an MP's yearly contributions, normalizing by the number of speeches

[19]Keywords include terms such as 'migra*', 'asylum', 'minorit*', 'traveller', 'ethnic*', 'racial*', and 'gypsy'.

[20]I utilize the SentimentIntensityAnalyzer from the VADER tool in the Natural Language Toolkit (NLTK) package, which leverages a sentiment-annotated lexicon to assess word polarity (positive, negative, neutral) and emotional intensity in various contexts. VADER's analysis, informed by grammatical and syntactical rules, effectively interprets modifiers like intensifiers, diminishers, and negations, impacting sentiment scores. The analyzer outputs four metrics: 'neg' (negative), 'neu' (neutral), 'pos' (positive), and 'compound'-an overall sentiment score. I focus on the 'compound' score for a concise summary of textual sentiment orientation.



(a) $MigrationTalk_{i,t}$

(b) $MigrationSentiment_{i,t}$

Figure 3.8. *Notes:* This plot shows the average of $MigrationTalk_{i,t}$ and $MigrationSentiment_{i,t}$ over time by party. $MigrationTalk_{i,t}$ for firm i at time t is normalized using the average $MigrationTalk_{i,t}$ in the sample; $MigrationSentiment_{i,t}$ for firm i at time t is normalized using the average $MigrationSentiment_{i,t}$ in the sample.

mentioning the keywords. The resulting metric, termed $MigrationSentiment_{i,t}$, reflecting an average sentiment score per relevant speech, quantitatively assesses the emotional and evaluative tone MPs adopt in their discourse on immigration.

Figure 3.8 illustrates the temporal trends of these measures across various parties. Panel A shows the frequency of mentions of "migration", while panel B shows the sentiment toward migration as measured by $MigrationSentiment_{i,t}$. As expected, the number of mentions of "migration" has increased until 2019, but there does not appear to be a significant difference between different parties. More notable is the trend shown in panel B, which reveals that MPs had the most negative tone toward migrants right before the referendum. Interestingly, there is no significant difference among parties in terms of their sentiment toward migrants at this time.

While these two measures have the advantage of looking at immigration directly, they don't necessarily capture the potential larger shift in party rhetoric along the cultural dimension. The following metric aims to capture this broader potential shift.

Universalism Measure: To this end, I again use Enke (2020) framework, which uses a simple word count that is based on keywords found in the Moral Foundations Dictionary (MFD) on the US Congressional Record. The used dictionary categorizes words into four dimensions: harm/care, fairness/reciprocity, in-group/loyalty, and authority/respect, totaling 215 words or word stems. The index of relative universalism, proxied by the relative frequency of universal terminology, is calculated as follows:

$$Universalism_{i,t} = \frac{Care_{it} + Fairness_{it} - Ingroup_{it} - Authority_{it}}{N_{it}} \quad (23)$$

Here, each term in the numerator represents the total count of words belonging to each category and the denominator, and N_{it} is the total number of non-stop words. According to this framework, individuals with a universalistic outlook tend to apply their value system broadly, often championing progressive civil rights and immigration policies. Thus, a decline in universalism might reflect a trend among right-wing politicians toward more culturally conservative rhetoric or a diminished propensity among left-wing politicians for progressive advocacy. Using this measure, Figure 3.6 shows a polarising at the national level between major parties. However, for a more granular analysis, I construct this measure for each MP and year to examine whether regions with higher exposure to immigration exhibit a shift toward more conservative or communal rhetoric, especially by right-wing parties.

Although ideal data would encompass the local stances of all parties across all constituencies, using parliamentary speeches provides a proxy for the sentiment at the constituency level only for the party currently holding the seat. This approach is particularly constrained in contexts like the UK, where the First-Past-The-Post (FPTP) electoral system is used. FPTP's winner-takes-all nature and its encouragement of strategic voting tend to amplify the voices of major parties. This system can result in a representation gap, leaving the viewpoints of some segments of the electorate, especially those backing smaller parties, underrepresented in Parliament. Therefore, it's crucial to interpret the forthcoming analysis as indicative of the impact of immigration on the rhetoric and positioning of the incumbent MPs, rather than a comprehensive reflection of the entire political landscape within constituencies.

To investigate the potential for the supply side of politics to respond to the level of immigration exposure at the location level, I estimate the following specifications:

$$y_{i,t} = \alpha_i + \eta_{r,t} + \beta \Delta IM_{it} + \epsilon_{i,r,t} \tag{24}$$

where $y_{i,t}$ represents either $MigrationTalk_{i,t}$, $MigrantSentiment_{i,t}$, or $Universalism_{i,t}$ for constituency i in year t. The term $\eta_{r,t}$ controls for region-year shocks.

The results are detailed in Table 3.10. Concentrating on the 2SLS estimates, the first column suggests a positive effect of immigration on the frequency of discussions about immigration by the region's MP, though this does not achieve statistical significance. This analysis was further refined in columns 2 to 4 by splitting the sample based on the party affiliation of the MPs throughout the observation period. Notably, the effect is more marked among Conservative MPs, as evidenced in column 3, the only column with a significant coefficient. The examination of $MigrantSentiment$ in columns 5 to 8 indicates that sentiment coefficients for Conservative MPs are negative and significant. In contrast, Labour MPs and other MPs show a positive coefficient, but these do not attain statistical significance. Together, these findings indicate that in areas with increased exposure to immigration, Conservative MPs are more likely to discuss immigration frequently and adopt a negative stance in their discussions. On the other hand, the data does not reveal a comparable pattern among Labour MPs, hinting at either a reluctance or an inability to engage with immigration issues.

Table 3.10's last three columns offer tentative evidence suggesting a divergence in responses to immigration exposure based on party lines. Labour MPs in constituencies with higher levels of immigration exposure exhibit a slight shift toward universalistic rhetoric. On the other hand, Conservative MPs have not markedly altered their rhetoric while MPs from other parties have shown a tendency to adopt a less universalistic stance. The apparent responsiveness of smaller parties' MPs suggests that these are the most agile ones to go beyond party lines and capitalize on these shocks.

It is important to note that the results in this table may reflect changes in rhetoric within individual MPs over time or shifts in the composition of MPs. That is, immigration shocks may alter the electoral landscape, making it more likely for certain candidates, who are perhaps more responsive or attuned to immigration issues, to be elected. Second, incumbent MPs may adjust their rhetoric to align more closely with the prevailing sentiments on immigration within their constituencies.

This section, by focusing on political party responses, complements the insights from the prior section, offering a more nuanced understanding of immigration's multifaceted impact. The previous section showed that immigration affects public attitudes and preferences toward immigration and subsequently influences voting patterns in alignment with parties' stances on immigration. This section highlights that political entities recalibrate their messaging and rhetoric in response to immigration shocks. This dual interaction-public sentiment evolving in response to immigration and political entities adjusting accordingly-suggests a transformative shift in political cleavage, moving away from traditional dichotomies toward a new axis centered on cultural dimensions, notably immigration. The ensuing chapter is devoted to a direct empirical investigation of this hypothesis, aiming to validate the proposed paradigm shift in the political landscape.

Table 3.10. Effects of Immigration on Parliamentary Speeches

	$MigrationTalk_{i,t}$				$MigrantSentiment_{i,t}$				$RelativeUniversalism_{i,t}$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(5)	(6)	(7)	(8)
Panel A. OLS												
Immigration Shock	0.13	0.12	0.31	0.01	-0.29	-0.15	-0.57	-0.21	0.10	0.08	0.18	-0.55
	(0.09)	(0.15)	(0.14)	(0.17)	(0.10)	(0.14)	(0.19)	(0.26)	(0.08)	(0.14)	(0.12)	(0.23)
Panel B. 2SLS												
Immigration Shock	0.26	-0.17	0.89	0.53	-0.11	0.10	-0.85	0.86	0.01	0.21	0.07	-2.07
	(0.21)	(0.29)	(0.42)	(0.42)	(0.25)	(0.45)	(0.39)	(0.92)	(0.18)	(0.27)	(0.26)	(0.84)
Constituencies	All	Labour	Con.	Others	All	Labour	Con.	Others	All	Labour	Con.	Others
Observations	6171	2479	2709	979	4249	1566	1947	704	6171	2479	2709	979

Notes: This table presents the impact of immigration on the respective sentiment measures. All dependent variables have been standardized. The regression includes controls for the share of the 2001 resident population lacking formal qualifications, the share of employees in routine occupations, and the shares of the working-age resident population employed in the manufacturing and retail sectors, interacted by year. Standard errors, adjusted for clustering at the constituency level, are provided in parentheses.

3.6 Cultural Realignment

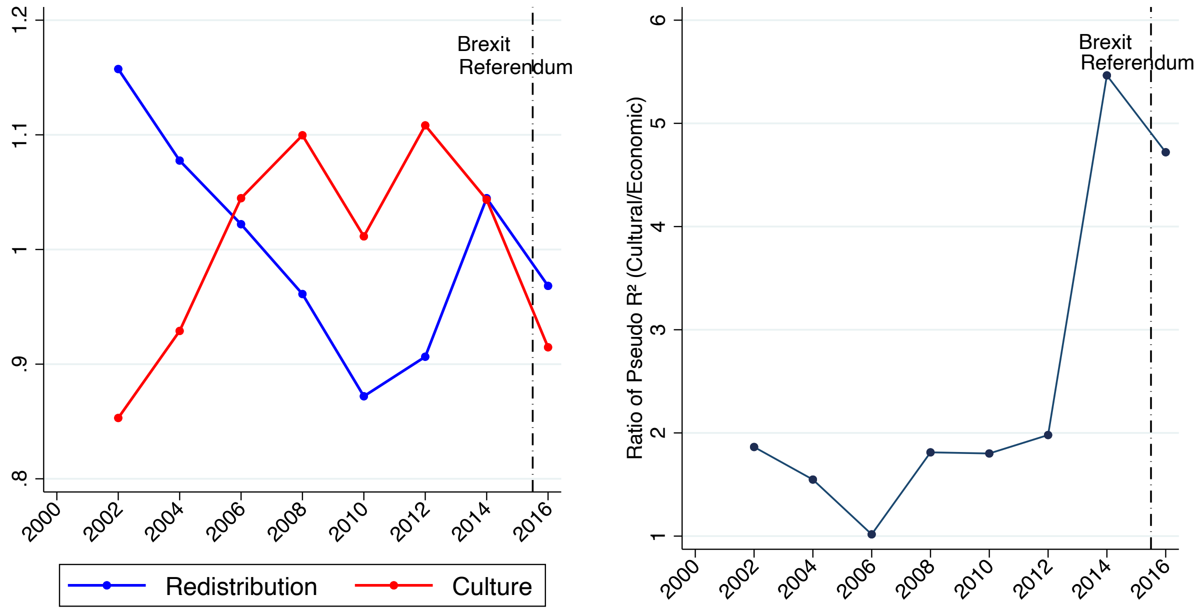
This section explores the dynamics driving the patterns of voter behavior and political responses observed, particularly the rise in support for right-wing parties amid anti-immigration sentiments. Section 3.4 demonstrated that immigration does not have a significant negative economic impact, suggesting that the underlying causes of this phenomenon are more cultural than economic. However, this raises the question of why those potentially economically disadvantaged who do not directly benefit from the right-wing agenda of minimal redistribution and reduced social welfare would support such parties in response to immigration. The evidence presented in this chapter suggests this puzzle can be explained by a pivotal shift: cultural alignment emerges as the primary cleavage, eclipsing traditional economic considerations in the voting calculus. This observation aligns with the theoretical concept of identity realignment as discussed in Bonomi *et al.* (2021), highlighting that economic incentives no longer encapsulate the main factors influencing political preferences. When electoral priorities change and cultural concerns predominate, the capacity and willingness of left-wing parties to adopt anti-immigration stances may find inherent limitations.

As already shown in Figure 3.2, concurrently with the rise of immigration in the UK there has been growing salience of immigration in public discourse, media, and politics. This chapter seeks to empirically validate the hypothesis that not only salience of immigration has increased but it has also led to a shift in how voters prioritize their political preferences. Specifically, it suggests that the visibility and frequent discussion of immigration may transform it into a critical point of political division as cultural considerations become more immediate and emotionally resonant. As a result, voters may begin to weigh cultural issues more heavily than economic policies, which could appear more abstract or distant. This dynamic suggests that immigration becomes a lens through which voters evaluate political parties and candidates, favoring those who reflect their cultural values.

First, I examine whether the heightened salience of immigration coincides with it evolving into a more contentious political cleavage issue. The shift in voters' disagreement over redistribution and culture and how these factors influence voting decisions is illuminated in Figure 3.9, which leverages data from the biennial European Social Survey (ESS) for the UK. Here, indices capturing the public's demand for redistribution and progressive cultural policies are constructed. The former is derived as the principal component from three questions related to public spending. Similarly, an index representing the demand for progressive cultural policies is formulated from opinions on immigration. I adjust both indices by estimating their residuals, conditioned on respondents' party affiliations and interacting with wave fixed effects to account for the dynamic nature of political party stances. Panel A of the figure delineates the variance of these indices from 2002 to 2016, where the last point refers to post-Brexit referendum data. The data presents a striking trend: while disagreements on redistribution show a general decline, the contention surrounding cultural policies intensifies notably during this period. This shift is not isolated to the UK context but resonates with similar trends observed in the US, as documented by Bonomi *et al.* (2021). Panel B shows the predictive power of redistributive and cultural attitudes in explaining voting behavior and further underscores this realignment, revealing the growing predominance of cultural issues in shaping voting patterns, a trend particularly pronounced around the Brexit referendum era. This evolving political landscape suggests a reshaping of the axes of political conflict, heralding a new era where cultural considerations increasingly dictate the electoral dynamics.

Building on the observation of increased cultural divisiveness and its growing role in voting dynamics, I explore voter realignment through cluster analysis, following the methodology outlined by Bonomi *et al.* (2021). Cluster analysis is a powerful approach for discerning shifts in voter alignment, particularly between cultural conflicts and economic dimensions. Utilizing the K-means algorithm, voters are classified into two distinct clusters within a bidimensional policy space that encompasses demands for progressive cultural policies and redistribution. As illustrated in Figure 3.10, a notable shift is observed in the 2015-14 period compared to 2002-2003. The primary distinctions between clusters have evolved, now more prominently based on cultural progressiveness versus conservatism, rather than pro- or anti-redistribution stances. This evidence supports the idea of voter realignment, indicating a transition in the political landscape where cultural issues, such as immigration, race, and national identity, increasingly influence political behavior, overshadowing traditional economic concerns like government spending and employment policies.

So far in this section, I have illustrated the shifts in political cleavages and a movement toward cultural clustering. The synchronicity of these shifts with the timing of immigration shocks suggests a causal relationship between immigration and political cleavages, which in turn causes the voting patterns discussed in earlier sections. To directly examine the existence of such a causal link, I utilize the cross-sectional variation in voter clustering. For this purpose, I turn to the British Election Study Internet Panel, initiated around the referendum period, which provides a broader sample size and finer geographical details for each respondent than the European Social Survey (ESS). Applying K-means clustering to individual local authorities enables an examination of whether immigration directly causes voter clustering around cultural issues. This is accomplished through Cluster Centroid Analysis, explained below.

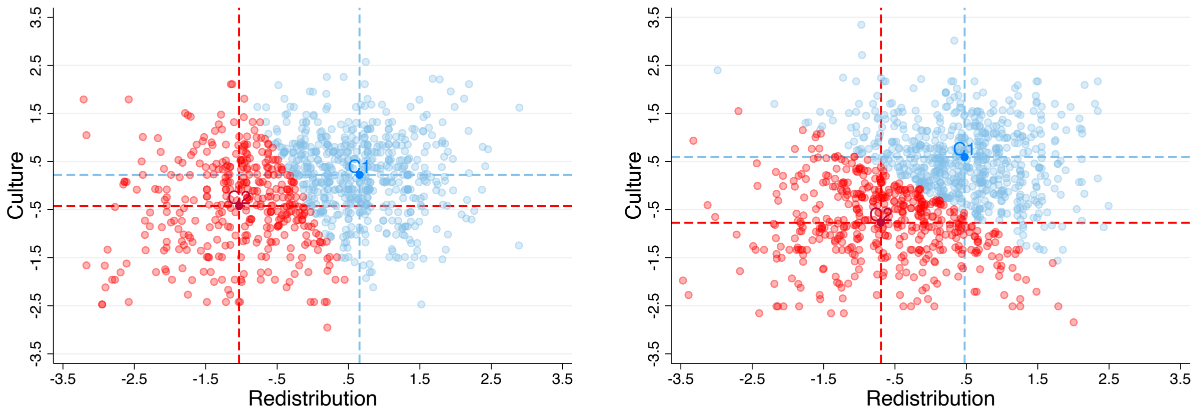


(a) Variance in Attitudes

(b) Culture/Redistribution in Voting Predictions

Figure 3.9. *Notes:* Panel A displays the variances of 'Redistribution' and 'Culture' concepts derived respectively from three questions about redistribution preferences/public spending and three questions concerning immigration, with responses standardized. The first principal component for each concept, calculated using polychoric principal component analysis, reflects higher values for more liberal views. Both 'Redistribution' and 'Culture' are then residualized based on party identity, factoring in interactions with wave-fixed effects from the European Social Survey (ESS). Across every survey wave, residuals have been standardized to achieve a mean of zero and a variance of one. Panel B illustrates the ratio of pseudo R-squared values. These values are obtained from separate multinomial logistic regressions, where party affiliation is regressed on 'Culture' and 'Redistribution' for each round of the ESS. This approach allows for an assessment of the relative explanatory power of cultural versus economic factors in predicting political party alignment across different periods covered in the ESS data.

Upon completing the K-means clustering in each local authority, I conduct a detailed examination of the centroids of the resulting clusters. A marked difference in centroids along the cultural dimension, coupled with minimal variance along the economic dimension, would suggest a primary influence of cultural factors. In contrast, if significant disparities are observed along the economic dimension, it would imply that economic factors are more influential. To quantify this distinction, I calculate the ratio of the differences in centroids along each axis. Specifically, I compute the following Culture-Redistribution Centroid Ratio (CRCR) measure for each local authority:



(a) 2002-3 **(b)** 2014-2015

Figure 3.10. *Notes:* This table illustrates UK respondents' attitudes toward cultural policies and redistribution for the years 2002–2003 (panel a) and 2014–2015 (panel b). The vertical axis represents cultural policy attitudes (with higher values indicating more open attitudes), while the horizontal axis reflects attitudes on redistribution (with higher values signifying a stronger preference for redistribution). These measures were derived by first extracting the principal polychoric component from two sets of questions, each addressing one of these political conflict dimensions. The principal component for cultural issues, labeled 'Culture', is based on questions regarding preferred immigration levels, abortion policy, and racial attitudes. The principal component for redistribution preferences, labeled 'Redistribution', is derived from questions about desired government spending levels and the government's role in ensuring citizens' employment and living standards. The residuals were then estimated after adjusting for respondents' party identity. Each marker in the graph represents an individual respondent. The color coding differentiates respondents according to the two clusters identified using the K-means method, applied to the aforementioned residuals for both periods separately with initial group means estimated using Ward's method. C1 and C2 mark the centroids of each cluster. Data Source: European Social Survey (ESS).

$$CRCR_i = \frac{C1_{i,culture} - C2_{i,culture}}{C1_{i,redist} - C2_{i,redist}} \tag{25}$$

In this formula, $C1$ represents the centroid of the cluster characterized by a stronger pro-redistribution stance. The subscript i refers to the specific local authority under analysis. This CRCR measure is then regressed against the immigration shock, with findings detailed in Table 3.11. While the results are somehow noisy in both OLS and 2SLS estimations, which is not surprising given the relatively small sample size in each local authority and resulting attenuation bias, they predominantly indicate that the immigration shock has led to a more pronounced realignment of voters along cultural lines, rather than redistribution lines. This trend persists even after adjusting for demographic variables and other industry shocks, suggesting a robust realignment of voter preferences along cultural lines in response to immigration.

This chapter's exploration sheds light on the nuanced influence of immigration on political cleavages and voter alignment, suggesting a gradual shift toward cultural considerations. The evidence points toward an emerging landscape where cultural factors outweigh economic factors in shaping voter decisions. In this evolving political context, not paying attention to the shift toward cultural issues in politics can lead to a range of adverse outcomes, from misreading the political landscape to exacerbating social divisions. Recognizing this shift is crucial for correctly interpreting electoral outcomes and the motivations behind voter behavior.

3.7 Conclusion

The increasing prevalence and political divisiveness of immigration in many Western countries coincide with a pivotal shift in political dynamics in these countries. Twentieth-century politics was largely shaped by economic divides, with the left advocating for workers and social welfare, and the right championing smaller government and the private sector. Contemporary politics, in contrast, pivots more on identity and cultural issues, with the left supporting various marginalized groups and the right focusing on protecting traditional national identity, often linked to race, ethnicity, or religion. This temporal juxtaposition raises a question: to what extent is immigration contributing to or influencing this profound political evolution?

To study this question, I began with an examination of how local exposure to immigration influences voting decisions, revealing a significant shift toward anti-immigrant right-wing parties. Employing a novel research design, this study tapped into previously unexplored variations in immigration exposure, utilizing migrant flows across industries and employment structures across regions. I instrument my measure using the industry-specific flow of migrants to other immigration destinations akin to the UK, i.e., pre-2004 EU countries. This approach uncovered immigration shock triggers a notable shift in political support, with individuals transitioning from the traditional left-leaning Labour Party toward the right-wing, anti-immigrant UK Independence Party (UKIP). Furthermore, this investigation extends to the domain of political rhetoric, highlighting an inclination among MPs from constituencies hit hard by immigration to discuss immigration issues negatively in their parliamentary speeches or to embrace a more localized discourse. Notably, such responses are markedly missing from Labour MPs, highlighting the complex, party-specific nature of reactions to the dynamics of immigration.

Table 3.11. Immigration Impact on Cultural and Redistribution Divides

	Culture-Redistribution Centroid Ratio			
	(1)	(2)	(3)	(4)
Panel A. OLS				
Immigration Shock	0.551	0.461	0.828	0.944
	(0.548)	(0.574)	(0.994)	(1.080)
Panel B. 2SLS				
Immigration Shock	0.579	0.493	0.940	1.020
	(0.370)	(0.397)	(1.218)	(1.288)
R-Squared	.00493	.00279	.0201	.0268
Observations	314	314	314	312
Region Fixed Effects	No	Yes	Yes	Yes
Demographics	No	No	Yes	Yes
Initial composition of immigrants	No	No	No	Yes
Routine Jobs	No	No	No	Yes
Import Competition Exposure	No	No	No	Yes

Notes: This table presents the results of analyses using data from the British Election Study Internet Panel (BES), specifically waves 8 (2015) and 14 (2017). The dataset was prepared by merging individual records based on unique identifiers. Responses marked as 'Don't know' were treated as missing values. Principal Component Analysis (PCA) was employed to construct composite indices for cultural attitudes and preferences for redistribution. Cultural attitudes were derived from views on immigration, racial equality, gender equality, and gay rights, while preferences for redistribution were based on attitudes toward government spending and taxation, as well as left-right self-placement. In both indices, higher values indicate more liberal stances. These principal components were normalized and residualized against political party identification. For each local authority, a clustering exercise was conducted in a policy space defined by two dimensions: demand for progressive cultural policies and redistribution demand. This process began with Ward's method to determine initial centroids, followed by refinement using K-means clustering. The final step involved calculating the ratio of the distances between two clusters' centroids along the cultural dimension versus the redistribution dimension for each local authority. This ratio was then used as the dependent variable in our analysis. The outcome variable is winsorised at 1% and 99%. Standard errors are clustered at the governmental region level.

Investigating various potential mechanisms, I provide evidence that regions undergoing immigration observe a subsequent reduction in unemployment rates and a boost in economic activity rates. Furthermore, these areas do not experience lower wage growth on average, although a slight decline in wage growth at the lower end of the wage distribution is noted. Additionally, these regions show a reduced burden on the welfare state. Thus, these economic factors, in isolation, cannot fully explain the emergence of anti-immigrant sentiments. The research then shifts to cultural dynamics, showing how immigration influences social attitudes and policy preferences, revealing a growing aversion to immigration.

Bringing these findings into a comprehensive perspective, I provide some suggestive evidence that can explain observed dynamics by voter realignment, transitioning from economic considerations to cultural factors, driven by immigration. Notably, the salience of immigration has surged significantly among voters, political discourse, and media narratives. This heightened prominence of immigration-related topics is concurrent with an increasing disagreement surrounding cultural issues and with cultural factors taking center stage as a pivotal force in shaping electoral choices. Moreover, it becomes evident that individuals tend to cluster along cultural dimensions as a response to immigration, thereby reshaping the political landscape away from traditional economic considerations.

However, this analysis is not without its limitations. The suggestive evidence on immigration's role in voter realignment, while illuminating, points to the need for further research to robustly establish causal links and grasp the full extent of this shift. Recognizing these limitations opens avenues for future inquiry into other potential shocks that might similarly influence political landscapes, such as economic downturns, technological changes, globalization, and environmental crises. Exploring these areas can enhance our grasp of political and social dynamics, informing the creation of responsive and inclusive policies.

These findings carry significant implications for the lens through which we should perceive the political landscape in recent years. We need to account for these dynamic political cleavages in both our theoretical and empirical analysis. Ignoring this evolution could result in a misreading of electoral outcomes, policies that fail to align with the public's needs, increased voter disenchantment, and potentially fueling the rise of populism and extremism.

In conclusion, this paper provides empirical insights that complement existing theoretical frameworks, underscoring the impact of shocks, such as immigration, on voter realignment from economic to cultural considerations. It provides an analysis of how immigration is reshaping the political landscape in the UK, underscoring the need for a more complex and multifaceted understanding of contemporary politics in the face of evolving cultural dynamics.

References

ABRAMITZKY, R., L. BOUSTAN, AND D. S. CONNOR (2024): "Leaving the enclave: Historical evidence on immigrant mobility from the industrial removal office," *The Journal of Economic History*, 84, 352–394. [*Cited on page 151.*]

ADAO, R., M. KOLESÁR, AND E. MORALES (2019): "Shift-share designs: Theory and inference," *The Quarterly Journal of Economics*, 134, 1949–2010. [*Cited on pages 99, 108, and 214.*]

AFESORGBOR, S. K. (2019): "The impact of economic sanctions on international trade: How do threatened sanctions compare with imposed sanctions?" *European Journal of Political Economy*, 56, 11–26. [*Cited on page 49.*]

AGHION, P., U. AKCIGIT, A. HYYTINEN, AND O. TOIVANEN (2017): "The social origins of inventors," Tech. rep., National Bureau of Economic Research. [*Cited on page 5.*]

AGHION, P., A. BERGEAUD, T. BOPPART, P. J. KLENOW, AND H. LI (2023): "A Theory of Falling Growth and Rising Rents," *The Review of Economic Studies*, 90, 2675–2702. [*Cited on page 4.*]

AHN, D. P. AND R. D. LUDEMA (2020): "The sword and the shield: The economics of targeted sanctions," *European Economic Review*, 130, 103587. [*Cited on page 49.*]

AKCIGIT, U., J. GRIGSBY, AND T. NICHOLAS (2017): "The rise of american ingenuity: Innovation and inventors of the golden age," Tech. rep., National Bureau of Economic Research. [*Cited on page 5.*]

ALESINA, A., A. MIANO, AND S. STANTCHEVA (2023): "Immigration and redistribution," *The Review of Economic Studies*, 90, 1–39. [*Cited on page 122.*]

ALESINA, A. AND M. TABELLINI (2024): "The political effects of immigration: Culture or economics?" *Journal of Economic Literature*, 62, 5–46. [*Cited on page 115.*]

ALMAGRO, M. AND T. DOMÍNGUEZ-IINO (2024): "Location sorting and endogenous amenities: Evidence from amsterdam," Tech. rep., National Bureau of Economic Research. [*Cited on page 5.*]

AMBEKAR, A., C. WARD, J. MOHAMMED, S. MALE, AND S. SKIENA (2009): "Name-ethnicity classification from open sources," in *Proceedings of the 15th ACM SIGKDD international conference on Knowledge Discovery and Data Mining*, 49–58. [*Cited on page 151.*]

AMITI, M., S. H. KONG, AND D. WEINSTEIN (2020): "The effect of the US-China trade war on US investment," Tech. rep., National Bureau of Economic Research. [*Cited on page 69.*]

Anginer, D., A. Donmez, H. N. Seyhun, and R. Zhang (2020): "Global economic impact of COVID-19: Evidence from insider trades," *Ray, Global Economic Impact of COVID-19: Evidence from Insider Trades (May 20, 2020)*. [*Cited on page 151.*]

Ashcroft, M. (2016): "How the United Kingdom voted on Thursday… and why," *Lord Ashcroft Polls*, 24, 1–14. [*Cited on page 91.*]

Autor, D. H., D. Dorn, and G. H. Hanson (2013): "The China syndrome: Local labor market effects of import competition in the United States," *American Economic Review*, 103, 2121–68. [*Cited on pages 84 and 96.*]

Azmat, F. (2013): "Opportunities or obstacles? Understanding the challenges faced by migrant women entrepreneurs," *International journal of gender and entrepreneurship*, 5, 198–215. [*Cited on page 31.*]

Baker, S. R., N. Bloom, and S. J. Davis (2016): "Measuring economic policy uncertainty," *The quarterly journal of economics*, 131, 1593–1636. [*Cited on page 75.*]

Barone, G., A. D'Ignazio, G. de Blasio, and P. Naticchioni (2016): "Mr. Rossi, Mr. Hu and politics. The role of immigration in shaping natives' voting behavior," *Journal of Public Economics*, 136, 1–13. [*Cited on page 85.*]

Bartlett, R., A. Morse, R. Stanton, and N. Wallace (2022): "Consumer-lending discrimination in the FinTech era," *Journal of Financial Economics*, 143, 30–56. [*Cited on page 30.*]

Becker, S. O. and T. Fetzer (2018): "Has Eastern European migration impacted UK-born workers?" *Working Paper*. [*Cited on page 116.*]

Becker, S. O., T. Fetzer, and D. Novy (2017): "Who voted for Brexit? A comprehensive district-level analysis," *Economic Policy*, 32, 601–650. [*Cited on pages 86 and 92.*]

Becker, S. O., T. Fetzer, *et al.* (2016): "Does migration cause extreme voting?" *Center for Competitive Advantage in the Global Economy and The Economic & Social Research Council*, 1–54. [*Cited on page 86.*]

Bell, A., R. Chetty, X. Jaravel, N. Petkova, and J. Van Reenen (2019): "Who becomes an inventor in America? The importance of exposure to innovation," *The Quarterly Journal of Economics*, 134, 647–713. [*Cited on page 5.*]

Besley, T. and T. Persson (2019): "The Rise of Identity Politics," . [*Cited on page 87.*]

Bhatiya, A. Y. (2023): "Do Enfranchised Immigrants Affect Politicians' Behaviour?" *'Working Paper'*. [*Cited on page 88.*]

Bologna Pavlik, J. and Y. Zhou (2023): "Are historic districts a backdoor for segregation? Yes and no," *Contemporary Economic Policy*, 41, 415–434. [*Cited on page 151.*]

Bonomi, G., N. Gennaioli, and G. Tabellini (2021): "Identity, beliefs, and political conflict," *The Quarterly Journal of Economics*, 136, 2371–2411. [*Cited on pages 87, 114, 123, 131, and 132.*]

BORUSYAK, K., P. HULL, AND X. JARAVEL (2022): "Quasi-experimental shift-share research designs," *The Review of Economic Studies*, 89, 181–213. [*Cited on pages 86, 97, 98, 99, 108, and 213.*]

BORUSYAK, K., X. JARAVEL, AND J. SPIESS (2024): "Revisiting event study designs: Robust and efficient estimation," *Review of Economic Studies*, rdae007. [*Cited on pages 18, 150, and 175.*]

BRYNJOLFSSON, E., Y. HU, AND D. SIMESTER (2011): "Goodbye pareto principle, hello long tail: The effect of search costs on the concentration of product sales," *Management science*, 57, 1373–1386. [*Cited on page 43.*]

BRYNJOLFSSON, E., Y. J. HU, AND M. D. SMITH (2006): "From niches to riches: Anatomy of the long tail," *Sloan management review*, 47, 67–71. [*Cited on page 43.*]

CALDER-WANG, S. (2021): "The distributional impact of the sharing economy on the housing market," *Available at SSRN 3908062*. [*Cited on page 5.*]

CALLAWAY, B. AND P. H. SANT'ANNA (2021): "Difference-in-differences with multiple time periods," *Journal of econometrics*, 225, 200–230. [*Cited on pages 18, 29, and 181.*]

CARBALLO, J., M. R. CHATRUC, C. S. SANTA, AND C. V. MARTINCUS (2022): "Online business platforms and international trade," *Journal of International Economics*, 137, 103599. [*Cited on page 4.*]

CARD, D., S. CHANG, C. BECKER, J. MENDELSOHN, R. VOIGT, L. BOUSTAN, R. ABRAMITZKY, AND D. JURAFSKY (2022): "Computational analysis of 140 years of US political speeches reveals more positive but increasingly polarized framing of immigration," *Proceedings of the National Academy of Sciences*, 119, e2120510119. [*Cited on page 87.*]

CARRERAS, M., Y. IREPOGLU CARRERAS, AND S. BOWLER (2019): "Long-term economic distress, cultural backlash, and support for Brexit," *Comparative Political Studies*, 52, 1396–1424. [*Cited on page 87.*]

CHEN, M. K., P. E. ROSSI, J. A. CHEVALIER, AND E. OEHLSEN (2019): "The value of flexible work: Evidence from Uber drivers," *Journal of political economy*, 127, 2735–2794. [*Cited on page 4.*]

ČIHÁK, M., A. DEMIRGÜÇ-KUNT, E. FEYEN, AND R. LEVINE (2012): "Benchmarking financial systems around the world," *World Bank policy research working paper*. [*Cited on page 55.*]

COLANTONE, I. AND P. STANIG (2018): "Global competition and Brexit," *American political science review*, 112, 201–218. [*Cited on pages 86 and 92.*]

COMBES, P.-P., B. DECREUSE, M. LAOUENAN, AND A. TRANNOY (2016): "Customer discrimination and employment outcomes: theory and evidence from the french labor market," *Journal of Labor Economics*, 34, 107–160. [*Cited on page 30.*]

CONG, L. W., B. LIU, X. YANG, AND X. ZHANG (2022): "Bridging the Gender Gap in Entrepreneurship and Empowering Women via Digital Technologies," . [*Cited on page 5.*]

COOK, C., R. DIAMOND, J. V. HALL, J. A. LIST, AND P. OYER (2021): "The gender earnings gap in the gig economy: Evidence from over a million rideshare drivers," *The Review of Economic Studies*, 88, 2210–2238. [*Cited on page 4.*]

COUTURE, V., B. FABER, Y. GU, AND L. LIU (2021): "Connecting the countryside via e-commerce: evidence from China," *American Economic Review: Insights*, 3, 35–50. [*Cited on page 5.*]

CROZET, M., J. HINZ, A. STAMMANN, AND J. WANNER (2021): "Worth the pain? Firms' exporting behaviour to countries under sanctions," *European Economic Review*, 134, 103683. [*Cited on page 49.*]

CROZET, M., J. HINZ, *et al.* (2016): *Collateral damage: The impact of the Russia sanctions on sanctioning countries' exports*, CEPII, Centre d'etudes prospectives et d'informations internationales. [*Cited on page 49.*]

DAILY MAIL (2019): "Deliveroo to expand to 50 new towns and cities in the UK," Accessed: 2024-08-06. [*Cited on page 20.*]

DANIELI, O., N. GIDRON, S. KIKUCHI, AND R. LEVY (2022): "Decomposing the Rise of the Populist Radical Right," *Available at SSRN 4255937*. [*Cited on page 87.*]

DE CHAISEMARTIN, C. AND X. D'HAULTFOEUILLE (2020): "Two-way fixed effects estimators with heterogeneous treatment effects," *American economic review*, 110, 2964–2996. [*Cited on pages 18 and 29.*]

DE RIDDER, M. (2024): "Market power and innovation in the intangible economy," *American Economic Review*. [*Cited on page 4.*]

DELIVEROO (2024): "Frequently Asked Questions - Deliveroo," `https://deliveroo.co.uk/faq`, accessed: 19-June-2024. [*Cited on page 10.*]

DOLEAC, J. L. AND L. C. STEIN (2013): "The visible hand: Race and online market outcomes," *The Economic Journal*, 123, F469–F492. [*Cited on page 30.*]

DRACA, M., J. GARRED, L. STICKLAND, AND N. WARRINNIER (2023): "On target? Sanctions and the economic interests of elite policymakers in Iran," *The Economic Journal*, 133, 159–200. [*Cited on pages 49, 50, 56, 66, 67, 68, 77, and 79.*]

DRORI, I. AND M. LERNER (2002): "The dynamics of limited breaking out: The case of the Arab manufacturing businesses in Israel," *Entrepreneurship & Regional Development*, 14, 135–154. [*Cited on page 31.*]

DUSTMANN, C. AND T. FRATTINI (2014): "The fiscal effects of immigration to the UK," *The economic journal*, 124, F593–F643. [*Cited on page 118.*]

DUSTMANN, C., T. FRATTINI, AND I. P. PRESTON (2013): "The effect of immigration along the distribution of wages," *Review of Economic Studies*, 80, 145–173. [*Cited on* page 116.]

DUSTMANN, C. AND I. P. PRESTON (2007): "Racial and economic factors in attitudes to immigration," *The BE Journal of Economic Analysis & Policy*, 7. [*Cited on page 115.*]

——— (2019): "Free movement, open borders, and the global gains from labor mobility," *Annual review of economics*, 11, 783–808. [*Cited on page 115.*]

DUSTMANN, C., K. VASILJEVA, AND A. PIIL DAMM (2019): "Refugee migration and electoral outcomes," *The Review of Economic Studies*, 86, 2035–2091. [*Cited on page 83.*]

EDELMAN, B., M. LUCA, AND D. SVIRSKY (2017): "Racial discrimination in the sharing economy: Evidence from a field experiment," *American economic journal: applied economics*, 9, 1–22. [*Cited on page 30.*]

ENKE, B. (2020): "Moral values and voting," *Journal of Political Economy*, 128, 3679–3729. [*Cited on pages 85, 88, 124, 126, and 127.*]

FAIRLIE, R., A. ROBB, AND D. T. ROBINSON (2022): "Black and white: Access to capital among minority-owned start-ups," *Management Science*, 68, 2377–2400. [*Cited on page 30.*]

FAIRLIE, R. W. AND A. M. ROBB (2007): "Why are black-owned businesses less successful than white-owned businesses? The role of families, inheritances, and business human capital," *Journal of Labor Economics*, 25, 289–323. [*Cited on page 5.*]

FAN, J., L. TANG, W. ZHU, AND B. ZOU (2018): "The Alibaba effect: Spatial consumption inequality and the welfare gains from e-commerce," *Journal of International Economics*, 114, 203–220. [*Cited on page 5.*]

FELBERMAYR, G., A. KIRILAKHA, C. SYROPOULOS, E. YALCIN, AND Y. V. YOTOV (2020): "The global sanctions data base," *European Economic Review*, 129, 103561. [*Cited on page 52.*]

FELBERMAYR, G. J., C. SYROPOULOS, E. YALCIN, AND Y. V. YOTOV (2019): "On the effects of sanctions on trade and welfare: New evidence based on structural gravity and a new database," *CESifo Working Paper*. [*Cited on page 49.*]

FIRPO, S., N. M. FORTIN, AND T. LEMIEUX (2009): "Unconditional quantile regressions," *Econometrica*, 77, 953–973. [*Cited on page 150.*]

GARCIA-LÓPEZ, M.-À., J. JOFRE-MONSENY, R. MARTÍNEZ-MAZZA, AND M. SEGÚ (2020): "Do short-term rental platforms affect housing markets? Evidence from Airbnb in Barcelona," *Journal of Urban Economics*, 119, 103278. [*Cited on page 5.*]

GAULÉ, P. AND M. PIACENTINI (2013): "Chinese graduate students and US scientific productivity," *Review of Economics and Statistics*, 95, 698–701. [*Cited on page 151.*]

GENNAIOLI, N. AND G. TABELLINI (2023): "Identity Politics," *Working Paper*. [*Cited on page 87.*]

GENTZKOW, M., B. KELLY, AND M. TADDY (2019): "Text as data," *Journal of Economic Literature*, 57, 535–74. [*Cited on pages 50 and 87.*]

GOLDSMITH-PINKHAM, P., I. SORKIN, AND H. SWIFT (2020): "Bartik instruments: What, when, why, and how," *American Economic Review*, 110, 2586–2624. [*Cited on page 97.*]

GORBACK, C. (2020): "Your uber has arrived: Ridesharing and the redistribution of economic activity," *Job Market Paper*. [*Cited on page 5.*]

GRAVES, A. AND J. SCHMIDHUBER (2005): "Framewise phoneme classification with bidirectional LSTM and other neural network architectures," *Neural networks*, 18, 602–610. [*Cited on page 151.*]

GROSSMAN, G. M. AND E. HELPMAN (2021): "Identity politics and trade policy," *The Review of Economic Studies*, 88, 1101–1126. [*Cited on page 87.*]

HAFNER, L., T. P. PEIFER, AND F. S. HAFNER (2023): "Equal accuracy for Andrew and Abubakar—detecting and mitigating bias in name-ethnicity classification algorithms," *AI & society*, 1–25. [*Cited on page 151.*]

HAIDAR, J. I. (2017): "Sanctions and export deflection: evidence from Iran," *Economic Policy*, 32, 319–355. [*Cited on page 49.*]

HALL, J. V. AND A. B. KRUEGER (2018): "An analysis of the labor market for Uber's driver-partners in the United States," *Ilr Review*, 71, 705–732. [*Cited on page 4.*]

HALLA, M., A. F. WAGNER, AND J. ZWEIMÜLLER (2017): "Immigration and voting for the far right," *Journal of the European Economic Association*, 15, 1341–1385. [*Cited on pages 83 and 85.*]

HASSAN, T. A., S. HOLLANDER, L. V. LENT, AND A. TAHOUN (2024): "The global impact of Brexit uncertainty," *The Journal of Finance*, 79, 413–458. [*Cited on pages 47, 50, 59, and 75.*]

HASSAN, T. A., S. HOLLANDER, L. VAN LENT, M. SCHWEDELER, AND A. TAHOUN (2023): "Firm-level exposure to epidemic diseases: Covid-19, SARS, and H1N1," Tech. rep., The Review of Financial Studies, Forthcoming. [*Cited on pages 47 and 50.*]

HASSAN, T. A., S. HOLLANDER, L. VAN LENT, AND A. TAHOUN (2019): "Firm-level political risk: Measurement and effects," *The Quarterly Journal of Economics*, 134, 2135–2202. [*Cited on pages 47, 50, 57, 59, 60, and 62.*]

HSIEH, C.-T., E. HURST, C. I. JONES, AND P. J. KLENOW (2019): "The allocation of talent and us economic growth," *Econometrica*, 87, 1439–1474. [*Cited on page 5.*]

HSIEH, C.-T. AND E. ROSSI-HANSBERG (2023): "The industrial revolution in services," *Journal of Political Economy Macroeconomics*, 1, 3–42. [*Cited on page 4.*]

HUFBAUER, G. C. (1990): *Economic sanctions reconsidered: Supplemental case histories*, vol. 2, Institute for International Economics. [*Cited on page 50.*]

Jackson, E. (2022): "Availability of the gig economy and long run labor supply effects for the unemployed," in *2021 APPAM Fall Research Conference*, APPAM. [*Cited on page 4.*]

Jaeger, D. A., J. Ruist, and J. Stuhler (2018): "Shift-share instruments and the impact of immigration," Tech. rep., National Bureau of Economic Research. [*Cited on pages 99, 100, and 106.*]

John Lewis Partnership (2023): "Waitrose Partners with Uber Eats to Build on Success of On-Demand Delivery," https://www.johnlewispartnership.media/pressrelease/jlp/details/17163, accessed: 19-June-2024. [*Cited on page 10.*]

Keeble, M., J. Adams, T. R. Bishop, and T. Burgoine (2021): "Socioeconomic inequalities in food outlet access through an online food delivery service in England: A cross-sectional descriptive analysis," *Applied geography*, 133, 102498. [*Cited on page 6.*]

Kerr, W. R. (2008): "Ethnic scientific communities and international technology diffusion," *The Review of Economics and Statistics*, 90, 518–537. [*Cited on page 151.*]

Kirikakha, A., G. J. Felbermayr, C. Syropoulos, E. Yalcin, and Y. V. Yotov (2021): "The Global Sanctions Data Base (GSDB): an update that includes the years of the Trump presidency," in *Research handbook on economic sanctions*, Edward Elgar Publishing, 62–106. [*Cited on page 52.*]

Kleven, H., C. Landais, and J. E. Søgaard (2019): "Children and gender inequality: Evidence from Denmark," *American Economic Journal: Applied Economics*, 11, 181–209. [*Cited on page 26.*]

Koustas, D. (2018): "Consumption insurance and multiple jobs: Evidence from rideshare drivers," *Unpublished working paper*, 1. [*Cited on page 4.*]

Lashkari, D., A. Bauer, and J. Boussard (2024): "Information technology and returns to scale," *American Economic Review*. [*Cited on page 4.*]

Leonard, J. S., D. I. Levine, and L. Giuliano (2010): "Customer discrimination," *The Review of Economics and Statistics*, 92, 670–678. [*Cited on page 30.*]

Lipton, A. (2022): "The Racial Wealth Gap and the Role of Firm Ownership," in *AEA Papers and Proceedings*, American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, vol. 112, 351–355. [*Cited on page 5.*]

Loughran, T. and B. McDonald (2011): "When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks," *The Journal of finance*, 66, 35–65. [*Cited on page 62.*]

Mayda, A. M., G. Peri, and W. Steingress (2022): "The political impact of immigration: Evidence from the United States," *American Economic Journal: Applied Economics*, 14, 358–389. [*Cited on page 85.*]

Melitz, M. J. (2003): "The impact of trade on intra-industry reallocations and aggregate industry productivity," *econometrica*, 71, 1695–1725. [*Cited on page 13.*]

MENDEZ, I. AND I. M. CUTILLAS (2014): "Has immigration affected Spanish presidential elections results?" *Journal of Population Economics*, 27, 135–171. [*Cited on page 85.*]

MUNSHI, K. (2003): "Networks in the modern economy: Mexican migrants in the US labor market," *The Quarterly Journal of Economics*, 118, 549–599. [*Cited on page 31.*]

NGUYEN, V.-A., J. BOYD-GRABER, P. RESNIK, AND K. MILER (2015): "Tea party in the house: A hierarchical ideal point topic model and its application to republican legislators in the 112th congress," in *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 1438–1448. [*Cited on page 87.*]

NIGMATULINA, D. *et al.* (2022): *Sanctions and Misallocation: How Sanctioned Firms Won and Russia Lost*, Centre for Economic Performance, London School of Economics and Political .... [*Cited on page 49.*]

OFFICE FOR NATIONAL STATISTICS (2015): "Population for postcode districts in England and Wales," Accessed: 2024-07-01. [*Cited on page 19.*]

OTTO, A. H. AND M. F. STEINHARDT (2014): "Immigration and election outcomes—Evidence from city districts in Hamburg," *Regional Science and Urban Economics*, 45, 67–79. [*Cited on page 85.*]

O'CONNELL, M., K. SMITH, AND R. STROUD (2022): "The dietary impact of the COVID-19 pandemic," *Journal of Health Economics*, 84, 102641. [*Cited on page 9.*]

PARASURAMA, P. (2020): *Gender and race preferences in hiring in the age of diversity goals: Evidence from Silicon Valley tech firms*, SSRN. [*Cited on page 151.*]

PATEL, K. AND F. VELLA (2013): "Immigrant networks and their implications for occupational choice and wages," *Review of Economics and Statistics*, 95, 1249–1277. [*Cited on page 31.*]

PERGELOVA, A., T. MANOLOVA, R. SIMEONOVA-GANEVA, AND D. YORDANOVA (2019): "Democratizing entrepreneurship? Digital technologies and the internationalization of female-led SMEs," *Journal of Small Business Management*, 57, 14–39. [*Cited on page 5.*]

POOLE, J. P. AND C. VOLPE (2023): "Can Online Platforms Promote Women-Led Exporting Firms?" *Journal of Globalization and Development*, 14, 357–384. [*Cited on page 5.*]

RAJ, M. AND D. CHOE (2023): "Digital Resilience and the Digital Divide: Broadband Availability and Brick-and-Mortar Retailer Survival," *Available at SSRN 4367290*. [*Cited on page 5.*]

RAJ, M. AND J. EGGERS (2023): "When Delivery Comes to Town: Heterogeneous Effects of Digital Distribution Platform Penetration on Establishment Exit," in *When Delivery Comes to Town: Heterogeneous Effects of Digital Distribution Platform Penetration on Establishment Exit: Raj, Manav— uEggers, JP*, [Sl]: SSRN. [*Cited on page 5.*]

RAJ, M., A. SUNDARARAJAN, AND C. YOU (2023): "COVID-19 and digital resilience: Evidence from Uber Eats," *Working Paper*. [*Cited on page 5.*]

SALEHI-ISFAHANI, D. (2023): "The impact of sanctions on household welfare and employment in Iran," *Middle East development journal*, 15, 189–221. [*Cited on page 48.*]

SAUTNER, Z., L. VAN LENT, G. VILKOV, AND R. ZHANG (2023): "Firm-level climate change exposure," *The Journal of Finance*, 78, 1449–1498. [*Cited on pages 47 and 50.*]

SCHAEFER, M. AND K. D. TRAN (2020): "Airbnb, hotels, and localized competition," *DIW Berlin Discussion Paper*. [*Cited on page 5.*]

SHAYO, M. (2009): "A model of social identity with an application to political economy: Nation, class, and redistribution," *American Political science review*, 103, 147–174. [*Cited on page 87.*]

SHORROCKS, A. (2013): "Decomposition procedures for distributional analysis: a unified framework based on the Shapley value," *The Journal of Economic Inequality*, 11, 99–126. [*Cited on page 163.*]

SICAT, M., A. XU, E. MEHETAJ, M. FERRANTINO, AND V. CHEMUTAI (2020): "Leveraging ICT technologies in closing the gender gap," *World Bank*. [*Cited on page 5.*]

SOOD, G. AND S. LAOHAPRAPANON (2018): "Predicting race and ethnicity from the sequence of characters in a name," *arXiv preprint arXiv:1805.02109*, 5. [*Cited on page 151.*]

STATISTA (2024): "Online Food Delivery - Worldwide," `https://www.statista.com/outlook/dmo/online-food-delivery/worldwide`, accessed: 2024-07-18. [*Cited on page 6.*]

STEIN, L. C. AND E. STONE (2013): "The effect of uncertainty on investment, hiring, and R&D: Causal evidence from equity options," *Hiring, and R&D: Causal Evidence from Equity Options (October 4, 2013)*. [*Cited on page 209.*]

STEINMAYR, A. (2021): "Contact versus exposure: Refugee presence and voting for the far right," *Review of Economics and Statistics*, 103, 310–327. [*Cited on page 85.*]

STONE, M. (2016): "The Response of Russian security prices to economic sanctions: policy effectiveness and transmission," *US Department of State Office of the Chief Economist Working Paper*. [*Cited on page 49.*]

SUN, L. AND S. ABRAHAM (2021): "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," *Journal of econometrics*, 225, 175–199. [*Cited on pages 18, 30, and 181.*]

TABELLINI, M. (2020): "Gifts of the immigrants, woes of the natives: Lessons from the age of mass migration," *The Review of Economic Studies*, 87, 454–486. [*Cited on pages 83, 85, and 123.*]

VISKANIC, M. (2017): "Fear and loathing on the campaign trail: did immigration cause Brexit?" *Available at SSRN 2941611*. [*Cited on page 86.*]

WILKERSON, J. AND A. CASAS (2017): "Large-scale computerized text analysis in political science: Opportunities and challenges," *Annual Review of Political Science*, 20, 529–544. [*Cited on page 87.*]

YANG, J., H. ASKARI, J. FORRER, AND H. TEEGEN (2004): "US economic sanctions: An empirical study," *The International Trade Journal*, 18, 23–62. [*Cited on page 49.*]

Appendix to Paper one

A1.1 Quantile Treatment Effects

In this section, I describe the application of a non-linear Difference-in-Differences (DiD) method to estimate the impact of the platform rollout on the distribution of the number of restaurants across different regions. The resulting Quantile Treatment Effects (QTEs) enable us to assess the impact on different parts of the outcome distribution.

Before delving into our results, it is helpful to briefly review the concepts of quantiles and QTEs. For any variable Y with a CDF function $F(y) \equiv \Pr[Y \leq y]$, the qth quantile of F is defined as the smallest value y_q such that $F(y_q) = q$. In a purely random treatment setting, we could compare two distributions, F_1 and F_0, representing the outcome variable in the treatment and control groups, respectively. The QTE at the qth quantile is then defined as $\Delta_q = y_q(1) - y_q(0)$, where $y_q(t)$ is the qth quantile of distribution F_t. This effect can graphically be represented as the horizontal distance between the graphs of F_1 and F_0 at the probability value q.

It is crucial to recognize that QTEs do not necessarily identify the treatment's impact on a specific locality or neighborhood. For example, if the platform leads to rank reversals in the distribution of restaurant numbers, simply knowing the median differences between the two distributions will not suffice to calculate the treatment effect for a locality that would have had the median number of restaurants either before or after the treatment. Nevertheless, the presence of a negative (positive) QTE indicates that the treatment effect is negative (positive) across some non-degenerate interval of the counterfactual restaurant distribution.

The inclusion of covariates and fixed effects complicates the analysis by necessitating a choice between conditional and unconditional quantile regression.[21] In conditional quantile regression (CQR), the inclusion of fixed effects controls for selection bias but also alters the definition of quantiles. That is, CQR estimates the treatment's impact on the nth conditional quantile of the outcome variable, indicating how the policy affects those at specific positions within the distribution of the outcome variable, conditional on the

[21]It is important to note that when no other covariates are involved, the conditional and unconditional treatment effects of a binary X are the same across all quantiles of Y. However, once additional covariates, such as fixed effects, are introduced-as in our study-the distinction between conditional and unconditional quantile regression becomes significant.

covariates or fixed effects. However, our primary interest lies in understanding the impact on units with low outcome levels unconditionally. Unconditional quantile regression (UQR) addresses this by estimating the effect of the policy on the overall distribution of the outcome variable, providing insights more relevant for policy evaluation. Unlike CQR, which focuses on within-group effects, UQR captures the impact of the independent variable on the entire distribution of the dependent variable, akin to OLS regression.

Unconditional quantile regression (UQR) offers the advantage of defining quantiles prior to model estimation, making it less susceptible to influence from right-hand-side variables. However, computational challenges arise when applying UQR to models with high-dimensional fixed effects. To address this, the recentered influence function (RIF) method is employed. This involves calculating Influence Function (RIF) for each observation and subsequently using these as the dependent variable in an OLS regression with the relevant independent variables. For a detailed methodological explanation, refer to Firpo *et al.* (2009). UQR estimates offer a more intuitive interpretation compared to conditional quantile regression, as they capture the impact of the treatment on specific quantiles of the outcome without conditioning on other variables or within groups, as in conditional quantile regression.

Figure A37 displays the QTE estimates derived from the RIF-DiD estimator. The point estimates are either zero or positive across the distribution up to the 92nd percentile. As we move to higher quantiles, particularly between the 70th and 80th percentiles, the QTE becomes more positive, peaking around the 80th percentile. However, confidence intervals widen significantly at higher quantiles, suggesting greater uncertainty or heterogeneity in treatment effects at the upper end of the distribution.

Overall, the QTE estimates suggest that the introduction of food apps had a positive impact across most of the distribution of the number of restaurants. However, it is important to interpret these results with caution. To the best of my knowledge, within the context of quantile regression in a Difference-in-Differences framework with staggered treatment rollout, there is no established package that fully addresses the complexities of treatment effect heterogeneity, as highlighted by Borusyak *et al.* (2024), and the challenges of conditional quantile regression with many fixed effects. Consequently, this analysis may not fully account for the concerns raised by the treatment heterogeneity literature. This is particularly important given the potential heterogeneity in treatment effects across markets of different sizes on the one hand and the relationship between treatment timing and the initial market conditions on the other.

A1.2 Name-Based Analysis to Infer Ethnicity and Gender

As described, the company house dataset lacks direct information on gender and ethnicity. However, these attributes can be inferred using name-based analysis. This approach, widely used in research economics and economic history (Kerr, 2008; Gaulé and Piacentini, 2013; Abramitzky *et al.*, 2024), employs an algorithm or machine learning methods to predict race and ethnicity based on names. For this purpose, the `ethnicolr` Python package is utilized, a tool increasingly common in academic literature (Anginer *et al.*, 2020; Parasurama, 2020; Bologna Pavlik and Zhou, 2023).

The `ethnicolr` package uses a long short-term memory (LSTM) neural network trained on US census data, Florida voter registration data, and Wikipedia data (Sood and Laohaprapanon, 2018). This study uses the model trained on Wikipedia, as it is less US-centric compared to other datasets. LSTM networks, a type of recurrent neural network (RNN) introduced by seminal work of Graves and Schmidhuber (2005), are particularly effective due to their unique memory cells that selectively remember and forget information, allowing for efficient incremental updates.

Using the Wikipedia training dataset compiled by Ambekar *et al.* (2009), `ethnicolr` predicts race and ethnicity based on first and last names. The package achieves higher accuracy when both first and last names are used together, as this provides more comprehensive information (Sood and Laohaprapanon, 2018). Although the training dataset is not specific to the UK, its global scope likely covers a wide range of immigrant backgrounds relevant to the UK. Hafner *et al.* (2023) showed Wikipedia-trained `ethnicolr` has shown a more balanced performance across ethnicities compared to other methods.

Technically, `ethnicolr` calculates the probability that a given name belongs to one of thirteen racial/ethnic groups: "Asian, Greater East Asian, East Asian", "Asian, Greater East Asian, Japanese", "Asian, Indian Subcontinent", "Greater African, Africans", "Greater African, Muslim", 'Greater European, British", "Greater European, East European", "Greater European, Jewish", 'Greater European, West European, French", "Greater European, West European, Germanic", "Greater European, West European, Hispanic", "Greater European, West European, Italian", and "Greater European, West European, Nordic". These categories are further classified into British, South Asia, East Asia, European, South American, Muslim, and African.

To infer genders from names in my dataset, I utilize the `gender-guesser` package. This package allows me to determine the likely gender associated with a given first name through a straightforward Python interface. By inputting names into the package, I can classify each as male, female, androgynous (andy), mostly male, mostly female, or unknown if the name is not found in the underlying database. For analysis, I treat "mostly male" as male and "mostly female" as female, as this does not significantly impact the results. The process involves creating a Detector object from the package, which uses a precompiled list of over 40,000 names and their associated genders and countries of origin. This dataset is designed to encompass the majority of first names used in European countries and several non-European countries, including China, India, Japan, and the US. By leveraging this tool, I can systematically infer and categorize the genders of individuals in my dataset, facilitating comprehensive demographic analysis.

A1.3 Matching Restaurants Across Datasets

Matching restaurants across different data sources, such as LDC, Google Maps, and delivery platforms, is complicated by inconsistencies in business names, address variations, and chain restaurants with multiple locations.

To achieve accurate matches, I employed a multi-step methodology combining exact matches with fuzzy matching techniques. Initially, I identified chain restaurants using a predefined list and matched them based on exact business names and postcodes. For non-chain or unmatched entries, I leveraged fuzzy matching based on restaurant names within the same postal district to account for minor discrepancies in naming conventions.

Specifically, I used the `fuzzywuzzy` library's `extractOne` function for this fuzzy matching. This tool compares a given restaurant name with a list of possible matches, calculating a similarity score based on Levenshtein distance, which measures how many single-character edits are needed to make the names identical. The function returns the closest match along with a similarity score ranging from 0 to 100.

If the similarity score was above 80, I accepted the match. For scores between 70 and 80, I further verified the match by checking exact postcode matches within the same postal district. This hybrid approach maximized accuracy and ensured a comprehensive understanding of market presence and business dynamics while accounting for variability in business records.

A1.4 Model Details

A1.4.1 Utility Function and Consumer Preferences

The first part of the model refers to consumer preferences and utility maximization under monopolistic competition:

$$\max\ U = \left[\int_{\Omega} \left(q(\omega)^{\frac{\sigma-1}{\sigma}} \right) d\omega \right]^{\frac{\sigma}{\sigma-1}}$$

Where:

- U is the utility of the representative consumer.
- Ω is the set of available varieties (goods).
- $q(\omega)$ is the quantity consumed of variety ω.
- σ is the elasticity of substitution between varieties (with $\sigma > 1$).

The first-order condition (FOC) of utility maximization leads to the demand for each variety:

$$q(\omega) = YP(\omega)^{-\sigma}P^{\sigma-1}$$

Where P is the aggregate price index, defined as:

$$P = \left(\int_{\Omega} P(\omega)^{1-\sigma} d\omega \right)^{\frac{1}{1-\sigma}}$$

There are L identical consumers:

$$y(\omega) = Lq(\omega) = LwP(\omega)^{-\sigma}P^{\sigma-1}$$

Where w is the wage rate which can be set to one.

A1.4.2 Production: Monopolistic Competition

The firm maximizes its profit under monopolistic competition:

$$\max_{\{q\}}\ \pi = pq - C(q)$$

From the firm's pricing rule and profit maximization:

$$P + \frac{dP}{dq}q - C = 0 \quad \Rightarrow \quad P(1 + \frac{dP}{dq}\frac{q}{P}) = C$$

This leads to the condition:

$$p(1 - \frac{1}{\epsilon_q}) = C \quad \text{or} \quad \frac{P}{C} = \frac{\epsilon_q}{\epsilon_q - 1}$$

Where ϵ_q is the price elasticity of demand. So, we can write:

$$P(\phi) = \frac{\sigma}{\sigma - 1}\frac{w}{\phi}$$

Where ϕ represents the firm's productivity. Revenue is given by:

$$r(\phi) = LP^{\sigma-1}\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}\left(\frac{\phi}{w}\right)^{1-\sigma}$$

The profit function is:

$$\pi(\phi) = y(\phi)r(\phi) - y(\phi)\frac{w}{\phi} - wf_d = \frac{r(\phi)}{\sigma} - wf_d$$

Which simplifies further into:

$$\pi(\phi) = LP^{\sigma-1}\frac{(\sigma-1)^{(\sigma-1)}}{\sigma^{\sigma}}\left(\frac{\phi}{w}\right)^{\sigma-1} - wf_d$$

Where f_d is the fixed cost of production.

A1.4.3 Zero-Profit Cutoff Condition

The profit condition for firms is given by:

$$\pi_d(\varphi^*) = YLP^{\sigma-1}\frac{(\sigma-1)^{(\sigma-1)}}{\sigma^{\sigma}}\frac{\varphi^{*\sigma-1}}{w^{\sigma-1}} - wf_d = B_p\varphi^{*\sigma-1} - wf_d = 0$$

Where $B_p = LP^{\sigma-1}\frac{(\sigma-1)^{(\sigma-1)}}{\sigma^{\sigma}w^{\sigma-1}}$

At the survival cutoff φ^*, profits are driven to zero:

$$B_p\varphi^{*\sigma-1} = wf_d$$

A1.4.4 Free Entry Condition

The free entry condition requires that the expected profit from entering the market equals the entry cost wf_e:

$$wf_e = \int_{\varphi^*}^{\infty}\pi(\phi)dG(\phi)$$

Where $G(\phi)$ is the distribution of firm productivities. Substituting the profit function:

$$wf_e = \int_{\varphi^*}^{\infty}\left(B_p\phi^{\sigma-1} - wf_d\right)dG(\phi)$$

If we substitute for B_p and set wage to one, this becomes:

$$wf_e = \int_{\varphi^*}^{\infty}\left(f_d\left(\frac{\phi}{\varphi^*}\right)^{\sigma-1} - f_d\right)dG(\phi) = f_d\int_{\varphi^*}^{\infty}\left(\left(\frac{\phi}{\varphi^*}\right)^{\sigma-1} - 1\right)dG(\phi)$$

Assuming a Pareto distribution for ϕ with scale parameter ϕ_{min} and shape parameter θ, the integral can be expressed in a closed-form solution. By substituting the probability density function (PDF) into the integral, we obtain:

$$I = \int_{\varphi^*}^{\infty} \left(\left(\frac{\phi}{\varphi^*} \right)^{\sigma-1} - 1 \right) \frac{\theta\, \phi_{\text{min}}^{\theta}}{\phi^{\theta+1}} \, d\phi$$

We simplify the integrand by separating the terms:

$$\begin{aligned}
I &= \theta\, \phi_{\text{min}}^{\theta} \left[\int_{\varphi^*}^{\infty} \left(\frac{\phi}{\varphi^*} \right)^{\sigma-1} \frac{1}{\phi^{\theta+1}} \, d\phi - \int_{\varphi^*}^{\infty} \frac{1}{\phi^{\theta+1}} \, d\phi \right] \\
&= \theta\, \phi_{\text{min}}^{\theta} \left[\varphi^{*-(\sigma-1)} \int_{\varphi^*}^{\infty} \phi^{\sigma-\theta-2} \, d\phi - \int_{\varphi^*}^{\infty} \phi^{-\theta-1} \, d\phi \right] \\
&= \theta\, \phi_{\text{min}}^{\theta} \left[\varphi^{*-(\sigma-1)} \left(-\frac{\varphi^{*\sigma-\theta-1}}{\sigma-\theta-1} \right) - \frac{\varphi^{*-\theta}}{\theta} \right] \\
&= \theta\, \phi_{\text{min}}^{\theta} \left[-\frac{\varphi^{*-(\sigma-1)} \varphi^{*\sigma-\theta-1}}{\sigma-\theta-1} - \frac{\varphi^{*-\theta}}{\theta} \right] \\
&= \theta\, \phi_{\text{min}}^{\theta} \left[-\frac{\varphi^{*-\theta}}{\sigma-\theta-1} - \frac{\varphi^{*-\theta}}{\theta} \right] \\
&= \phi_{\text{min}}^{\theta} \varphi^{*-\theta} \left[-\frac{\theta}{\sigma-\theta-1} - 1 \right] \\
&= \left(\frac{\varphi^*}{\phi_{\text{min}}} \right)^{-\theta} \left(\frac{\sigma-1}{1+\theta-\sigma} \right)
\end{aligned}$$

setting wage equal to one, this would imply:

$$\varphi^* = \phi_{\text{min}} \left(\frac{f_d}{f_e} \right)^{\frac{1}{\theta}} \left(\frac{\sigma-1}{1+\theta-\sigma} \right)^{\frac{1}{\theta}} \tag{26}$$

A1.4.5 Proofs

Proposition I Proof The labor market equilibrium can be expressed as:

$$L = M_a \int_{\phi_a}^{\infty} \left[\frac{Lq(\phi)}{\phi} + f_d \right] \frac{g(\phi)}{1-G(\phi_a)} d\phi + M_e f_e$$

We know that $M_e = \frac{M_a}{1-G(\phi_a)}$

$$L = M_a \int_{\phi_a}^{\infty} \left[\frac{Lq(\phi)}{\phi} + f_d \right] \frac{g(\phi)}{1-G(\phi_a)} d\phi + M_a \frac{f_e}{1-G(\phi_a)}$$

Using the relationship for labor demand:

$$\frac{Lq(\phi)}{\phi} + f_d = \frac{Py(\phi)}{W} - \frac{\pi(\phi)}{W} = (\sigma - 1)\frac{\pi(\phi)}{W} + \sigma f_d$$

Substituting this into the labor market equilibrium condition, we have:

$$L = M_a \int_{\phi_a}^{\infty} \left[(\sigma - 1)\frac{\pi(\phi)}{W} + \sigma f_d\right] \frac{g(\phi)}{1 - G(\phi_a)} d\phi + M_a \frac{f_e}{1 - G(\phi_a)}$$
$$L = M_a \left[(\sigma - 1)\frac{\bar{\pi}}{W} + \sigma f_d\right] + M_a \frac{f_e}{1 - G(\phi_a)}$$
$$L = M_a \left[(\sigma - 1)\frac{f_e}{(1 - G(\phi_a))W} + \sigma f_d\right] + M_a \frac{f_e}{1 - G(\phi_a)}$$

Finally, solving for M_a:

$$M_a = \frac{L}{\sigma \left(\frac{f_e}{1 - G(\phi_a)} + f_d\right)}$$

This expression gives the equilibrium number of active firms M_a as a function of total labor, the fixed costs, and the productivity distribution cutoff ϕ_a.

Assuming Pareto distribution for ϕ, we have $1 - G(\phi_a) = \left(\frac{\phi_{\min}}{\varphi^*}\right)^{\theta}$. Substituting the value of φ^* from equation 26, this becomes $\frac{f_e}{f_d}\frac{1+\theta-\sigma}{\sigma-1}$. Using this relationship, we can express the equilibrium mass of active firms as:

$$M_a = \frac{L}{\sigma\ f_d \frac{\theta}{1+\theta-\sigma}} = \frac{L\,(1 + \theta - \sigma)}{\sigma f_d \theta} \tag{27}$$

From this expression, we observe that M_a increases with θ. This implies that if a technology enhances the productivity of superstar firms—resulting in a decrease in θ (i.e., making the productivity distribution more fat-tailed)—the mass of active firms in the market decreases. □

Proposition II Proof From the labor market equilibrium equation (Equation 27), it is evident that a decrease in f_d leads to an increase in the mass of firms. □

Proposition III Proof Starting from the expression for the equilibrium mass of active firms:

$$M_a = \frac{L(1 + \theta - \sigma)}{\sigma f_d \theta}$$

Treating L and σ as constants, we can write M_a as a function of θ and f_d:

$$M_a = \frac{C(1 + \theta - \sigma)}{f_d \theta} \quad \text{where} \quad C = \frac{L}{\sigma}$$

To find how M_a changes with θ and f_d, we compute the total differential dM_a:

$$dM_a = \frac{\partial M_a}{\partial \theta} d\theta + \frac{\partial M_a}{\partial f_d} df_d$$

Calculating the partial derivatives:

$$\frac{\partial M_a}{\partial \theta} = \frac{C\left[(\theta)(1) - (1 + \theta - \sigma)(1)\right]}{f_d \theta^2} = \frac{C(\sigma - 1)}{f_d \theta^2}$$
$$\frac{\partial M_a}{\partial f_d} = -\frac{C(1 + \theta - \sigma)}{f_d^2 \theta}$$

Substituting back into the total differential:

$$dM_a = \frac{C(\sigma - 1)}{f_d \theta^2} d\theta - \frac{C(1 + \theta - \sigma)}{f_d^2 \theta} df_d$$

To express the changes in proportional terms, we divide both sides by M_a:

$$\begin{aligned}\frac{dM_a}{M_a} &= \left(\frac{1}{M_a}\right)\left(\frac{\partial M_a}{\partial \theta} d\theta + \frac{\partial M_a}{\partial f_d} df_d\right) \\ &= \left(\frac{f_d \theta}{C(1 + \theta - \sigma)}\right)\left(\frac{C(\sigma - 1)}{f_d \theta^2} d\theta - \frac{C(1 + \theta - \sigma)}{f_d^2 \theta} df_d\right) \\ &= \frac{(\sigma - 1)}{\theta(1 + \theta - \sigma)} d\theta - \frac{df_d}{f_d}\end{aligned}$$

Recognizing that $d\theta$ and df_d are negative due to decreases in θ and f_d, we let:

$$d\theta = -\Delta\theta \quad \text{and} \quad df_d = -\Delta f_d$$

Substituting back:

$$\frac{dM_a}{M_a} = -\frac{(\sigma - 1)}{(1 + \theta - \sigma)} \frac{\Delta\theta}{\theta} + \frac{\Delta f_d}{f_d}$$

For the equilibrium number of firms to increase ($dM_a > 0$), we require:

$$\frac{\Delta f_d}{f_d} > \frac{(\sigma - 1)}{(1 + \theta - \sigma)} \frac{\Delta\theta}{\theta}$$

This condition ensures that the positive effect of reduced fixed costs outweighs the negative effect of a more unequal productivity distribution. □

The Impact on the Price index:

We have:

$$YLP^{\sigma-1}\frac{(\sigma-1)^{(\sigma-1)}}{\sigma^{\sigma}}\varphi^{*\sigma-1}=f_d$$

We can see that if φ^* increases, P has to decrease.

A1.5 Analysis of Restaurateurs' Experiences on Reddit

To gain insights into the motivations and experiences of restaurant owners regarding the adoption of food delivery applications, I collected data from two Reddit subreddits: `r/restaurateur` and `r/restaurantowners`. These forums serve as platforms where restaurant owners and prospective owners discuss industry-related topics, share experiences, and seek advice.

I extracted all posts from these subreddits that contained keywords indicative of food delivery applications, such as "UberEats", "Deliveroo", "food delivery app", "Grubhub", and "DoorDash." The time frame for data collection spanned from January 2015 to December 2023, capturing the period during which food delivery applications became prominent in the industry.

Traditional text analysis methods, such as Latent Dirichlet Allocation (LDA), often require extensive preprocessing and may not capture nuanced language used in informal online discussions. To address these limitations, I employed Large Language Models (LLMs) to analyze the collected Reddit posts. Specifically, I utilized OpenAI's GPT series models, known for their advanced natural language understanding capabilities.

The analysis proceeded in two main stages:

A1.5.1 Topic Identification and Classification Scheme Development

I randomly sampled 100 posts from the dataset to serve as a representative subset for initial analysis. Using the GPT-4 model (version `OpenAI's GPT-o1`), I performed a qualitative content analysis to identify recurring themes and topics within these posts. From this analysis, I identified several key topics, which formed the basis of the classification scheme. The primary categories included:

(1) **Expanding Customer Reach**
(2) **Marketing and Visibility**
(3) **Operational Efficiency and Workflow Infrastructure**
(4) **Reducing On-Premise Delivery Costs**
(5) **Reducing Premises Costs**
(6) **Competitive Pressure**
(7) **Customer Convenience**
(8) **Data and Analytics Access**

A1.5.2 Automated Classification of the Full Dataset

With the classification scheme established, I proceeded to classify the entire dataset of posts using the GPT-4 model (`gpt-4o-2024-08-06`). The process involved feeding each post into the LLM with instructions to assign one or more of the predefined categories based on the content of the post. Posts were allowed to be assigned multiple categories if they touched on several topics. To assess the reliability of the LLM's classifications, I compared the automated classifications with manual annotations on a random subset of 50 posts, achieving an agreement rate of over 90%, which indicates high consistency. Discrepancies between the automated and manual classifications were analyzed to refine prompts and improve the model's performance.

The prompt used was:

> *"Read the following post from a restaurant owner discussing food delivery applications. If the post discusses the possibility of joining a food delivery application, expresses interest in joining, or mentions benefits of using such platforms, identify and list the main reasons they mention for adopting or considering the use of these platforms from the following list..."*

A1.5.3 Results

The classification results are summarized in Figure A5. The most frequently cited reason for adopting food delivery applications was **Expanding Customer Reach**, highlighting the importance restaurant owners place on accessing a broader market.

Other significant motivations included:

- **Marketing and Visibility (18.2%)**: Restaurant owners appreciated the promotional benefits provided by the platforms, which reduce the need for independent marketing efforts.
- **Operational Efficiency and Workflow Infrastructure (17.6%)**: The platforms' integrated order management and delivery logistics streamline operations, lowering the burden on in-house staff.
- **Reducing On-Premise Delivery Costs (13.1%)**: Outsourcing delivery services to the platforms eliminates the need to maintain a fleet of delivery personnel.
- **Reducing Premises Costs (7.4%)**: Some restaurant owners noted that partnering with delivery apps allows them to operate in smaller physical spaces or less expensive locations, as dine-in facilities become less critical.

A1.6 Determinants of Platform Rollout Dates

I employ a basic machine learning approach to identify the subsets of regional factors that most effectively predict the platform's rollout dates accross UK postal districts. Although my goal is not to establish a causal explanation due to the multifaceted nature of platform decisions, an in-depth examination of various socio-economic variables can shed light on the elements influencing the system's rollout, which serves as the identifying variation in this study.

More concretely, I conduct a feature selection procedure to determine the strongest predictors of the rollout date. For this, I apply Best Subset Selection (BSS), a machine learning method used for feature selection, aimed at reducing the dimensionality of the feature space. The concept behind BSS is to test all possible models, considering every combination of control variables, and produce the statistically best-fit model that minimizes an information criterion. The detailed steps are as follows:

A1.6.1 Covariates Selection

I consider covariates that pass a first plausibility test. If this test is not satisfied, the model may include variables lacking theoretical justification, practical relevance, or empirical support, leading to several issues. These issues include compromised interpretability, reduced predictive accuracy and reliability due to noise, and overlooked multicollinearity causing unstable coefficients.

The covariates I choose include variables from groups such as indicators of the area's restaurant industry, variables reflecting the trend in demographic and human capital characteristics, and metrics that capture the region's economic structure and its evolution. More specifically, there are more than 30 variables used in this analysis, including both level and trend variables. These variables represent aspects such as population size, number of restaurants, GDP, urbanization levels, age demographics, hourly pay statistics, migration growth, unemployment rates, and economic dependence on various sectors, migration growth, unemployment rates, and sectoral employment shares in agriculture, mining, manufacturing, construction, retail, hotel and restaurant, transport, and finance. These variables capture both the current state and the changes in regional socio-economic conditions.

A1.6.2 Best Subset Identification

BSS involves evaluating all possible combinations of predictors to find the subset that best fits the data for different numbers of parameters. Initially, models containing a single predictor ($p = 1$) are evaluated, with each model assessed for its fit using metrics like the residual sum of squares (RSS). Next, all possible models containing exactly two predictors ($p = 2$) are evaluated. This step involves assessing the fit of models with pairs of predictors. The process continues for models with three predictors, four predictors, and so on, until all combinations of predictors have been considered. This exhaustive search ensures that the best subset of predictors is identified for each possible number of parameters.

$$\min_{\beta} \sum_{c=1}^{C} \left(y_c - \beta_0 - \sum_{j=1}^{p} x_{cj}\beta_j \right)^2 \quad \text{Residual sum of squares}$$

Once I have the total set of covariates, BSS evaluates all possible combinations of predictors and selects the subset that minimizes a specific criterion. In this analysis, I use the commonly employed Akaike Information Criterion (AIC).

A1.6.3 Information Criterion

The previous step helps to find the best predictors for each number of predictors. The information criterion refines the model selection by providing a criterion for choosing the best model among the subsets of predictors. The AIC balances model fit and complexity, ensuring that the selected model is not only accurate but also parsimonious.

More formally, the objective is to minimize the AIC for each subset of predictors S:

$$\min_{S \subseteq \{1,2,\ldots,p\}} \left\{ n \ln \left(\frac{\text{RSS}(S)}{n} \right) + 2|S| \right\}$$

where:

- $\text{RSS}(S) = \sum_{i=1}^{n} \left(y_i - \sum_{j \in S} \beta_j X_{ij} \right)^2$
- $|S|$ is the number of predictors in the subset S
- n is the number of observations

Using AIC in BSS ensures that the selected model not only fits the data well but also remains parsimonious, avoiding the pitfalls of overfitting.

This statistically optimal approach can quickly become impractical as the number of potential regressors, p, increases. In BSS, the process involves estimating models for every possible combination of regressors using Ordinary Least Squares (OLS). Initially, models with one regressor are evaluated, followed by models with two regressors, and so on, until all combinations are considered. This results in evaluating 2^p models in total. As p grows, the computational burden becomes immense, making the process infeasible for large datasets. While our model had just enough potential features to remain feasible, larger sets of features necessitate the use of regularization methods like LASSO and Ridge Regression. These methods solve convex optimization problems efficiently, making them suitable for high-dimensional data.

One should bear in mind that the BSS method can generate models with varying levels of complexity, which are not necessarily nested. I outline the sequence of 'best' models for each set of predictors p and assess how including additional covariates enhances the model's fit. A drawback of this approach is that highly correlated variables may be excluded. This implies that even if a predictor x_i provides a unique contribution when conditioned on x_j, it might be left out of the analysis if its signal isn't strong enough.

A1.6.4 Results

Table A9 presents the results of the BSS analysis. The first column reports the model that includes only the best predictor. The second column adds the best when we can have two predictors, and so forth, with each subsequent column incorporating an additional permissible predictor. As evident from the results, rurality emerges as the most significant predictor, followed by population size and educational attainment. Notably, trend variables do not seem to play a significant role in predicting platform rollout dates. This indicates that while certain static socio-economic factors are critical, rather than underlying trends. It is worth highlighting that the R^2 is overall high.

A1.6.5 Shorrocks-Shapley Decomposition:

After using BSS to select the best subset of predictors, the Shorrocks-Shapley decomposition (Shorrocks, 2013) can be applied to the final model to understand the relative contribution of each selected predictor to the R^2. The Shorrocks-Shapley decomposition works by considering all possible permutations of the predictors and calculating the marginal contribution of each predictor to the R^2 of the model. This marginal contribution is the change in R^2 when a predictor is added to a model that includes a subset of the other predictors. By averaging these marginal contributions across all possible orderings of the predictors, the Shorrocks-Shapley value for each predictor is obtained.

Figure A38 shows the results. As you can see Urbanisation, population and GDP level are the most important contributors to predicting the rollout date.

A1.7 Extra Graphs

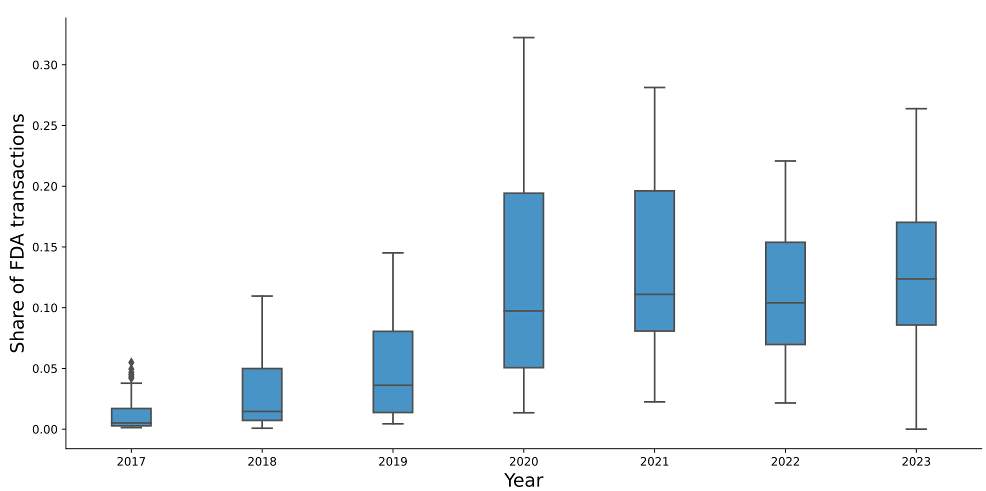


Figure A1. *Notes:* The figure shows a box-and-whisker plot depicting the penetration of UberEats and Deliveroo across ONS subgroups over time. The y-axis represents the share of Food App transactions, while the x-axis shows the years from 2017 to 2023. Data reflects the distribution and trends in the adoption of these platforms over the specified period.

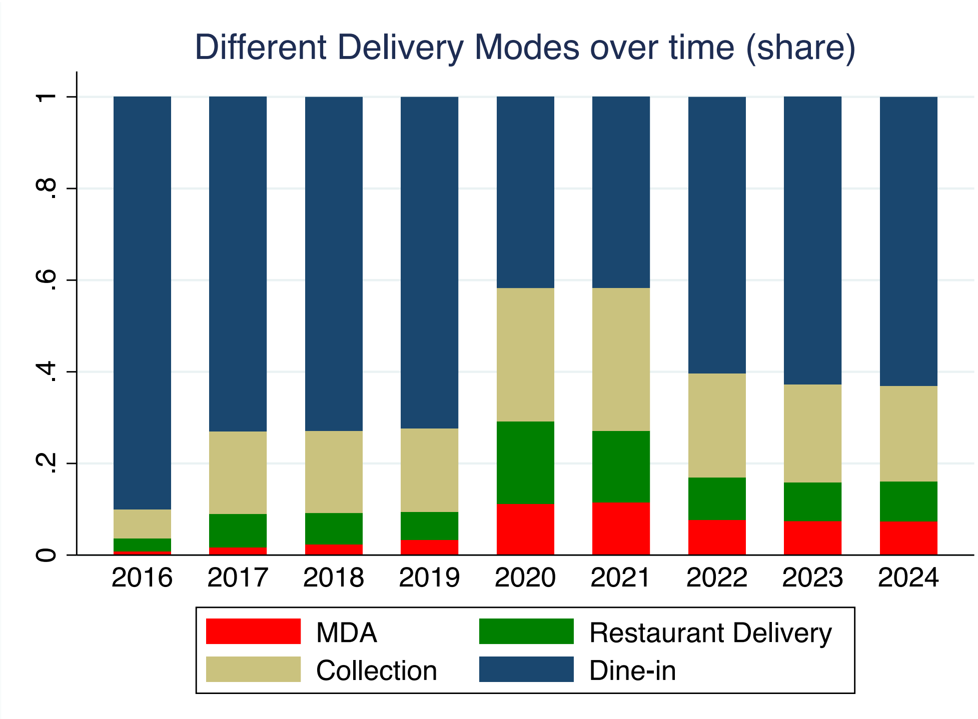
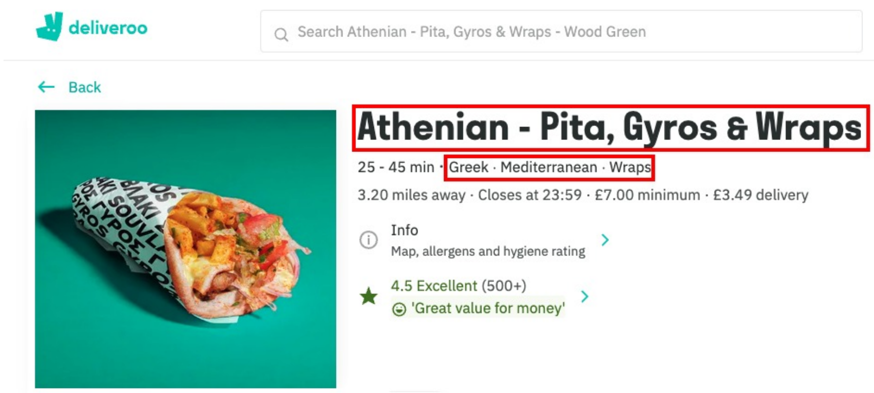
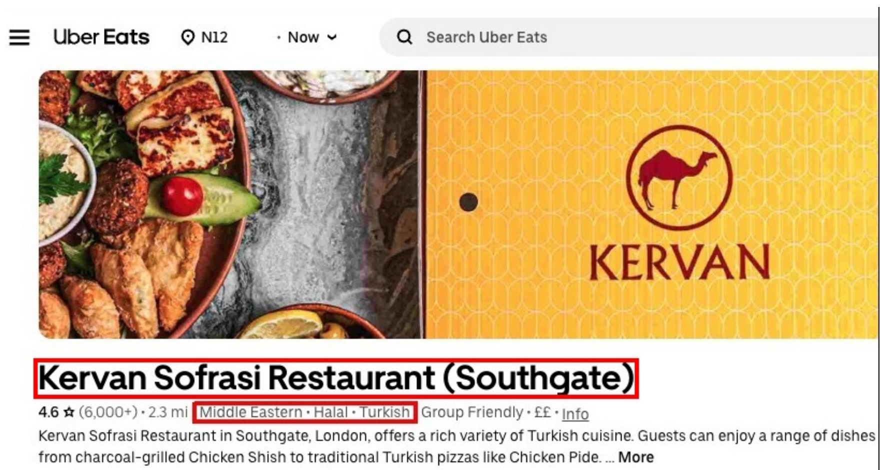


Figure A2. *Notes:* The graph depicts the distribution of delivery modes over time. The first category includes dine-in orders. The second category encompasses orders made through food delivery applications from services Just Eat, Uber Eats, Deliveroo, Amazon Restaurants, and Hungry House. This category also includes orders placed through these platforms for personal collection. The third category represents orders delivered by the restaurant's own fleet, placed either through the restaurant's application, website or via phone. The final category is for customers who personally visit the restaurant to pick up their food. The data is from Kantar's Worldpanel Out of Home Panel for the years 2016 to 2024 Q1. It is important to note that orders labeled a "Restaurant's Own Website" (approximately 0.19% of observations) are assumed to involve delivery, though this label does not explicitly distinguish between delivery and collection. The graph is constructed based on observations where mealcomponent==1 thus excluding drinks and side dishes only transactions. It excludes data from years before 2017, as there are no recorded deliveries for those years.

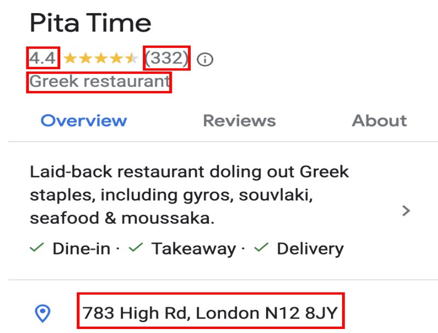


(a) Deliveroo

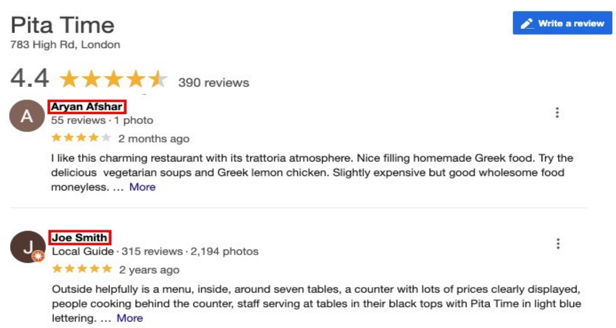


(b) UberEats

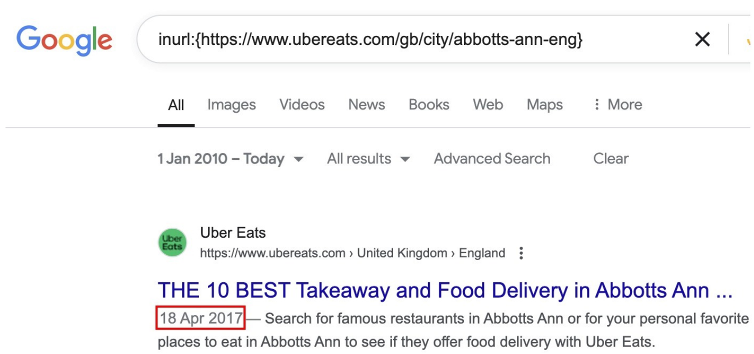
Figure A3. *Notes:* Panel (a) shows a sample restaurant on Deliveroo along with key information extracted from it, such as the name, rating, cuisine type, and address. Panel (b) displays the same procedure for a restaurant on UberEats.



(a) Google Maps



(b) Google Reviewers



(c) Google Indexed date

Figure A4. *Notes:* This figure presents a sample of information extracted from Google Maps. Panel (a) shows a restaurant listing with key details, such as rating, cuisine type, and address. Panel (b) displays sample reviewers for the restaurant. In practice, all reviews for each restaurant are scraped, and reviewers' names are extracted to infer their backgrounds. Panel (c) shows a Google search result for a location's UberEats URL, with the indexed date indicated.

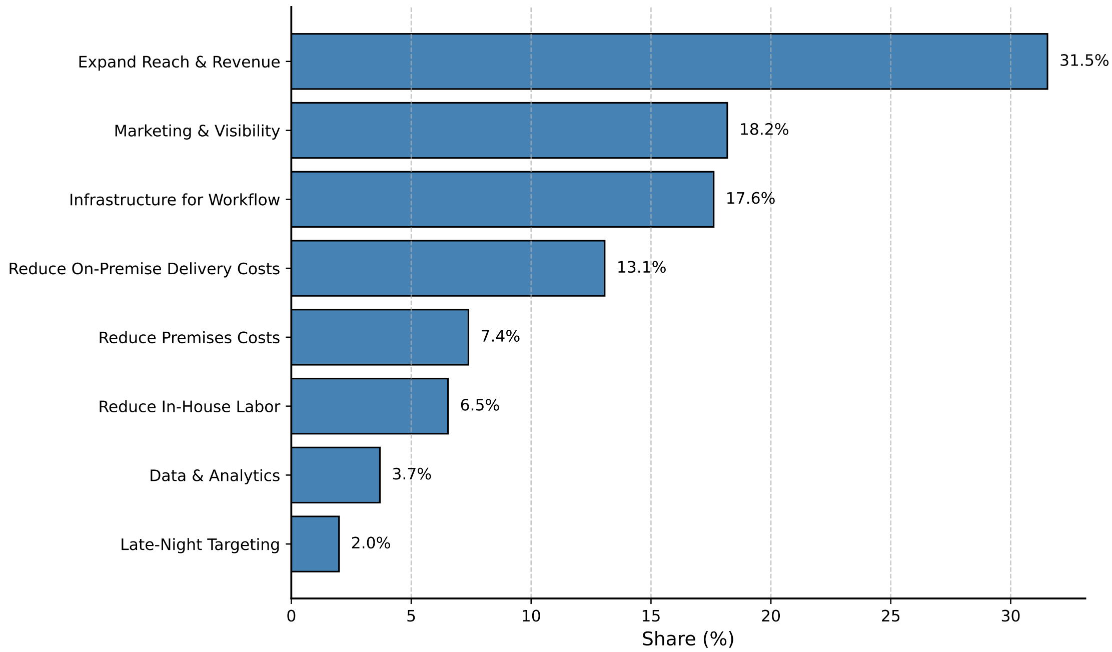


Figure A5. *Notes:* This figure illustrates the percentage distribution of identified reasons why restaurateurs consider partnering with food delivery apps, based on data from Reddit's `r/restaurateur` and `r/restaurantowners` subreddits. Posts mentioning relevant keywords indicative of using food apps (661 posts) were aggregated and analyzed using OpenAI's GPT-4 language model (`gpt-4o-2024-08-06`), which classified each post according to a set of predefined benefit categories.

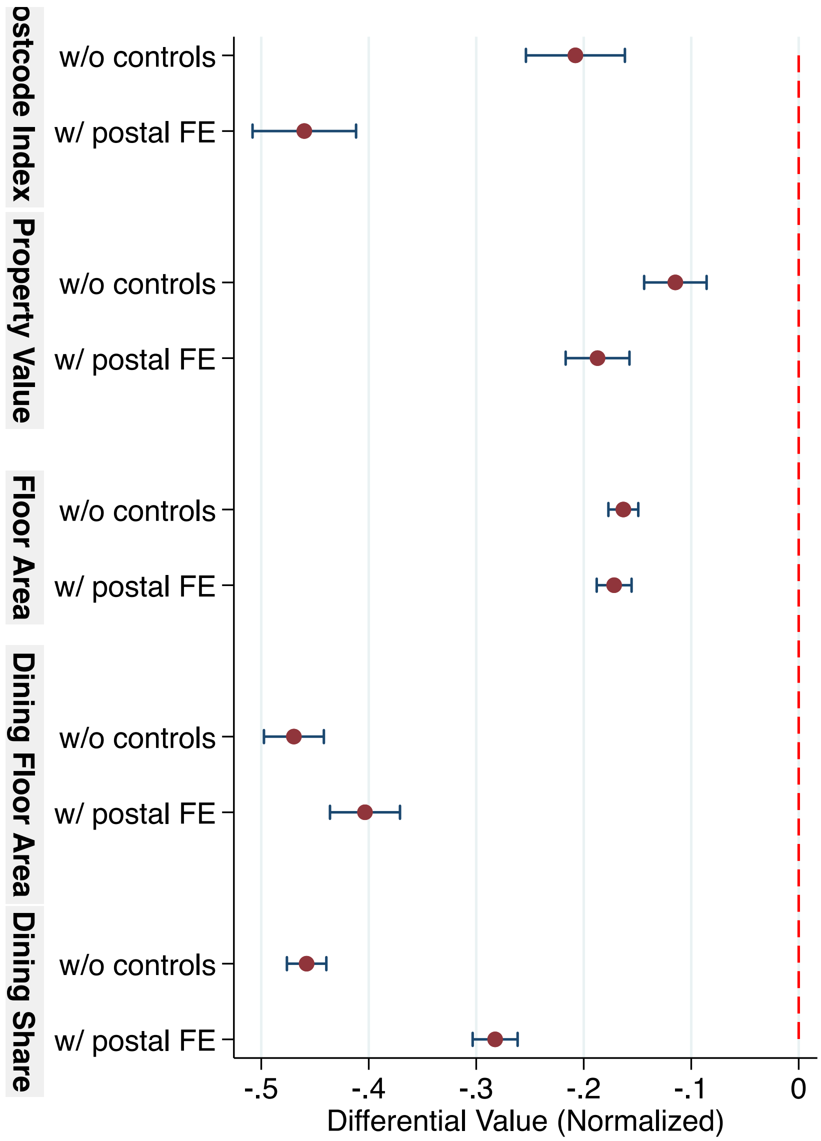
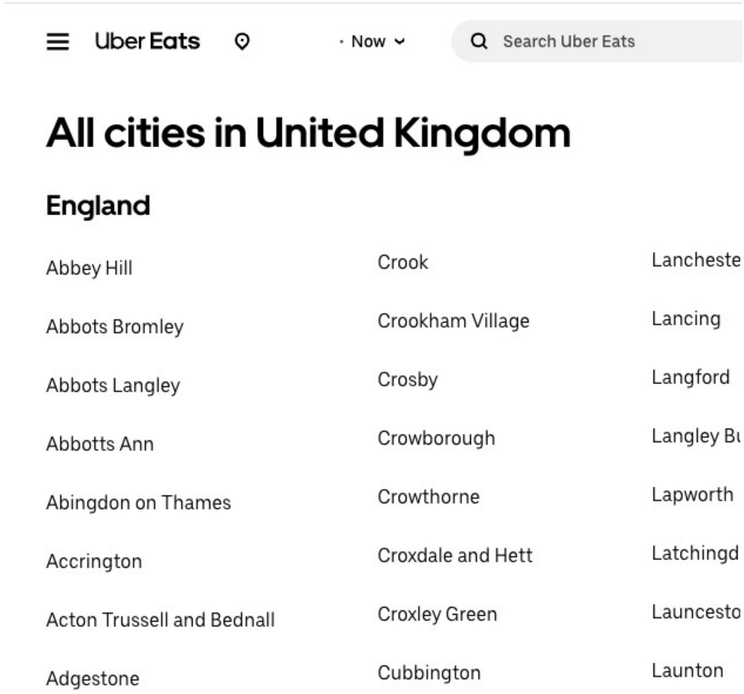
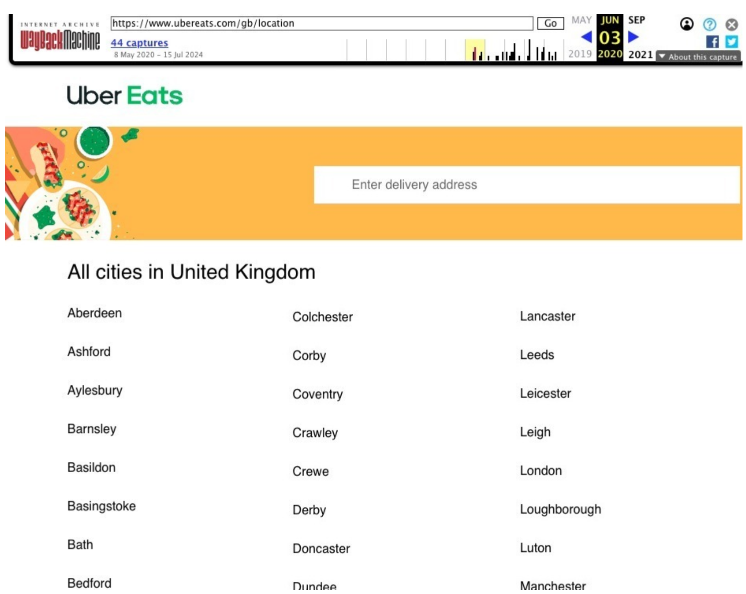


Figure A6. *Notes:* This graph presents the differences in property characteristics between restaurants listed on food delivery applications (Food App) and those not listed, among non-chain restaurants. The outcome variables include Price Index (estimated average property price within the restaurant's postcode), Property Value (total market value of the property), Total Area (overall floor area), Dining Area (floor area dedicated to dining), and Dining Share (percentage of the property's area used for dining). Each outcome is regressed on the Food App indicator both without controls and with postal district fixed effects, with estimates normalized using the dependent variable's mean. The price index is derived from Price Paid data to measure postcode-level land price, while other outcomes are sourced from the Valuation Office Agency (VOA) 2023 dataset.

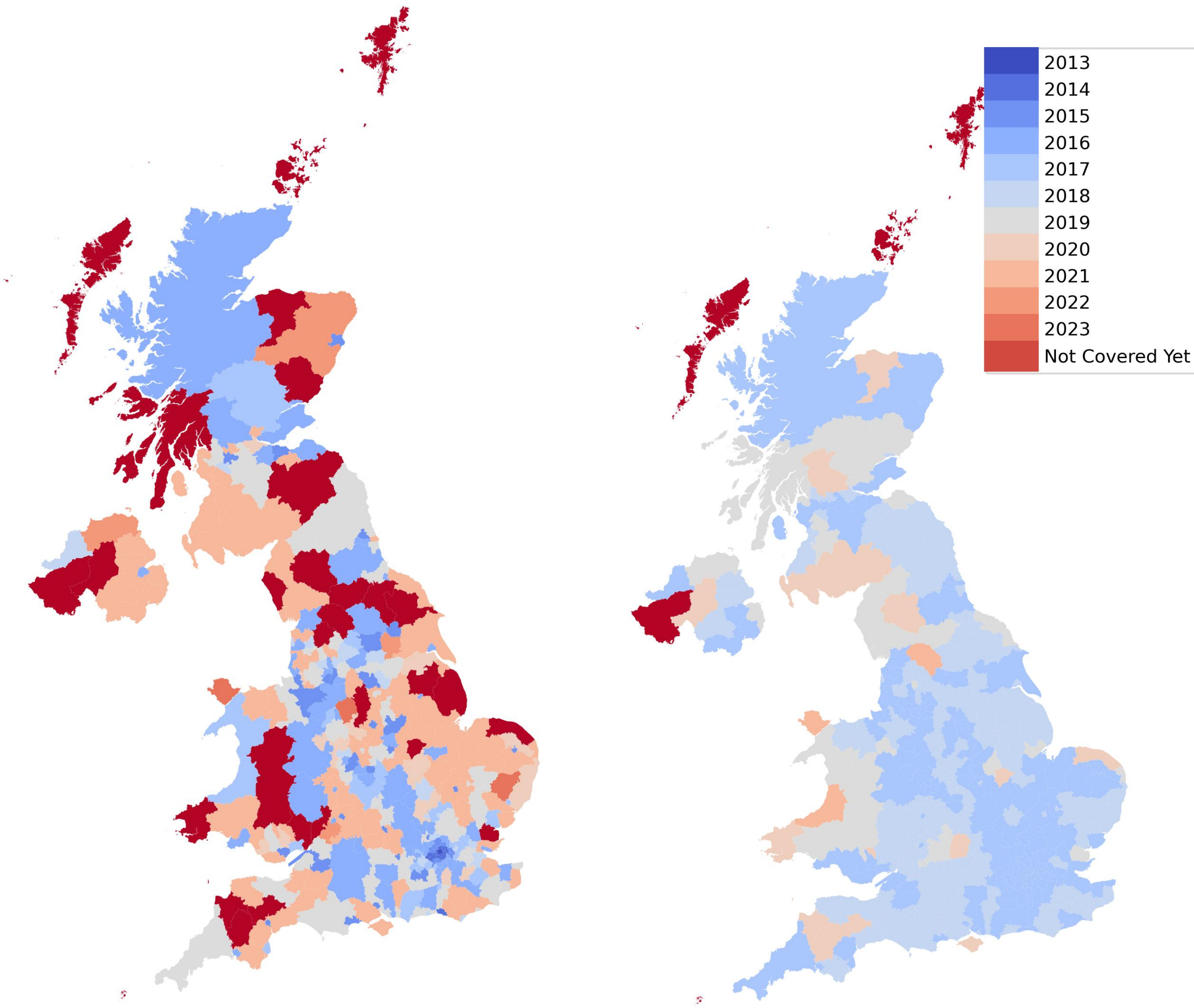


(a) UberEats Coverage Page



(b) Internet Archive

Figure A7. *Notes:* Panel (a) displays a webpage from UberEats showcasing their UK coverage area. Panel (b) illustrates the archived version of this page, retrieved from the Wayback Machine (Internet Archive) on June 3, 2020.



(a) Deliveroo Rollout

(b) UberEats Rollout

Figure A8. *Notes:* This map displays UK local authorities that have a minimum of one restaurant featured on Google Maps. Panel (a) depicts the introduction of the Deliveroo application, and Panel (b) indicates the introduction of the UberEats application. The UK local authority boundary file is sourced from: `https://geoportal.statistics.gov.uk/datasets/196d1a072aaa4882a50be333679d4f63/explore`. A small number of postal districts could not be directly mapped due to updates in postal district definitions. These unmatched districts were associated with the closest matching district from the boundary file.

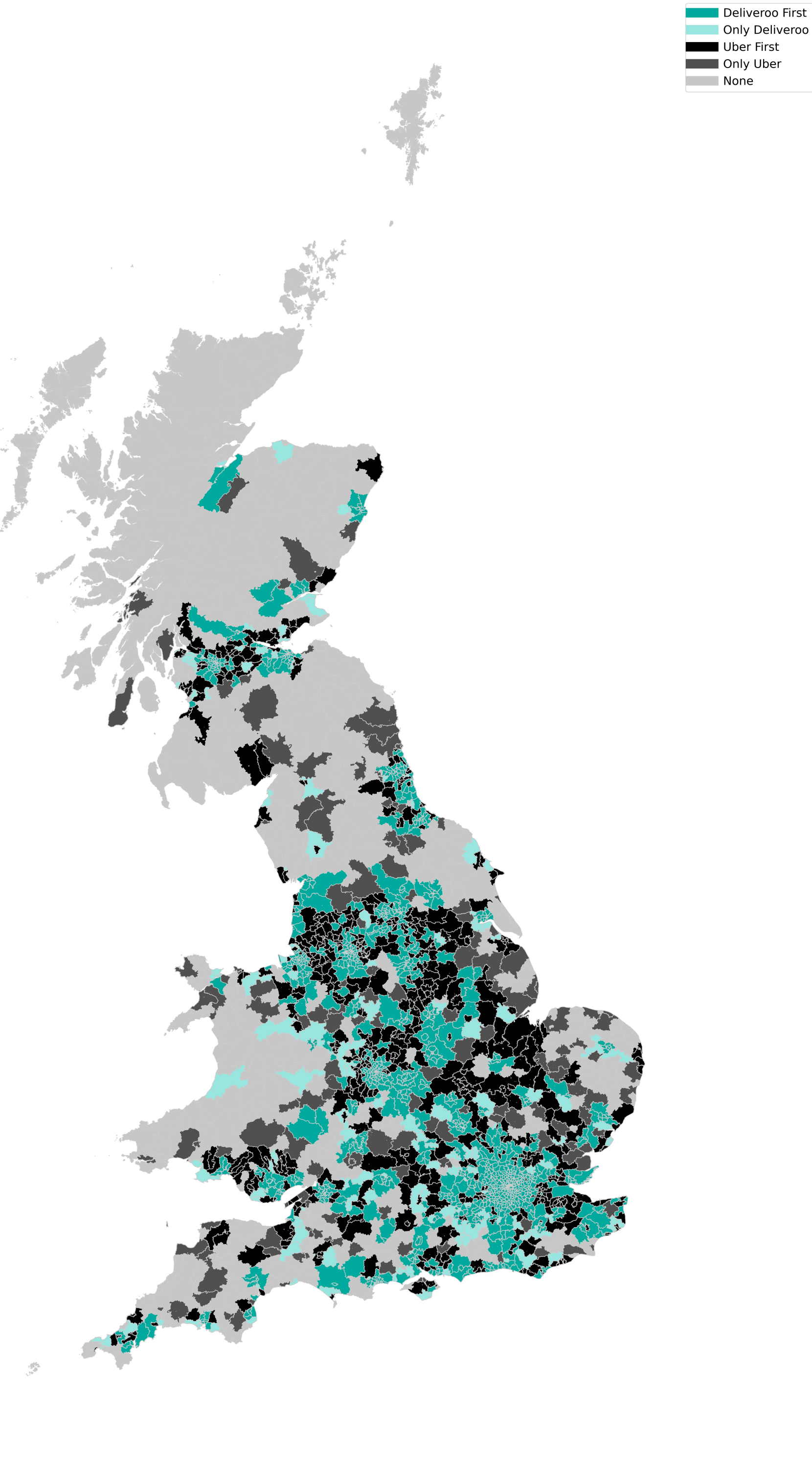
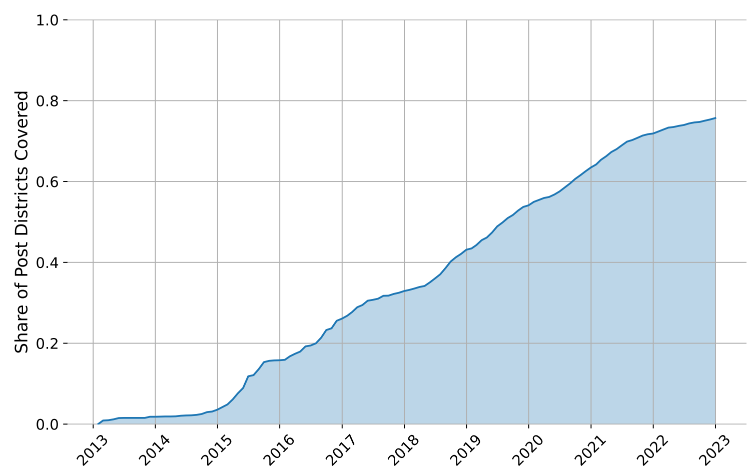
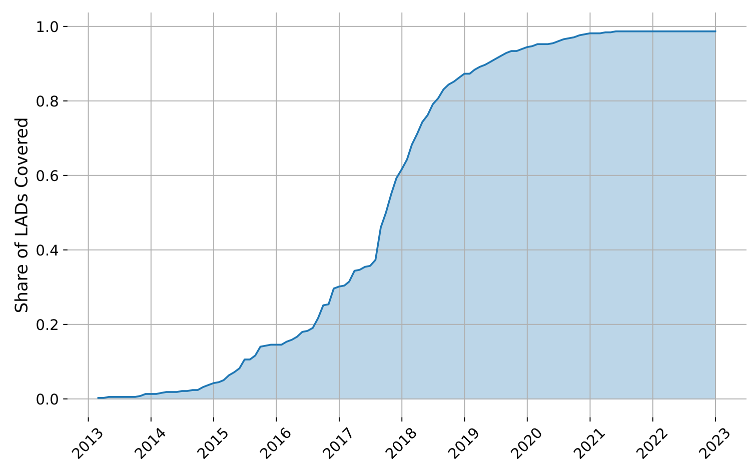


Figure A9. *Notes:* The map illustrates the rollout dynamics of Deliveroo and UberEats in UK postal districts, showing which platform entered first, which districts are only served by Deliveroo or Uber, and which districts are not served by either platform. Data is sourced from the author's scraping data collection.



(a) Share of Postal Districts Covered



(b) Share of LADs Covered

Figure A10. *Notes:* The figures illustrate the proportion of postal districts (panel a) and local authorities (panel b) covered by either UberEats or Deliveroo. The definition of penetration, i.e., rollout for each platform in each spatial unit, is discussed in section 1.3.

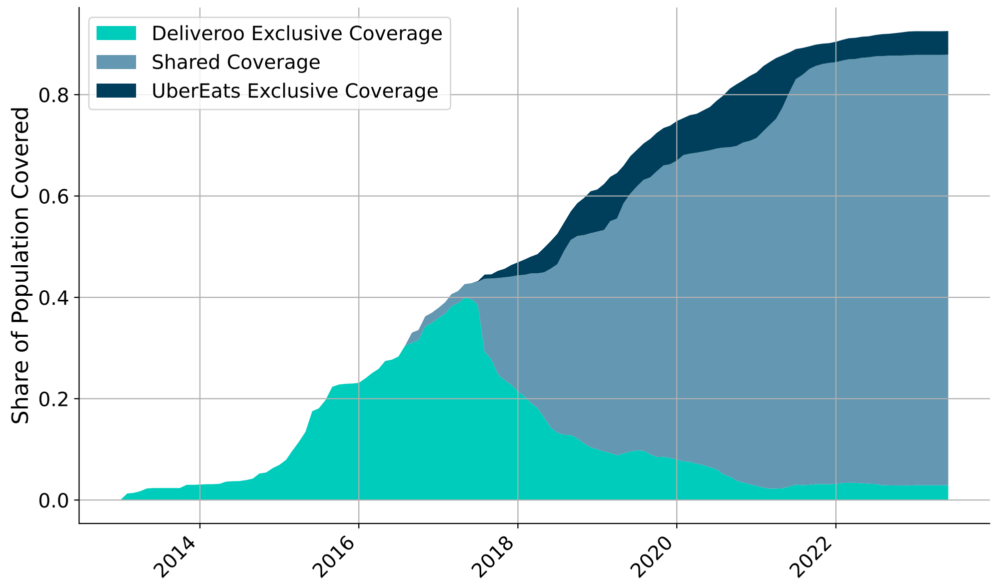


Figure A11. *Notes:* The figure illustrates the share of the population with access to Deliveroo (teal), UberEats (dark blue), and both services (light blue) over time. 'Having access' is defined for each postal district as detailed in the accompanying text, with population figures derived from the 2021 census.

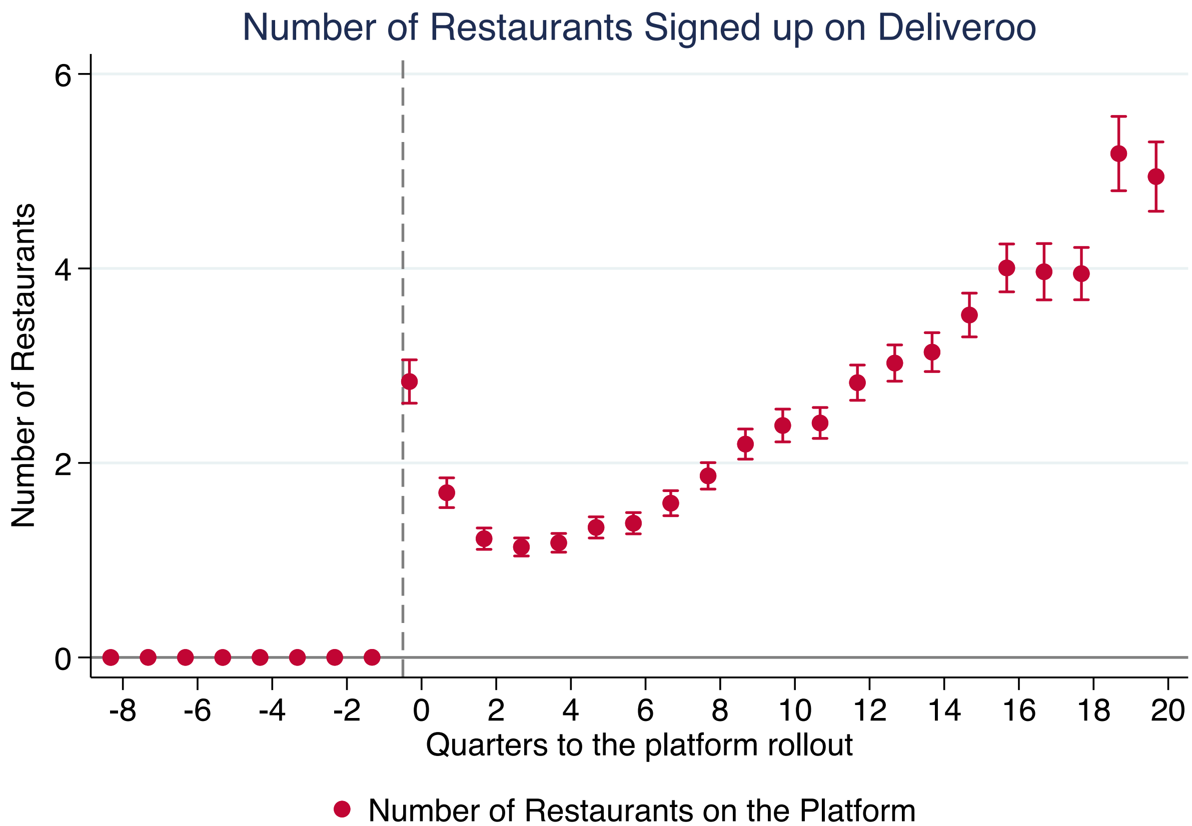
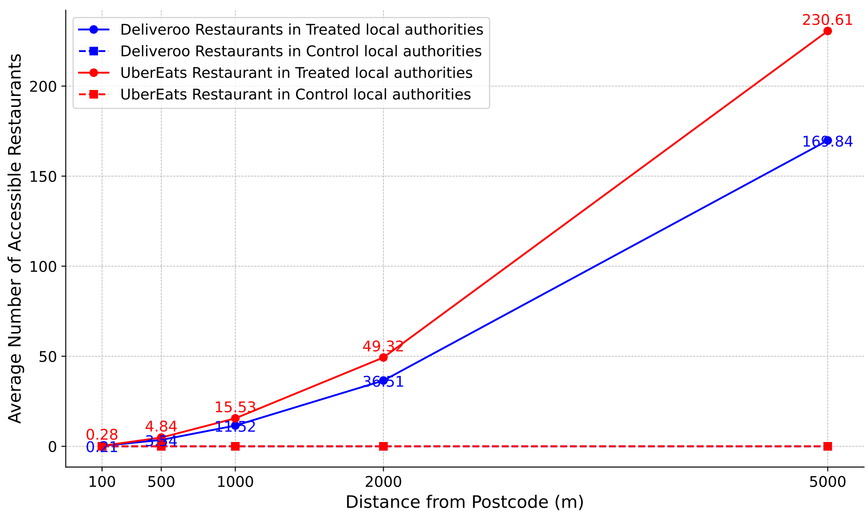


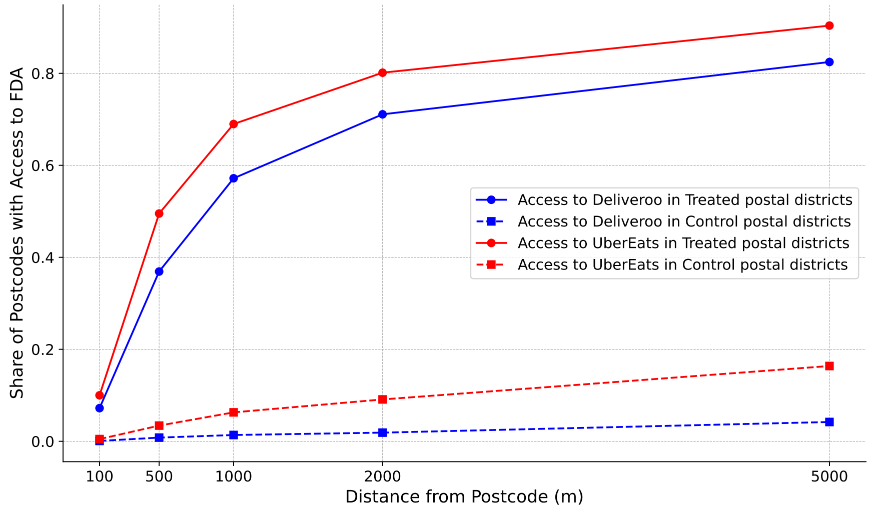
Figure A12. *Notes:* This figure presents the event study results for Deliveroo, where the outcome variable is the number of new restaurants on the platform. Borusyak *et al.* (2024) estimator is used. Postcodes with Deliveroo rollout before 2017-03 (768 postcodes) are dropped since imputation is impossible for these units as they are treated in all periods in the sample. The graph represents a fully dynamic regression incorporating all leads and lags, though only the first 12 leads and 48 lags are visually depicted.

(a) Postal District

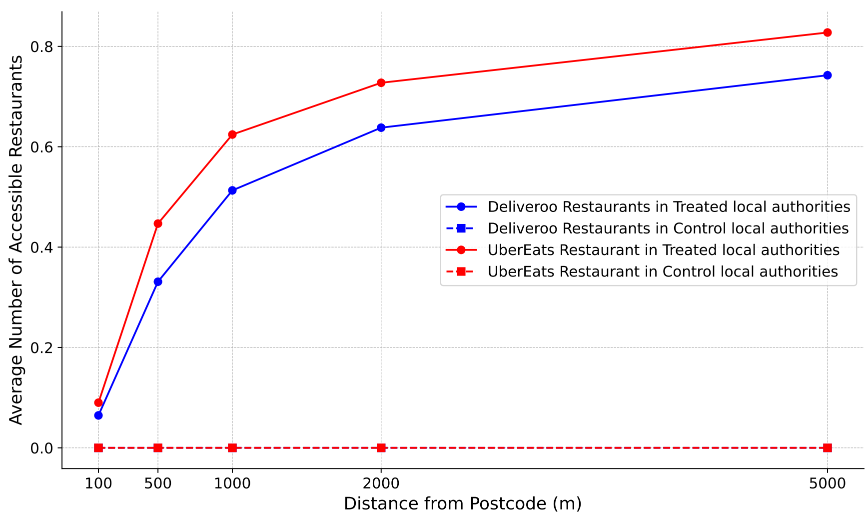


(b) Local Authority

Figure A13. *Notes:* The figure illustrates the average number of accessible restaurants on each platform as a function of distance from postcodes, comparing postcodes in treated and control spatial units as of 2023. Panel (a) shows postcodes in treated and control postal districts, while Panel (b) shows postcodes in treated and control local authorities. The definition of how units are classified into these control and treated groups is detailed in Section 1.3. The dataset includes 1.6 million postcodes, and the geodesic distance between a given postcode's coordinates and those of nearby restaurants is calculated. Data for this analysis were derived from the National Statistics Postcode Lookup (NSPL) and the author's calculation of Deliveroo restaurant entries.

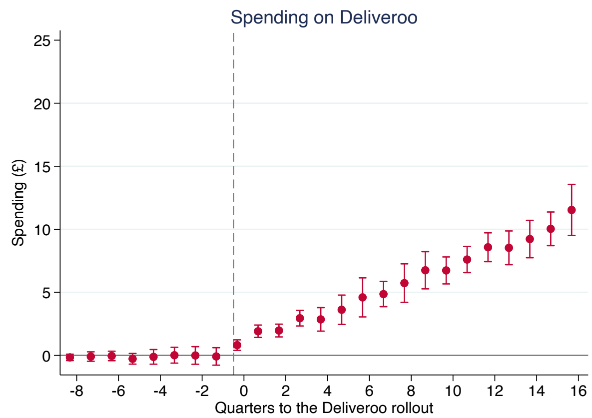


(a) Postal District

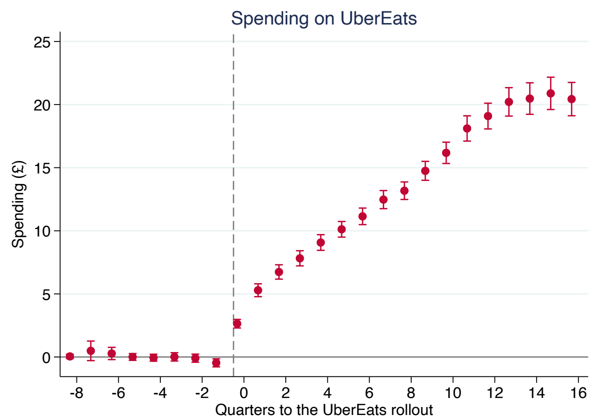


(b) Local Authority

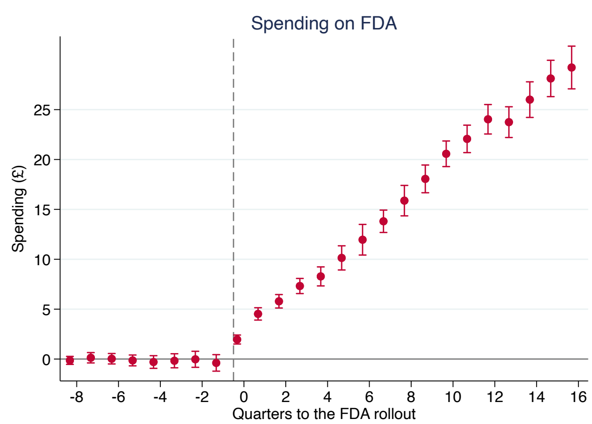
Figure A14. *Notes:* The figure illustrates the share of postcodes having access to Food App as a function of distance from postcodes, comparing postcodes in treated and control spatial units as of 2023. Panel (a) shows postcodes in treated and control postal districts, while panel (b) shows postcodes in treated and control local authorities. The definition of how units are classified into these control and treated groups is detailed in Section 1.3. The dataset includes 1.6 million postcodes, and the geodesic distance between a given postcode's coordinates and those of nearby restaurants is calculated. Data for this analysis were derived from the National Statistics Postcode Lookup (NSPL) and the author's calculation of Deliveroo restaurant entries.



(a) Deliveroo Rollout Analysis

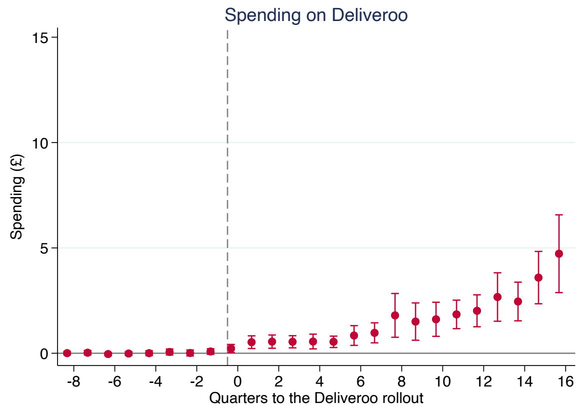


(b) UberEats Rollout Analysis

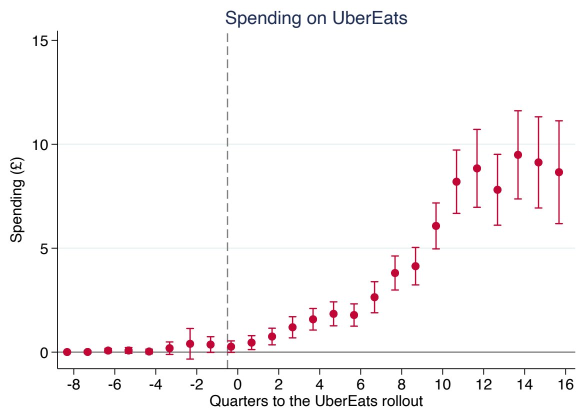


(c) Combined Platform Rollout Analysis

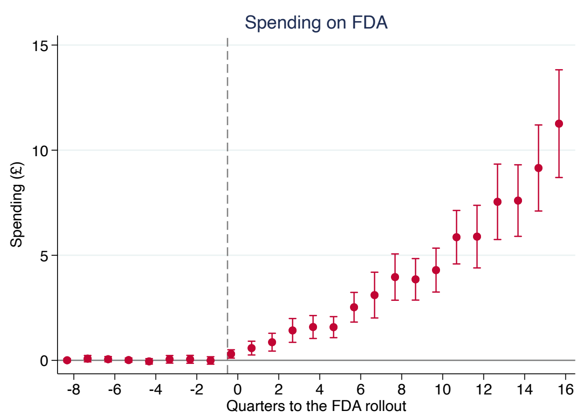
Figure A15. *Notes:* Panel (a) presents the event study results for the individual spending on Deliveroo following its rollout, panel (b) details the results for UberEats, and panel (c) shows the combined spending for Deliveroo and UberEats, based on the earliest rollout date of either platform. The outcome variable analyzed is expenditure, measured using the Fable dataset.



(a) Deliveroo Rollout Analysis



(b) UberEats Rollout Analysis



(c) Combined Platform Rollout Analysis

Figure A16. *Notes:* Panel (a) presents the event study results for the individual spending on Deliveroo following its rollout, panel (b) details the results for UberEats, and panel (c) shows the combined spending for Deliveroo and UberEats, based on the earliest rollout date of either platform. The outcome variable analyzed is expenditure, measured using Kantar's Worldpanel Take Home Purchase Panel for years 2017 to 2023.

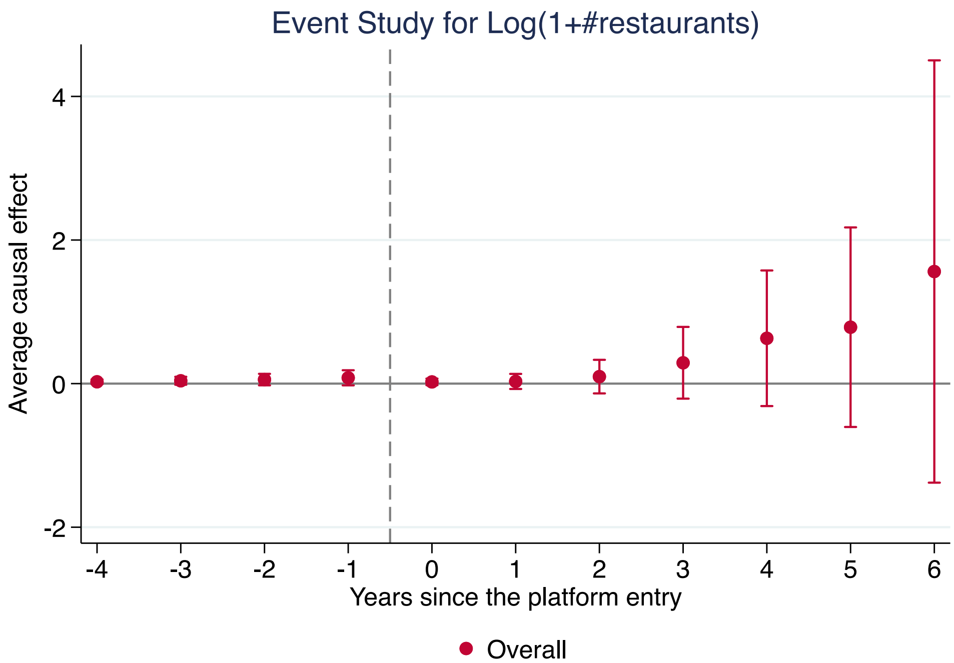
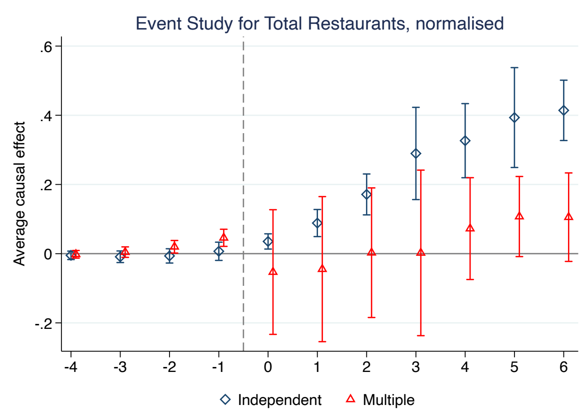


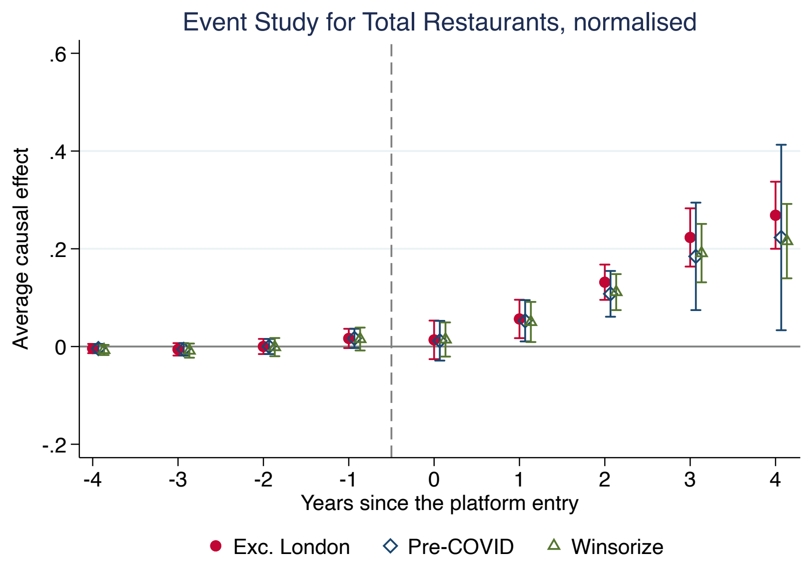
Figure A17. *Notes:* The figure shows the impact of food delivery applications on the log transformation of the number of restaurants for different cuisine types, indicating changes in the number of establishments across various culinary categories. Cuisine types are categorized as outlined in Table A8. Data is sourced from the Local Data Company (LDC).

(a) Overall

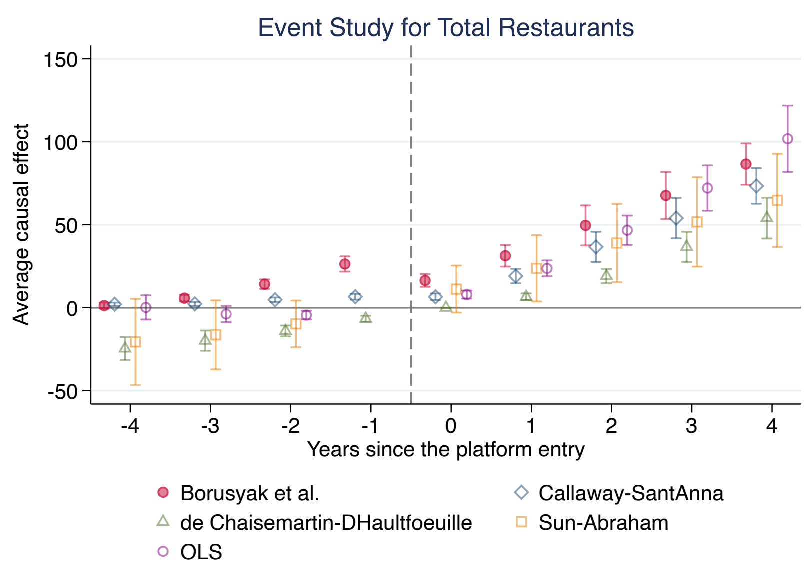


(b) Independent vs Multiple

Figure A18. *Notes:* Panel (a) presents the average causal effect of food delivery application rollout on the total number of restaurants over time as a percentage of the counterfactual outcome, absent food delivery application (i.e., $P_j \equiv \hat{\beta}_j / E[\hat{y}_{st} | t = E_s + j]$ as defined in Section 1.4.1). Panel (b) shows the average causal effect on the number of independent versus multiple establishment restaurants normalized in the same way. The x-axis represents the years since the platform rollout, and the y-axis shows the average causal effect. Data is sourced from the Local Data Company (LDC).



(a) Robustness Checks



(b) Different Estimators

Figure A19. *Notes:* The top panel displays robustness checks where we first exclude local authorities associated with London, then remove the COVID-19 years, and finally winsorize the data at the 5th and 95th percentiles. The bottom panel presents the main analysis using different Difference-in-Differences estimators. Specifically, the Sun and Abraham (2021) estimator is calculated using last-treated units as the control group, while the Callaway and Sant'Anna (2021) estimator uses not-yet-treated units as the control group. The data is sourced from the Local Data Company, and the outcome variable is the number of restaurants in each local authority.

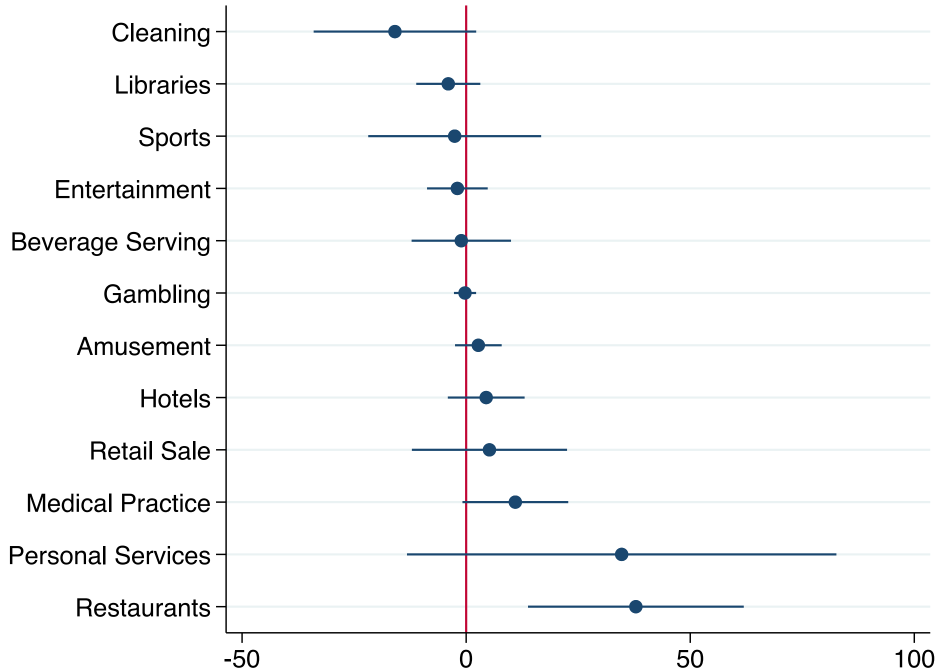


Figure A20. *Notes:* This figure displays the estimated effect of food apps rollout on the number of businesses in various placebo industries. Each coefficient represents the result of a separate regression, where the outcome variable is the number of businesses in a local authority, controlling for local authority and year-fixed effects, along with local economic indicators and population interacted by time. Data are sourced from the UK Business Counts from Nomis, which is an extract from the Inter-Departmental Business Register (IDBR), for the years 2010-2023.

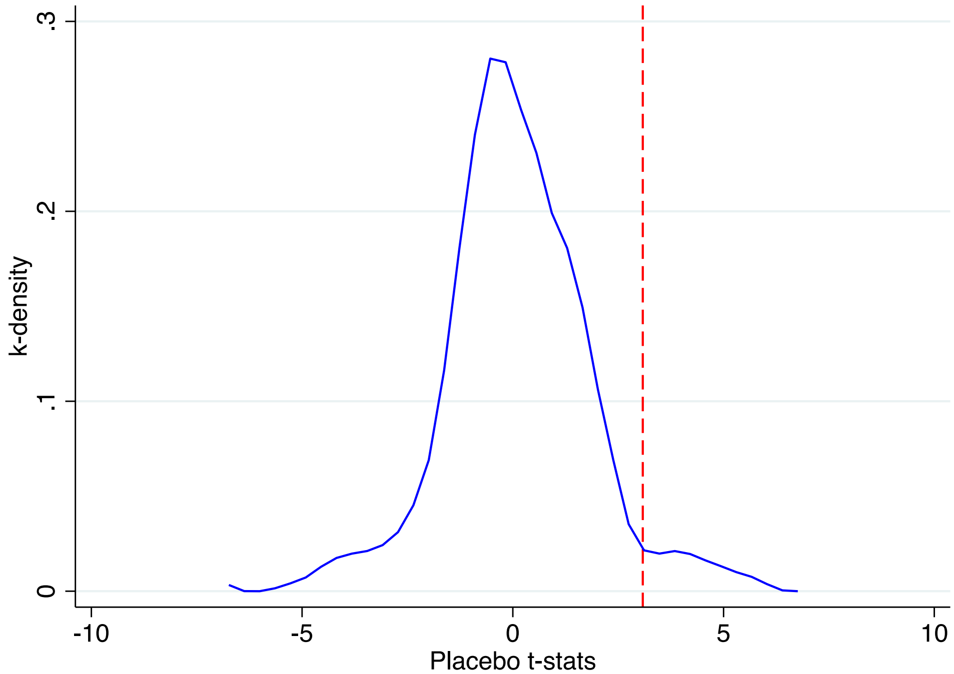


Figure A21. *Notes:* This figure presents the kernel density function of t-statistics for the effect of food apps rollout on the number of businesses across all three-digit SIC 2007 industries. The vertical line indicates the true point estimate for the restaurant industry. Data are sourced from the UK Business Counts from Nomis, which is an extract from the Inter-Departmental Business Register (IDBR), for the years 2010-2023.

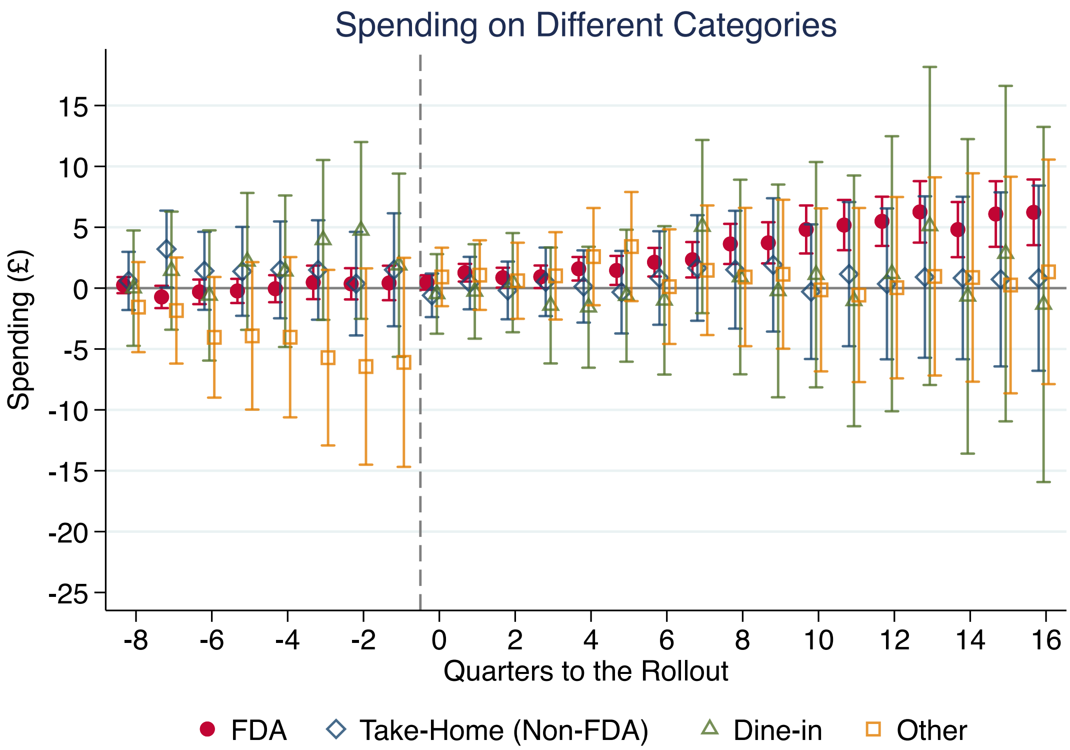
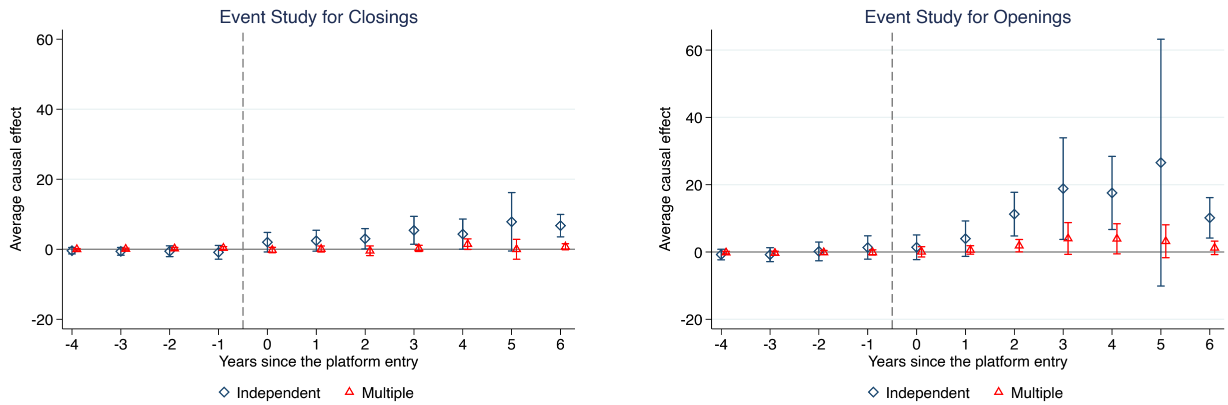
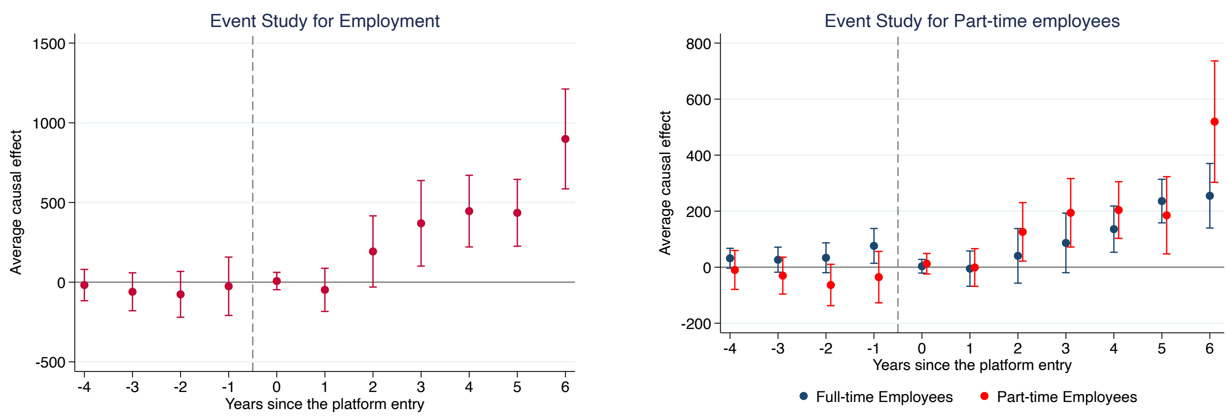


Figure A22. *Notes:* This graph shows the impact of the rollout of food delivery applications on different spending categories. Data is from Kantar's Worldpanel Out of Home Panel for the years 2017 to 2023.



(a) Closings **(b)** Openings

Figure A23. *Notes:* Panel (a) presents the average causal effect of food delivery application rollout on restaurant closings, distinguished by independent and multiple establishment types. Panel (b) shows the average causal effect on restaurant openings, also categorized by independent and multiple establishment types. The x-axis represents the years since the platform rollout, and the y-axis shows the average causal effect. Data is sourced from the Local Data Company and covers the period from 2010 to 2020.



(a) Overall Employment **(b)** Part-time vs Full-time Employment

Figure A24. *Notes:* Panel (a) presents the average causal effect of food delivery application rollout on overall employment levels in Local Authority Districts (LADs). Panel (b) shows the average causal effect on part-time versus full-time employment within the same districts. The x-axis represents the years since the platform rollout, and the y-axis shows the average causal effect. Data is sourced from the Business Register and Employment Survey (BRES) covering the period from 2015 to 2023. Full-time employees work more than 30 hours per week, while part-time employees work 30 hours or less per week. Employment includes employees plus working owners, covering self-employed workers registered for VAT or PAYE but excluding those not registered, HM Forces, and Government Supported trainees.

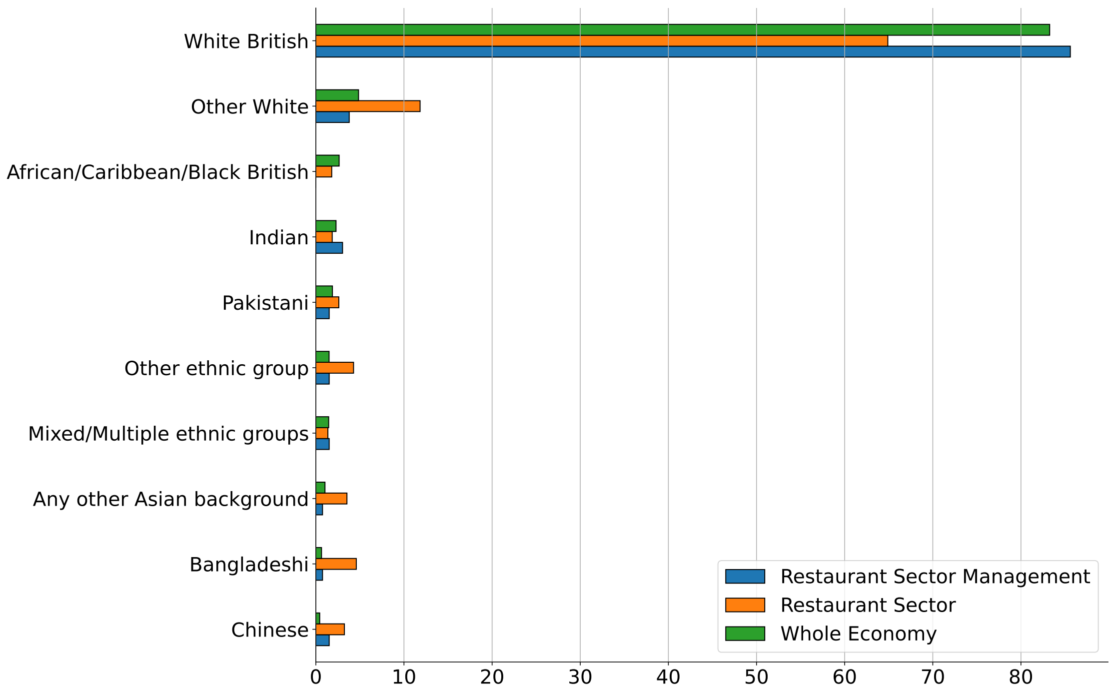
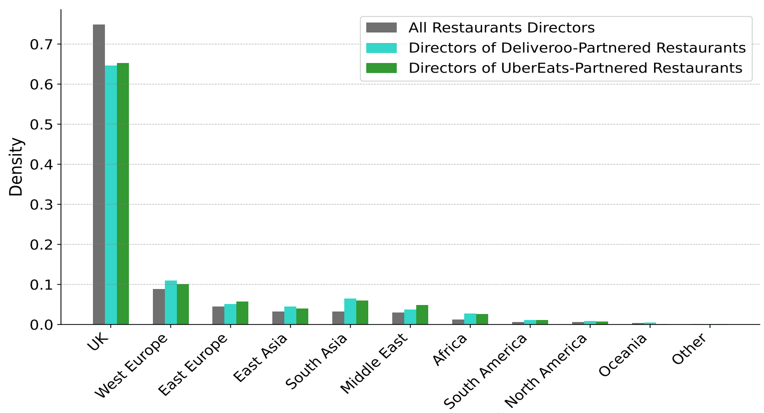
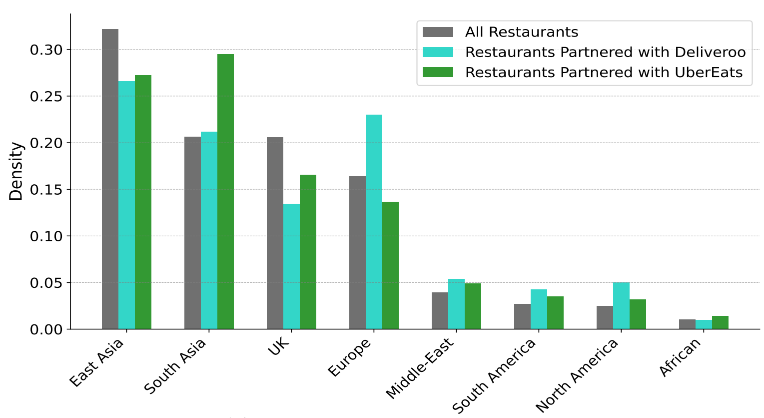


Figure A25. *Notes:* This graph shows the distribution of ethnic groups across the overall economy, the restaurant sector, and managerial positions within the restaurant sector. The data are drawn from the UK Labour Force Survey (2013Q1–2015Q4). The restaurant industry corresponds to SIC code 561, and managerial positions are based on the SOC category ”Higher managerial and professional.” ”White Irish” are grouped under ”Other White.”

(a) Directors' Ethnicity



(b) Directors' Nationality



(c) Restaurant Cuisine Type

Figure A26. *Notes:* Panel (a) shows the distribution of restaurant directors by ethnic background across three categories: all restaurant directors, Deliveroo-partnered, and UberEats-partnered restaurant directors, where directors' backgrounds are inferred from their names. Panel (b) displays the distribution by nationality across the same categories, with nationalities classified as in Table A7. Panel (c) presents the distribution of cuisine types across three categories: all restaurants in the LDC dataset, LDC restaurants matched with Deliveroo listings, and LDC restaurants matched with UberEats listings. The matching process, detailed in Section 1.1, utilizes fuzzy matching algorithms based on restaurant names. It focuses on geographic cuisine and excludes generic restaurants. Data for Panels (a) and (b) come from Companies House and Data for Panel (c) come from LDC.

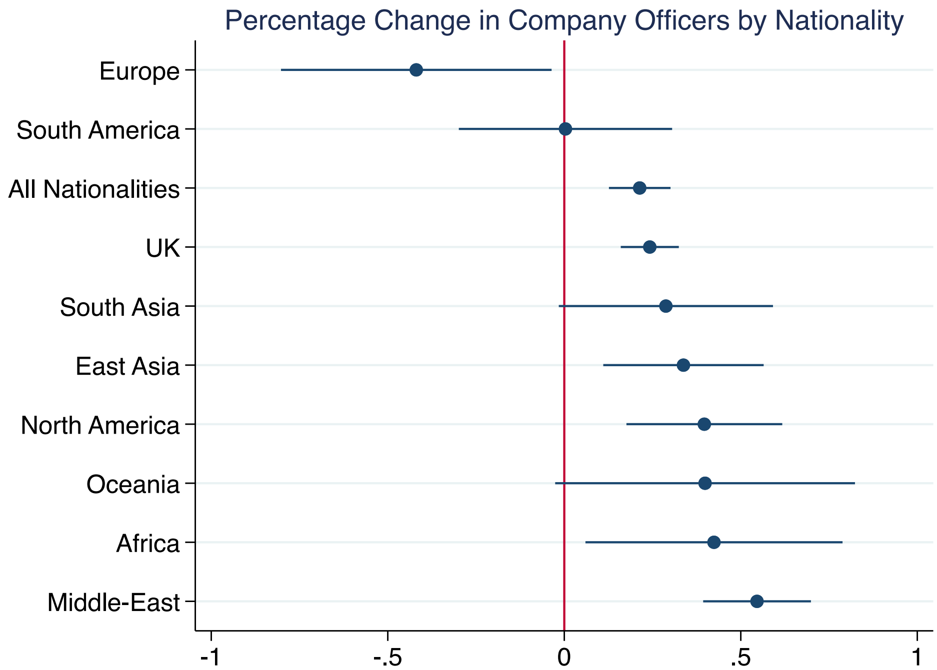
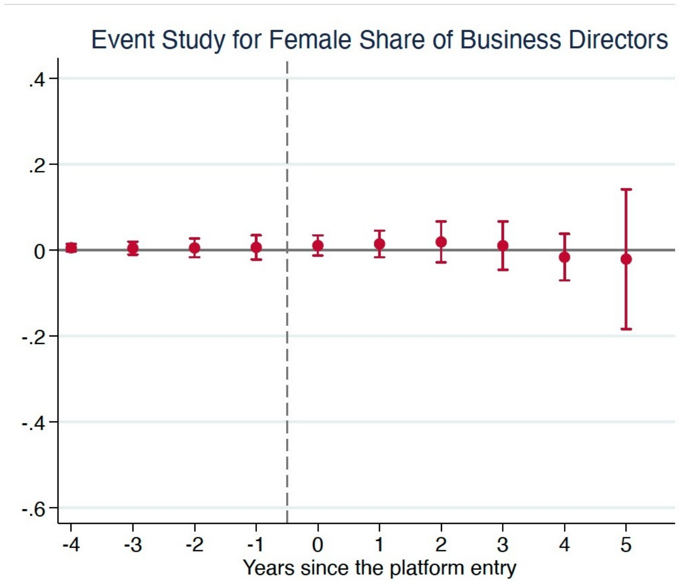
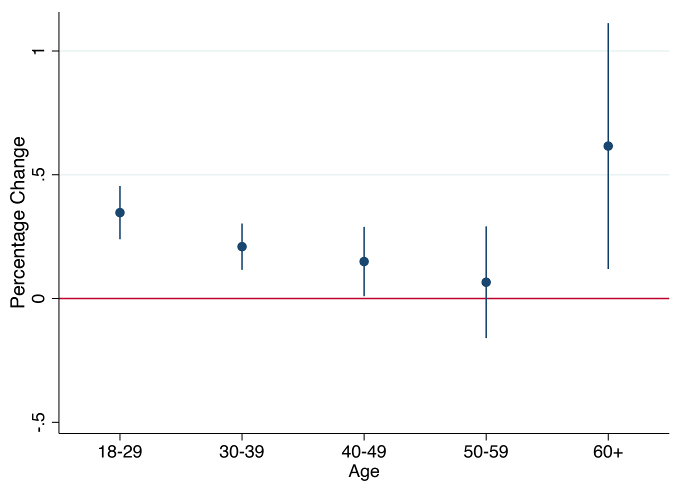


Figure A27. *Notes:* The figure shows the impact of the food apps on different entrepreneur nationalities, reported as the percentage changes by computing $\Delta\hat{y}_m = \hat{\beta_m}/E(\hat{y_m}|D_{it} = 1)$, where $E(\hat{y_m}|D_{it} = 1)$ is the average predicted number of entrepreneurs from nationality m after the rollout of the platform when omitting the contribution of the treatment variable for the presence of the platform. The analysis controls for location and year-fixed effects, as well as local economic indicators and population interacted by time. Data is sourced from Companies House.



(a) Impact on Female Share



(b) Impact on Age

Figure A28. *Notes:* Panel (a) depicts the impact of food apps rollout on the share of female entrepreneurs, estimated using an event study design. Panel (b) illustrates the impact on different age groups as a percentage change, where each coefficient represents the average effect of all lags in the event study for comparability. Data is sourced from Companies House.

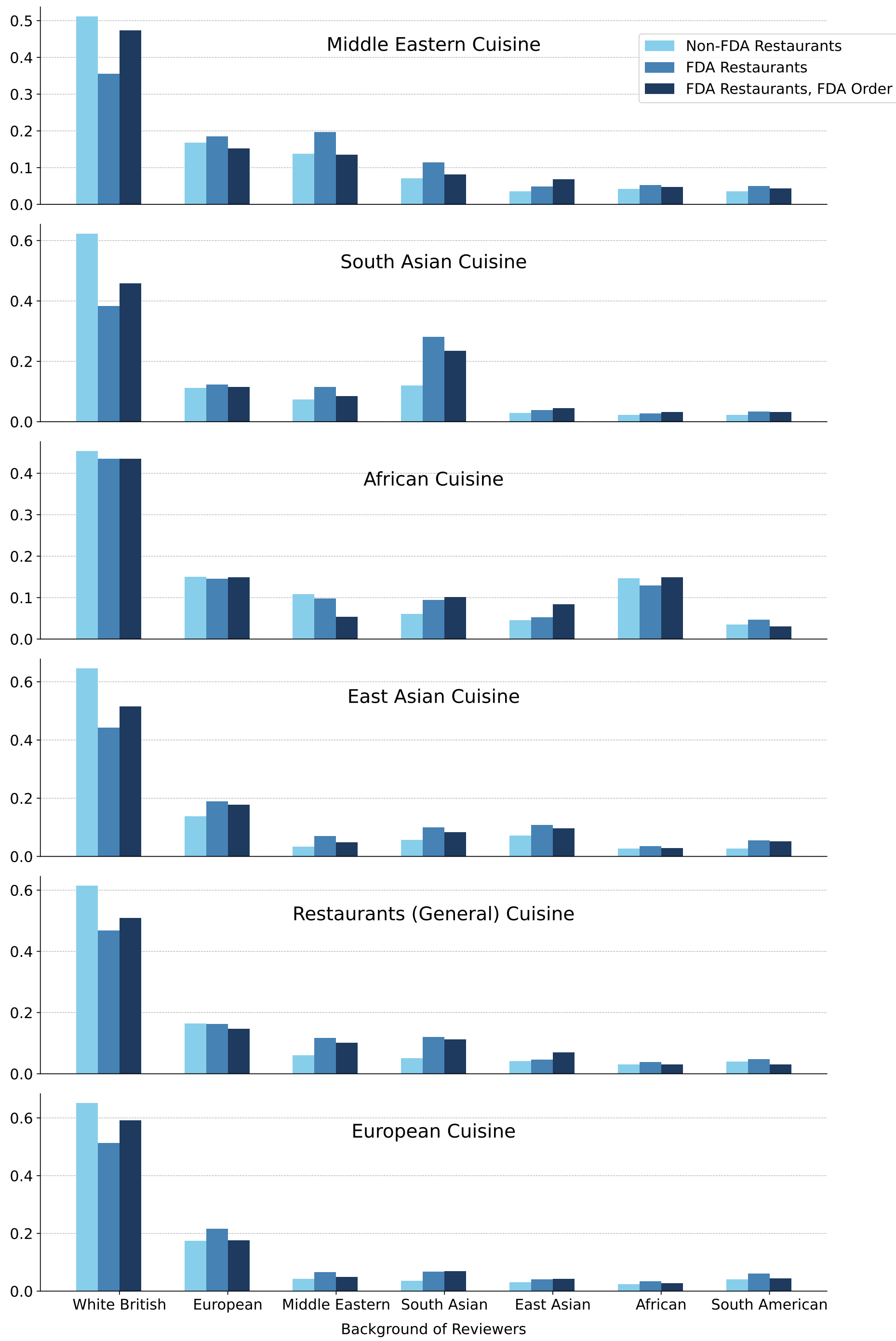


Figure A29. *Notes:* The figure illustrates the distribution of ethnicities of reviewers for different cuisines across three categories: non-partnered restaurants, app-partnered restaurants, and a subset of app-partnered restaurant customers confirmed to have placed orders through Food Apps. The data is based on restaurants on Google Maps. Cuisine types are categorized as outlined in Table A8. Ethnicities were inferred using a predictive algorithm based on first and last names.

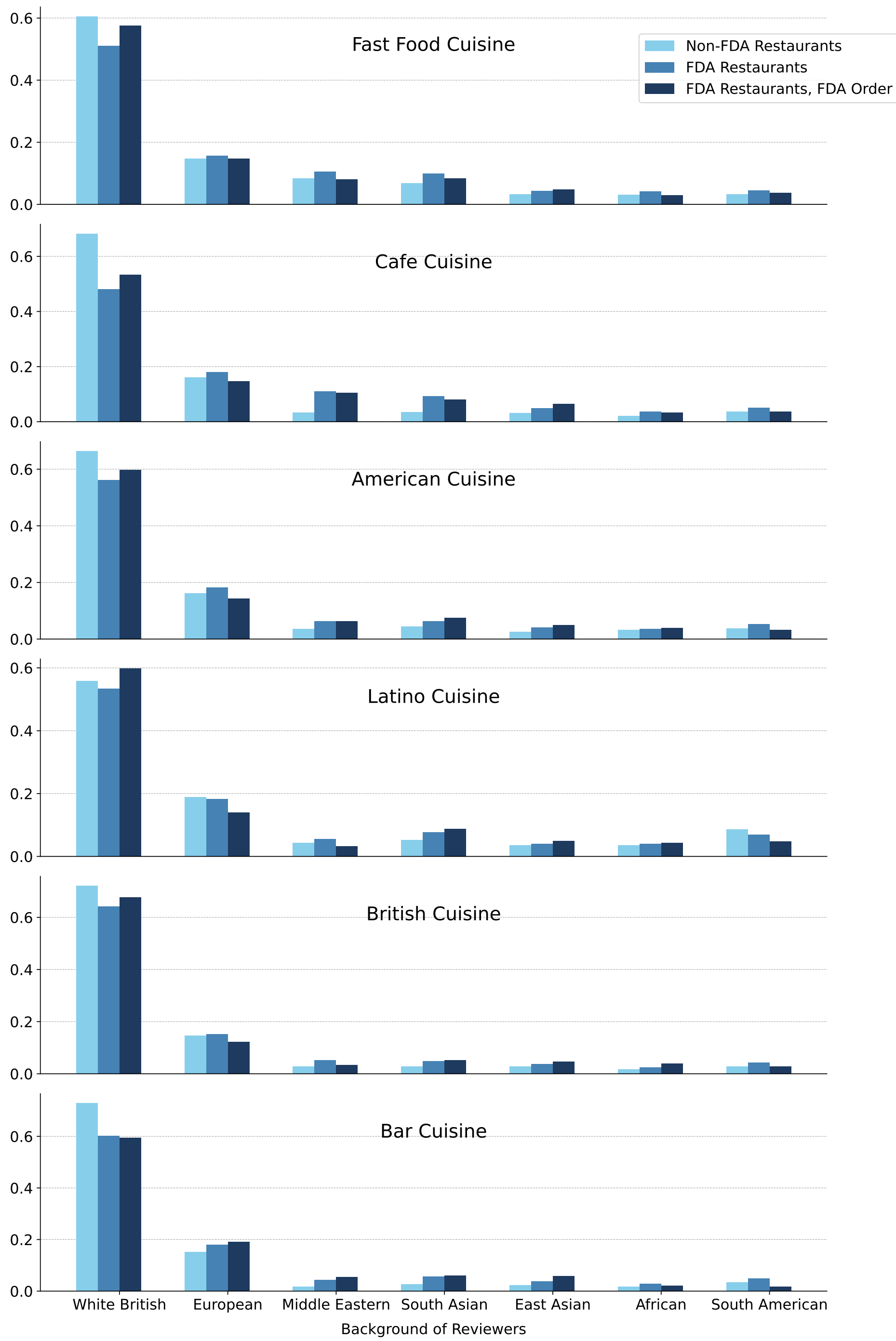


Figure A30. *Notes:* The figure illustrates the continuation of Figure A29, which shows the distribution of ethnicities of reviewers for additional cuisines. The data is sourced from app-partnered restaurants.

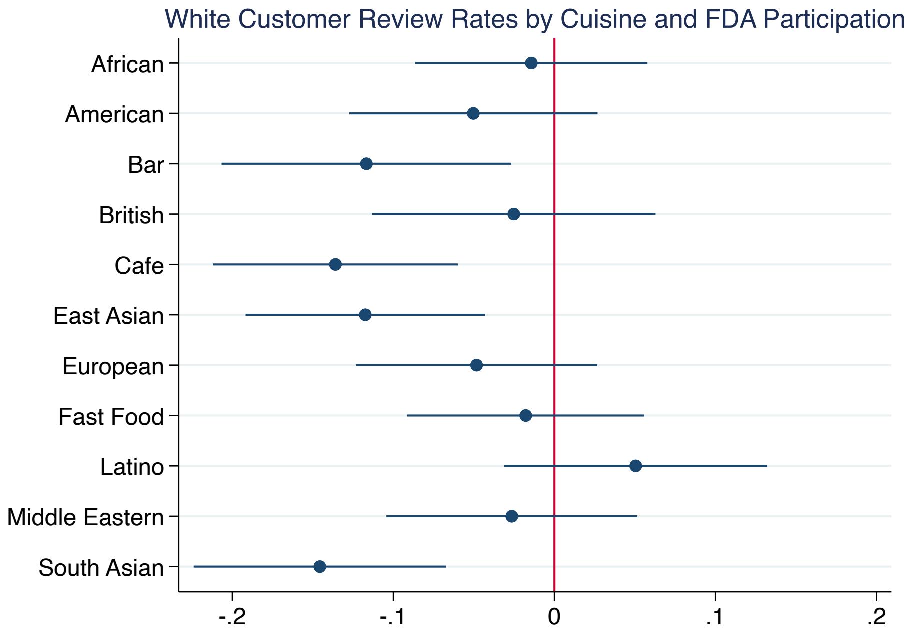
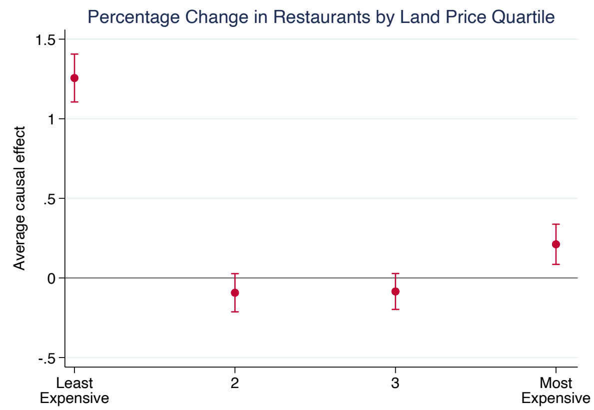
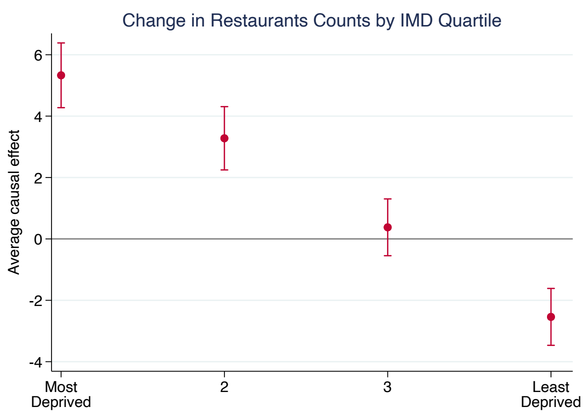


Figure A31. *Notes:* This figure reports the coefficient estimates from equation 3, which capture whether the share of white reviewers (i.e., the probability that a reviewer is white) is higher for platform orders compared to orders at non-platform restaurants, across various cuisine types. The background of the reviewer and the cuisine type is inferred from Google Maps.

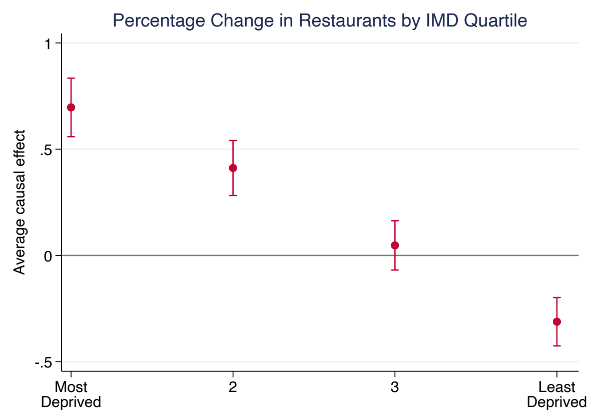
(a) Total Restaurants



(b) Total Restaurant, normalized

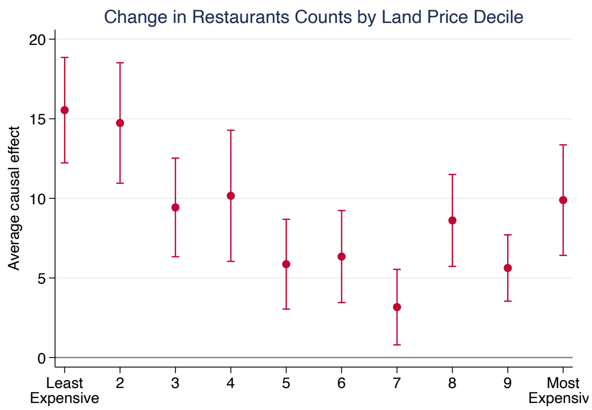


(c) Total Restaurants



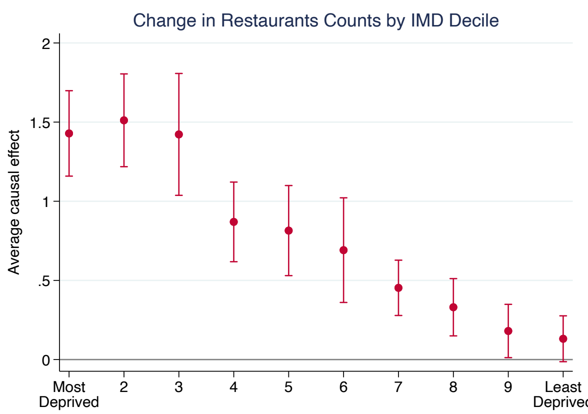
(d) Total Restaurant, normalized

Figure A32. *Notes:* Panel (a) presents the average causal effect of food delivery application rollout on the count of the total number of restaurants, segmented by price quartile within the postal district. Panel (b) presents the average causal effect of food delivery application rollout on the total number of restaurants segmented by price quartile within postal district as a percentage of the counterfactual outcome, absent food delivery application (i.e., $P_j \equiv \hat{\beta}_j / \mathbb{E}[\hat{y}_{st} | s \in$ decile j of IMD] as defined in Section 1.4.1). Panels (c) and (d) show the same thing for IMD deciles. The analysis controls for postal district and year-fixed effects, as well as local economic indicators and population interacted by time. Data is sourced from the Local Data Company and covers the period from 2010 to 2023.

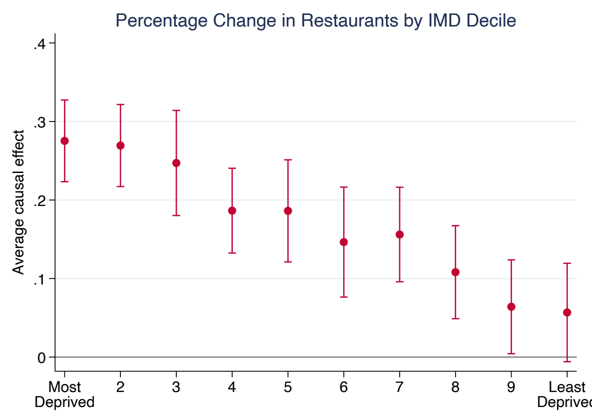


(a) Postal District Price Decile

(b) Postal District Price Decile, normalized



(c) Postal District IMD



(d) Postal District IMD, normalized

Figure A33. *Notes:* Panel (a) presents the average causal effect of food delivery application rollout on the count of the total number of restaurants, segmented by postal district physical space price deciles. Panel (b) shows in percentage terms of the counterfactual outcome, absent food delivery application (i.e., $P_j \equiv \hat{\beta}_j / \mathbb{E}[\hat{y}_{st}|s \in \text{decile } j \text{ of IMD}]$ as defined in Section 1.4.1). Panel (c) presents the average causal effect of food delivery application rollout on the total number of restaurants segmented by postal district IMD deciles. Panel (d) shows it as a percentage change of the counterfactual outcome. The analysis controls for postal district and year-fixed effects, as well as local economic indicators and population interacted by time. Data is sourced from the Local Data Company and covers the period from 2010 to 2023.

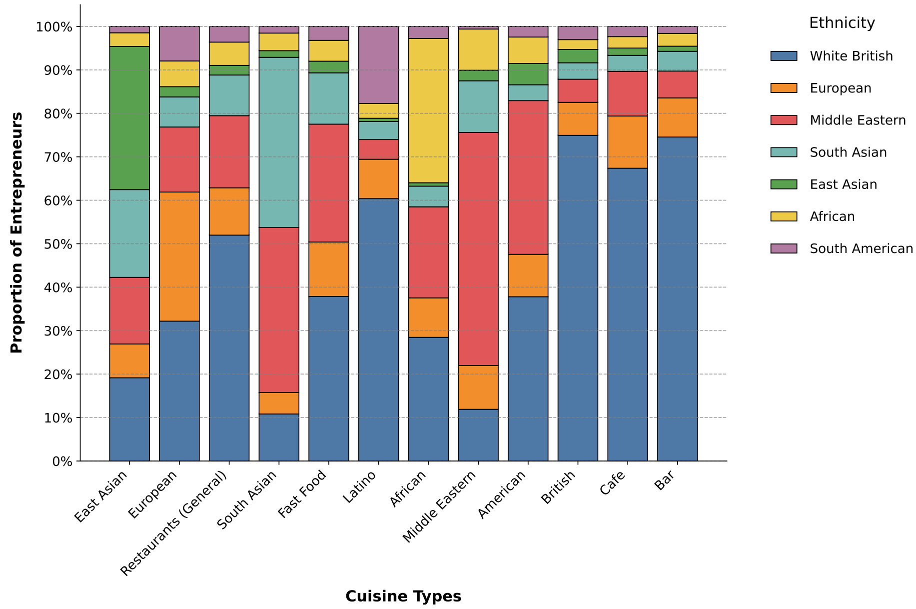


Figure A34. *Notes:* The figure displays the ethnic composition of entrepreneurs across various cuisine types. Cuisine classifications are sourced from Google Maps data, which are then matched with Companies House records containing entrepreneurs' names. As detailed in the text, ethnicity is inferred based on name analysis to estimate the representation of different ethnic backgrounds within each cuisine category..

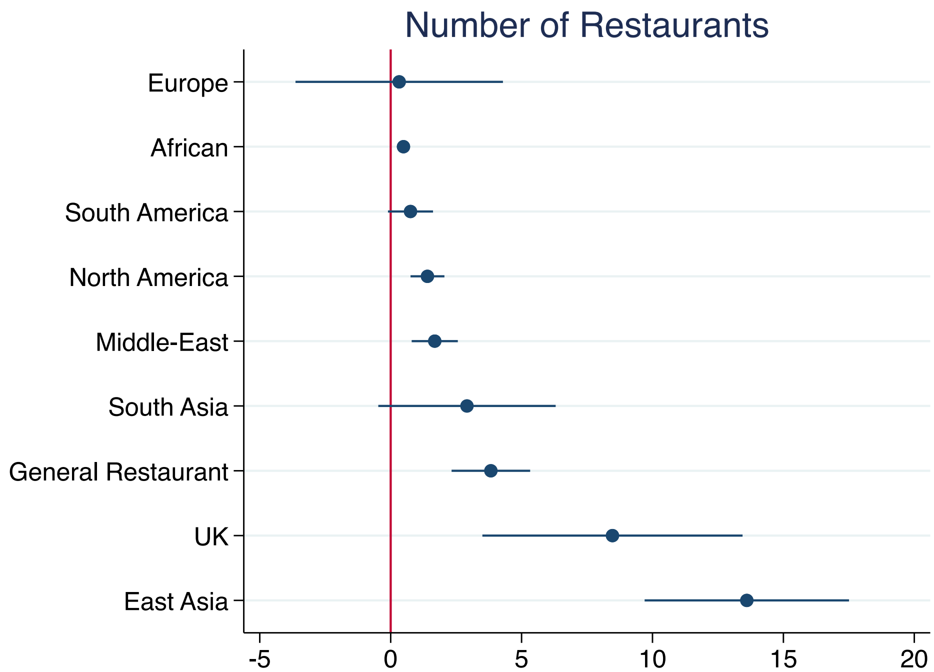


Figure A35. *Notes:* The figure shows the impact of food delivery applications on the number of restaurants for different cuisine types, indicating changes in the number of establishments across various culinary categories. Cuisine types are categorized as outlined in Table A8. Data is sourced from the Local Data Company (LDC).

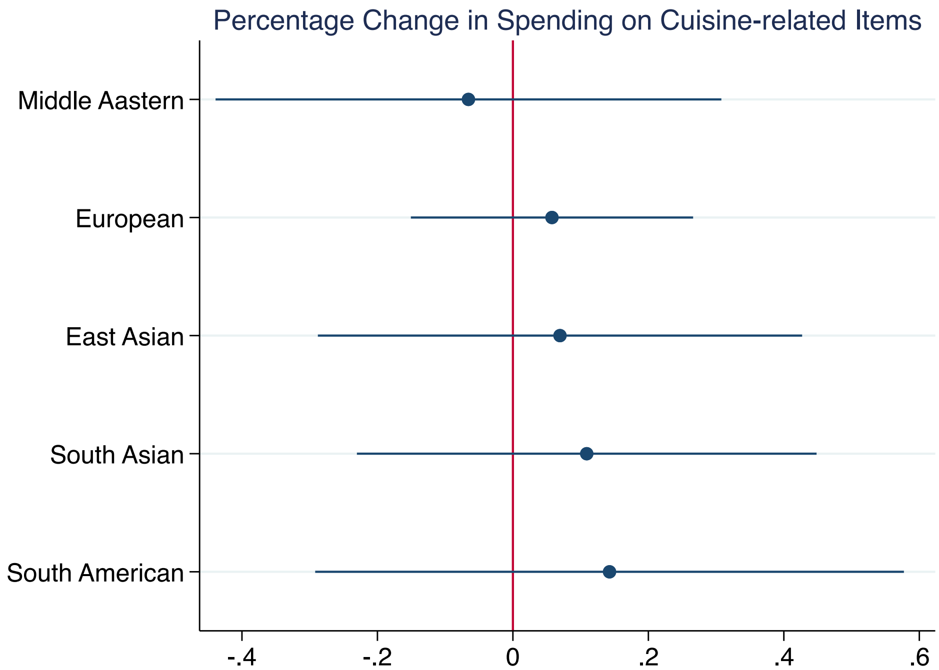


Figure A36. *Notes:* This graph displays the percentage impact of food delivery app rollouts on consumer spending from grocery stores for items representative of specific cuisines, used as placebo tests. The data, sourced from Kantar's Worldpanel Out of Home Panel, and for years 2017 to 2023, includes items like pizza, pasta, and sauerkraut for European cuisine; curry, samosa, and biryani for South Asian cuisine; burritos, nachos, and tapas for South American cuisine; falafel, hummus, and shawarma for Middle Eastern cuisine; sushi, miso, and tofu for East Asian cuisine; and cornbread, buffalo sauce, and clam chowder for North American cuisine.

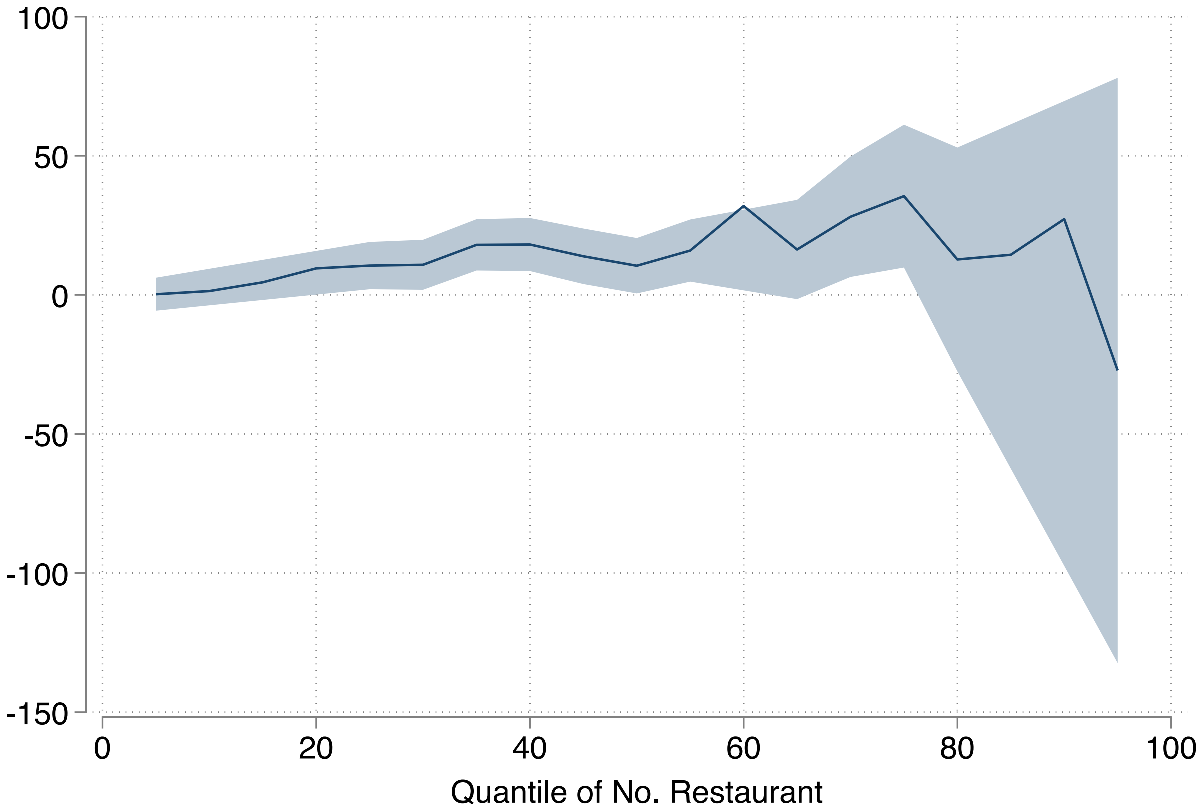


Figure A37. *Notes:* This figure graphs Quantile Treatment Effect (QTE) estimates from the RIF-DiD estimator, including a 90% confidence interval. The outcome variable is the number of restaurants in each local authority and all specifications include postal district and year-fixed effects, as well as local economic indicators and population interacted by time. Data is sourced from the Local Data Company and covers the period from 2010 to 2023.

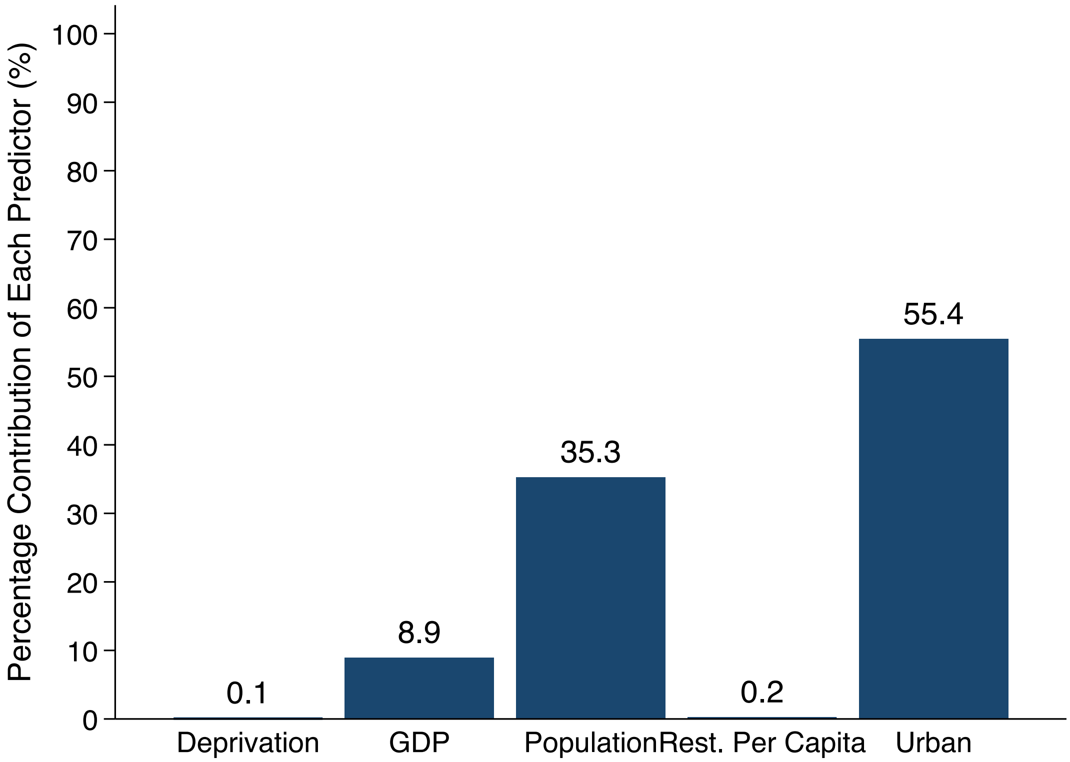


Figure A38. *Notes:* This graph shows the percentage contribution of each predictor to the R-squared value of the regression model assessing the impact of various factors on the rollout dates of food delivery applications in different postal districts. Predictors were selected using the Best Subsets Selection (BSS) method. The Shorrocks-Shapley decomposition method was used to determine the relative importance of each predictor.

A1.8 Extra Tables

Table A1. Summary of Data Sources

Variable	Fable	Kantar's Worldpanel	ONS
Income (£)			
Mean	4,380	3,236	3,083
25th Percentile	2,212	2,083	1,667
50th Percentile	3,184	2,917	2,500
75th Percentile	4,811	4,583	3,667
Age			
20-39	0.53	0.41	0.33
40-59	0.38	0.42	0.34
60+	0.09	0.17	0.33
Female Share	0.54	0.61	0.52
FDA Consumption (£)			
Mean	139.48	30.85	
Mean (conditional on using)	294.90	113.36	
50th Percentile	0.00	0.00	
75th Percentile	92.52	10.98	
90th Percentile	344.56	69.98	
Share of Restaurant Spending	0.22	0.07	
Proportion of FDA Users	0.51	0.32	

Notes: This table compares key variables across the Fable (2021-2022), Kantar (2022-2023), and ONS datasets. Fable income was inferred from likely income transactions, excluding refunds and those under £250, and restricted to individuals with at least 5 months of consistent inflows. ONS income data is sourced from the ONS Average Household Income (UK: financial year 2020), while Kantar's Worldpanel reports income in bands, with values mapped to the midpoint of these bands. Age distribution is divided into three groups (20-39, 40-59, and 60+) for comparison across datasets. The gender share reflects the proportion of females, excluding unknown entries in Fable but included in Kantar's Worldpanel and ONS. Food App consumption data, representing household expenditure on platforms like Deliveroo, Uber Eats, and Just Eat, is shown for both Fable and Kantar's Worldpanel at the 50th, 75th, and 90th percentiles. Age and gender data for ONS are sourced from Population Estimates by the Office for National Statistics, National Records of Scotland, and the Northern Ireland Statistics and Research Agency.

Table A2. Restaurants on Platform vs Non-Platform Restaurants

	Q1	Q2	Q3	Q4
Price Level	1.14	0.92	0.67	0.53
Average Reveiw	1.42	1.17	0.74	0.53
Number of Reviews	0.41	0.67	1.08	1.99
Opening Year	0.95	0.86	1.02	1.24
Number of Nearby Sales	0.37	0.60	1.07	2.13
Number of Nearby Properties	0.38	0.59	1.09	2.12
Nearby Property Price	0.38	0.57	1.11	2.11
Within-District Property Price	1.11	1.09	0.96	0.86

Notes: This table conducts a comparative analysis between restaurants listed on Deliveroo and those not affiliated with the platform. Quartiles (Q1 through Q4) are calculated based on the distributions within the complete dataset of restaurants for each measure. Each quartile in the table represents the ratio of Deliveroo to non-Deliveroo restaurants, calculated by comparing the proportion of Deliveroo restaurants within each quartile to the proportion of non-Deliveroo restaurants in the same quartile. This analysis aims to highlight potential differences in the geographical and economic landscapes between Deliveroo-participating restaurants and the wider restaurant sector.

Table A3. Summary of UK Business Counts

Employment Sizeband	Company	Private Non-Company	Non-Private
	Panel A: Local Units		
Total	95,200	22,465	540
Micro (0 to 9)	64,655	19,775	400
Small (10 to 49)	27,790	2,660	130
Medium-sized (50 to 249)	2,690	30	0
Large (250+)	65	0	0
	Panel B: Enterprises		
Total	79,155	22,215	380
Micro (0 to 9)	59,935	19,540	280
Small (10 to 49)	17,635	2,645	90
Medium-sized (50 to 249)	1,215	30	5
Large (250+)	365	0	0

Notes: The data is derived from 'UK Business: Activity, Size and Location', utilizing an extract from the Inter-Departmental Business Register (IDBR) on businesses with a restaurant code that were live at a reference date in March 2023. An 'enterprise' refers to the entire business, encompassing all individual sites or workplaces. It is defined as the smallest aggregation of legal units (usually based on VAT and/or PAYE records) that possesses a degree of autonomy within an enterprise group. A 'local unit' represents an individual site (e.g., a factory or shop) linked to an enterprise, also known as a workplace. In this context, 'Private Non-Company' includes Partnerships and Sole Proprietorships, while 'Non-Private' encompasses Non-Profit Bodies or Mutual Associations, Public Corporations, and entities under Central Government or Local Authority ownership.

Table A4. Summary of Data Sources

Dataset	Source	#Restaurants	Time Period	Information
Deliveroo	Scraping	50,000	2021-2024	Name, cuisine, postcode
UberEats	Scraping	63,000	2021-2024	Name, cuisine, postcode
Company House	Official (Scraped)	30,000	2010-2024	Name, age, and nationality of directors, registered address postcode
Local Data Company	Proprietary Data	80,000	2010-2024	Name, cuisine, postcode, entry and exit dates
IDBR	Official	70,000	2010-2024	Size, independent vs multiple
Google Maps	API call Scraping	180,507	2024	Name, cuisine, geolocation, average ratings, total number of reviews, names' of reviewers, price indicators

Notes: This table provides a summary of the data sources used in the analysis. The data for Deliveroo and UberEats was webscraped, while the data for Google Maps was partly scraped and partly fetched using API calls. The "No. of Restaurants" column indicates the number of restaurants included in the dataset, with the figures for UberEats and Deliveroo corresponding to the most recent batch of data. The "Time Period" column specifies the coverage period for each dataset. The "Information" column describes the types of data collected from each source.

Table A5. Official Announcements of Rollout

Region	Date	#Link
London Central	16 Jun 2016	uber.com/en-GB/newsroom/ubereats-9
London Zone 2	29 Sep 2016	uber.com/en-GB/newsroom/ubereats-zone2
Manchester	8 Feb 2017	uber.com/en-GB/newsroom/whos-hungry-manchester-introducing-ubereats
Bromley	15 Feb 2017	uber.com/en-GB/newsroom/ubereats-london-coverage-area
Birmingham	9 Mar 2017	uber.com/en-GB/newsroom/whos-hungry-birmingham-introducing-ubereats
Edinburgh	25 Apr 2017	uber.com/en-GB/newsroom/ubereats-edinburgh-is-here
Glasgow	4 May 2017	uber.com/en-GB/newsroom/serving-up-ubereats-in-glasgow
Leeds	11 May 2017	uber.com/en-GB/newsroom/leeds-whos-hungry
Nottingham	12 May 2017	uber.com/en-GB/newsroom/nottingham-whos-hungry
Liverpool	30 May 2017	uber.com/en-GB/newsroom/liverpool-whos-hungry
Southampton	30 May 2017	uber.com/en-GB/newsroom/southampton-whos-hungry
Leicester	1 Jun 2017	uber.com/en-GB/newsroom/leicester-whos-hungry
Sheffield	7 Jun 2017	uber.com/en-GB/newsroom/ubereats-launches-in-sheffield
Cardiff	7 Jun 2017	uber.com/en-GB/newsroom/ubereats-launches-in-cardiff
Swansea	7 Jul 2017	uber.com/en-GB/newsroom/ubereats-launches-in-swansea
Bristol	26 Jul 2017	uber.com/en-GB/newsroom/ubereats-launches-in-bristol
Guildford	26 Jul 2017	uber.com/en-GB/newsroom/ubereats-launches-in-guildford
Bath	26 Jul 2017	uber.com/en-GB/newsroom/ubereats-launches-in-bath
Derby	3 Aug 2017	uber.com/en-GB/newsroom/ubereats-launches-in-derby/
Chelmsford	3 Aug 2017	uber.com/en-GB/newsroom/ubereats-launches-in-chelmsford
Norwich	10 Aug 2017	uber.com/en-GB/newsroom/ubereats-launches-in-norwich
Windsor	10 Aug 2017	uber.com/en-GB/newsroom/ubereats-launches-in-windsor
Portsmouth	10 Aug 2017	uber.com/en-GB/newsroom/ubereats-launches-in-portsmouth

Notes: This table provides a summary of the UberEats rollout dates across various regions in the UK. The "Region" column lists the specific areas where UberEats was launched, and the "Date" column indicates the respective launch dates. The "Link" column contains shortened URLs to the official announcements on the Uber Newsroom website.

Table A6. Productivity of Minority-own and Platform-affiliated and Other Restaurants

	Google Average Review		
	(1)	(2)	(3)
FDA Restaurant	-0.10		-0.13
	(0.01)		(0.01)
Minority-run		-0.14	-0.16
		(0.01)	(0.01)
Minority $\times$ FDA			0.11
			(0.02)
Mean of dep. variable	4.46	4.46	4.46
Observations	8084	8084	8084

Notes: This table compares productivity levels, measured by the average Google review for different restaurants. The outcome variable is the Google Average Review, extracted from Google Maps in Q1 2024. These listings are then matched with Company House, Deliveroo, and UberEats listings. The matching process uses fuzzy algorithms based on restaurant names and postcodes, and only observations with a high likelihood of a successful match are retained. Minority-owned is a binary variable equal to one if the most common background of the restaurant's directors is inferred to be "African," "Muslim," "East Asian," or "South American." Food App equals one if the restaurant is also listed on either Deliveroo or UberEats.

Table A7. Classification of Nationalities

Group	**Countries**
UK	United Kingdom, England, Wales, Scotland, Northern Ireland
North America	United States, Canada
Europe	Austria, Belgium, Cyprus, Denmark, Finland, France, Germany, Greece, Guernsey, Hungary, Iceland, Ireland, Italy, Jersey, Luxembourg, Malta, Monaco, Netherlands, Norway, Portugal, San Marino, Spain, Sweden, Switzerland, Ireland, Albania, Armenia, Azerbaijan, Belarus, Bosnia, Bulgaria, Croatia, Czech Republic, Estonia, Georgia, Kosovo, Latvia, Lithuania, Macedonia, Moldova, Montenegro, Poland, Romania, Russia, Serbia, Slovakia, Slovenia, Ukraine
Middle East	Iran, Iraq, Israel, Jordan, Kuwait, Lebanon, Oman, Palestine, Qatar, Saudi Arabia, Syria, United Arab Emirates, Yemen, Turkey, Afghanistan, Armenia, Bahrain,
South Asia	Bangladesh, Bhutan, India, Nepal, Pakistan, Sri Lanka
East Asia	Brunei, Burma, Cambodia, China, Indonesia, Japan, Kazakhstan, Korea, Laos, Macau, Malaysia, Maldives, Mongolia, Myanmar, North Korea, Philippines, Singapore, South Korea, Taiwan, Thailand, Turkmenistan, Uzbekistan, Vietnam, Kyrgyzstan, East Timor
Africa	Algeria, Angola, Botswana, Burkina Faso, Burundi, Cameroon, Congo, Djibouti, Egypt, Eritrea, Ethiopia, Gabon, Gambia, Ghana, Guinea-Bissau, Guinea, Ivory Coast, Kenya, Lesotho, Liberia, Libya, Madagascar, Malawi, Mali, Mauritania, Morocco, Mozambique, Namibia, Niger, Nigeria, Rwanda, Senegal, Seychelles, Sierra Leone, Somalia, South Africa, South Sudan, Sudan, Swaziland, Tanzania, Togo, Tunisia, Uganda, Zambia, Zimbabwe
Oceania	Australia, Vanuatu, Fiji, Nauru, New Zealand, Papua New Guinea, Samoa, Tonga
South America	Antigua, Argentina, Bahamas, Barbados, Belize, Bermuda, Bolivia, Brazil, Chile, Colombia, Costa Rica, Cuba, Dominican Republic, Ecuador, El Salvador, Grenada, Guatemala, Guyana, Haiti, Honduras, Jamaica, Mexico, Nicaragua, Paraguay, Peru, Saint Kitts and Nevis, Saint Lucia, Trinidad and Tobago, Uruguay, Venezuela, Panama
Other	Stateless, Stateless Refugee

Notes: This table shows how different nationalities, as recorded in the Company House database, are classified into various groups.

Table A8. Classification of Cuisine Types

Category	**Cuisine Types**
UK	Irish, British, Fish & Chip Shops, English, Scottish, Welsh
North America	American
Europe	Austrian, Belgian, French, German, Greek, Hungarian, Italian, Mediterranean, Polish, Portuguese, Russian, Scandinavian, Spanish, Swedish, Swiss, Brasserie, European, Continental, Eastern European, Danish
Middle-East	Lebanese, Iranian, Iraqi, Israeli, Turkish, Middle Eastern, Moroccan, Afghan
South Asia	Indian, Indian Takeaway, Nepalese, Bangladeshi, Pakistani
East Asia	Asian, Chinese, Japanese, Korean, Thai, Vietnamese, Chinese Fast Food, Oriental, Malaysian, Philippine, Indonesian, Mongolian, Tibetan, Burmese, Southwestern
African	African, Sudanese, Mauritian, Egyptian
South America	Argentinian, Brazilian, Colombian, Mexican/Tex Mex, South American, Caribbean, Jamaican, Cuban
Specialty Cuisine	Oceanic, International, Seafood, Vegan, Vegetarian, Kosher
Fast Food	Pizzeria, Fast Food Takeaway, Fast Food Delivery, Pizza Takeaway, Take Away Food Shops, Sandwich Delivery Service
Cafe & Casual Dining	Cafe & Tearoom, Coffee Shops, Juice Bars, Creperie, Internet Cafes
General Restaurant	Restaurant, Bar, Cruises, Other
Culinary Services	Cake Makers, Decorators & Supplies, Caterers

Notes: This table shows the classification of different cuisine types, as recorded in Local Company House or Google Maps, into the broader categories used in our study.

Table A9. Best Subset Selection Results for Platform Rollout Dates

	Food App Rollout Date				
	(1)	(2)	(3)	(4)	(5)
Urban	-52.867	-39.249	-40.297	-36.359	-34.889
	(1.423)	(1.508)	(1.458)	(1.342)	(1.329)
Population 60 older (2001)		404.358	342.134	319.944	298.382
		(17.731)	(17.817)	(15.399)	(15.300)
Share of res. pop. qualification 1 (2001)			298.937	314.579	296.891
			(23.934)	(20.278)	(20.352)
Population				-0.000	-0.000
				(0.000)	(0.000)
GDP					-0.032
					(0.004)
Best Subset				X	
Observations	2307	2088	2088	2021	2012
R-Squared	.375	.491	.526	.654	.665

Notes: This table reports results from OLS regressions. The dependent variable is the rollout date of the earliest platform in months (months since January 1960) for each postal district. Empirical models were selected using BSS. The best subset marked by 'X' indicates the top models selected using BSS on the set of predictors, based on the AIC information criterion. Column 1 shows the best subset across all variables, Column 2 the best subset with two predictors, Column 3 the best subset with three predictors, and so on. Robust standard errors are presented in parentheses.

Appendix to Paper Two

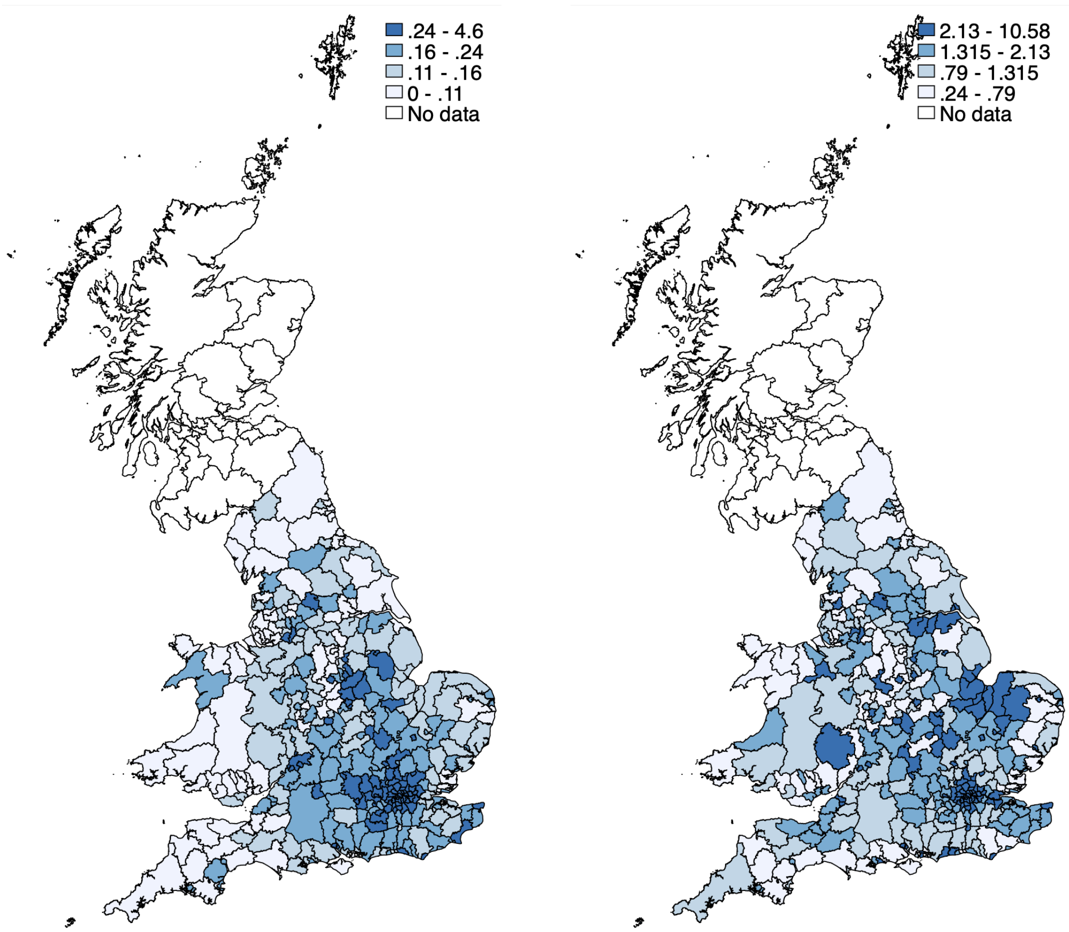
A2.1 Variable Construction

The firm-level data and sales information used in this analysis are obtained from the income statement, cash flow statement, and balance sheet published on the official Codal outlet. I calculate the change in sales, represented as $\Delta Sales_{i,t}/Sales_{i,t-1}$, by determining the difference in sales between two periods and dividing the result by the sales from the previous year. Subsequently, I perform winsorization on this variable at the 1st and 99th percentile.

The capital expenditure measure utilized in this study, denoted as $I_{i,t}/K_{i,t-1}$, is recursively determined using a perpetual inventory approach. This is necessitated due to the financial statements presenting capital values at book value rather than replacement value. This method draws upon established methodologies, such as the one demonstrated by Stein and Stone (2013). The computation begins with the initial observation for each company spell available within the dataset. More specifically, the capital expenditure measure for $t = 2$ is calculated as $\frac{I_{i,2}}{PPE_{i,1}}$. For instances where $t > 2$, the measure is computed as $\frac{I_{i,t}}{K_{i,t-1}}$. For periods $t > 2$, the denominator is recursively determined as $K_{i,t} = \frac{\pi_t^K}{\pi_{t-1}^K}(1 - \delta)K_{i,t-1} + I_{i,t}$, where $I_{i,t}$ represents capital expenditure (CapEx), PPE denotes the net value of property, plant, and equipment, $\frac{\pi_t^K}{\pi_{t-1}^K}$ is the ratio of the current period's Producer Price Index to that of the previous period, and δ symbolizes depreciation, which is set at 10%. Additionally, the variable is winsorized at the first and 99th percentiles to minimize the effect of extreme values.

Appendix to Paper Three

A3.1 Extra Graphs



(a) 2001 **(b)** 2010

Figure A1. *Notes:* This map displays the spatial distribution of immigrants from the New Member States (NMS) in 2001 (left panel) and 2010 (right panel) as a share of the total population in England and Wales. The data used for this visualization is derived from the 2001 and 2011 census, which quantifies the resident population in each local authority area according to the country of birth.

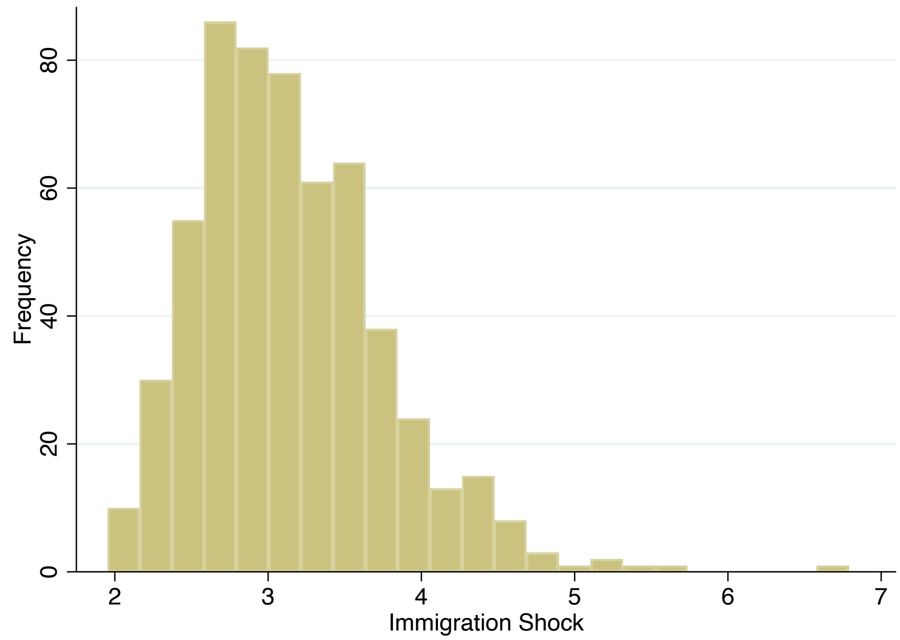
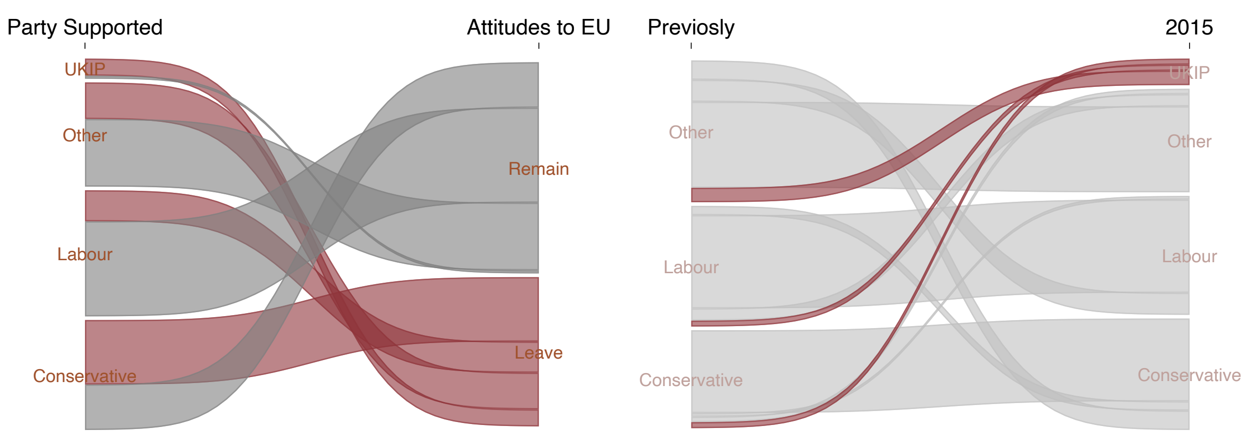


Figure A2. *Notes:* This graph shows the distribution of immigration shock in 2016 across constituencies in the UK. The immigration shock is a measure of the impact of immigration on each constituency, with higher values indicating a greater impact, as defined in equation 12.



(a) Attitude toward EU and Party Support

(b) Party Support transition

Figure A3. *Notes:* This graph shows the transition of voters over time. Panel A is a Sankey diagram showing the supported party of those who prefer either leaving or remaining in the UK, based on their attitudes toward the EU. The attitude to EU variable is constructed using answers to several questions. Panel B displays how respondents moved between parties from 2015. Each respondent is matched to the last party they supported, with UKIP supporters matched to their previous non-UKIP party.

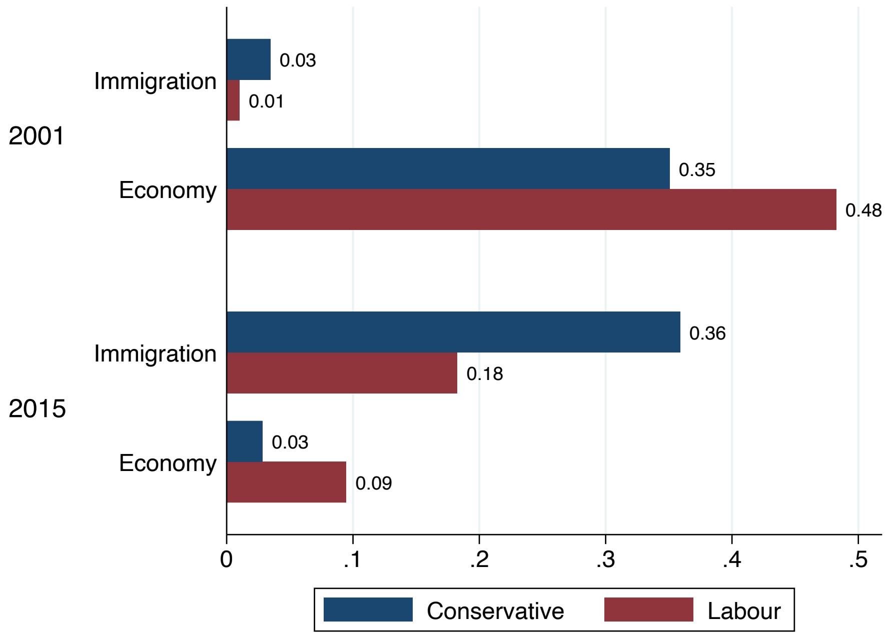


Figure A4. *Notes:* This graph depicts the proportion of individuals identifying either economic factors or immigration as the Most Important Issue (MII), segmented by party support and year. The 2001 data stems from the BES Panel's post-election aggregation, while the 2015 data is sourced from the BES Internet Panel (Wave 8). In this context, 'Immigration' represents the fraction of respondents who consider immigration/asylum the most pressing issue facing the country, whereas 'Economy' aggregates the shares of individuals prioritizing health (NHS), education, or taxation.

A3.2 Extra Tables

Table A1. Shock Distribution

	Over years	In 2016
Mean	.02	.047
Standard deviation	.038	.067
Interquartile range	.026	.048
Effective sample size (1/HHI)	389	24
Largest average exposure	.0068	.11
Number of shocks	1344	84

Notes: This table presents distributional statistics of the shift-share instrument, constructed based on migration from NMS to EU10 countries. Statistics are weighted by average industry exposure shares and are based on employment share at the start of the period. Column 1 includes all shocks over time, while Column 2 only includes shocks in 2016. Effective sample size (inverse renormalized Herfindahl index of exposure weights, as suggested by Borusyak *et al.* (2022)), is also reported.

Table A2. Revised Analysis of Table 3.2 with Alternative Inference Approaches

	UKIP Vote Share Change		
	European 2014-2004 (1)	General 2015-2005 (2)	Local (2012-15)-(2000-3) (3)
Current Immigration Shock	2.045	2.919	3.032
	(0.612)	(0.394)	(0.941)
Alternative Standard Errors:			
Robust	0.587	0.407	0.806
Adao et al (2019)	0.917	0.956	1.085
Wild cluster bootstrap	0.612	0.394	0.941
Estimator	IV	IV	IV
Region FE	Yes	Yes	Yes
Observations	347	573	346
Outcome mean	12.68	12.69	12.91
Adj. R^2	0.0415	0.0404	0.0506
F-statistic	77.92	260.9	75.30

Notes: This table presents a re-estimation of Table 3.2, employing various inference methods in addition to the conventional approach of clustered standard errors, which are denoted in parentheses. It includes robust standard errors, standard errors clustered at the regional level-with adjustments for potential biases arising from a limited number of clusters via the wild-cluster bootstrap method-and adjusted standard errors for shift-share designs as suggested by Adao *et al.* (2019).

Table A3. Effects of Immigration on the Electoral Performance of Labour and Conservative

	European Elections		Local Elections		General Elections	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Labour Party						
Immigration Shock	-3.210	-2.009	-2.982	-2.353	-2.817	-2.694
	(0.638)	(0.487)	(0.920)	(0.709)	(0.554)	(0.442)
Panel B. Conservatives Party						
Immigration Shock	0.060	0.006	1.594	0.354	0.561	0.285
	(0.382)	(0.328)	(0.937)	(0.676)	(0.453)	(0.351)
Method	2SLS	OLS	2SLS	OLS	2SLS	OLS
LA/Constituency FE	Yes	Yes	Yes	Yes	Yes	Yes
Region-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1041	1041	3263	3263	2283	2283

Notes: This table analyzes the effects of immigration on the electoral performance of the Labour and Conservative parties across European, local, and general elections. The analysis is structured into two panels: Panel A focuses on the Labour Party, while Panel B is dedicated to the Conservative Party. For each party, the table presents both Ordinary Least Squares (OLS) and Two-Stage Least Squares (2SLS). The analysis is conducted using data from local authorities and constituencies, excluding Scotland. Standard errors, adjusted for clustering at the local authority or constituency level, are shown in parentheses.

Table A4. Individual-level Pre-trend Analysis

	(1) UK membership of EU a bad thing	(2) UK benefited from being in EU	(3) UK longterm policy wr. EU	(4) EURO currency
OLS Estimates:				
2016 Imm. Shock	0.004	0.027	0.004	0.009
	(0.008)	(0.009)	(0.007)	(0.015)
2SLS Estimates:				
2016 Imm. Shock	0.013	0.042	0.007	0.011
	(0.010)	(0.011)	(0.009)	(0.016)
Observations	19,113	21,585	17,796	13,990

Notes: This table presents the results of a pre-trend analysis examining the relationship of immigration and historical individual attitudes toward various aspects of the UK's relationship with the EU before the Brexit referendum. The analysis uses questions from previous waves of the survey to construct outcome variables related to UK membership, benefits of being in the EU, long-term policy toward the EU, and opinions on the EURO currency. It employs both Ordinary Least Squares (OLS) and Two-Stage Least Squares (2SLS) methods. All regressions include region-wave-time fixed effects and control for individual qualification, age, economic activity statute, income decile and employment sector. Data are from surveys conducted in 1999, 2000, 2001, and 2002. Standard errors, adjusted for clustering at the local authority or constituency level, are shown in parentheses.

Table A5. Effects of Immigration on the Annual Pay Distribution

log(Annual Pay):	(1) Avg	(2) 90th Pct	(3) 75th Pct	(4) Med	(5) 25th Pct	(7) 10th Pct
Panel A. OLS						
Immigration Shock	-0.007	0.041	-0.004	-0.004	-0.002	0.002
	(0.004)	(0.024)	(0.003)	(0.004)	(0.006)	(0.009)
Average effect	-.73%	4.42%	-.43%	-.39%	-.20%	.215%
Standard deviation	.826	4.99	.494	.449	.236	.243
Panel B. 2SLS						
Immigration Shock	-0.006	0.054	0.001	-0.008	-0.011	-0.007
	(0.007)	(0.034)	(0.005)	(0.006)	(0.009)	(0.013)
F-stat	214	212	170	236	191	209
Average effect	-.61%	5.79%	.120%	-.88%	-1.2%	-.71%
Standard deviation	.691	6.54	.135	1.00	1.36	.803
Pre-log mean of DV	2705	3881	3316	2245	1384	7377
LA FE	Yes	Yes	Yes	Yes	Yes	Yes
Region-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Spatial units	348	332	336	345	346	339
Observations	7286	719	5548	7042	6431	4211

Notes: This table presents the estimated impacts of immigration shocks on wage distribution, with the dependent variable being the log of annual wages at the mean and also various percentiles within the earnings distribution of a local authority, as derived from the Annual Survey of Hours and Earnings. Some data points are excluded due to the Office of National Statistics determining insufficient precision in the statistics. The term ”F-stat” refers to the Kleibergen-Paap rk Wald F-statistic, which is used to test for weak instruments. The table presents robust standard errors, which are clustered by local authority, in parentheses.