Essays on Political Economy and the European Union

Alexandru Ionescu

PhD

University of York

Economics

September 2020

Abstract

This thesis is a collection of essays about political economy, in particular to the relationship between the United Kingdom and the European Union.

The second chapter aims to analyse whether and to what extent negative newspaper coverage of topics relating to immigration concerns influenced the percentage of leave votes in the 2016 European Union (EU) membership referendum. Using data at constituency level, I find that negative newspaper coverage had a small, statistically significant and positive effect on the percentage of EU leave votes mainly in constituencies in the bottom quarter of the education qualification distribution.

This work in the third chapter generalises the work of Besley and Prat (2006) analysing political media capture by bribing some or all media outlets in the event of an election. The most crucial item of generalisation which I introduce is that the challenger politician can observe any bribing by the incumbent politician to the media outlets and has an opportunity to make counter-offers. The most important novel finding is that under sufficiently general conditions, there exists one Perfect Bayesian Equilibrium for which total media capture never arises in equilibrium.

The final chapter decomposes differentials in wages and propensities to work a second job of nationals of countries which joined the European Union (EU) in 2004 relative to natives to analyse their integration in the British Labour Market, before and after their accession. This paper finds that at the mean migrants who arrived at least 5 years prior being better integrated in all periods. The Quantile decomposition and finds that joining the European Union improved the labour market outcomes of A10 nationals at the bottom of the wage distribution at the cost of creating a glass ceiling. The logit Oaxaca decomposition results reveal that migrants are significantly more likely to work second jobs after 2004.

Extended Abstract

This thesis is a collection of essays about political economy, in particular to the relationship between the United Kingdom and the European Union.

The second chapter aims to analyse whether and to what extent negative newspaper coverage of topics relating to immigration concerns influenced the percentage of Leave votes in the 2016 European Union (EU) membership referendum. Using data at constituency level, I find that negative newspaper coverage had a small, statistically significant and positive effect on the percentage of EU Leave votes mainly in constituencies in the bottom quarter of the education qualification distribution, where a one unit increase in the coverage variable increased the Leave vote percentage between 0.003 and 0.005 percentage points depending on the analysed type of coverage. The effect decreases the more educated the electorate in the constituency is, and the overall effect is small, positive, and statistically insignificant. Heterogenous results are observed when separating the coverage by publication type and position taken with respect to the then upcoming referendum. A unit increase in the weighted articles variable that tabloid coverage supporting Remain decreased the Leave vote between 0.672 and 0.9 percentage points, while an equivalent increase in the coverage variable for Leave supporting broadsheets increased the Leave vote between 0.1 and 0.3 percentage points. The coverage effects are stronger in the last six months before the referendum. Anti-EU coverage from Leave supporting tabloids also increased the Leave vote by around 0.8 percentage points per one unit increase in the coverage variable.

The work in the third chapter generalises the work of Besley and Prat (2006) analysing political media capture by bribing some or all media outlets in the event of an election. The most crucial item of generalisation which I introduce is that the challenger politician can observe any bribing by the incumbent politician to the media outlets and has an opportunity to make counter-offers. The most important novel finding is that under sufficiently general conditions, there exists one Perfect Bayesian Equilibrium for which total media capture never arises in equilibrium. The results also indicate that political interference in media markets is welfare reducing only if it is a monopsony.

The final chapter decomposes differentials in wages and propensities to work a second job of nationals of countries which joined the European Union (EU) in 2004 relative to natives to analyse their integration in the British Labour Market, before and after their accession. This paper finds that at the mean migrants who arrived at least 5 years prior being better integrated

in all periods. In the 1994-2003 period, no statistically significant log wage gap is observed for migrant men, irrespective of whether they arrived in the UK at least five years prior. However, in the period immediately after accession, this study finds a statistically significant gap of 50%, whereas in the case of migrants present for at least 5 years the gap is statistically insignificant at 4%. In 2009-2013 this gap increases and becomes statistically significant for both migrant groups, indicating wage gaps relative to natives of 49% and 36% in the case of migrants present for less and more than 5 years respectively. At the bottom of the income distribution, the observed wage gap reduced from 59% before accession to 0.14% and 0.12% in the two periods after 2004, out of which 50% and 83% are explained by differences in characteristics. At the top, the gap increases from 10% and statistically insignificant before 2004 to 60% in both post-accession periods, whereas the proportion explained by differences in characteristics decreases from 53% to 40%.The Quantile decomposition finds that joining the European Union improved the labour market outcomes of A10 nationals at the bottom of the wage distribution at the cost of creating a glass ceiling. The logit Oaxaca decomposition results reveal that migrants are significantly more likely to work second jobs after 2004.

Contents

List of Tables

Acknowledgement

I owe a debt of gratitude to my supervisors Professor Cheti Nicoletti, Professor Jo Swaffield and Dr. Anindya Bhattacharya. Their invaluable knowledge and experience, as well as their limitless patience and continued feedback were crucial in creating this research. They have encouraged a stimulating environment which fostered the analytical and organised frame of mind required to complete this thesis. I also wish to thank Professor Thomas Cornellisen and Dr. Makoto Shimoji for their insightful observations and suggesting new avenues of thinking. Without their help, this would not have been possible.

I am also indebted to my fiancée, Jen, whose unwavering support was crucial through this programme. Her unconditional belief in me during the good and bad times of this adventure made this thesis possible and I can't wait to start our next one.

This work was supported by the Economic and Social Research Council [grant number ES/J500215/1.].

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Chapter 1: Introduction

This thesis is composed of three papers in the spheres of political economy and media economics. The first and third papers are concerned with the relationship between the United Kingdom (UK) and the European Union, while the first and second are concerned with media economics.

The first paper is concerned with one of the most significant recent political events, namely the 2016 UK European union membership referendum, colloquially known as the Brexit Referendum. There is already a rich literature on the plausible factors which caused the Brexit vote. Particular attention was paid to economics and social factors. The study of Becker et al. (2017) uses a machine-learning method to determine what covariates best predict the Leave vote and the results underline that underlying socio-economic factors were important in predicting the vote to Leave. On the other hand, the local area's exposure to the European Union was not a significant predictor, but Colantone and Stanig (2016) suggest that import competition from China had a causal effect on a region's propensity to vote Leave in the referendum. Fetzer (2018) underlines the influence of austerity policies undertaken by previous governments

Becker et al. (2017) note that there was a 'disconnect between the key correlates of the vote outcome and the topics dominating the political debate'. Considerable attention was awarded to the role of media coverage in the period leading up to the referendum in influencing individual decisions to vote to Leave. To my knowledge, this paper is the first attempt to estimate the effect of media coverage intensity on the local level decision to vote Leave in the 2016 referendum.

The important question then becomes: Did the rhetoric used in newspapers by the Leave campaign in the six to twelve months before the EU referendum have an effect on the referendum result? In this paper I answer this question by evaluating the effects of the two main types of newspapers coverage, mainly used by the Leave campaign to portray the EU in an unfavourable light, coverage on migration with a focus on either crime or public resources, and more in general coverage on anti-EU topics.

The main hypothesis examined in the first paper is that higher exposure to negative coverage about the EU makes the electorate more Eurosceptic and thus more likely to vote

Leave. Of particular interest is whether the negative coverage incited higher levels of Leave votes in areas more affected by government cuts and/or with lower levels of education.

To the best of my knowledge, this is the first paper that evaluates the effect of written media on the outcome of the EU membership referendum in 2016. People do not generally access official statistics on crime and economic conditions related to EU and they usually form their beliefs on EU issues through the media. Newspapers are one of the main media sources providing such information. Data from the National Readership Survey (NRS) indicates that newspapers, although now less popular relative to other news sources, are still a significant source of news, with an estimated monthly reach of 47 million adults in the UK, which exceeds that of Google. Further literature indicates that newspapers are still one of the main sources of information about political and economic developments. Additionally, as argued in Murphy (2014), a measure of newspaper coverage can be a good proxy of other types of media coverage.

Newspaper coverage can bias the beliefs of voters by reporting incorrect information, or by deliberately over-reporting news and information that support a viewpoint and under-reporting or omitting to cover specific news and information that support a different viewpoint (see Williams and Dickinson (1993), Bjorvatn et al. (2015), Grigorieff et al. (2016), and Larreguy et al. (2016)). If newspapers had some information bias against or in favour of the EU Leave decision, then this bias might have had an effect on the final referendum result. Moore and Ramsay (2017) provide evidence of slanted news coverage in the context of the 2016 EU referendum. In particular, they note that "both sides engage in mutual accusations of lying" and "highly partisan" reporting.

Additionally, the information presented by news sources can strengthen a reader's previously held expectations via confirmation bias if the message presented matches their prior belief. Importantly, theoretical evidence outlined in Rabin and Schrag (1999) indicates that an overconfident reader could dismiss evidence that is too different from their prior expectation. A voter who has a strong enough preconception could dismiss unbiased news entirely. On the other hand, a discerning reader would be able to realise the misleading effect present in some publications and update their expectation accordingly, suggesting voter sophistication in the news material they consume.

The first main contribution of this paper is to provide an estimate of the effects of newspaper coverage of news about migrant crime, the economic impact of migration, and anti-

EU topics on the percentage of Leave votes by exploiting variation in the newspapers' coverage across constituencies. I find that newspaper over-coverage (under-coverage) of negative migrant related news lead to a higher (lower) Leave vote share, but only in constituencies with a low level of education. This seems to suggest a potential effect of over-coverage (under-coverage) of news in favour of the Leave campaign on voters' beliefs in low-educated areas.

The second main contribution of this paper is to provide new insights on the effect of availability bias on the EU referendum results. I find evidence that newspaper coverage of migration related news in the 6 months before the EU referendum had a larger effect on voters than newspaper coverage in the 12 months before the EU referendum. This seems to suggest that voters formed their decision about what to vote in the EU referendum by giving more weight to most recent news and articles.

The third main contribution of this paper is to evaluate whether the effect of newspaper coverage varies by publication type and stance (in support of or against leaving the EU). I find that coverage originating in broadsheets which favoured Remain did not have significant effects on the referendum vote. On the other hand, tabloids which supported Remain were successful at dissuading voters from voting Leave. I find also that coverage originating in broadsheets which supported Leave increased the Leave vote percentage. Anti-EU coverage is also found to be positive and statistically significant if originating in Leave-supporting tabloids, but the other two definitions are not.

Given the results obtained in the first paper and those seen in the literature concerning the effect of media on electoral outcomes, the second paper in this thesis is focusing with the effects of political competition on the effectiveness of media capture. Given that media can influence elections, does political competition help or hinder the probability that media is captured in the first instance?

Beginning at the end of the last century, the world experienced a wave of democratisation, and, in the context of former Iron Curtain countries, freer media markets. The democratisation trend was believed to lead to less censorship and biased media, as they were seen as the purview of autocratic regimes. However, Besley and Prat (2006, page 720) note that in a modern context, 'despite the lack of old-fashioned pre-emptive censorship' economic means are increasingly used to restrict media freedom in order to gain favourable coverage. Furthermore, Corduneanu-Huci and Hamilton (2018) show that media restriction via both censorship and economic means is an increasing trend, 'affecting both democracies and

autocracies, with negative consequences for the overall quality of political institutions and civil liberties'. Given these results and the backsliding on media freedom outlined empirically in Corduneanu-Huci and Hamilton (2018), it becomes interesting to analyse whether the media and political competition together are sufficient to increase media transparency and voter welfare. A study analysing the role of potentially biased media in the presence of imperfectly informed electorates is a relevant and interesting avenue of research given the recent attention to the perceived role of fake news in affecting the outcomes of democratic elections and referenda.

With respect to the origin of bias, two main categories are distinguished: demand side and supply side. The overall message is that media market competition hinders supply driven bias and exacerbates demand driven bias. There are two types of bias distinguished in the literature: (outright) distortion of fact and information filtering. Focusing on media slant, Mullainathan and Shleifer (2005), Baron (2006), Rudiger (2013), Sobrio (2013), and Várdy and Oliveros (2015) find that media competition does not alleviate bias, which in turn can lead to a suboptimal election result. The results of Besley and Prat (2006) and Anand et al. (2007) indicate that electoral outcomes improve with competition. Within this category, media slant can occur either via the demand side due to audience bias, or via the supply side due to the beliefs of the reporter or the reputational concerns of the editor. The closest papers to this study are those of Besley and Prat (2006) and Trombetta and Rossignoli (2020), who analyse political games of endogenous media capture. Both papers analyse the conditions under which media capture is feasible and desirable from the perspective of the incumbent, indicating that political media capture can occur despite media freedom. However, neither paper accounts for political competition and electoral spending laws, which could help or hinder media freedom.

How can increased competition, both politically and within media market, lead to media suppression, media bias and dual information systems? Is voter welfare negatively affected? This paper proposes a novel theoretical approach that analyses the incentives of politicians and media outlets to misrepresent information to the electorate.

This paper analyses whether political competition in capture and electoral standards help mitigate these effects and looks at the theoretical incentives of media suppression in more democratic settings, with novel results. This study aims to contribute to the burgeoning literature on the interplay between media, political outcomes, and the flow of information. The proposed model builds on the seminal paper of Besley and Prat (2006) and that of Trombetta

and Rossignoli (2020) in two ways. Firstly, it introduces an active challenger in the political game, effectively emulating a simple majority democratic election or referendum. Secondly, it extends the analysis to account for binding and non-binding electoral spending laws. However, unlike Trombetta and Rossignoli (2020) and Besley and Prat (2006), the results indicate that in pure strategy Perfect Bayesian Equilibria, media and political competition and electoral standards ensure that the true state of politicians is always revealed to a subset of the electorate. However, if the proportion of voters who actively seek political news is sufficiently large, full disclosure occurs in the media market even if both politicians are bad. The present model can be seen as a generalisation of the previous models as it collapses to that of Trombetta and Rossignoli (2020) if the politicians' budgets are highly asymmetrical and there are no reputation concerns. With the additional assumption that the share of informed voters is zero, it collapses to the analysis of Besley and Prat (2006).

One of the most contentious issues in the media in the run up to the 2016 referendum is whether incoming EU migrants are able to integrate in the British society and whether they are a strain on the public infrastructure. The final paper analyses the effect of EU membership on the integration of EU migrants, with a particular focus on the 2004 enlargement episode. Labour market discrimination and integration of immigrants are contentious issues in the UK, which has experienced a migrant growth from 6.4% to 13.4% between 1990 and 2017, according to the 2017 edition of the UN's trend in immigrant stock. Previous literature such as Ottaviano and Peri (2012) suggests that migrant and native labour are not perfect substitutes and as a result different outcome in the labour market are observed for the two groups, whereas Dustmann, Glitz and Vogel (2010) and Clark and Drinkwater (2008) indicate that non-OECD migrants in the UK are relatively higher educated and better remunerated than their native counterparts. It is therefore important to ask to what extent is it possible for migrants to integrate in the UK labour market, and to analyse whether their labour market performance is fair given their characteristics. As migrants accumulate experience and social capital in the receiving country, their labour market outcomes are also expected to improve. Is this observed in the UK?

In the final paper, I examine differences in labour market outcomes between UK natives and immigrants from 10 countries that joined the European Union in 2004 (hereafter also referred to as EU10 countries). In this context, the absence of wage gaps can be interpreted as a measure of the UK's efficiency at attracting and integrating successful immigrants, whereas

the presence of a gap can indicate discrimination. The presence of wage gaps between natives and EU migrants can also give an indication of the extent to which the UK is integrated in the EU labour market.

As a member of the European Union until 2020, the UK was bound by legislation surrounding the freedom of movement that outlaws discrimination of workers from member states. Therefore, it becomes interesting to analyse how changes in the EU membership in 2004 for 10 European countries affected the degree of discrimination and integration in the UK labour market of immigrants from these countries. To assess such changes, I focus on the 2004 EU membership enlargement and I carry out an empirical analysis on whether EU migrant-native wage gaps exist and how they evolve after a country joins the EU.

. In order to test this hypothesis, I propose focusing on the 2004 EU enlargement episode and analyse three periods, namely 1994-2003, 2004-2008 and 2009-2013. In particular, I focus on the observed wage differentials between native and EU10 migrants, both before and after accession and whether the integration of these migrants improves with time spent in the UK. I also analyse whether there is heterogeneity in observed wage differentials at different quantiles of the wage distribution in all three periods and whether they are affected by EU membership. I also focus on the probability of working a second job, which can be seen as an indication of short-term stays and unwillingness to fully integrate in the British labour market.

Given these questions, the third paper makes the following contributions. Firstly, this study adds to the literature of the integration of migrant workers in the British labour market by using the Blinder-Oaxaca decomposition to estimate log wage differentials and ascertain to what extent they are motivated by differences in characteristics between the native and migrant populations. It is therefore important to analyse the explained part of the decomposition to understand to what extent the differentials are attributable to different characteristics distributions across nationalities. Furthermore, it analyses whether EU10 migrant men integration improves as they spend time in the UK, both before and in two periods after accession. Secondly, this study proposes using the Logit-Oaxaca decomposition proposed by Fairlie (2005) to split the differences the probability of working a second job into explained and unexplained parts. Thirdly, it adopts an extension proposed by Firpo et al. (2007) to decompose the wage differentials at different quantiles on the wage distribution to analyse which subgroups are better integrated in the British labour market.

Chapter 2: An Assessment of the Effect of Migration Media Coverage on the European Union Referendum

Section 2.1: Introduction

Did the rhetoric used in newspapers by the Leave campaign in the six to twelve months before the EU referendum have an effect on the referendum result? In this paper I answer this question by evaluating the effects of the two main types of newspaper coverage, mainly used by the Leave campaign to portray the EU in an unfavourable light, coverage on migration with a focus on either crime or public resources, and more in general coverage on anti-EU topics.

The main hypothesis examined in this study is that higher exposure to negative coverage about the EU makes the electorate more Eurosceptic and thus more likely to vote Leave. Of particular interest is whether the negative coverage incited higher levels of Leave votes in areas more affected by government cuts and/or with lower levels of education.

To the best of my knowledge, this is the first paper that evaluates the effect of written media on the outcome of the EU membership referendum in 2016. People do not generally access official statistics on crime and economic conditions related to EU and they usually form their beliefs on EU issues through the media. Newspapers are one of the main media sources providing such information. Data from the National Readership Survey (NRS) indicates that newspapers, although now less popular relative to other news sources, are still a significant source of news, with newsbrands and magazines still having a monthly reach of 94% of individuals aged 15 or over in 2016.[1] Newsworks, a marketing body for national newspapers, reports that written media has a monthly reach of 47 million adults in the UK, which exceeds that of Google. The prevalence of written media as a news source is also highlighted in studies such as Blinder and Krueger (2004) and Gerber et al. (2009) who suggest that newspapers are still one of the main information sources on economic issues such as wage growth and unemployment. Additionally, as argued in Murphy (2014), a measure of newspaper coverage can be a good proxy of other types of media coverage.

Newspaper coverage can bias the beliefs of voters by reporting incorrect information, or by deliberately over-reporting news and information that support a viewpoint and under-reporting or omitting to cover specific news and information that support a different viewpoint (see Williams and Dickinson (1993), Bjorvatn et al. (2015), Grigorieff et al. (2016), and Larreguy et al. (2016)). If newspapers had some information bias against or in favour of the

[1] Source: NRS PADD Apr 2016 – Mar 2017.

EU Leave decision, then this bias might have had an effect on the final referendum result. Moore and Ramsay (2017) provide evidence of slanted news coverage in the context of the 2016 EU referendum. In particular, they note that "both sides engage in mutual accusations of lying" and "highly partisan" reporting.

The first main contribution of this paper is to provide an estimate of the effects of newspaper coverage of news about migrant crime, the economic impact of migration, and anti-EU topics on the percentage of Leave votes by exploiting variation in the newspapers' coverage across constituencies. I find that newspaper over-coverage (under-coverage) of negative migrant related news lead to a higher (lower) Leave vote share, but only in constituencies with a low level of education. This seems to suggest a potential effect of over-coverage (under-coverage) of news in favour of the Leave campaign on voters' beliefs in low-educated areas.

Additionally, the information presented by news sources can strengthen a reader's previously held expectations via confirmation bias if the message presented matches their prior belief. Importantly, theoretical evidence outlined in Rabin and Schrag (1999) indicates that an overconfident reader could dismiss evidence that is too different from their prior expectation. A voter who has a strong enough preconception could dismiss unbiased news entirely. On the other hand, a discerning reader would be able to realise the misleading effect present in some publications and update their expectation accordingly, explaining why some voters are less prone to media persuasion. Furthermore, voters in the EU referendum might have formed their beliefs about the EU based on 'what comes to mind', for example overweighing information and events covered by newspapers in the period just before the referendum and underweighing more distant news and events. This type of bias is known in the literature as availability bias (see Tversky and Kahneman (1973) Kosse (2013), Ferraz and Finnan (2008), and Grigorieff et al. (2016)) and it may imply that newer coverage may offset the effect of past coverage. This would imply that the effect of news would be much stronger the closer the date of publication is to the EU referendum.

The second main contribution of this paper is to provide new insights on the effect of availability bias on the EU referendum results. I find evidence that newspaper coverage of migration related news in the 6 months before the EU referendum had a larger effect on voters than newspaper coverage in the 12 months before the EU referendum. This seems to suggest that voters formed their decision about what to vote in the EU referendum by giving more weight to most recent news and articles.

A third main contribution of this paper is to evaluate whether the effect of newspaper coverage varies by publication type and stance (in support of or against leaving the EU). I find that coverage originating in broadsheets which favoured Remain did not have significant effects on the referendum vote. On the other hand, tabloids which supported Remain were successful at dissuading voters from voting Leave. I find also that coverage originating in broadsheets which supported Leave increased the Leave vote percentage. Anti-EU coverage is also found to be positive and statistically significant if originating in Leave-supporting tabloids, but the other two definitions are not.

The empirical analysis is based on area (constituency) level data from the 2011 census, the British Election Study (BES), the Office of National Statistics (ONS), Local Media Works which provides data on newspaper circulation and distribution, and an international database of national and local newspapers articles called Nexis. By combining the last two datasets I derived new measures of newspaper coverage which better capture the type of published articles that might have affected the EU beliefs of readers and which better measure the coverage intensity at level of constituency by considering the circulation of each newspaper and the population present in each constituency. This improves upon the article count measures utilised in previous literature (see e.g. Lamla and Lein (2008) and van der Wiel (2009)).

The results indicate that overall, differences in media coverage intensities had a small, positive, and statistically insignificant effect on the constituency level decision to vote Leave. However, the effects are heterogeneous, with a small, statistically significant and positive effect on the percentage of EU leave votes mainly in constituencies in the bottom quarter of the education qualification distribution, where a one unit increase in the coverage variable increased the Leave vote percentage between 0.003 and 0.005 percentage points depending on the analysed type of coverage. The effect decreases the more educated the electorate in the constituency is, and the overall effect is small, positive and statistically insignificant. Heterogeneous results are observed when separating the coverage by publication type and position taken with respect to the then upcoming referendum. A unit increase in the weighted articles variable that tabloid coverage supporting Remain decreased the Leave vote between 0.672 and 0.9 percentage points, while an equivalent increase in the coverage variable for Leave supporting broadsheets increased the Leave vote between 0.1 and 0.3 percentage points. The coverage effects are stronger in the last six months before the referendum. Anti-EU coverage from Leave supporting tabloids also increased the Leave vote by around 0.8 percentage points per one unit increase in the coverage variable.

One potential drawback of this study is the possible presence of endogeneity. This can occur if the share of leave votes is correlated with unobserved characteristics that also affect the level of weighted (negative) coverage. One potential factor is Euroscepticism, the inherent dissatisfaction of the electorate towards the EU, which would in turn affect the public demand for negative coverage. For example, if the electorate has inherent Euroscepticism, it would affect both the resulting vote leave percentage and the demand for negative coverage prior to and at the time of the referendum. As a result, this affects not only the overall coverage variables, but also how this coverage is split across publication types. This creates a potential endogeneity issue, which in turn causes a bias in the OLS estimates. A measure of local Euroscepticism, detailed in the data section, is included in all regressions as a robustness check.

The remainder of the chapter is structured as follows: Section 2 provides a review of previous relevant literature, Section 3 describes the dataset and media variable construction details, Section 4 outlines the empirical strategy, while Section 5 presents a discussion of the results and concludes.

Section 2.2: Literature Review

Two main methodologies are employed in the economic literature concerned with how exposure to the media affects decisional outcomes. The first involves using survey data to ascertain the effects of media exposure on a plethora of outcomes such as the respondents' fear of crime, attitudes towards immigration, or the effects of corruption uncovered by the media on voting outcomes. Studies that use this methodology include those by Williams and Dickinson (1993), Banerjee et al. (2011), Kasper, Kogler and Kirchler (2015), Bjorvatn et al. (2015), and Facchini et al. (2016), among others. The second methodology involves quasi-experiments using variation in media availability over time. This strand of literature explores mainly exogenous variation in television or radio signal strength created by geographical features, as seen in Enikopolov et al. (2011) Durante et al. (2009, 2015), Bursztyn and Cantoni (2014), or due to staggered introduction, as observed in Jacobsen (2011) or Durante and Knight (2009).

Noteworthy variations of the latter methodology consist of the studies of Ferraz and Finan (2008) and Larreguy et al. (2014). The former study uses variation created by the 2003 Brazilian anti-corruption programme in which municipalities were randomly selected for expenditure auditing, alongside the presence of radio in the municipality. Their results indicate that media coverage diminished the chances that a corrupt mayor would be elected, as well as

increasing the chances that an 'honest' mayor would be re-elected. Interestingly, the results here also indicate voter discretion, inasmuch mayors with populistic platforms were more negatively affected by the presence of the media.

Larreguy et al. (2014) find similar results in Mexico, outlining the relative importance of local media in changing voters' perception of the credibility and competence of politicians seeking re-election. In addition to the findings of Ferraz and Finan (2008), Larreguy et al. (2014)'s results indicate that redirecting resources that are meant for the poor has especially large effects for the Populist Party, attesting to voter sophistication in the form of punishing ideological dishonesty. However, media coverage of the malfeasance of a candidate who was based in another municipality is found to have no effect, indicating the relative importance of local media in changing voter perception. These two studies indicate that the media influences political outcomes through the distribution of information and that the electorate has the capacity to filter media bias to a certain extent. This becomes relevant in this context since the analysis of Moore and Ramsay (2017) reveals media slant in the coverage leading up to the EU referendum particularly on migration and economic concerns.

A similar study was conducted by DellaVigna and Kaplan (2007) in the United States. The authors analyse the impact of the introduction of Fox News on the Republican vote share and vote switch in the 1996 and 2000 US Presidential elections. The analysis suggests that the pro-Republican news station had a positive effect on the vote share for Republican candidates as well as on voter turnout. Additionally, DellaVigna and Kaplan (2007) estimate that Fox News convinced 5% to 30% of the non-Republican audience to vote for the party. The authors put forward three possible explanations. Their first contention is that towns were becoming more conservative in areas where Fox News was first introduced, picking up political trends; however, the authors supply evidence to discredit this explanation. Secondly, they propose that initially, voters were unsure about the bias of Fox News, resulting in a temporary effect on voting. The observed effect does not disappear in 2004, which indicates that the availability of Fox News is an insufficient justification. Their third explanation is that viewers are subject to nonrational persuasion, altering their beliefs and voting behaviour; this is of particular interest as it is also a hypothesis present in this study. Another plausible explanation is that, like in Durante et al. (2015), exposure to Fox News has been continuous until re-evaluation, which does not allow the effects to dissipate.

Durante et al. (2015) examine the staggered introduction in 1985 of Mediaset, a commercial station owned by Silvio Berlusconi, on citizens' attitudes and ultimately on their voting decisions. The media variable is Mediaset availability, based on the location and strength of the transmitters. The results indicate that areas exposed to Mediaset displayed higher electoral support for Berlusconi's party in 1994, an effect which lasts for two decades. Using survey data, the authors bring evidence that this effect was stronger for individuals who watched TV more, with older people becoming more likely to watch pro-Berlusconi news via increased attachment to the network. Conversely, younger individuals became less interested in politics and more prone to populist messages. The continuous coverage present in this study precludes the possibility of availability bias attenuation but highlights the electoral effect of information.

In line with the work of DellaVigna and Kaplan (2007), Banerjee et al. (2011) indicate that media coverage is correlated with higher voting turnout and a higher propensity to cast an 'informed' vote. However, the results of Dutta et al. (2013) underline that the media induced a difference between actual and perceived outcomes, indicating that unfavourable media coverage may compound the potential effects of migrant shares and net migration. The resulting bias indicates a possible disparity in effects between the expectations formed due to media exposure and official figures on unemployment, crime, or growth figures.

An interesting study on the persistence of availability bias and the effects of coverage of economic issues is that of Doms and Morin (2004). Their results indicate that consumers update their expectations about the economy more frequently in times of high media coverage, that the resulting expectation stickiness is countercyclical, and that expectation changes are short-lived. The results not only indicate the presence of availability bias, but also that the effects of the media depend on reporting intensity.

The studies of Williams and Dickinson (1993) and Tversky and Kahneman (1973) are relevant to this study in the sense of both information bias and availability bias. Williams and Dickinson (1993) argue that the media is a contributor in consumers' expectation formation. More precisely, reports of crime and increased criminality increased individuals' fear of crime, although the effect was dependent upon the reporting style of the paper. This indicates both a fundamental information bias as well as an underlying availability bias, as more recent papers had a greater effect on consumers, as introduced in Tversky and Kahneman (1973). The analyses in both studies indicate that the media has a strong effect in influencing the population,

and that media tone rather than media coverage is relatively more important when analysing its effects. In this context, it suggests that electorates may be prone to media persuasion that would alter voting intentions.

More recently, Facchini et al. (2016) used a large-scale Japanese experiment to demonstrate that exposing citizens to material pertaining to the positive social and economic effects of immigration combats the often-antagonistic public sentiment towards immigrants and leads to increased support for a more open immigration policy. The effects were heterogeneous across the population and, in line with the results of Williams and Dickinson (1993), heterogeneous in tone. These effects were also present at re-evaluation, indicating a potential policy measure that could quell scepticism towards immigration and potentially populism. The results from the study by DellaVigna et al. (2011) propose similar results with the opposite effect, indicating that Serbian radio transmissions in Croatia were correlated with higher incidences of nationalism. This outlines a potential channel through which a media induced distortionary effect on decision making can occur.

A counterpoint to these papers is the study of Gerber et al. (2009), whose findings suggest that exposure to the media is more important than media tone in changing readers' perceptions. The authors' survey results indicate that despite the 'media slant', the qualitative results were similar for the two newspapers that expressed significant differences in reporting style. The results also show that voters filtered out the media bias, effectively negating the non-informative role of media coverage, providing evidence of voter sophistication in filtering biased information.

The studies of Burstzyn and Cantoni (2014), Eisensee, and Stromberg (2007) also find media effects on various outcomes to be short lived. The latter contends that the media effect may be transitory since their results indicate that natural disasters are more likely to receive monetary relief when the competition for US news time is low, due to the American public being more likely to notice the contributing organisations' charitable behaviour. Since media coverage influences public outcry, this result is in alignment with the availability bias hypothesis.

Bjorvatn et al. (2015) suggest that exposure to certain media can make individuals more financially literate and more interested in entrepreneurship but provides mixed evidence on the persistence of media effects. The study of Hovland et al. (1949), cited in Prat and Stromberg (2013) notes the presence of a 'sleeper effect', which manifests itself in the form of gradually

reverting to one's original position after being exposed to a persuasive message. This result is similar to the hypothesis of Tversky and Kahneman (1973) who argue that more recent and more common information is more relevant in opinion formation.

Economic literature indicates that exposure to the media has an impact on decision making and on individual behaviour, but that this effect can be transitory and offset by newer coverage.

Section 2.3: Data

Section 2.3.1: Sources

For the present study, I use data from Version 2.1 of the 2015 British Election Study Constituency Results. This dataset is also enhanced with demographic data from the 2011 census, including total population, ethnicity, country of birth, religion, and socio-economic classification. The population data from the census is used to define variables in the constituency controls and migrant controls categories defined below. Vote Leave shares in the EU membership referendum were constructed and added to the BES dataset. In addition, this analysis uses data from the Office of National Statistics (ONS) to add constituency level data on, net migration, gross value added, wage median and unemployment levels, and regional levels of public spending and investment from the EU. Net migration data, which is used to construct a variable detailing the net migration flow, constitutes the final variable in the migration control set, while the remainder set of variables are used to construct economic indicator variables. The latter variable group accounts for relative economic performance between the constituencies, which may have contributed to determining the percentage of Leave votes. Intuitively, a relatively 'worse off' constituency is expected to have a higher propensity to vote to Leave as suggested in Hobolt (2016). In this context, changes in unemployment and wage median are at constituency level and are defined as changes from the year prior to the referendum. The change in public spending variable is defined in the same period, but is only available at NUTS-1 aggregation level, which comprises of several constituencies. Changes in EU funding are defined at the same aggregation level but are defined as changes in the allocated regional budgets from between the 2007-2014 and 2014-2020 periods. Gross Value Added (GVA) is a measure of the level of economic activity in a region, available at the level of NUTS-3 regions.

To construct the media variables, this study uses Nexis and the Newspaper Database (JICREG), offered by Local Media Works. The former is an online database of regional and

national newspaper articles from a set of countries, including the United Kingdom. It allows the retrieval of all newspaper articles published in the UK fitting some specified search criteria. Relevant article searches depend on using keywords linked by logical operators, country of publication, whether regional newspapers are included, and period. To improve precision, relevant search terms can be constricted to be in either the headline or the main text, and all articles pertaining to events or reports outside the UK are eliminated. The second database used in the media construction variable is the Newspaper Database (JICREG), offered by Local Media Works, which offers information about both the titles and geographic circulation of a newspaper. Based on information from these two databases, it is possible to construct variables detailing the number of applicable articles to which constituencies were exposed during the chosen period. Due to circulation data unavailability for Northern Ireland, the present study uses the remaining 632 General Election (GE) constituencies in England, Scotland and Wales. A constituency is defined as an electoral area that elects one Member of Parliament (MP) at each General Election.

Section 2.3.2: Media variable construction

For the analysis, this study uses the behavioural economic assumption of myopic expectation building motivated by previous literature; in other words, a disproportionate emphasis on information encountered nearer the present. Since the average voting citizen does not have access to the official police records of crime, they form subjective expectations based on available data. More recent articles are expected to be relatively more important in shaping individual decisions, and as such articles published closer to the 23rd of June 2016 are expected to have a stronger effect on electors' voter decisions, as suggested by Burstzyn and Cantoni (2014), Tversky and Kahneman (1973), and DellaVigna and Kaplan (2007). Because of myopic expectation building and the fact that the referendum was announced on the 27th of May 2015, the present analysis focuses mainly on media coverage in the year prior to the referendum.

The count variable, as used in previous literature, is given by:

$$Mediavariable_i = \sum_{j \in N_i} articles_j \qquad \text{[Eq. 1]}$$

where i denotes the constituency, N_i denotes the set of newspapers distributed in constituency i during the analysed period, j denotes a newspaper, $articles_j$ is the number of articles published by newspaper j in a specific topic found by Nexis according to a search criteria, and $Mediavariable_i$ measures the total number of articles published in constituency i in a particular topic. I will consider articles published on topics portraying the EU in a negative manner and

more in particular articles about migrant crime and about economic consequences of migration. Similar article count variables have been used by Lamla and Lein (2008), van der Wiel (2009) and Murphy (2014).

A drawback with this $Mediavariable_i$ measure is that it allocates the same weight to all newspapers distributed in a constituency regardless of their relative local circulation figures. This overestimates the effect of regional, smaller but spatially fragmented publications. Additionally, some regional newspapers have a wide geographical distribution despite the focus on their coverage being relatively small. While the majority of relevant articles originate in national newspapers with high distribution, the majority of the variation in constituency coverage derives from regional newspapers. For this reason, I introduce a new $Mediavariable_i$ measure that take account of different level of circulation of a newspaper by constituency defined as follows:

$$WeightedMediavariable_i = \sum_{j \in N_i} \frac{articles_j * circ_{i,j}}{elector_{i,1000s}} \qquad \text{[Eq. 2]}$$

where $circ_{i,j}$ denotes the total number of copies of newspaper j distributed in constituency i, which can be thought as a proxy for readership in the constituency, and $elector_{i,1000s}$ denotes the size of the electorate (in units of 1000) in the constituency i in 2015. The size of the electorate is used instead of the total population for two reasons. Firstly, eligibility criteria for the 2015 general election are identical to that of the referendum one year later. This, by definition, excludes the share of the population for whom the effect of media coverage is irrelevant due to institutional constraints which prevent them from voting. Secondly, due to data constraints, total population size data originates in the 2011 census, but the electorate size data matches populations in 2015. These two reasons render the size of the electorate in 2015 a better approximation of the size of the constituency population who could have voted in the referendum than the total population measured at the 2011 census. One caveat on the second measure is to account for the populations of different constituencies to be relatively better informed or more up to date relative to others, in which case the first version of the weighted variable would partly pick up positive correlation between more qualified electorates and remain votes. Analysis using the unweighted media variables can be found in Appendix A2, with results listed in Tables 2.10 and 2.11.

To explain the share of Leave votes at constituency level I will use the $WeightedMediavariable_i$ as my main explanatory variable to capture the effect of

newspaper coverage on specific topics but always controlling also for the level of readership at constituency level using the following measure:

$$Constituencyreadership_i = \sum_{j \in N_i} \frac{circ_{i,j}}{elector_{i,1000s}} \ . \qquad \text{[Eq.3]}$$

I will consider three different specific definitions of media variables using equation (2) to derive a weighted count of the number of articles on the following three different topics:

- migrant crime and crime involving migrants;
- migrant induced resource constraints, including alleged pressure on public services such as schooling and the NHS, 'benefit scrounging', etc.;
- anti-EU sentiment, due to socio-economic reasons or concerning alleged loss of border control/sovereignty associated with continued EU membership.

The delimitation outlined above stems from the main arguments used by the Leave campaign in the period leading up to the referendum and is based on the review presented in Moore and Ramsay (2016). For the construction of the variables, migration concerns are split into two categories: crime, and public resource constraints. Moore and Ramsay (2016) point out that migration coverage prior to the referendum was overwhelmingly negative in all categories. Separating coverage in this fashion allows the analysis of the effect of variation in differing types of content and their relative effects. The definitions for each variable are given below and alternative definitions for the migrant crime category along with corresponding results are found in the appendix A1.

The methodology employed in devising the keywords is based on a linguistic approach in which relevant keywords, or combinations thereof, are used to identify the articles of interest as well as the style of content. Here, as in the other work exemplified below, the negative focus of the article is ensured by the combination of negative keywords and the subject focus of the study. This method is applicable to other fields such as health economics, as observed and exemplified in Nagelhout et al. (2011) and Niederdeppe (2010), law (Garber et Bower (1999), and others. The approach used in this study also appears in economic literature such as Lamla and Lein (2008), van der Wiel (2009), Murphy (2014), and McAuliffe et al. (2017) who employ specific wordings in order to capture the sentiment of their relevant articles. In particular, Lamla and Lein (2008) use terminology such as 'expensive euro' or 'teuro' in order to search for and distinguish articles in which the introduction of the euro currency was blamed for rising

inflation, with a specifically negative outlook. The study of van der Wiel (2009) employs a similar strategy, constructing variables indicating the number of relevant articles published weekly containing pertinent terms in the headline or introduction respectively as well as combinations in the headline, but does not make distinctions regarding the tone of the articles.

The study of Murphy (2014) also uses the LexisNexis database and employs a similar strategy to this paper in order to identify news reporting cases in which teachers were either accused of or investigated for misconduct. The author uses the word 'teacher' to pinpoint the focus of the coverage, and then various keyword alternatives suggesting differing formulations in which the idea of being accused and/or under investigation could be expressed. Furthermore, due to the focus of his paper, further keywords are added to identify whether the teacher was in a union, but other relevant information such as the gender of the teacher and what level they were teaching was inferred from reading the articles.

For the present study, a list of relevant keywords was drafted covering each issue as exhaustively as possible based on the examples of headlines provided in Moore and Ramsay (2016). In order to avoid false positives in the search results, the main keywords were linked with identifiers found in the first string. In the case of all variables, the first set of keywords aims to ensure that the article is specifically referring to migrants or migration in general and refers to differing formulations in which newspapers may refer to a non-native.

One potential weakness of this definition is that the nationality of the culprit may be used instead of a variation of the word migrant, however, the search criteria used still uncovers articles which match this description. The second keyword string in the first definition is more restrictive, requiring the keywords to be found in the headline, ensuring that the focus of the article is indeed the required topic and it includes variation of nouns and phrases commonly associated with the subject. In the case of the definition of the definition of resource constraints resulting from migration, the second keyword string attempts to capture formulations which portray migrants or migration in general as a strain on public services, in particular the National Health Service (NHS). The third string of keywords used in this definition follows a similar intuition. . The United Kingdom index is added to ensure that the articles uncover coverage of economic pressure of migration on the British public infrastructure only. The keyword codebook exemplified above was obtained through iterative modifications aimed at encompassing all formulations in which the negative news could be phrased, based on the report of Moore and Ramsay (2016). The codebook may not necessarily be comprehensive, but

it was expanded to account for a majority of possible combinations and phrases. In line with previous literature, the results were checked to ensure the validity of the search.

The keywords used to construct the main set of media coverage variables are detailed in Appendix A1. The keywords for the crime are included in the keywords for the anti-EU sentiment because crime is one of the reasons for which anti-EU sentiment is expressed.

Two alternative keyword specifications are tested in the case of the migrant crime coverage to check the validity of the main keyword codebook. The first alternative definition aims to narrow the search to articles in which migrants are tried and found guilty, and to filter out mere allegations of criminal activity. This is in line with previous literature involving coverage of crime, such as work by van der Wiel (2009), Nagelhout et al. (2014), and Murphy (2014), who focus on creating relevant categories. Therefore, the second set of search strings is aimed at narrowing down coverage of migrant crime to only cases in which a migrant is convicted and/or imprisoned. Since in these articles the guilt of the migrant has been established, a reader whose perception would be affected negatively by such coverage would find such articles more relevant. The first keyword string refers to differing formulations in which newspapers may refer to a non-native, while the last second string aims to capture coverage in which crime or criminality may be reported. The second search string in the first alternative definition string attempts to narrow down the search to instances in which the guilt of the culprit is verified, as opposed to the first definition which does not make this distinction. The first two strings are relatively less stringent and can be found in the main text of the article. It can be expected that an article may omit the guilty status and or the nationality of the culprit in the headline. The third keyword string is more restrictive, requiring the keywords to be found in the headline, ensuring that the focus of the article is indeed fixed on crime and it includes variation of possible crimes and nomenclatures for persons engaging in them. This variation has two potential advantages. Firstly, if negative coverage of migrant crime affects voting intentions in respect of the EU membership referendum through fear of crime, its increased relevance would better capture this effect, similar to the literature listed above. Secondly, its formulation is aimed specifically at ensuring that the origin of the crime is not a native person. As in the original keyword formulation, a plausible weakness of this definition is that the nationality of the culprit may be used instead a variation of the word migrant, but the search still uncovers articles which match this description. These alternative specifications and their results can be found in Appendix A.3.

On the other hand, the third search imitates the concept of a 'brute force search' from computer science. The aim is to list all plausible word combinations which express the link between a crime and migrants or migration. In particular, the second alternative definition systematically tallies all the elements in the Cartesian product of the set of keywords from the first two strings of previous definitions. The logical link between the products ensures the search returns articles which feature at least one element in the Cartesian product. The strength of this definition is that it ensures, by construction, the exclusion of articles in which the migrant is the victim of the crime. Although intuitively appealing, this approach has the disadvantage of being inefficient in the context of higher complexity queries such as written language where certain combination would be superfluous due to unnatural speech. Additionally, the complexity of the resulting 'brute-force' keyword search increases exponentially with the complexity of the previous two keyword searches. Since generating all the plausible word combinations in which migrant crime may be referred to is implausible, the resulting variable is more likely an imprecise proxy of total coverage. Appendix A3 presents alternative formulations for the media variables on the coverage of migrant crimes and their similar) corresponding results.

Furthermore, overall media coverage variables can be split into broadsheet, tabloid, and regional. This division is motivated by the results of Williams and Dickinson (1993), who find that the self-reported fear of crime effects was stronger in the case of 'tabloid' newspapers, which also displayed a greater propensity to cover crime and to cover it in a more sensationalist fashion. Since the original variables included all three categories, their estimated coefficient is likely capturing an average effect, with a stronger effect for tabloids and weaker for broadsheets. This paper uses the classification between tabloids and broadsheets of Bednarek (2005) and Moore and Ramsay (2017).

Broadsheets	Tabloids	Regionals
Daily Telegraph	Daily Mail	Publications with limited spatial distribution
Sunday Telegraph	The Mail on Sunday	
The Times	Daily Express	
Sunday Times	Sunday Express	
Financial Times	The Sun	
The Guardian	The Sun on Monday	
The Observer	Daily Mirror	
The Independent (until the 26th of March 2016)	Sunday Mirror	
	Sunday People	
	Daily Star	
	Morning Star	

The delimitation outlined above accounts only for reporting style, but not for stated positions. The newspapers can be further split by their editorial position in relationship to the referendum choice. According to Firmstone (2016), The Guardian, The Mirror, The Independent, The Times and Financial Times have stated support for remain; while Leave was supported by The Sun, The Daily Mail, Daily Express, and Daily Telegraph. The Sunday editions followed the same line as the main publication with two notable exceptions: The Mail on Sunday which supported remain, and the Sunday Times which supported Leave. Further to the publications mentioned in Firmstone (2016), The Observer, Morning Star and Daily Star also explicitly stated their position, with the former backing remain and the latter two supporting Leave. This allows grouping tabloids and broadsheet newspapers in four categories depending on their position on the referendum, accounting for reporting style heterogeneity.

Section 2.3.3: Summary statistics

Summary statistics for the main variables are detailed in Table 2.1. All coverage variables presented here are weighted.

Table 2.1: Summary Statistics of media coverage variables and Leave votes					
Variable	Obs	Mean	Std. Dev.	Min	Max
Vote Leave percent	632	52.11407	11.40347	18.48123	74.96078
Weighted Migrant crime coverage					
Migrant crime articles 1 year	632	738.8511	386.172	126.1288	1944.349
Migrant crime articles 6 months	632	477.2277	261.5479	79.6209	1316.427
Weighted Resource constraint coverage					
Resource constraint articles 1 year	632	493.5794	251.5361	85.85534	1284.373
Resource constraint articles 6 m	632	318.463	172.124	45.24339	847.438
Weighted Aggregate anti-EU coverage					
Aggregate coverage articles 1 year	632	427.7436	238.7579	58.54731	1170.73
Aggregate coverage articles 6 months	632	252.9924	139.8663	34.95857	685.8712
Coverage in the last six months before the 2015 General Election					
Weighted Migrant crime coverage					
Migrant crime	632	242.82	135.5237	39.24323	685.7206
Regional coverage	632	.6372705	.5487861	.2638319	5.234414
Weighted Resource constraint coverage					
Resource constraint coverage	632	117.4623	66.09668	18.98097	335.5742
Tabloid coverage	632	52.18262	17.35128	12.51709	108.3701

Table 2.1.1: Summary Statistics of media coverage variables and Leave votes (continuation)					
Variable	Obs	Mean	Std. Dev.	Min	Max
Broadsheet supporting Leave					
Migrant crime articles 1 year	632	115.6089	68.18404	7.819592	290.6104
Migrant crime articles 6 months	632	62.86246	37.61009	2.879936	155.9359
Aggregate coverage articles 1 year	632	69.19222	40.79668	6.185204	181.4025
Aggregate coverage articles 6 months	632	38.3419	22.34291	3.948822	100.1305
Resource constraint articles 1 year	632	51.41028	30.49804	3.388902	130.8966
Resource constraint articles 6 m	632	35.26233	20.9185	2.284552	89.69765
Broadsheet supporting Remain					
Migrant crime articles 1 year	632	303.3806	221.9983	42.08141	1068.551
Migrant crime articles 6 months	632	201.9173	163.713	26.87609	773.1436
Aggregate coverage articles 1 year	632	168.9584	133.3664	18.51411	629.3856
Aggregate coverage articles 6 months	632	120.1814	91.81594	12.23505	432.6766
Resource constraint articles 1 year	632	159.5925	136.2441	15.92887	635.2037
Resource constraint articles 6 m	632	109.8156	90.55837	10.29046	424.5381
Tabloids supporting Leave					
Migrant crime articles 1 year	632	305.8619	112.8783	73.17387	661.2477
Migrant crime articles 6 months	632	203.7952	73.49953	49.42139	435.9189

Aggregate coverage articles 1 year	632	181.7865	75.42665	33.27156	412.6098
Aggregate coverage articles 6 months	632	126.5895	53.49541	22.28644	289.9303
Resource constraint articles 1 year	632	257.489	99.9897	57.01041	566.3782
Resource constraint articles 6 m	632	173.2691	72.11565	32.54597	391.4078
Tabloids supporting Remain					
Migrant crime articles 1 year	632	14.43221	5.154451	3.300129	29.97733
Migrant crime articles 6 months	632	9.154418	4.032874	0.632659	19.98489
Aggregate coverage articles 1 year	632	7.606239	3.400746	0.451899	16.65408
Aggregate coverage articles 6 months	632	132.6745	55.72814	22.64796	303.2535
Resource constraint articles 1 year	632	11.00576	4.219103	1.830824	23.31571
Resource constraint articles 6 m	632	7.606239	3.400746	0.451899	16.65408

Table 2.2: Summary statistics of other variables					
Variable	Obs	Mean	Std. Dev.	Min	Max
Constituency Controls					
Constituency Readership	632	0.673264	0.317043	0.273459	2.201388
Share of people aged 30 to 44	632	20.27771	2.751368	13.99852	31.75177
Share of people aged 45 to 59	632	19.71651	2.203759	11.45902	24.13966
Share of people age over 60	632	20.69523	4.755848	7.20494	35.01723
Population density	632	20.22218	25.65448	0.055717	146.3846
Male share	632	49.07457	0.804575	46.8641	53.04909
Share of married people	632	33.34806	5.772054	14.63279	46.32761
Share individuals with qualification level at least 4, 2011 Census	632	26.74864	8.319327	12.06501	57.39058
Migration Controls					
Share of Pre-2004 EU immigrants	632	1.503952	1.482931	0.325282	14.04585
Share of Post 2004 EU immigrants	632	1.818318	1.688463	0.16075	11.86967
Share of non-EU immigrants	632	7.853686	8.611992	0.981044	47.39812
Net migration in the last 5 years	632	5.304484	5.938312	-1.989	30.983
Economic indicators					
Unemployment rate, census 2011	632	4.373037	1.421444	1.837399	9.526676
Decrease in unemployment rate (2015-16)	632	0.060177	0.292878	-1.35075	1.251096

Decrease in EU funding (% 2007-14 to 2014-2020)	632	-14.4383	30.81162	-90	29.16667
Change in public spending (2015-16 %)	632	1.157129	1.660864	-0.55899	5.880651
GVA per head (£ 000s)	632	24.30064	18.84479	13.411	292.855
Change in wage median (2015-16 %)	632	-2.51502	4.963389	-22.1532	14.85149

Data in Table 2.1 above suggests the presence of variability within the constructed media coverage variables and the constituency level Leave vote shares. The Leave vote displays a mean of 52.1% and a standard deviation of 11.4 percent, while overall 401 constituencies had a majority of Leave votes. In absolute terms, media coverage relating to migration is the most prominent. In the case of weighted migrant crime variables, the measure of article coverage involving migrants and crime displays a smaller mean of 738.85 but a considerably larger standard deviation of 386.17 is observed. As expected, coverage about migration pressure on the British resources is also extensive, but less so than that of crime, with a mean of 493.57 weighted articles per constituency in the year prior to the referendum. Shortening the time span to 6 months reduces means and standard deviations in all variable categories. It is worth noting that an increase in the intensity of coverage was observed in the last 6 months compared to the 6 to 12-month period prior to the referendum, which indicates an increase in the relative intensity of relevant coverage. Summary statistics of the weighted media variables and constituency readership by newspaper can be found in Table 2.15 in the appendix..

Data in Table 2.2 presents the summary statistics for the main explanatory variables used in the analysis below. People aged 15 or older at the 2011 census are considered for this study since they would have reached the minimum age of 18 at the time of the referendum. All age groups display similar means of roughly 20% of the population and 2% standard deviations, but the oldest band display higher standard deviations of nearly 5%. Three measures of migrant population are used based on whether the migrant is a member of an EU country, which in turn is split based on whether the country joined prior the 2004 EU enlargement. The shares of both categories of EU migrants display similar mean and standard deviations of roughly 1.8 and

1.5% respectively, while the shares of non-EU migrants display a higher mean of 7.85% and a standard deviation of 8.61%. It can also be argued that individuals' migration concerns stemmed not from local stocks of migrants, but from perceptions of net increases in one's local area. Net migration indicators data is appended from the ONS to test this hypothesis, where net migration is defined as the difference between the inflow and outflow of migrants in the area. A set of plausible determinant of a constituency's vote are changes in local economic conditions. To this end, changes in unemployment, EU funding, public spending and local economic activity data appended from the ONS are included in the analysis. Constituency level data is only available in the case of unemployment, which is considered in the year prior to the referendum, while changes in the other variables are available at NUTS-1 and NUTS-3 level. NUTS-1 level are the largest statistical areas of the UK, including nine statistical regions for England, with Scotland and Wales as separate entities. NUTS-3 level is defined by groups of unitary authorities, counties or council areas in England, Scotland or Wales, which can be aggregated into constituencies. To account for differing levels in economic activity, measures of Gross Value Added (GVA) are added from the ONS. Table 2.1 displays the summary statistics the GVA per head in 2015. The change in public spending and EU funding are only available at NUTS-1 level, which coincide with the main UK administrative divisions.

Since graduates and student shares were found to be inversely related to Leave votes, the share of individuals whose highest qualification is at least level four according to the Regulated Qualifications Framework (RQF) at the time of the 2011 census was included as a proxy. The summary statistics suggest that at the time of the 2011 Census, a mean 26.74 percent of a constituency's population satisfied this qualification requirement, with a standard deviation of 8.31. An RQF qualification level four indicates that the person has completed some further education beyond college, but less than a bachelor's degree. On this scale, Bachelor and Master degrees are levels six and seven respectively, which renders this an imprecise measure of share of tertiary education. For the purpose of the present analysis, we use the share of individuals whose highest qualification is at least four and create a dummy which equals one if the constituency's share is above the sample average, as well as four distinct quartile range dummies.

The Eurobarometer surveys have collected historical levels of Euroscepticism in the UK, but they are only available at country level. Since this is constant across constituencies in a cross-section, it is unusable in this context. An alternative measure of Euroscepticism proposed in Godwin and Milazzo (2015) comes from the sixth wave of the Internet Panel of

the British Election study (BES), collected 13 months before the referendum and before the start of the media coverage variable used in this study. The respondents were asked: "If there was a referendum on Britain's membership of the European Union, how do you think you would vote?". The share of respondents in the constituency who responded Leave would be a reasonable proxy of inherent local Euroscepticism, especially since their responses were collected before start of the 1-year media coverage variables used in the study. Given the number of respondents at the constituency level ranges from 17 to 154, it is unlikely the share of Leave responders is a very accurate representation of the true value. This value is added in all regressions as a robustness check, and all results subsequently account for this measure of local Euroscepticism.

Section 2.4: Empirical Strategy

The estimated model is:

$$leaveshare_i = \alpha_1 + \beta_1 mediavariable_i + \mu_1 \boldsymbol{V_i} + \boldsymbol{\delta_1 migrantcontrol_i} + \boldsymbol{\rho_1 Economicindicators_i} + \boldsymbol{\varepsilon_{1,i}} \quad \text{[Eq.4]}$$

In equation 4, $leaveshare_i$ is the percentage vote to Leave in the EU membership referendum, α_1 is the intercept, $mediavariable_i$ denotes the weighted media variable, V_i is a vector of constituency specific characteristics, $migrantcontrol_i$ is a vector containing the migration controls previously mentioned, $Economicindicators_i$ is a vector containing constituency and region level economic controls and ε_i is the error term. The vector of economic indicators includes changes in economic variables such as public spending, unemployment and wage medians, while the migrant controls vector includes the percentage of migrants from pre-2004 expansion EU, post-2004 expansion and non-EU countries in the constituency recorded in the census. Similarly**, μ, ρ** and **δ** are vectors of coefficients.

One extension to the analysis is to check whether the effect of media coverage is heterogeneous across constituencies with differing shares of 'qualified' individuals. To this end, equation 5 below outlines the new model, where the qualification shares vector includes dummy variables equalling 1 if the constituency is in one of the top three quartiles of the qualification distribution. The media variable interactions vector includes the interaction between the media variable of interest and the qualification dummies denoting constituencies in the top three quartiles of the high qualifications share distribution:

$$leaveshare_i = \alpha_2 + \beta_2 mediavariable_i + \boldsymbol{\theta_2 mediavariableinteractions_i} + \mu_2 \boldsymbol{V_i} + \boldsymbol{\tau_2 QualificationDummy_i} + \boldsymbol{\rho_2 Economicindicators_i} + \boldsymbol{\delta_2 migrantcontrol_i} + \boldsymbol{\varepsilon_{2,i}}$$

[Eq.5]

Since the first set of variables includes article counts irrespective of the credibility, reputation and writing style of publication, the estimated effect will lie between the effect of exposure to reputable newspapers and that of tabloid-type publications. Furthermore, national newspapers and their regional counterparts may enjoy different levels of trustworthiness, so a further split is proposed for further research. Further analysis can be done using the weighted index described in the previous section, differentiating between national article coverage and regional, and analysing whether the effects were differentiated across publication types and referendum stances. The estimated model is outlined in equation 6 below:

$$leaveshare_i = \alpha_3 + \beta_3 nationalbroadsheetvariable_i + \beta_4 nationaltabloidvariable_i + \beta_5 regionalvariable + \boldsymbol{\theta_3 mediavariableinteractions_i} + \mu_3 \boldsymbol{V_i} + \boldsymbol{\tau_3 QualificationDummies_i} + \boldsymbol{\rho_3 Economicindicator_i} + \boldsymbol{\delta_3 migrantcontrol_i} + \varepsilon_{3,i}$$

[Eq.6]

Another interesting extension is analysing whether variation in coverage closer to the referendum date is relatively more important, as outlined in the literature by Tversky and Kahneman (1973) or Doms and Morin (2004). For this analysis, variants of the variables defined above are created to account for coverage in the 6 months prior to the referendum and the six months prior to the General Election in 2015. This extension can be analysed alone or in conjunction with the credibility split extension. This empirical strategy exploits two sources of exogenous variation: coverage intensity across parliamentary constituencies as well as differences in reporting propensities in the six months before the referendum.

One potential drawback of this study is the possible presence of endogeneity. This can occur if the share of Leave votes is correlated with unobserved characteristics that also affect the level of weighted (negative) coverage. One potential factor is Euroscepticism, the inherent dissatisfaction of the electorate towards the EU, which would in turn affect the public demand for negative coverage. For example, if the electorate has inherent Euroscepticism, it would affect both the resulting vote Leave percentage and the demand for negative coverage prior to

and at the time of the referendum. This creates a potential endogeneity issue, which in turn causes a bias in the OLS estimates.

Denote local Euroscepticism by Z_i and assume the true model, in matrix form, is given by:

$$leaveperc_i = \lambda \boldsymbol{X_i} + \gamma \boldsymbol{Z_i} + e$$

Where $\boldsymbol{X_i}$ is a matrix containing the included variables, media and the controls, λ is the associated vector of coefficients, $\boldsymbol{Z_i}$ is a column vector containing 'omitted' Euroscepticism and *e* is the residual. The normal form equations will be:

$$\begin{bmatrix} \boldsymbol{X_i'X_i} & \boldsymbol{X_i'Z_i} \\ \boldsymbol{Z_i}'\boldsymbol{X_i} & \boldsymbol{Z_i}'\boldsymbol{Z_i} \end{bmatrix} \begin{bmatrix} \hat{\lambda} \\ \hat{\gamma} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X_i'}leaveperc_i \\ \boldsymbol{Z_i}'leaveperc_i \end{bmatrix}$$

Solving for $\hat{\lambda}$:

$$\boldsymbol{X_i'X_i}\hat{\lambda} + \boldsymbol{X_i'Z_i}\hat{\gamma} = \boldsymbol{X_i'}leaveperc_i$$

$$\boldsymbol{X_i'X_i}\hat{\lambda} = \boldsymbol{X_i'}leaveperc_i - \boldsymbol{X_i'Z_i}\hat{\gamma}$$

$$\hat{\lambda} = (\boldsymbol{X_i'X_i})^{-1}\boldsymbol{X_i'}leaveperc_i - (\boldsymbol{X_i'X_i})^{-1}\boldsymbol{X_i'Z_i}\hat{\gamma}$$

$$\hat{\lambda} = \lambda - (\boldsymbol{X_i'X_i})^{-1}\boldsymbol{X_i'Z_i}\hat{\gamma}$$

The second right hands side term denotes the correction arising from omitted variable bias. Note that if $(\boldsymbol{X_i'X_i})^{-1}\boldsymbol{X_i'Z_i}\hat{\gamma} = \boldsymbol{0}$ the OLS estimate is unbiased and preferable. This can happen under two conditions. First, if the elements in $\boldsymbol{X_i}$ are orthogonal to 'unobserved' Euroscepticism, then $(\boldsymbol{X_i'X_i})^{-1}\boldsymbol{X_i'Z_i} = \boldsymbol{0}$ and the estimate is unbiased. However, this may not be true in the case of the coverage variables if Euroscepticism is a newspaper demand determinant. For example, if Euroscepticism increases demand for negative coverage in an area, publications may increase the amount of negative coverage aimed at such areas. Inversely, publications which feature more negative coverage may circulate more in areas with higher Euroscepticism due to increased demand. This in turn increases coverage intensity as measured by the weighted variables, rendering the orthogonality assumption invalid.

Secondly, the estimates would be unbiased if $\hat{\gamma} = 0$, or if Euroscepticism had no impact on a constituency's propensity to vote Leave, which is unlikely to hold. Since higher levels of Euroscepticism are expected to be positively correlated with both local propensities to vote

Leave and with the amount of negative coverage the constituency is exposed to, the resulting OLS coefficients would be biased.

The endogeneity issue is not restricted to the general model shown above, but also to the case in which coverage is split by publication type. More precisely, the difference in publication type readership may be endogenous as well. According to previous literature and the split observed above, most of the tabloid publications declared their support for the Leave vote while the opposite is true for broadsheets. Therefore, it may be the case that while both corr($tabloidvariable_i, \varepsilon_i$) $\neq 0$ and corr($broadsheetvariable_i, \varepsilon_i$) $\neq 0$, Euroscepticism has a higher impact on tabloid coverage. The editorial stance split indicated at the end of the data section suggests tabloid coverage has a stronger correlation with Euroscepticism.

Section 2.5: Results

Section 2.5.1: Benchmark Results

Data in Table 2.3 displays the regression results when the variable of interest is the total coverage of crime involving migrants in the year prior to the referendum. The first column indicates the OLS results obtained with only constituency controls, whereas the second and the third successively add migration and economic indicators. If neither migration nor economic indicators are accounted for, an article increase in average elector exposure in the year prior to the referendum increased Leave vote by 0.003 percentage points or by 1.159 per standard deviation increase (see first column in Table 2.3). Interestingly, the estimated effect on the Leave vote increases to 0.004 percentage point increase per unit increase in the weighted coverage once migration controls are considered. While this does not suggest more cosmopolitan constituencies were more susceptible to media framing, migrant shares are significant predictors of constituency level decisions to vote to Leave. Interestingly, the estimated effects of constituency readership, although negative in all three specifications, become statistically significant once differences in migrant shares are included. The estimated effect of minus 2.46 percentage points per circulated newspaper. The measure may be interpreted as a proxy for overall 'informedness', indicating that more 'informed' or up-to-date electorates were less likely to vote to Leave the European Union. Only the migrant stock from the post-2004 accession countries is estimated to have had a positive statistically significant effect of 0.54 percentage points on the vote Leave percentage per percentage point increase. On the other hand, the estimated effect on the pre-2004 accession EU migrants is strongly statistically significant and negative, indicating a 3.56 percentage points reduction in Leave votes per 1 percentage point increase, indicating the vote leaving propensity is lower in more

cosmopolitan areas. This is further strengthened by minus 0.37 percentage points decrease observed as the estimated effect of the share of non-EU migrants or 3.7 percentage points per 10 percentage points increase, which is significant. These findings support the hypothesis that the stock of migrants is a relevant predictor of the constituency level Leave vote or due to pre-2004 EU migrants selecting more affluent areas.

The third column reports the regression results when economic indicators are considered. The estimated effect of the media coverage on vote Leave percentage in the year up to the referendum becomes statistically insignificant, positive and smaller in magnitude at 0.001 percentage points (or 0.38 per standard deviation increase), indicating the negative estimates observed in the first two columns were underlined by economic disparities. Constituency newspaper readership is still negative and statistically significant even after accounting for migration and economic factors, although its estimated effect on the Leave vote decreased to -2.28 percentage points . All economic indicator variables are statistically significant and have the expected sign, further indicating that economic grievances were a more accurate predictor of the decision to vote to Leave the EU. A one percentage point increase in the unemployment rate at the 2011 census is estimated to have increased the propensity to vote to Leave by approximately 4.09 percentage points. Furthermore, a constituency level one percent decrease in unemployment in the year before the referendum is estimated to have decreased the Leave vote by 2.44 percentage points. This suggests that constituencies with better labour market conditions were less likely to vote to Leave in the 2016 referendum. Exposure to the EU programs was also significant, although small in magnitude. Here, a one percentage point decrease in EU funds allocated to the region lead to 0.046 percentage points increase in the Leave vote, or 1.41 per standard deviation. However, in some regions, the decrease may be attributed to the region becoming too wealthy to be eligible for some EU funds, but the estimated effects are small in magnitude. A one percent decrease in public spending in the year before the referendum has a similar effect as a one standard deviation decrease in media coverage, indicating a decrease in the Leave vote of2.22 percentage points. Given the politics of austerity begun by Conservative governments in 2010, cuts on public spending may be a proxy for economic deprivation. If the pattern of reduction is proportional, the estimates indicate that pre-existing patterns of deprivation at the time of the referendum, were relevant in predicting the propensity to vote Leave. All four estimates are statistically significant at 99% confidence level. Furthermore, the analysis suggests that once public spending patterns are accounted for, the estimates for total media coverage become statistically

insignificant. However, the estimated coefficient of the media coverage effect in column three might hide heterogeneous effects of media type coverage. All control variable groups are also jointly statistically significant.

The results on the migration indicators are similar to those observed in the second column. The estimated coefficient for the share of pre-2004 migrants is negative and statistically significant, albeit reduced in magnitude to a 1.17 percentage point reduction in the Lave vote per percentage point increase in this migrant share. This change between the second and the third column indicates possible migrant self-selection as the majority of the negative effect observed in column 2 was driven by economic factors. A similar observation can be made about the estimated coefficient of the share of migrants from post 2004 accession EU countries. The estimated coefficient is strongly statistically significant and strengthened relative to that in column two, indicating that a percentage point increase in the migrant share led to a 0.99 percentage points increase in Leave votes.

Table 2.4 below contains the results of equivalent regressions when the main explanatory variable is the weighted coverage of migration induced resource constraints in the year prior to the referendum. Results are similar to those found in Table 2.3. In the absence of migration and economic controls, a unit increase in weighted media coverage is estimated to increase Leave vote by 0.004 percentage points, or one percentage point per standard deviation increase. This result is statistically significant at 99% confidence interval and statistically higher than the migrant crime equivalent ($p=0.002$). When migration controls are taken into account, the estimated effect of media coverage remains statistically significant and increases to indicate a 1.5 percentage points increase in the Leave vote per standard deviation increase in this media coverage variable. As in the previous table, the pattern of results may be indicative of migrant self-selection in less or more affluent areas. When both migration and economic indicators are taken into account, the estimated effect of media coverage decreases and becomes statistically insignificant, indicating the previous estimates were driven by differences in economic determinants. However, this is the effect of aggregate coverage which does not distinguish reporting style and assumed position on EU membership. Therefore, the coefficient may mask heterogeneities in effect, which are analysed in subsequent subsections. Similarly, the effect of overall constituency is negative in all specifications and statistically significant in the last two, indicating that more informed constituencies were less likely to vote to Leave.

Looking at the effects of migrant shares in the second column, only the share of EU migrants from the post-2004 accession countries is estimated to have had a positive statistically

significant effect of 0.53 percentage points on the vote Leave percentage per percentage point increase in this share. On the other hand, the estimated effect on the pre-2004 accession EU migrants is strongly statistically significant and negative at minus 3.5 percentage points, indicating the vote leaving propensity is lower in more cosmopolitan areas. This is further strengthened by a minus 0.37 percentage points decrease observed as the estimated effect of the share of non-EU migrants or 3.7 percentage points per 10 percentage points increase. As before, these results may indicate migrant self-selection, with non-EU and pre-2004 EU nationals settling in areas that are more affluent. This is reinforced by the results of the migration coefficients in the last column. The estimated coefficient of a percentage point increase in the pre-2004 EU migrant stock decreased in magnitude to suggest a 1.17 percentage points decrease in Leave votes. However, the share of EU migrants from countries which joined the EU after 2004 is estimated to have increased the Leave vote by one percentage point per increase in this migrant share. This may be indicative of newer EU member state nationals being more prevalent in areas with more individuals 'targeted' by the Leave campaign and/or with worse economic conditions.

The estimated results on economic factors again indicate that it is economic factors rather than variation in media coverage which matters in explaining constituency level propensities to vote Leave. A one percentage point increase in the unemployment rate at the 2011 census is estimated to have increased the propensity to vote to Leave by approximately 4.09 percentage points, the equivalent of 2.5 standard deviations increase in the media variable from the previous specification. Furthermore, a constituency level 1 percent decrease in unemployment in the year before the referendum is estimated to have decreased the Leave vote by 2.43 percentage points. Exposure to the EU programmes was also significant, although small in magnitude. Here, a 1 percentage point decrease in EU funds allocated to the region lead to 0.045 percentage points increase in the Leave vote, or 1.67 per standard deviation increase, while a one percent decrease in public spending in the year before the referendum has a similar effect as a one standard deviation decrease in media coverage at minus 2.23 percentage points. Furthermore, previous analyses suggest that once public spending patterns are accounted for, the estimates for total media coverage become insignificant. All four estimates are statistically significant at 99% confidence level**.**

This underlines that economic grievances were, as predicted, important determinants of constituency level propensity to vote Leave, as also observed in Becker et al. (2017). Furthermore, the overall message emerging from the estimates of the effects of the economic indicators suggest that areas with worse labour market conditions and declining public finances

were more likely to vote Leave. The analysis of Fetzer (2018) comes to similar results, outlining that austerity measures taken by Conservative governments after 2010 were significant determinants of Leave vote. Given that the media effect estimates are insignificant once economic indicators considered all benchmark regressions, it can be ascertained that economic conditions have had a stronger effect on the propensity to vote Leave than written media coverage of crime involving migrants.

The signs and significance patterns of the constituency controls variables are similar to those observed in the previous table and in agreement with the study of Becker et al. (2017). In the migration controls category, it can be observed that only the stock of migrants from post-2004 EU countries was positively correlated with an increase in Leave votes. Since the skill distribution of this group is predominantly low skilled, they are competing with similarly skilled natives and putting pressure on their wages. This assertion, aimed at low-skilled natives, was used by Leave campaign rhetoric and noticed in previous economic literature (see Borjas (2003) and Borjas and Monras (2017)), which may cause the positive coefficient observed in Tables 2.3 and 2.4. A side result in the next subsection also outlines that a lower share of skilled natives correlates with higher Leave vote shares, but the estimated coefficient for the post-2004 EU migrant stock remains statistically significant and positive. Becker et al (2017) also finds that the proportion of post 2004 EU Nationals increased the predicted Leave vote, which brings evidence in favour of this hypothesis.

One concern with the results from third specifications in section 2.5.1 is the possibility that variation in media coverage is entirely explained by the remaining vector of covariates. This hypothesis is tested and the results are reported in Tables 2.8 and 2.9. It can be observed that for all weighted media variables there is still significant variation which is not explained by the remainder of the vector of covariates. The addition of interaction effects as seen in section 2.5.2 yields similar results.

Table 2.3: Estimation of the effect of weighted newspapers coverage of migrant crime in the year before the EU referendum:

Dependent variable: percentage of votes to Leave EU	(1)	(2)	(3)
Coverage variables			
Total weighted coverage of migrant crime	0.003***	0.004***	0.001
	(0.001)	(0.001)	(0.001)
Constituency controls			
Constituency readership	-0.818	-2.465**	-2.285***
	(1.054)	(0.967)	(0.882)
Leave voting intention	0.573***	0.484***	0.288***
	(0.031)	(0.031)	(0.025)
Share of people aged 30 to 44	-1.247***	-0.046	0.006
	(0.255)	(0.299)	(0.205)
Share of people aged 45 to 59	0.578*	-0.610*	-0.774***
	(0.314)	(0.358)	(0.251)
Share of people age over 60	-0.250	0.390**	0.749***
	(0.159)	(0.172)	(0.124)
Population density	-0.058**	0.093**	-0.017
	(0.027)	(0.037)	(0.028)
Male share	1.362***	1.831***	1.264***
	(0.479)	(0.527)	(0.340)
Share of married people	-0.044	0.148	0.614***
	(0.099)	(0.114)	(0.091)
Migration Controls			
Share of Pre 2004 EU immigrants		-3.560***	-1.172**
		(0.624)	(0.537)
Share of Post 2004 EU immigrants		0.540*	0.999***

		(0.286)	(0.255)
Share of non-EU immigrants		-0.372***	-0.478***
		(0.090)	(0.074)
Net migration in the last 5 years		-0.027	-0.016
		(0.054)	(0.044)
Economic Indicators			
Unemployment rate, census 2011			4.097***
			(0.301)
Decrease in unemployment rate (2015-16)			-2.445***
			(0.837)
Decrease in EU funding (% 2007-14 to			0.046***
2014-20)			(0.007)
Change in public spending (2015-16 %)			-2.228***
			(0.180)
GVA per head (£ 000s)			0.046**
			(0.020)
Change in wage median (2015-16 %)			0.092**
			(0.041)
Constant	-14.934	-50.734*	-51.758***
	(26.842)	(29.409)	(19.675)
Observations	632	632	632
R-squared	0.533	0.639	0.783

Notes to Table 2.3: All variables at constituency level except for EU funding and public spending variables. Dependent variable defined as 'Constituency level percentage voting 'Leave''. White robust standard errors in parentheses and significance level denoted by *** p<0.01, ** p<0.05, * p<0.1. 'Coverage type' defined as: 'migrant crime 1 year before the referendum'.

Table 2.4: Estimation of the effect of weighted newspapers coverage of migrant induced resource constraints in the year before the EU referendum:

Dependent variable: percentage of votes to Leave EU	(1)	(2)	(3)
Coverage variables			
Total weighted resource constraint coverage	0.004***	0.006***	0.002
	(0.002)	(0.002)	(0.001)
Constituency controls			
Constituency readership	-0.783	-2.427**	-2.282**
	(1.053)	(0.966)	(0.883)
Leave intention	0.572***	0.483***	0.288***
	(0.031)	(0.031)	(0.025)
Share of people aged 30 to 44	-1.249***	-0.048	0.009
	(0.255)	(0.299)	(0.205)
Share of people aged 45 to 59	0.578*	-0.616*	-0.774***
	(0.314)	(0.358)	(0.251)
Share of people age over 60	-0.249	0.389**	0.751***
	(0.159)	(0.172)	(0.124)
Population density	-0.060**	0.091**	-0.017
	(0.027)	(0.037)	(0.028)
Male share	1.365***	1.832***	1.262***
	(0.479)	(0.526)	(0.340)
Share of married people	-0.047	0.148	0.614***
	(0.099)	(0.114)	(0.091)
Migration Controls			
Share of Pre 2004 EU immigrants		-3.548***	-1.172**
		(0.624)	(0.536)
Share of Post 2004 EU immigrants		0.539*	1.002***
		(0.286)	(0.255)
Share of non-EU immigrants		-0.375***	-0.476***

		(0.090)	(0.074)
Net migration in the last 5 years		-0.028	-0.016
		(0.054)	(0.044)
Economic Indicators			
Unemployment rate, census 2011			4.094***
			(0.302)
Decrease in unemployment rate (2015-16)			-2.439***
			(0.837)
Decrease in EU funding (% 2007-14 to			0.045***
2014-20)			(0.007)
Change in public spending (2015-16 %)			-2.230***
			(0.180)
GVA per head (£ 000s)			0.046**
			(0.020)
Change in wage median (2015-16 %)			0.092**
			(0.041)
Constant	-15.063	-50.722*	-51.761***
	(26.842)	(29.368)	(19.680)
Observations	632	632	632
R-squared	0.533	0.640	0.783

Notes to Table 2.4: All variables at constituency level except for EU funding and public spending variables. Dependent variable defined as 'Constituency level percentage voting 'Leave''. White robust standard errors in parentheses and significance level denoted by *** p<0.01, ** p<0.05, * p<0.1. 'Coverage type' defined as: 'resource constraint articles 1 year before the referendum'.

Section 2.5.2: Analysis of heterogeneous effects by relative shares of qualified residents

Whereas the previous section outlined that the overall effect of negative media coverage preceding the EU referendum becomes insignificant while economic indicators are considered, this overall result may obfuscate heterogeneity in effects. For this purpose, this section analyses whether constituencies with different vectors of population qualifications differ in their predisposition to internalise media framing and uncover whether media coverage was more successful in constituencies with fewer highly qualified inhabitants.

Precise educational attainment data is not available in this context. The qualification variables used in this sub-section come from the 2011 census data, which includes information on highest qualification achieved, which in this case encompasses the share of individuals whose qualification level is at least level four. More details are available in the data section. Using the quartile distribution of this variable, four binary variables are created which equal one if the constituency lies in the 0 to 25^{th}, 25^{th} to 50^{th}, 50^{th} to 75^{th} or above the 75^{th} quantile respectively. These variables together with their interactions with the six months and one-year weighted media variables are included in the model to account for possible heterogeneity in effects, with the reference category set as the bottom 25^{th} percentile of population whose highest qualification level is at least four. Since the analysis in the previous section outlined that the initial estimated effects of media coverage were due to differences in economic and migration indicators, subsequent analysis focuses solely on the third specification from the previous section. In this specification, the coefficient for weighted media coverage estimates its effect on a reference constituency, while the following three interactions capture the additional effect of media coverage relative to a reference constituency.

In this section, the first column in a table will outline the regression results where the media variable is concerned with crime involving migrants, while the second will display the estimates coverage of migrant induced resource constraints in the year before the referendum. The results in column 3 are the ones obtained when the media coverage is focused on anti-EU coverage. Table 2.5 contains the results for regressions when the media duration is the year before the referendum, while Table 2.6 contains the results for regressions for when the media duration is in the last six months before. The reference category denotes constituencies whose share of ‘qualified’ inhabitants is below the first quartile.

The first column in Table 2.5 outlines the regression results when the media variable is weighted coverage of crime including migrants, outlining a relatively small increase of 0.003 percentage points per media article an average elector in a reference constituency is exposed

to in the year prior to the referendum. Similar results can be observed in the case of resource constraint variable in column 2, indicating an increase in Leave votes of 0.004 percentage points per article increase an average voter in a reference constituency 'observes'. The third regression suggests that anti-EU coverage was also associated with an increase in Leave votes of 0.005 percentage points per weighted media variable increase. All three results are statistically significant at the 95% confidence level, providing weak evidence that constituencies whose 'educated' electorate share is below the national 25th percentile are prone to media framing. Given the results for the first two interactions are statistically insignificant in all regressions while their combined effect is still positive, this argument can be extended to include all constituencies whose share of qualified inhabitants is below the 75th percentile. Additionally, the one-year estimates of media coverage for a reference constituency are significantly higher in the case of resource constraint and anti-EU than migrant crime (p=0.02), indicating the former coverages were influential on Leave votes. This may be indicative of sovereignty and economic issues being relatively more important in the decision to vote Leave.

The pattern of results indicates result reversal for constituencies in the top quartile of the qualification distribution. In the case of migrant crime, the estimate of the top quartile interaction term suggests a decrease relative to a reference constituency of 0.005 percentage points per unit increase in weighted media coverage. This effect is larger than the one observed in reference constituencies, but the total effect is not significantly different from zero (p=0.052), suggesting no effect of media at the top of the qualification distribution. Similar results can be observed in the second column in the case of resource constraint coverage. Here, the estimated coefficient of the third interaction term indicates a 0.007 and 0.008 percentage points relative decrease in Leave votes relative to benchmark constituencies per article an average elector is exposed to. This effect is again strong enough to reverse the sign observed in the benchmark constituency (p=0.016), and higher than the one observed in the case of migrant crime (p=0.033). The same pattern holds in the case of anti-EU coverage, where a reversal of effect is observed at the top of the qualification distribution.

Further evidence of voter sophistication arises from the estimated results of the first three variables in the constituency controls categories. The results for the qualification quartile dummies are all negative and strongly statistically significant. Furthermore, the pattern observed in all three specifications indicates that the estimated effect of being in a higher qualification quartile band is progressively higher in magnitude, relative to the reference case. A constituency situated in the second quartile is estimated to cast 4.76, 4.62 and 4.56 percentage points fewer votes relative to reference case in the three regressions respectively,

ceteris paribus. The results indicate that being the immediately higher band reduces the estimated Leave vote by roughly 8 percentage points for all coverage types, while being in the top quarter of the qualification quartile distribution reduces Leave vote by 12.09, 11.8 and 11.9 percentage points respectively relative to the reference category. These results bring evidence in support of the theory that constituencies that have a higher share of highly qualified inhabitants were less likely to vote to Leave. This is reinforced by the estimated effect of constituency readership, which is negative and statistically significant at 90% significance level. The results on the age bands indicate a Leave voting prevalence amongst the over-60 electors, whose estimated coefficient suggests a 0.53 percentage point increase in Leave votes per percentage point increase in the band.

The estimates of the migrant controls are qualitatively similar to those observed in the previous section. Only the migrant stock from the post-2004 accession countries is estimated to have had a positive statistically significant effect of roughly 0.75 percentage points on the vote Leave percentage per percentage point increase in the share. This is lower than and statistically different from the benchmark estimate for all types of coverage (p=0.09, p=0.09 and p=0.09 respectively). This and the estimated results on the qualification quartile binary variables suggest the positive effect estimated for post-2004 EU migrants is not exclusively due to self-selection. On the other hand, the estimated effect on the pre-2004 accession EU migrants is still strongly statistically significant but significantly decreased in magnitude, indicating an approximate 0.74 percentage points decrease in Leave votes per percentage increase in the pre-2004 EU migrant population. However, the results suggest I am unable to reject the null hypothesis that the coefficients are different from the benchmark case (p=0.27, p=0.28 and p=0.27 respectively). On the other hand, the equivalent null is rejected in the case of non-EU migrants, whose estimated coefficients increase relative to the benchmark case in all specifications. A one-percentage point increase in the non-EU migrant stock is estimated to have decreased Leave votes by 0.139, 0.137 and 0.141 percentage points in the three cases respectively.

Although there is some evidence of significant media effects in this section, the results on economic factors are again significant predictors of the constituency level Leave votes. A one percentage point increase in the unemployment rate at the 2011 census is estimated to have increased the propensity to vote to Leave by approximately 0.86, 0.85 and 0.87 percentage points in the migrant crime, resource constraint and aggregate coverage specifications respectively. All three results are statistically significant and significantly different then their benchmark counterparts, indicating a relative prevalence of ‘qualified’ individuals in

constituencies with less unemployment.

The results for the other economic indicators are similar to their benchmark equivalents and not statistically different from them. A constituency level 1 percent decrease in unemployment in the year before the referendum is estimated to have decreased the Leave vote by roughly 2.2 percentage points in all specifications. Exposure to the EU programs was also significant, although small in magnitude. Here, a 1-percentage point decrease in EU funds allocated to the region lead to 0.05 percentage points increase in the vote, or 1.67 per standard deviation in the EU funding variable. On the other hand, a one percent decrease in public spending in the year before the referendum correlates with a minus 2.3 percentage points decrease in Leave votes in all specifications. In magnitude, this is similar to roughly two standard deviations decrease in media coverage.

The results in Table 2.6 present the results obtained when the focus of the media variables is the last six months before the referendum. The first column in Table 2.6 outlines the regression results when the media variable is weighted coverage of crime including migrants, outlining a relatively small increase of 0.005 percentage points in the Leave vote per media article an average elector in a reference constituency is exposed to in the year prior to the referendum. This result is weakly statistically significant and higher than the equivalent estimate from Table 2.5 (p=0.004). Similar results can be observed in the case of resource constraint variable in column 2, indicating an increase in the Leave vote of 0.007 percentage points per article increase an average voter in a reference constituency 'observes'. The third regression suggests that anti-EU coverage was also associated with an increase in Leave votes, of 0.009 percentage points per weighted article increase. All three results are statistically significant at the 95% confidence level, providing weak evidence that constituencies whose 'educated' electorate share is below the national 25th percentile are prone to media framing.

Given the results for the first two interactions are statistically insignificant in all regression and that the combined effect is still positive, this argument can be extended to include all constituencies whose share of qualified inhabitants is below the 75th percentile. As with the estimate of migrant crime coverage, the 6 months estimates of weighted coverage in a reference constituency of both resource constraints and anti-EU are significantly higher than their one-year equivalents (p=0.009 and p=0.007 respectively). These results provide evidence of the availability bias hypothesis.

As in Table 2.5, the second and third column estimates of media coverage for a reference constituency are significantly higher in the case of resource constraint and anti-EU than migrant crime (p=0.006), indicating the former coverages were more associated with

Leave votes. This may be indicative of sovereignty and economic issues being relatively more important in the decision to vote Leave in both the year and six months before the referendum.

The qualification interaction estimates yield similar results to the one-year regressions, with the results becoming increasingly negative in each successive quartile. However, in all specifications, the interaction with the top quartile is negative and strongly statistically significant. In the case of migrant crime, the estimate of the top quartile interaction term suggests a decrease of 0.007 percentage points per unit increase in weighted media coverage. This effect is larger than the one observed in reference constituency and the total effect is weakly significantly different from zero (p=0.057), suggesting an opposite effect of media at the top of the qualification distribution. The top quartile interaction in the second column indicates a 0.012 percentage points decrease in Leave votes per article an average elector is exposed to. The overall media effect in a constituency in the top qualification quartile is statistically insignificant (p=0.08). The same pattern holds in the case of anti-EU coverage, where a negation of effect is observed at the top of the qualification distribution. Plausible explanations for the effect observed with the third interaction term in all six regressions may be voter discretion or a higher propensity to vote remain amongst 'qualified' individuals which negative coverage was unable to overturn. Additionally, constituencies with a higher share of qualified individuals were less likely targeted by the Leave campaign. The increasingly negative pattern of effect observed with the interaction terms in for all coverage supports the first theory.

The estimates of the migrant controls are qualitatively similar to those observed in the previous section and in Table 2.5. Only the migrant stock from the post-2004 accession countries is estimated to have had a positive statistically significant effect of roughly 0.75 percentage points on the vote Leave percentage per percentage point increase in the share, which is lower than and weakly statistically different from the benchmark estimate for all types of coverage (p=0.09 in all cases). This and the estimated results on the qualification quartile binary variables suggest the positive effect estimated for post-2004 EU migrants is not exclusively due to self-selection. On the other hand, the estimated effect on the pre-2004 accession EU migrants is still strongly statistically significant but significantly decreased in magnitude, indicating decreases of 0.79 or 0.8 percentage points decrease in Leave votes. However, the results suggest I am unable to reject the null hypothesis that the coefficients are different from the benchmark case (p=0.23, p=0.23 and p=0.24 respectively).

Although there is some evidence of significant media effects in this section, the results on economic factors are again significant predictors of the constituency level Leave votes. A

one percentage point increase in the unemployment rate at the 2011 census is estimated to have increased the propensity to vote to Leave by approximately 0.87, 0.85 and 0.86 percentage points in the migrant crime, resource constraint and aggregate coverage specifications respectively. All three results are statistically significant and significantly different then their benchmark counterparts, indicating a relative prevalence of 'qualified' individuals in constituencies with less unemployment. A constituency level 1 percent decrease in unemployment in the year before the referendum is estimated to have decreased the Leave vote by roughly 2.21 percentage points in all specifications. Exposure to the EU programs was also significant, although small in magnitude. Here, a 1-percentage point decrease in EU funds allocated to the region lead to 0.049 percentage points increase in the vote, or 1.67 per standard deviation in the EU coverage variable. On the other hand, a one percent decrease in public spending in the year before the referendum correlates with a minus 2.29 percentage points decrease in Leave votes in all specifications. In magnitude, this is similar to roughly two standard deviations decrease in media coverage.

The notable difference with respect to the benchmark results is the evidence of positive media effects in the bottom three quartiles of the qualification distribution, which persist once public spending patterns are considered. The estimated effects of economic indicators are still strongly statistically significant and important in magnitude. Overall, the effects of economic coefficients are largely identical between the regressions involving all types of negative coverage considered in this paper. Appendix A3 presents alternative formulations for the media variables on the coverage of migrant crimes and their (similar) corresponding results in Tables 2.13 and 2.14.

Table 2.5: Estimation of the effect of weighted newspapers coverage in the year before the EU referendum: heterogeneity by education quartiles

Dependent variable: percentage of votes to Leave EU	Coverage type: Migrant crime	Coverage type: Resource constraint	Coverage type: Aggregate anti EU
Coverage variables			
Total weighted media coverage	0.003**	0.004**	0.005***
	(0.001)	(0.002)	(0.002)
Interactions with education quartile dummy			
Interaction with Education Second quartile dummy	-0.000	-0.001	-0.001
	(0.002)	(0.003)	(0.003)
Interaction with Education Third quartile dummy	-0.002	-0.003	-0.003
	(0.002)	(0.002)	(0.002)
Interaction with Education Top Quartile dummy	-0.005***	-0.007***	-0.008***
	(0.002)	(0.002)	(0.002)
Constituency controls			
Second quartile dummy	-4.761***	-4.626***	-4.569***
	(1.216)	(1.252)	(1.150)
Third quartile dummy	-8.281***	-8.099***	-7.990***
	(1.257)	(1.288)	(1.203)
Top Quartile dummy	-12.095***	-11.871***	-11.945***
	(1.328)	(1.352)	(1.272)
Leave intention	0.175***	0.175***	0.175***
	(0.020)	(0.020)	(0.020)
Constituency readership	-1.302**	-1.307**	-1.333**
	(0.590)	(0.590)	(0.589)
Share of people aged 30 to 44	0.068	0.072	0.072
	(0.158)	(0.158)	(0.157)
Share of people aged 45 to 59	-0.080	-0.080	-0.081

	(0.207)	(0.207)	(0.206)
Share of people age over 60	0.531***	0.534***	0.534***
	(0.112)	(0.112)	(0.112)
Population density	0.014	0.015	0.014
	(0.018)	(0.018)	(0.018)
Male share	0.632**	0.631**	0.640**
	(0.275)	(0.275)	(0.275)
Share of married people	0.456***	0.455***	0.455***
	(0.069)	(0.069)	(0.068)
Migration Controls			
Share of Pre-2004 EU immigrants	-0.800***	-0.804***	-0.795***
	(0.267)	(0.267)	(0.267)
Share of Post 2004 EU immigrants	0.759***	0.761***	0.755***
	(0.151)	(0.151)	(0.151)
Share of non-EU immigrants	-0.139**	-0.137**	-0.141***
	(0.054)	(0.054)	(0.054)
Net migration in the last 5 years	-0.011	-0.012	-0.012
	(0.034)	(0.034)	(0.034)
Economic Indicators			
Unemployment rate, census 2011	0.865***	0.856***	0.877***
	(0.282)	(0.282)	(0.282)
Decrease in unemployment rate (2015-16)	-2.229***	-2.223***	-2.215***
	(0.637)	(0.637)	(0.635)
Decrease in EU funding (% 2007-14 to 2014-20)	0.049***	0.049***	0.049***
	(0.006)	(0.006)	(0.006)
Change in public spending (2015-16 %)	-2.317***	-2.315***	-2.289***
	(0.137)	(0.137)	(0.139)

GVA per head (£ 000s)	0.033***	0.033***	0.033***
	(0.012)	(0.012)	(0.012)
Change in wage median (2015-16 %)	0.049	0.048	0.048
	(0.034)	(0.034)	(0.034)
Constant	-4.925	-5.087	-5.650
	(15.870)	(15.863)	(15.827)
Observations	632	632	632
R-squared	0.869	0.869	0.869

Notes to Table 2.5: All variables at constituency level except for EU funding and public spending variables. Dependent variable defined as 'Constituency level percentage voting 'Leave''. White robust standard errors in parentheses and significance level denoted by *** p<0.01, ** p<0.05, * p<0.1.

Table 2.6: Estimation of the effect of weighted newspapers coverage in the 6 months before the EU referendum: heterogeneity by education quartiles

Dependent variable: percentage of votes to Leave EU	Coverage type: Migrant crime	Coverage type: Resource constraint	Coverage type: Aggregate anti EU
Coverage variables			
Total weighted media coverage	0.005**	0.007**	0.009**
	(0.002)	(0.003)	(0.004)
Interactions with dummy quartile variables			
Interaction with Second quartile dummy	-0.000	-0.001	-0.002
	(0.002)	(0.004)	(0.004)
Interaction with Third quartile dummy	-0.003	-0.005	-0.006
	(0.002)	(0.003)	(0.004)
Interaction with Top Quartile dummy	-0.007***	-0.012***	-0.015***
	(0.002)	(0.003)	(0.004)
Constituency controls			
Second quartile dummy	-4.724***	-4.425***	-4.491***
	(1.197)	(1.173)	(1.156)
Third quartile dummy	-8.187***	-7.854***	-7.925***
	(1.237)	(1.223)	(1.209)
Top Quartile dummy	-12.013***	-11.802***	-11.889***
	(1.304)	(1.290)	(1.276)
Leave intention	0.175***	0.175***	0.175***
	(0.020)	(0.020)	(0.020)
Constituency readership	-1.311**	-1.333**	-1.337**
	(0.590)	(0.589)	(0.589)
Share of people aged 30 to 44	0.072	0.073	0.072
	(0.158)	(0.157)	(0.157)
Share of people aged 45 to 59	-0.087	-0.075	-0.077

	(0.206)	(0.206)	(0.206)
Share of people age over 60	0.536***	0.533***	0.533***
	(0.112)	(0.112)	(0.112)
Population density	0.014	0.015	0.015
	(0.018)	(0.018)	(0.018)
Male share	0.647**	0.630**	0.634**
	(0.275)	(0.275)	(0.275)
Share of married people	0.454***	0.453***	0.454***
	(0.068)	(0.068)	(0.068)
Migration Controls			
Share of Pre-2004 EU immigrants	-0.791***	-0.806***	-0.800***
	(0.267)	(0.267)	(0.267)
Share of Post 2004 EU immigrants	0.753***	0.759***	0.757***
	(0.151)	(0.150)	(0.151)
Share of non-EU immigrants	-0.143***	-0.137**	-0.139***
	(0.054)	(0.054)	(0.054)
Net migration in the last 5 years	-0.012	-0.013	-0.012
	(0.034)	(0.034)	(0.034)
Economic Indicators			
Unemployment rate, census 2011	0.874***	0.857***	0.869***
	(0.282)	(0.282)	(0.282)
Decrease in unemployment rate (2015-16)	-2.246***	-2.219***	-2.221***
	(0.636)	(0.636)	(0.635)
Decrease in EU funding (% 2007-14 to 2014-20)	0.049***	0.049***	0.049***
	(0.006)	(0.006)	(0.006)
Change in public spending (2015-16 %)	-2.308***	-2.293***	-2.291***
	(0.136)	(0.139)	(0.139)

GVA per head (£ 000s)	0.033***	0.033***	0.033***
	(0.012)	(0.012)	(0.012)
Change in wage median (2015-16 %)	0.048	0.048	0.048
	(0.034)	(0.034)	(0.034)
Constant	-5.778	-5.233	-5.433
	(15.847)	(15.825)	(15.826)
Observations	632	632	632
R-squared	0.869	0.869	0.869

Notes to Table 2.6: All variables at constituency level except for EU funding and public spending variables. Dependent variable defined as 'Constituency level percentage voting 'Leave''. White robust standard errors in parentheses and significance level denoted by *** p<0.01, ** p<0.05, * p<0.1.

Section 2.5.3: Analysis of newspaper style

A third extension stems from the fact that original variables included all relevant articles regardless of the source and their respective reputation and credibility. Consequently, both sets of variables can be further split into national broadsheet, tabloid, and regional. Since the original variables included all three categories, their estimated coefficient is likely capturing an average effect, with a positive effect for tabloids and expected negative for broadsheets. Although a distinction exists between tabloid and broadsheet style reporting at the level of national publications, a clear distinction does not exist for regional newspapers. For this reason, the latter regional coverage remains a separate category.

For the purpose of this analysis, the year before the referendum and the 6 months prior are considered as separate cases for all coverage categories defined above. Four media variables are considered based on the type of the publication and their assumed position on the referendum. In Table 2.7, the first two columns contain the results obtained when the media coverage variable is migrant crime in the twelve and six-month intervals before the referendum respectively. The results from the resource constraint regressions are outlined the third and fourth columns, while the last two contain the results of the OLS regressions with anti-EU coverage media variables. Since the analysis indicates that the estimated effects of constituency, migrant, and economic controls are largely similar in both sign and magnitude to those observed in previous sections, their interpretation is omitted here.

Table 2.7: Estimation of the effect of newspapers coverage of migrant crime in the year and in the 6 months before the EU referendum: Heterogeneous effect by share of highly educated people in the constituency and by type of newspaper

VARIABLES	(1) Migrant Crime 1 year	(2) Migrant Crime 6 months	(3) Resource Constraint 1 year	(4) Resource constraint 6 months	(5) Anti-EU 1 year	(6) Anti-EU 6 months
Broadsheet coverage Remain	0.00771	0.00779	0.0118	0.00636	-0.0112	-0.0121
	(0.00477)	(0.00590)	(0.00806)	(0.0116)	(0.0180)	(0.0183)
Tabloid coverage Remain	-0.672***	-0.793***	-0.895***	-0.941***	-0.896**	-0.856**
	(0.144)	(0.181)	(0.220)	(0.259)	(0.366)	(0.367)
Broadsheet coverage Leave	0.0217	0.101***	0.0379	0.187**	0.286**	0.309**
	(0.0253)	(0.0316)	(0.0802)	(0.0793)	(0.139)	(0.138)
Tabloid coverage Leave	0.00722	-0.0199	0.0143	-0.0155	0.840**	0.791*
	(0.0139)	(0.0146)	(0.0231)	(0.0216)	(0.404)	(0.404)
Regional coverage	-0.0147	0.363	-0.170	-2.530**	-3.434***	-3.465**

weighted						
	(0.280)	(0.474)	(0.169)	(0.981)	(1.012)	(1.484)
Constituency Controls	X	X	X	X	X	X
Migration Controls	X	X	X	X	X	X
Economic Indicators	X	X	X	X	X	X
Observations	632	632	632	632	632	632
R-squared	0.761	0.760	0.761	0.763	0.767	0.763

Notes to Table 2.7: All variables at constituency level except for EU funding and public spending variables. Dependent variable defined as 'Constituency level percentage voting 'Leave''. White robust standard errors in parentheses and significance level denoted by *** p<0.01, ** p<0.05, * p<0.1. 'Coverage type' defined in the column heading. Tabloid, Broadsheet and Regional coverage represents coverage from the respective type of publication on the specified period. Qualification quartiles based on the share of individuals in the constituency whose maximum qualification level is at least four.

The estimation results suggest that broadsheet coverage from publications that endorsed Remain did not have a statistically significant effect in any specification while estimates have the expected sign in the last two specifications. On the other hand, broadsheet coverage originating in publications that declared their support for Leave is found to have had positive effects on the constituency level Leave vote, particularly in the last six months before the referendum. A unit increase in the weighted media coverage of migrant crime is estimated to have increased the Leave vote by 0.1 percentage points, while an equivalent increase in the resource constraint and anti-EU coverages led to increases of 0.18 and 0.3 percentage points respectively. Both the resource constraint and the anti-EU are weakly significantly larger than their migrant crime equivalent in column 2, with p-values of 0.08 and 0.054 respectively. However, the null hypothesis they are not significantly different from each other is not rejected (p=0.14), suggesting that, in the six months before the referendum, coverage of economic issues and sovereignty originating in broadsheets were relatively more important determinants of the Leave vote, even after relevant economic and migration indicators are considered. Broadsheet coverage from Leave supporting publications in the year prior to the referendum was a significant predictor of the vote only in the case of aggregate coverage. The estimated effect is a 0.28 percentage points increase per unit increase in the weighted media variable, which is statistically significant at 90% confidence interval and significantly lower than the coefficient in column 6 (p=0.02), providing evidence in support of availability bias.

On the other hand, the results suggest heterogeneity in the pattern of effects of tabloid coverage. Unlike broadsheet publications which supported Remain, the equivalent estimates of tabloid coverage are all negative and statistically significant for all types and both periods. A one article increase in weighted migrant crime coverage in the year before the referendum is estimated to have decreased Leave votes by 0.67 percentage points, and 0.79 percentage points in the six months equivalent. The six months estimate is significantly higher than the one in column 1 (p=0.05), bringing further evidence in support of the availability bias theory. The results for resource constraint coverage suggest that a unit increase weighted media coverage in the year before the referendum decreased the Leave vote by 0.89 percentage points. This result is strongly statistically significant and weakly statistically different from the one-year migrant crime equivalent (p=0.06), but the same does not hold in the six months case (p=0.15). The estimate of tabloid coverage about resource constraints in the last six months before the referendum suggest a decrease of .94 percentage points per unit increase in the media variable. However, I am unable to conclude that it is significantly higher than the yearly resource

constraint equivalent in column 3. A similar pattern is observed in columns 5 and 6 where the media focus is on anti-EU coverage, where the estimates suggest decreases of 0.89 and 0.85 percentage points in Leave votes per article an average elector is exposed to, for the year and six months regressions respectively. but I am unable to conclude the two estimates are significantly different (p=0.31). Given that the coverage originated in publications which supported Remain, the negative signs suggest successful media persuasion. The results in the last two columns suggest that in the case of Leave supporting tabloids, anti-EU coverage the only coverage found to be statistically significant.

Overall evidence on availability bias in this section is mixed since the estimate of the migrant crime coverage of is higher in the six months compared to the one-year case for Remain backing tabloids and, in the case of Leave backing broadsheets, for both migrant crime and resource constraint. However, this effect does not hold in the other pairwise comparisons.

Overall, the evidence suggests that in the six months before the referendum, the tabloid results suggest that resource constraint coverage was the most successful at persuading the electorate to vote remain if originating from a tabloid publication which supported Remain. Consequently, if a tabloid supported Leave, their most significant coverage type was anti-EU coverage. In the case of Leave supporting broadsheets, the estimates suggest that an extra article increase per average elector in the last six months was estimated to significantly increase Leave vote more for the resource constraint and anti-EU cases relative to migrant crime.

Section 2.5.4: Determinants of overall coverage: is there evidence of bias?

Table 2.8 below provides the results obtained when estimating the effects on the weighted media coverage variables of the other regressors used in the benchmark specification. In each column, the dependent variable is the coverage of type and period displayed at the top. Table 2.9 below displays the equivalent results when the dependent variables are weighted tabloid coverage variables. The estimation results indicate that in all specifications, for both types of coverage and both time periods, coverage intensity increases with population density and in all shares for all age bands, indicating a possible newspaper preference for individuals living in cities and those aged 30 or more. A one percentage point increase in the share of individuals aged between 30 to 44 is estimated to increase the number of circulated articles on migrant crime an average elector is exposed to by 33.5 in the year prior to the referendum and 22.98 in the 6 months prior, both of which are statistically significant at the 99% confidence interval. The estimated effects of the age bands are progressively lower as the age bands

increase, which indicates heterogeneity in newspaper demand across age groups. A similar pattern is observed in the case of resource constraint coverage and the magnitudes observed in columns 2 and 4 are consistent with the fact that the majority of relevant articles occurring in the last 6 months before the 23rd June 2016.

The estimated coefficients in the migration controls sections provide evidence that the migrant stock from the pre-2004 accession countries did not have a statistically significant effect on the level of coverage of any type. On the other hand, a 1-percentage point increase in the share of both non-EU and post-2004 EU migrants is found to have increased coverage intensity for all coverage types and in both periods. Of particular interest the case of migrant crime, for which a percentage point increase in the post-2004 EU migrant stock is estimated to have increased migrant crime coverage by 27.5 and 19.1 articles per average elector in the year and 6 months before the referendum respectively. Furthermore, the estimated coefficients of the non-EU migrant stocks are also positive and statistically significant, indicating increases of 22.4 and 15.6 articles an average elector is exposed to in the year and six months before the referendum respectively. The results in Table 2.9 indicate the same pattern of significance and sign in the case of tabloid coverage. We do not have access to local crime figures, but it can reasonably be expected that migrant crime is positively correlated with local migrant populations. The positive sign observed on these estimates provides weak evidence of the absence of bias in migrant crime reporting volume since the results validate the expected direction.

This result however does not preclude bias in reporting style, but this analysis is beyond the scope of this paper. The results in the last four columns indicate similar results of migrant stocks on the weighted resource constraint variables. Since these variables were constructed with a migrant focus in mind, the results strengthen the validity of the keyword choice. Additionally, these results do not provide evidence of bias in resource constraint reporting since the coverage is increasing with the migrant stock. The results of the economic indicators suggest that a 1-percentage point increase in public spending significantly decreased the weighted coverage variable in all four specifications. In the case of resource constraint coverage, this suggests that the overall media coverage is not biased in terms of volume, as the amount of coverage is inversely proportional to the increase in local funds.

The estimated results for tabloid coverage are qualitatively similar with respect to constituency and migration controls. One crucial difference with respect to the overall coverage regressions is that the estimates suggest that a 1-percentage point decrease in unemployment

rate increased by the weighted resource constraint coverage by 8.7 and 6.1 articles per average elector in the 1 year and 6-month cases respectively. This indicates an increase in tabloid coverage of resource constraints despite improving economic conditions.

Table 2.8: Estimation of the effect on total weighted media coverage type of the regressors

VARIABLES	(1) Migrant crime 1year	(2) Migrant crime 6 months	(3) Resource constraint 1 year	(4) Resource Constraint 6 months	(5) Anti-EU 1 year	(6) Anti-EU 6 months
Constituency controls						
Constituency readership	-55.241	-41.704*	-39.928*	-28.801*	-37.699*	-22.483*
	(36.215)	(23.807)	(23.498)	(15.891)	(21.739)	(12.811)
Share of people aged 30 to 44	33.566***	22.998***	21.486***	14.232***	20.345***	11.834***
	(9.498)	(6.244)	(6.163)	(4.168)	(5.702)	(3.360)
Share of people aged 45 to 59	27.783**	19.888**	19.246**	12.852**	17.809**	10.308**
	(12.752)	(8.383)	(8.274)	(5.595)	(7.655)	(4.511)
Share of people age over 60	15.357**	10.590**	9.692**	6.553**	9.295**	5.465**
	(6.721)	(4.418)	(4.361)	(2.949)	(4.034)	(2.377)
Population density	5.023***	3.554***	3.307***	2.203***	3.113***	1.808***
	(1.101)	(0.724)	(0.714)	(0.483)	(0.661)	(0.389)
Male share	-36.648**	-27.587**	-23.994**	-15.998**	-23.497**	-13.400**
	(17.108)	(11.246)	(11.100)	(7.507)	(10.269)	(6.052)
Share of married people	9.351**	5.981**	5.996**	4.090**	5.453**	3.238**
	(4.204)	(2.763)	(2.728)	(1.845)	(2.523)	(1.487)
Migration Controls						

Share of Pre- 2004 EU immigrants	-6.385	-4.217	-4.153	-2.110	-3.173	-1.865
	(16.621)	(10.926)	(10.785)	(7.293)	(9.977)	(5.880)
Share of Post 2004 EU immigrants	27.554***	19.123***	17.725***	12.043***	16.949***	9.889***
	(9.280)	(6.100)	(6.021)	(4.072)	(5.571)	(3.283)
Share of non-EU immigrants	22.244***	15.645***	14.394***	9.659***	13.899***	8.027***
	(3.222)	(2.118)	(2.090)	(1.414)	(1.934)	(1.140)
Net migration in the last 5 years	0.596	0.217	0.487	0.366	0.243	0.192
	(2.110)	(1.387)	(1.369)	(0.926)	(1.266)	(0.746)
Economic Indicators						
Unemployment rate, census 2011	-2.422	-2.890	0.549	0.848	-2.089	-0.562
	(17.546)	(11.534)	(11.385)	(7.699)	(10.532)	(6.207)
Decrease in unemployment rate (2015-16)	40.904	26.014	25.572	17.881	21.002	14.017
	(39.548)	(25.998)	(25.661)	(17.354)	(23.740)	(13.990)
Decrease in EU funding (% 2007-14 to	-0.368	-0.201	-0.201	-0.084	-0.125	-0.075
2014-20)	(0.380)	(0.250)	(0.247)	(0.167)	(0.228)	(0.134)
Change in public spending	-58.466***	-37.450***	-39.258***	-31.677***	-42.211***	-25.115***

(2015-16 %)						
	(7.709)	(5.068)	(5.002)	(3.383)	(4.628)	(2.727)
GVA per head (£ 000s)	-0.773	-0.533	-0.509	-0.353	-0.490	-0.287
	(0.728)	(0.479)	(0.473)	(0.320)	(0.437)	(0.258)
Change in wage median (2015-16 %)	1.138	0.711	0.705	0.467	0.663	0.368
	(2.149)	(1.413)	(1.394)	(0.943)	(1.290)	(0.760)
Constant	495.148	428.394	329.418	210.326	341.616	185.497
	(980.887)	(644.820)	(636.455)	(430.413)	(588.804)	(346.993)
R-squared	0.549	0.575	0.552	0.563	0.575	0.57

Notes to Table 2.8: All variables at constituency level except for EU funding and public spending variables. Dependent variable is the total weighted coverage of type and duration outlined in the column heading. White robust standard errors in parentheses and significance level denoted by *** p<0.01, ** p<0.05, * p<0.1

Table 2.9: The effects on weighted tabloid coverage of the qualification quartiles and controls.

VARIABLES	(1) Tabloid Migrant crime 1 year	(2) Tabloid Migrant crime 6 months	(3) Tabloid Resource constraint 1 year	(4) Tabloid Resource constraint 6 months	(5) Tabloid Anti-EU 1 year	(6) Tabloid-EU 6 months
Constituency controls						
Constituency readership	-9.016	-6.210	-9.740	-8.515	-7.516	3.406**
	(13.65)	(8.877)	(11.87)	(8.300)	(8.634)	(1.605)
Share of people aged 30 to 44	7.504**	4.924**	6.461**	4.467**	4.806**	2.129
	(3.212)	(2.089)	(2.785)	(1.946)	(2.265)	(2.155)
Share of people aged 45 to 59	4.907	3.162	4.483	3.271	2.989	1.582
	(4.246)	(2.766)	(3.679)	(2.573)	(3.040)	(1.136)
Share of people age over 60	3.484	2.248	3.087	2.185	2.228	0.409**
	(2.318)	(1.509)	(2.002)	(1.395)	(1.602)	(0.186)
Population density	0.910***	0.591***	0.812***	0.575***	0.579**	-0.402
	(0.338)	(0.220)	(0.295)	(0.207)	(0.262)	(2.891)
Male share	0.446	0.174	0.0292	-0.213	-0.598	1.910***
	(5.730)	(3.733)	(4.964)	(3.467)	(4.079)	(0.710)
Share of married people	4.212***	2.756***	3.629***	2.506***	2.717***	3.238**
	(1.247)	(0.813)	(1.082)	(0.757)	(1.002)	(1.487)

Migration Controls						
Share of Pre- 2004 EU immigrants	-2.072	-1.313	-1.489	-0.748	-0.316	-0.218
	(5.557)	(3.642)	(4.863)	(3.436)	(3.963)	(2.808)
Share of Post 2004 EU immigrants	6.416*	4.186*	5.709*	4.084*	4.157*	2.944*
	(3.610)	(2.348)	(3.112)	(2.164)	(2.213)	(1.568)
Share of non-EU immigrants	3.541***	2.323***	3.181***	2.312***	2.402***	1.715***
	(1.006)	(0.656)	(0.866)	(0.603)	(0.768)	(0.544)
Net migration in the last 5 years	1.210	0.786	1.056	0.736*	0.744	0.520
	(0.746)	(0.485)	(0.641)	(0.442)	(0.503)	(0.356)
Economic Indicators						
Unemployment rate, census 2011	9.808**	6.238**	8.729**	6.117**	5.716	3.962
	(3.955)	(2.575)	(3.412)	(2.376)	(4.183)	(2.965)
Decrease in unemployment rate (2015-16)	25.40**	16.51**	22.64**	16.00**	16.059*	11.209*
	(12.75)	(8.299)	(11.00)	(7.660)	(9.429)	(6.682)
Decrease in EU funding (% 2007-14 to	-0.193*	-0.115	-0.149	-0.0629	-0.043	-0.023
2014-20)	(0.116)	(0.0753)	(0.101)	(0.0720)	(0.091)	(0.064)

Change in public spending (2015-16 %)	-19.39***	-13.29***	-19.68***	-17.62***	-18.670***	-13.702***
	(2.358)	(1.533)	(2.007)	(1.365)	(1.838)	(1.303)
GVA per head (£ 000s)	-0.185	-0.120	-0.169	-0.126	-0.128	-0.091
	(0.396)	(0.259)	(0.347)	(0.244)	(0.174)	(0.123)
Change in wage median (2015-16 %)	0.492	0.328	0.405	0.274	0.291	0.205
	(0.814)	(0.530)	(0.703)	(0.490)	(0.512)	(0.363)
Constant	-237.2	-142.8	-199.6	-131.3	-112.434	-80.958
	(328.8)	(214.2)	(284.5)	(198.4)	(233.865)	(165.740)
R-squared	0.298	0.306	0.327	0.375	0.375	0.382

Notes to Table 2.9: All variables at constituency level except for EU funding and public spending variables. Dependent variable defined as 'Weighted tabloid coverage' of type and period outlined in the column heading. White robust standard errors in parentheses and significance level denoted by *** p<0.01, ** p<0.05, * p<0.1.

Section 2.6: Discussion and Conclusion

The fundamental result of this analysis is that economic indicators were the main predictors of Leave vote percentages, whereas total weighted media coverage had a relatively small and overall statistically insignificant effect on aggregate. The use of the weighted media variables proposed in this paper mitigates the drawbacks of the relevant article count variable (defined in Equation 1) used in previous literature such as Lamla and Lein (2008) and van der Wiel (2009), leading to more accurate estimates and thus eliminating bias resulting from measurement error.

The benchmark weighted media results indicate that newspaper coverage of migrant crime had an initial small, positive, and statistically insignificant effect on the constituency level percentage Leave votes regardless of the version of coverage used. Further analysis reveals heterogeneities of effects both from relative population qualifications and from coverage type. The estimates suggest that negative coverage had a small, statistically significant and positive effect in constituencies in the bottom three quartiles of the qualification distribution. This provides evidence in favour of information bias since overall media coverage had a statistically significant effect which persists after accounting for migration and economic indicators. However, this effect dissipates in the top quartile of the qualification distribution for all coverage types and all periods. Furthermore, the estimated results of the dummies indicating the relative qualification levels of the electorates are all negative and strongly statistically significant in all specifications. This suggests that the percentage of Leave votes is significantly lower in constituencies with higher proportions of highly educated people.

Analysis on publication type and stance reveals a heterogeneous pattern of effects. Coverage originating in broadsheets which favoured Remain was not found to have significantly altered the referendum vote. On the other hand, tabloids which supported Remain were successful at dissuading voters from voting Leave, which is significant in all specifications. Coverage originating in broadsheets which supported Leave is estimated to have increased the Leave vote percentage in the last six months before the referendum for all coverage types and in both periods in the case of anti-EU coverage. Anti-EU coverage is also found to be positive and statistically significant if originating in Leave-supporting tabloids, but the not if found in the other two definitions.

Generally, constituencies which experienced a decrease in public spending and EU funding and those which experienced increases in unemployment in the year preceding the

referendum were estimated to vote more in favour of exiting the EU. The estimated effect of an EU funding decrease was relatively small in magnitude, indicating a 0.045 to 0.049 percentage point increase in Leave vote percentages. A one percentage point decrease in public spending in the year before the referendum has strong effect in all specifications, larger in magnitude than 2.3 percentage points. A one percentage point increase in the unemployment rate at the 2011 census is estimated to have significantly increased the propensity to vote to Leave by approximately four percentage points in the benchmark specifications. The estimated effect drops to indicate an increase in Leave vote percentages of 0.8 points in subsequent analyses. Furthermore, a constituency level one percentage point decrease in unemployment in the year before the referendum is estimated to have decreased the Leave vote by roughly 2.2 to 2.4 percentage points in all specifications.

Looking at the effects of migrant shares, only the migrant stock from the post-2004 accession countries is estimated to have had a positive statistically significant effect on the propensity to vote Leave, which is consistent across specifications. On the other hand, the estimated effects of the pre-2004 accession EU migrants and non-EU migrants is still strongly statistically significant and negative, indicating the propensity to vote Leave is lower in areas that are more cosmopolitan and possible migrant self-selection. On the other hand, recent net migration is not a significant predictor of the percentage of Leave votes. Overall, the effects of economic coefficients are largely identical between the regressions involving all types of negative coverage considered in this paper. Although there is some evidence of significant media effects in this study, the results on economic factors are significant predictors of the constituency level Leave votes in all specifications. The results on migrant shares are similar to those obtained by Becker et al. (2017).

This study has two main limitations. The first is the endogeneity issue outlined above, since the resulting OLS estimates would be biased upwards. However, given that the results are already relatively small, this reinforces the finding that the overall media effect is insignificant in magnitude. The Leave voting intention variable obtained from the May 2015 British Election Study survey is added to account for potential differences in latent Euroscepticism and to alleviate concerns about endogeneity.

Although the main results of the paper indicated that media coverage had no significant overall effect, they suggest that coverage is more effective if it originates in a broadsheet publication supporting Leave or a tabloid supporting Remain. This finding is similar to that of

Chiang and Knight (2011) in the sense that credibility affects the effect of media coverage. These results are in line with the studies of Williams and Dickinson (1993), Banerjee et al. (2013), and Larreguy et al. (2014), among others, where being exposed to information can alter both the perception and voting actions of individuals in line with the message being relayed. Furthermore, this study finds evidence that not only are more qualified electorates less prone to media persuasion, but also that voters are aware of the reputation of a newspaper. This result is strengthened by Larreguy et al. (2014) and Chiang and Knight (2011), who also find evidence of voter sophistication along the lines of reputation and expected bias. A further analysis of this manner in the context of the 2014 Scottish independence referendum would be an interesting avenue for further research, with a specific focus on pro-independence coverage.

Chapter 3: An Analysis on Political Competition and Media Suppression

Section 3.1: Introduction

Beginning at the end of the last century, the world experienced a wave of democratisation, and, in the context of former Iron Curtain counties, freer media markets. The democratisation trend was believed to lead to less censorship and biased media, as they were seen as the purview of autocratic regimes. However, Besley and Prat (2006, page 720) note that in a modern context, 'despite the lack of old-fashioned pre-emptive censorship' economic means are increasingly used to restrict media freedom in order to gain favourable coverage. Furthermore, Corduneanu-Huci and Hamilton (2018) show that media restriction via both censorship and economic means is an increasing trend, 'affecting both democracies and autocracies, with negative consequences for the overall quality of political institutions and civil liberties'. Given these results and the backsliding on media freedom outlined empirically in Corduneanu-Huci and Hamilton (2018), it becomes interesting to analyse whether the media and political competition together are sufficient to increase media transparency and voter welfare. A study analysing the role of potentially biased media in the presence of imperfectly informed electorates is a relevant and interesting avenue of research given the recent attention to the perceived role of fake news in affecting the outcomes of democratic elections and referenda.

How can increased competition, both politically and within media market, lead to media suppression, media bias and dual information systems? Is voter welfare negatively affected? This paper proposes a novel theoretical approach that analyses the incentives of politicians and media outlets to misrepresent information to the electorate.

Previous literature suggests that media competition is generally an insufficient deterrent for media capture. Theoretical inquiry into the effect of the biased media market on electoral outcomes is a relatively recent academic development. There is ample literature regarding the perception of media bias from the perspective of the reader, as observed in Burke (2008), Besley and Prat (2006), and Sobrio (2013), etc. Of interest in the previous theoretical literature was whether an independent media could prevent or limit the incentives of elected politicians to engage in dishonest practices, yielding mixed results. With respect to the origin of bias, two main categories are distinguished: demand side and supply side. The overall message is that media market competition hinders supply driven bias and exacerbates demand driven bias. There are two types of bias distinguished in the literature: (outright) distortion of fact and information filtering. Focusing on media slant, Mullainathan and Shleifer (2005), Baron

(2006), Rudiger (2013), Sobrio (2013), and Várdy and Oliveros (2015) find that media competition does not alleviate bias, which in turn can lead to a suboptimal election result. The results of Besley and Prat (2006) and Anand et al. (2007) indicate that electoral outcomes improve with competition. Within this category, media slant can occur either via the demand side due to audience bias, or via the supply side due to the beliefs of the reporter or the reputational concerns of the editor. The closest papers to this study are those of Besley and Prat (2006) and Trombetta and Rossignoli (2020), who analyse political games of endogenous media capture. Both papers analyse the conditions under which media capture is feasible and desirable from the perspective of the incumbent, indicating that political media capture can occur despite media freedom. However, neither paper accounts for political competition and electoral spending laws, which could help or hinder media freedom.

This paper analyses whether political competition and electoral standards, here defined as having equal campaign budgets, help mitigate these effects and looks at the theoretical incentives of media suppression in more democratic settings, with novel results. This study aims to contribute to the burgeoning literature on the interplay between media, political outcomes, and the flow of information. The proposed model builds on the seminal paper of Besley and Prat (2006) and that of Trombetta and Rossignoli (2020) in two ways. Firstly, it introduces an active challenger in the political game, effectively emulating a simple majority democratic election or referendum. Secondly, it extends the analysis to account for binding and non-binding electoral spending laws. However, unlike Trombetta and Rossignoli (2020) and Besley and Prat (2006), the results indicate that in pure strategy Perfect Bayesian Equilibria, media and political competition and electoral standards ensure that the true state of politicians is always revealed to a subset of the electorate. However, if the proportion of voters who actively seek political news is sufficiently large, full disclosure occurs in the media market even if both politicians are bad. The present model can be seen as a generalisation of the previous models as it collapses to that of Trombetta and Rossignoli (2020) if the politicians' budgets are highly asymmetrical and there are no reputation concerns. With the additional assumption that the share of informed voters is zero, it collapses to the analysis of Besley and Prat (2006).

The remainder of this chapter follows the following structure. Section 3.2 provides a brief literature review. Section 3.3 outlines the actors and features of the model. I then analyse a limited game to pinpoint which features are paramount for the results. Section 3.5 analyses

the full games with media market credibility. Section 3.6 includes an outline of the results and discussions.

Section 3.2: Literature

The relationship between the media and electoral or political outcomes has been widely researched in the theoretical political economy literature, as outlined comprehensively in the reviews of Prat and Stromberg (2013) and Gentzkow et al. (2016). Two main strands of literature can be distinguished, namely supply driven bias and demand driven bias respectively. The former arises when media market bias arises from rational consumers who may prefer confirmatory news, as outlined in Chan and Suen (2008) and Duggan and Martinelli (2011) or gain psychological utility from reading reports whose stance is closer to the reader's prior beliefs, as in Mullainathan and Shleifer (2005). Alternatively, reporting closer to what the readers believe to be the true state of the world could be used as a signal of accuracy and trustworthiness by outlets in an attempt to increase their (future) demand, as outlined in the model of Gentzkow and Shapiro (2006). The overall message is that media bias persists despite increased media market competition, as misreporting may become a profit maximising decision for the outlets. Supply side bias arises when the outlets themselves have a political preference, a mobilisation incentive, or act under pressure from third parties such as politicians and advertisers. This is outlined in the models of Baron (2006), Besley and Prat (2006), Anderson and McLaren (2012) and Trombetta and Rossignoli (2020). Whereas these incentives can lead to bias even with rational consumers, increased competition tends to diminish bias and increase consumer welfare.

The issue of side payments or 'bribes' from the politicians to the media is motivated by evidence of the government ownership of media outlets and political courting of favourable coverage, as outlined empirically in Durante and Knight (2012) in Italy and Enikopolov et al. (2011, 2018) in Russia, and theoretically by Gehlbach and Sonin (2014). In explaining the effect of the media on political outcomes, two important aspects arise in the literature, namely the ideological motive and the profit motive. In this context, it may be a profit maximisation decision for outlets to misreport the correct 'state of the world' to cater to their customer base, even if they received the correct signal. An example of the latter appears in Anderson and McLaren (2012), but their main focus is on the market structure in the media market. An interesting feature is the behaviour of voters in this framework, especially since the voters do not know whether the signal the media reports is genuine or deceitful, or whether the outlet obtained the wrong signal themselves.

Chakraborty and Ghosh (2016) enhance probabilistic modelling with a Hotelling-Downs structure to construct an electoral model with potentially biased media, in which voters care about policy and a politician's character. The Hotelling-Downs model is not optimal in the context of this analysis due its focus on political position rather than competency. While some degree of ideological solidarity appears in this model, the ultimate determinant of relative second period voter utility is the competency of the elected politician, since not providing any public good is his optimal strategy. The probabilistic model employed in Várdy and Oliveros (2015) allows for the possibility that voters abstain. Here, voters have a pre-existing belief of the state of the world, which acts both as a measure of ideological conviction and as an input in a voter's decision on whether to consult the media. Their results suggest that the possibility of abstention and media competition lead to diminish the intensity of bias, but not remove it.

The closest models to the one analysed in this paper are those of Besley and Prat (2006) and Trombetta and Rossignoli (2020). Besley and Prat (2006) analyse a political game of endogenous media capture, in which an incumbent politician can attempt to 'capture' outlets such that they misreport their signal to the electorate. In this context, media capture is defined by a politician, who by economic or coercive means induces one outlet to misreport their signal. The authors analyse the conditions under which media capture is feasible and desirable from the perspective of the politician, indicating that the incumbent captures either all or none of the outlets. In this framework, the cost of capture is strictly increasing with competition as each outlet needs to be paid monopoly profits in equilibrium. Trombetta and Rossignoli (2020) extends the analysis by introducing heterogeneous voters, who can either intentionally follow political content or only see it when their outlet reports it. This results in heterogeneous levels of political capture if the share of informed voters is small enough, but the cost of capture becomes non-monotonic.

Other papers extending the model of Besley and Prat (2006) are Drufuca (2014) who endogenizes the information choice voters face, and Corduneanu-Huci and Hamilton (2018) who analyse the substitutability of censorship and capture via economic inducements. Neither paper considers both political and media market competition nor reputation concerns for the outlets. The results shown below show that heterogeneous media capture can arise in pure strategy Perfect Bayesian Equilibria in the presence of both political and media market competition, but total media suppression never occurs.

Section 3.3: Model

In this model, I attempt to retain as many features of the models of Besley and Prat (2006) and Trombetta and Rossignoli (2020) in order to provide a contrast to the result obtained therein. The proposed two period electoral model includes three actors, namely: two politicians, $n \in \mathbb{N}$ media outlets, and voters with mass normalised to 1. In order to win, a politician must acquire at least a 50% share of the total votes. In this respect, the following model is approximating a second run of a majority vote system. The voters have restricted access to information and are expected utility maximisers. This model further develops the work of Besley and Prat (2006), enhancing the model by not only adding heterogeneous voters, as seen in Trombetta and Rossignoli (2020), but also adding an active challenger politician who has identical incentives as the incumbent and heterogeneous media agents. Additionally, this model proposes an extension concerning reputation in the media market, which should decrease the probability of 'capture'. In their model, Besley and Prat (2006) analyse a two-period game in which an incumbent politician attempts to 'capture' media agents in order to determine what facilitates media capture as well as its welfare effects. Trombetta and Rossignoli (2020) extends their model by allowing voters to be heterogeneous in their attention and demand for political news. Unlike Besley and Prat (2006), this feature allows for heterogeneous levels of media capture to arise in pure strategy Perfect Bayesian Equilibria.

The game takes place in three stages, which occur successively: the political competition stage, the media market competition stage, and the voting stage. The political competition occurs successively, while media competition and voting are simultaneous.

Section 3.3.1: Politicians

There are two risk-neutral politicians in this model, namely the incumbent and the challenger; for exogenous reasons, the incumbent is in office at the beginning of the first period. In the rest of the paper, subscripts *I* and *C* will denote an incumbent and challenger specific variables respectively. The politicians can be either good or bad. Let $\theta_i = \{g, b\}$, denote the type of politician, where i indexes the politicians. Both politicians desire to win the election and be in power in the second period. Assume that the prior probability a politician is good is γ, i.e. Pr $\{ \theta_i = g\} = \gamma, i \in \{I, C\}$, and is known to all agents. This assumption is similar to that observed in Adachi and Hizen (2014).

The utility of either politician is independent of the administration they provide, but a good politician always choses the policy which maximises voter's welfare, while a bad

politician chooses a policy which yield no utility to the voters. The incumbent is in office at the start of the first period and has a positive and finite budget $B \in \mathbb{R}$ to spend on campaigning. The politicians know their type as well as their opponent's, both of which are unknown to voters, who need to rely on the media to obtain such information. Therefore, politicians can pay a vector of 'bribes' to the *n* media outlets, where $n \in \mathbb{N}$ is a finite number, in an attempt to persuade media outlets to report on them in a beneficial manner.[2] Unlike the studies of Besley and Prat (2006) and Trombetta and Rossignoli (2020), an honest incumbent would also have an incentive to offer a vector of payments towards media outlets to prevent a dishonest challenger from influencing these media outlets and potentially winning the election. Politicians observe their type and what signals the media agents received, as explained in detail below. Let $\tau_J, J \in \{I, C\}$ denote the vector of bribes offered by politician *J*, and $\tau_{Ji} \geq 0$ the proposed transfer from politician *J* to outlet *i*, such that

$$\tau_J = [\, \tau_{J1,}\, \tau_{J2},\, \tau_{J3}, \ldots,\, \tau_{Jn}]$$

The resulting utility of the politicians is given by:

$$U_I = \begin{cases} B - \sum_{i \in M} \tau_{Ii}\,, if\ he\ loses \\ 2B - \sum_{i \in M} \tau_{Ii}\ \ if\ he\ wins \end{cases},$$

Where *B* is the budget a politician can spend on electoral campaigning, *M* is the set of news agents, and τ_{Ii} is the monetary transfer aimed at outlet *i* from the incumbent. For the benchmark analysis, it is assumed that both politicians have the same budget. This assumption is relaxed in Section 6. Therefore, the utility of the challenger is given by:

$$U_C = \begin{cases} B - \sum_{i \in M} \tau_{Ci}\,, if\ he\ loses \\ 2B - \sum_{i \in M} \tau_{Ci}\ \ if\ he\ wins \end{cases}$$

For tractability reasons, politicians pay by discrete units of 1.[3] Additionally, there is an incumbency advantage in the sense that the incumbent wins in the event of a tied vote and he is preferred by both types of voters should voting for either politician yield the same expected utility.

[2] This will be defined below.

[3] This assumption helps in tiebreaking optimal decisions.

Section 3.3.2: Media

Media outlets can receive signals about the true nature of the politicians and are revenue maximisers. They are risk neutral and have two potential sources of revenue: advertising, which is proportional to their audience, or a possible bribe from either politician. This assumption is similar to those of Besley and Prat (2006) and Trombetta and Rossignoli (2020), where the bribe could originate only from the incumbent. They are revenue maximisers, and should outlets receive offers from both politicians, they would accept only the contribution that maximises their revenue. Should accepting either offer lead to the same expected revenue, they prefer the one that allows the outlet to remain honest. Their audience revenue is dependent on their readership, which in turn is determined by the actions of all media outlets. The revenue structure will be detailed in the next subsection.

Should an outlet accept the bribe, they must report favourably towards that politician. Let $s_k \in \{\Phi_k, b_k\}$ denote the signal received by media outlets about politician k. The Φ_k, signal is uninformative while b_k indicates the politician is bad. The uninformative signal is always received in the case of a good politician. The bad signal can only occur if the politician is bad, i.e. $\Pr(s_k = b_k|\theta_k = g) = 0$. Furthermore, let $\Pr(s_k = b|\theta_i = b) = q$, denote the probability that the outlets receive a bad signal about a bad politician and $\Pr(s = \Phi|\theta_i = b) = 1 - q$. Correspondingly, the signal set is given by $S_k = \{\Phi_I\Phi_C, b_I b_{C,} b_I\Phi_C, \Phi_I b_{C,}\}$. Given their signal and potential transfers received from politicians, outlets decide their optimal action. The outlets' actions can be either to accurately report the nature of the politicians or misreport the state of one of them. To simplify the analysis, the outlets are assumed to only be capable of reporting negative information about politicians after receiving a bad signal about them.

A novel feature of an extension of this proposed model is the introduction of reputation concerns in the media market. For each outlet, with probability λ, nature reveals at the beginning of the second period an outlet which suppressed a bad signal as dishonest. In this case, the guilty outlet loses all their audience and the associated revenue. Since it is optimal for the elected politician to keep the entire second period budget to himself, this revelation eliminates second period revenue from the media outlets revealed to be dishonest. The assumption that the $\lambda=0$ across media outlets types is made for the benchmark analysis. However, this assumption is not necessary to obtain the main results and will be revisited in a later section.

Section 3.3.3: Voters

Voters are risk-neutral and have an interest in an honest politician being elected, receiving no utility otherwise. Hence, voter utility is:

$$U_V = \begin{cases} 1 \; if \; a \; good \; politician \; is \; in \; office \\ 0 \; otherwise \end{cases}$$

For tractability reasons, the voters cannot observe their utility until after having cast their votes in the first period. This assumption is taken from the papers of Besley and Prat (2006) and Trombetta and Rossignoli (2020) and is introduced to provide a contrast to their results by adapting their models as little as possible. Recall that voters have heterogeneous interest in following political news. The first group, called attentive voters, are like those modelled in Besley and Prat (2006) in the sense that they will seek media that covers political news. Rationally inattentive voters are split equally between the competing outlets and get politically informed if their media outlet publishes political news in the form of a bad signal about either politician. Let α denote the share of attentive voters. Consequently, the share of rationally inattentive voters is 1-α. Voters know their own type and α is common knowledge.

Section 3.3.4: Benchmark case: No audience related revenues

In this section, I analyse the benchmark case in which audience related revenues are normalised to 0 irrespective of the signals received by media outlets and their actions. All other features of the model are maintained. This is the most difficult case, as the addition of audience revenues will reinforce the results since suppression becomes more difficult. For this section, we assume that the probability that a media outlet is discovered as dishonest in the second period, λ, is zero. Therefore, the only way in which outlets receive positive revenue is if they receive a positive transfer from the politician. In this section, I determine the main drivers of the equilibrium result. These are the inability of outlets to fabricate a bad signal, the consecutive movement in the politician game and constant budget political competition, rather than media market competition. I begin by analysing the Perfect Bayesian Equilibrium arising in this case, and then prove that under these general conditions, total media suppression never arises as an equilibrium outcome if politicians have equal budgets. This result cannot be extended to the unequal budgets case as sufficiently asymmetrical budgets approximate the absence of political competition in media capture. This collapses the analysis to that of Besley and Prat (2006) if no voter heterogeneity is considered, or Trombetta and Rossignoli (2020) if it is, under both of which total media suppression can occur.

Under this baseline scenario, the minimal acceptable bribe from any bad politician is 1. Formally, the payoffs for the outlets becomes:

Then the payoff structure of outlet i in the first period, P_{1i}, is the following:

$$P_{1i} = \begin{cases} \tau_{J,i} \; if \; it \; accepts \; the \; offer \; from \; politician \; j \\ 0 \; otherwise \end{cases}$$

$$P_{2i} = 0, \forall i$$

One inherent tie breaking assumption is made, namely that outlets prefer to report when they receive no offer from the politician. This renders truth telling the optimal action in the absence of political interference in the media market, which is in line with what is observed in the remainder of the paper, where, in the absence of political intervention, reporting audience revenues exceed non-reporting audience revenues.

The purpose of this section is to show that the absence of total media suppression result is driven by two factors: firstly, that media outlets can only report a politician is bad when they receive a signal and secondly, that political competition increases the cost of total media capture.

The timing of the game is:

- Nature determines the types of the politicians θ_i, where $i \in \{I, C\}$ and $\theta_i \in \{g, b\}$. Only the politicians know their type and their opponent's type prior to the signal. The prior probability that a politician is good is γ for both the incumbent and the challenger and is known by all agents. The shares of informed voters α is determined and is known to all agents.
- Outlets receive a composite signal about the state of the politicians, with each bad politician being revealed with probability q. The signal set outlets may receive is $\{\Phi_I \Phi_C, b_I b_{C,} b_I \Phi_C, \Phi_I b_{C,}\}$. The voters do not know that all outlets received the signal, and their prior belief that a politician is bad is γ. If a politician is revealed as bad, all media agents know it, but the voters do not. Additionally, both politicians know which has been exposed as bad
- Politicians decide successively on two vectors of bribes τ_I and τ_C towards the media outlets and they specify a reporting strategy towards the outlets they are bribing, with the incumbent moving first. These vectors are unknown only to the voters.
- Media outlets decide simultaneously whether to accept the bribe and report accordingly. The voters do not observe this stage and do not know which, if any, outlets have accepted

compensation from the politician. If a news outlet is approached by both politicians, its best response is to adopt the strategy which maximises the outlet's revenue. Should both politicians make equal acceptable offers, it will report the truth. Should the expected payoff after receiving donations from either politician match, they prefer the incumbent.

- Voters simultaneously observe the signal from the media according to their type and vote to maximise their expected utility.

Second period begins.

- The election results are announced along with politician types and are known to all agents.
- Voters simultaneously consume media and the game ends.

In the following sections, I identify pure strategy Perfect Bayesian Equilibria (PBE), which are outlined through a series of lemmas.

Proposition 1

In the Perfect Bayesian Equilibrium, the incumbent wins in cases $\Phi_I\Phi_C$, $\Phi_I b_C$, $b_I b_C$, while the incumbent wins if case $b_I\Phi_C$. Consequently, a good politician always wins the election if he is present and a bad incumbent defeats a bad challenger. Outlets report honestly if at in all cases. The minimum share of outlets a bad challenger needs to 'influence' to win the election is $\Gamma = \frac{\mathbf{0.5}}{\mathbf{1-\alpha}}$, unless both politicians are revealed as bad. If both politicians are bad, the number of outlets the incumbent needs to capture $\Gamma^* = \frac{\mathbf{0.5-\alpha}}{\mathbf{1-\alpha}} \leq \Gamma$.

Assume that outlets receive no readership revenues and that politicians have constant budgets.

Histories/Signal received by media	$\Phi_I\Phi_C$	$\Phi_I b_C$	$b_I\Phi_C$	$b_I b_C$
Prior Probability	$\gamma^2 + 2(1-\gamma)\gamma q + (1-\gamma)^2(1-q)^2$	$\gamma(1-\gamma)q + (1-\gamma)^2 q(1-q)$	$\gamma(1-\gamma)q + (1-\gamma)^2 q(1-q)$	$(1-\gamma)^2 q^2$
Agents and strategies				
Incumbent strategy	$\boldsymbol{\tau_{I,i}} = 0, \forall i,$ $\forall B, \forall\, E(\tau_C)$	$\boldsymbol{\tau_{I,i}} = 0, \forall i,$ $\forall B\ \forall\, E(\tau_C)$	$\boldsymbol{\tau_{I,i}} = 0, \forall i \quad \forall B, and\ \forall\, E(\tau_C)\ such\ that\ \boldsymbol{\tau_{C,i}} =$ **0** for $n + 1 - \lceil \Gamma n \rceil$ outlets Otherwise, a) if $B \geq \lceil \Gamma n \rceil$ $\boldsymbol{\tau_{I,i}} = \begin{cases} 1\ to\ \lceil \Gamma n \rceil \text{ outlets} \\ 0\ otherwise \end{cases}$ b) if $B < \lceil \Gamma n \rceil \tau^*\ \tau_{I,i} = 0, \forall i$	If α=0 and n>2 or if α>0 and $\frac{\lceil \Gamma^* n \rceil (\lceil \Gamma n \rceil + 1)}{\lceil \Gamma^* n \rceil - 2} < B$ $\tau_{I,i} = 0, \forall i$ If α=0, B≥ 1 and n=2 $\tau_{I,i} = B - 1\ to\ one\ outlet$ If $\frac{\lceil \Gamma^* n \rceil (\lceil \Gamma n \rceil + 1)}{\lceil \Gamma^* n \rceil - 2} \geq B$ and $B \geq \lceil \Gamma n \rceil$ $\boldsymbol{\tau_{I,i}} = \begin{cases} 1 + \frac{(l+k)1}{\lceil \Gamma^* n \rceil}\ to\ \lceil \Gamma^* n \rceil \text{ outlets} \\ 0\ to\ the\ rest \end{cases}$ If $\frac{\lceil \Gamma^* n \rceil (\lceil \Gamma n \rceil + 1)}{\lceil \Gamma^* n \rceil - 2} \geq B$ and $B \in [\lceil \Gamma^* n \rceil, \lceil \Gamma n \rceil)$ $\tau_{i,I} = \begin{cases} 1\ to \lceil \Gamma^* n \rceil\ outlets \\ 0\ to\ the\ rest \end{cases}$ Otherwise $\boldsymbol{\tau_{I,i}} = 0, \forall i$

Challenger strategy	$\boldsymbol{\tau_{C,i}} = 0, \forall i, \forall B, \forall\, \tau_I$	$\boldsymbol{\tau_{C,i}} = 0, \forall i \forall B, \forall\, \tau_I$	If $\tau_{I,i} < 1$ for $\lceil \Gamma n \rceil$ outlets, then $\boldsymbol{\tau_{C,i}} = 0, \forall i\ \forall B$. If $\tau_{I,i} \geq 1$ for x+$\lceil \Gamma n \rceil$ outlets, pick the x+1 outlets with the smallest $\tau_{I,i} \geq 1$, τ_{I,i^*}, and $$\boldsymbol{\tau_{C,i}} = \begin{cases} \boldsymbol{\tau_{I,i}} + 1\ for\ i \in \tau_{I,i^*} \\ 0\ otherwise \end{cases}$$	If α=0 or if α>0 and $\frac{\lceil \Gamma^* n \rceil (\lceil \Gamma n \rceil + 1)}{\lceil \Gamma^* n \rceil - 2} < B$ $$\boldsymbol{\tau_{C,i}} = \begin{cases} 1\ for\ \lceil \Gamma n \rceil\ outlets \\ 0\ otherwise \end{cases}$$ Otherwise $\boldsymbol{\tau_{C,i}} = 0\ , \forall i$ if $B < \lceil \Gamma n \rceil$ If $B \geq \lceil \Gamma n \rceil$, let x denote the number of outlets for which $\tau_{I,i}$=0 and order the remaining outlets ascendingly. Then if 1) x≥$\lceil \Gamma n \rceil$ or 2) x<$\lceil \Gamma n \rceil$ and $$B - x \geq \sum_{j=x+1}^{\lceil \Gamma n \rceil} \tau_{I,j} + (\lceil \Gamma n \rceil - x)$$ $$\tau_{C,i} = \begin{cases} 1\ to \min(x, \lceil \Gamma n \rceil)\ \ to\ outlets\ s.t\ \tau_{I,i} = 0 \\ \tau_{I,i} + 1\ to \text{ the first } \max(\lceil \Gamma n \rceil - x, 0)\ \ outlets \\ 0\ otherwise \end{cases}$$ Otherwise $\boldsymbol{\tau_{C,i}} = 0$
Electoral result	Incumbent wins	Incumbent wins	Challenger wins	Depends on α, n and B
Media Agents	Report honestly,$\Phi_I \Phi_C$, for $\forall\, \tau_I\ and\ \forall\, \tau_C$. If $\tau_{i,I} \neq 0$ and or $\tau_{C,i} \neq$	Report honestly $\Phi_I b_C$ if $\tau_{C,i} < 1, \forall\, \tau_I$	Report honestly $b_I \Phi_C$ unless	Report honestly $b_I b_C$ if either set of conditions holds

	0, they would accept the higher offer and still report honestly	Or if $\tau_{C,i} \geq 1$ and $\tau_I \geq \tau_{C,i} + 1$ Report $\Phi_I\Phi_C$ if $\tau_{C,i} \geq 1$ and $\tau_I < \tau_{C,i} + 1$	$\begin{cases} \tau_{I,i} \geq 1 \textbf{and} \\ \tau_{C,i} < \tau_{I,i} + 1 \end{cases}$ then report $\Phi_I\Phi_C$	$\tau_{I,i} < 1$ **and** $\tau_{C,i} < 1$ ***OR*** $\tau_{I,j} > 1$ ***and*** $\tau_{I,-j} > \tau_{I,j} + 1$ For $j \in \{I, C\}$ Report $\Phi_I b_C$ if $\tau_{I,i} \geq 1$ *and* $\tau_{I,i} \geq \tau_{C,i}$ Report $b_I\Phi_C$ if $\tau_{C,i} \geq 1$ *and* $\tau_{I,i} < \tau_{C,i}$
Informed Voters	Vote incumbent unless observe $b_I\Phi_C$ in the media	Vote incumbent unless observe $b_I\Phi_C$ in the media	Vote challenger if they observe $b_I\Phi_C$ in the media. Vote incumbent if they observe $\Phi_I\Phi_C$ in the media	Vote challenger if they observe $b_I\Phi_C$ in the media. Vote incumbent if they observe b_Ib_C or Φ_Ib_C in the media
Uninformed Voters	Vote incumbent unless observe $b_I\Phi_C$ from their outlet	Vote incumbent unless observe $b_I\Phi_C$ from their outlet	Vote challenger if they observe $b_I\Phi_C$ from their outlet. Vote incumbent if they observe $\Phi_I\Phi_C$ from their outlet	Vote challenger if they observe $b_I\Phi_C$ from their outlet. Vote incumbent if they observe b_Ib_C or Φ_Ib_C from their outlet
Equilibrium	$\boldsymbol{\tau_{I,i}} = 0, \forall i$, $\boldsymbol{\tau_{C,i}} = 0, \forall i$ Media outlets report honestly $\Phi_I\Phi_C$. Both types of voters vote incumbent.	$\boldsymbol{\tau_{I,i}} = 0, \forall i$, $\boldsymbol{\tau_{C,i}} = 0, \forall\, i$ Media outlets report honestly $\Phi_I b_C$. Both types of voters vote incumbent.	$\boldsymbol{\tau_{I,i}} = 0, \forall i, \boldsymbol{\tau_{C,i}} = 0, \forall\, i$ Media outlets report honestly $b_I\Phi_C$. Both types of voters vote challenger.	If α=0 and n>2 or if α>0 and $\frac{\lceil \Gamma^* n \rceil (\lceil \Gamma n \rceil + 1)}{\lceil \Gamma^* n \rceil - 2} < B$ $\tau_{I,i} = 0, \forall i\, \boldsymbol{\tau_{C,i}} = 0, \forall\, i$ If α=0, B≥ 1 and n=2 $_{I,i} = B - 1$ *to one outlet*, $\boldsymbol{\tau_{C,i}} = 0, \forall\, i$ If $\frac{\lceil \Gamma^* n \rceil (\lceil \Gamma n \rceil + 1)}{\lceil \Gamma^* n \rceil - 2} \geq B$ and $B \geq \lceil \Gamma n \rceil$

				$\boldsymbol{\tau_{I,i}} = \begin{cases} 1 + \frac{(l+k)1}{\lceil \Gamma^* \mathrm{n} \rceil} \ to \ \lceil \Gamma^* \mathrm{n} \rceil \text{ outlets} \\ 0 \ to \ the \ rest \end{cases}$ $\boldsymbol{\tau_{C,i}} = 0, \forall \ i$ If $\frac{\lceil \Gamma^* \mathrm{n} \rceil (\lceil \Gamma \mathrm{n} \rceil + 1)}{\lceil \Gamma^* \mathrm{n} \rceil - 2} \geq B$ and $B \in [\lceil \Gamma^* \mathrm{n} \rceil, \lceil \Gamma \mathrm{n} \rceil)$ $\tau_{i,I} = \begin{cases} 1 \ to \lceil \Gamma^* \mathrm{n} \rceil \ outlets \\ 0 \ to \ the \ rest \end{cases}$ $\boldsymbol{\tau_{C,i}} = 0, \forall \ i$ Media outlets report honestly $b_I b_C$. Both types of voters vote incumbent.

Lemma 1:

In any Perfect Bayesian Equilibrium with weakly undominated strategies, voters will prefer to vote for the incumbent unless he is revealed as bad and the challenger is not.

Proof:

For the purposes of this proof, a distinction between how informed and uninformed voters update their beliefs needs to be made.

From the perspective of any voter:

$$\Pr(\theta_I = g | s_i = b_I b_C) = 0 < \gamma$$

$$\Pr(\theta_I = g | s_i = b_I \Phi_C) = 0 < \gamma$$

$$\Pr(\theta_C = g | s_i = b_I b_C) = 0 < \gamma$$

$$\Pr(\theta_C = g | s_i = \Phi_I \mathrm{b}_C) = 0 < \gamma$$

which results from the structure of the game for both politicians and for both voter types and all histories. Therefore, upon observing a bad signal, a voter knows the politician is bad and that he would get utility of 0 if said politician is in office. It is therefore an optimal decision for the voter to not vote for a politician revealed as bad. If both politicians are revealed as bad, the voters prefer the incumbent due to the incumbency advantage.

Voters are aware that bad politicians could attempt to 'silence' outlets, but they update their beliefs differently. An attentive voter acts like one in Besley and Prat (2006) and can be misled only by total media suppression.

Applying Bayes' rule for both the incumbent and the challenger from the perspective of an informed voter based on the politician strategies outlined above:

$$\Pr(\theta_I = g | s_i = \Phi_I \Phi_C) = \frac{\Pr(\theta_I = g)\Pr(s_i = \Phi_I \Phi_C | \theta_I = g)}{\Pr(s_i = \Phi_I \Phi_C)}$$

$$= \frac{\gamma[\gamma + (1-\gamma)(1-q)]}{\gamma^2 + (1-\gamma)^2(1-q)^2 + 2\gamma(1-\gamma)(1-q)}$$

Similarly, about the challenger:

$$\Pr(\theta_C = g | s_i = \Phi_I \Phi_C) = \frac{\Pr(\theta_I = g)\Pr(s_i = \Phi_I \Phi_C | \theta_I = g)}{\Pr(s_i = \Phi_I \Phi_C)}$$

$$= \frac{\gamma[\gamma + (1-\gamma)(1-q)]}{\gamma^2 + (1-\gamma)^2(1-q)^2 + 2\gamma(1-\gamma)(1-q)}$$

This is because an informed voter observes a pair of null signals in the following three situations only: both politicians are good, both politicians are good and undiscovered, or one politician is good and the other is bad and undiscovered. Therefore, given politician *j* is good, a null signal can be observed only if a bad signal is not observed about politician *-j*.

Now consider Bayesian updating for an informed voter when they observe signal $b_j \Phi_{-j}$

$$\Pr(\theta_j = g | s_i = b_j \Phi_{-j}) = 0 < \gamma$$

$$\Pr(\theta_j = g | s_i = \Phi_j b_{-j}) = \frac{\Pr(\theta_j = g)\Pr(s_i = \Phi_j b_{-j} | \theta_j = g)}{\Pr(s_i = \Phi_j b_{-j})}$$

$$= \frac{\gamma q(1-\gamma)}{\gamma q(1-\gamma) + q(1-\gamma)^2(1-q)} > 0$$

This happens because upon observing signal $\Phi_j b_{-j}$, informed voters know with certainty that politician *-j* is bad, while there is a strictly positive probability that politician *j* is good. Consequently, they would prefer politician *–j* as $\Pr(\theta_{-j} = g | s_i = \Phi_j b_{-j}) = 0$.

Because of the incumbency advantage assumption, informed voters prefer to vote for the incumbent should they observe $b_I b_C$. Thus, the only scenario in which an informed voter prefers the challenger to the incumbent is if the expected utility from electing the challenger exceeds that of electing the incumbent, i.e. after observing $b_i \Phi_c$

I now analyse how an uninformed voter updates their belief. As above, after observing a bad signal, the posterior belief that the politician is bad becomes one. The difference with respect to the informed voters is that in case $b_I b_C$, some rationally inattentive voters can be subscribed to an outlet which misreports one of the politicians. The winning politician depends on the realisation of B and α, which in turn determine the fraction of outlets suppressed in equilibrium.

Let n denote the total number of outlets. Therefore, in both histories, the probability that an inattentive voter reads a captured outlet is $\frac{\lceil \Gamma n \rceil}{n}$ or $\frac{\lceil \Gamma^* n \rceil}{n}$, where m_i and m_c denote the number of outlets silenced by the incumbent and challenger respectively. For the incumbent:

$$\Pr(\theta_I = g | s_i = \Phi_I \Phi_C) = \frac{\Pr(\theta_I = g)\Pr(s_i = \Phi_I \Phi_C | \theta_I = g)}{\Pr(s_i = \Phi_I \Phi_C)}$$

$$= \frac{\gamma[\gamma + (1-\gamma)(1-q)]}{\gamma^2 + (1-\gamma)^2(1-q)^2 + 2\gamma(1-\gamma)(1-q)}$$

Similarly, for the challenger:

$$\Pr(\theta_C = g | s_i = \Phi_I \Phi_C) = \frac{\Pr(\theta_I = g)\Pr(s_i = \Phi_I \Phi_C | \theta_I = g)}{\Pr(s_i = \Phi_I \Phi_C)}$$

$$= \frac{\gamma[\gamma + (1-\gamma)(1-q)]}{\gamma^2 + (1-\gamma)^2(1-q)^2 + 2\gamma(1-\gamma)(1-q)}$$

As with informed voters, we can observe that $\Pr(\theta_I = good | s_{i,I} = \Phi) = \Pr(\theta_C = good | s_{i,I} = \Phi)$, and due to the incumbency advantage assumption, both informed and rationally inattentive voters prefer to vote for the incumbent after observing a pair of null signals.

Now consider Bayesian updating for a rationally inattentive voter when they observe signal $b_I \Phi_C$. Since outlets cannot fabricate a bad signal, it follows that

$$\Pr(\theta_I = g | s_i = \mathrm{b}_I \Phi_C) = 0$$

In the case of the challenger:

$$\Pr(\theta_C = g | s_i = \mathrm{b}\ \ \Phi_C) = \frac{\Pr(\theta_C = g)\Pr(s_i = \mathrm{b}_I \Phi_C | \theta_C = g)}{\Pr(s_i = \mathrm{b}_I \Phi_C)}$$

$$= \frac{\gamma(1-\gamma)q}{\gamma(1-\gamma)q + (1-\gamma)^2 q(1-q) + \omega(1-\gamma)^2 q^2 \frac{\lceil \Gamma n \rceil}{n}} > 0$$

Where ω denotes the probability that α>0 and $\frac{\lceil \Gamma^* n \rceil (\lceil \Gamma n \rceil + 1)}{\lceil \Gamma^* n \rceil - 2} < B$. This expression is strictly larger than 0 since $\gamma > 0$, $\omega > 0$, and $q > 0$. Therefore, an inattentive voter will prefer to vote for the challenger upon observing $\mathrm{b}_I \Phi_C$. A similar argument occurs for the $\Phi_I b_C$ signal. Since

outlets cannot fabricate a bad signal, it follows that Pr $(\theta_C = g|s_i = \Phi_I b_C) = 0$. In the case of the incumbent, we have:

$$\Pr(\theta_I = g|s_i = \Phi_I b_C) = \frac{\Pr(\theta_I = g)\Pr(s_i = \Phi_I b_C|\theta_I = g)}{\Pr(s_i = \Phi_I b_C)}$$

$$= \frac{\gamma(1-\gamma)q}{\gamma(1-\gamma)q + (1-\gamma)^2 q(1-q)} > 0$$

This occurs because voters only observe $\Phi_I b_C$ in three situations. These are: the incumbent is good while the challenger is bad and discovered; both politicians are bad and only the challenger is discovered; or if they are both bad and discovered and the voters read one of the outlets bribed by the incumbent when neither α=0 and n>2 nor α>0 and $\frac{\lceil \Gamma^* n \rceil (\lceil \Gamma n \rceil + 1)}{\lceil \Gamma^* n \rceil - 2} < B$ hold. Because of the incumbency advantage assumption, rationally inattentive voters prefer to vote for the incumbent should they observe $b_I b_C$. Thus, the only scenario in which a rationally inattentive voter prefers the challenger to the incumbent is if the expected utility from electing the challenger exceeds that of electing the incumbent, i.e. after observing $b_i \Phi_c$.

QED Lemma 1■

Lemma 2:

If the budget is constant across politicians and a bad signal is received about only one politician, then total media suppression will not occur in equilibrium.

Proof:

In cases $\boldsymbol{g_i g_c}$ there is no scope for media interference. In case $\boldsymbol{g_i b_c}$, following Lemma 1, media interference from the challenger is ineffective since both types of voters prefer the incumbent if no signal is observed. It remains to show that total media suppression will not occur in case $\boldsymbol{b_i g_c}$.

Suppose ∃ is a vector of payment from the incumbent to the media outlets τ_I such that achieves total media suppression. Therefore, it must be the case that it is a winning strategy for the incumbent (but not necessarily optimal) and:

$$\sum_{i=1,..,n} \tau_{i,I} \leq B \; and \; \tau_{Ii} \geq 1 > 0 \; \forall i$$

Three cases are distinguished depending on the realisations of B and α:

1. $B \leq n$

Under this realisation of B, total media suppression is impossible as the incumbent does not have enough funds to achieve it even if the challenger is inactive.

2. $B \geq n \; and \; \alpha < 0.5$

If the challenger persuades at least one outlet to report honestly, he gains a share of α votes from attentive voters and rationally inattentive votes proportional to the share of honest outlets. In order for the challenger to win the election, he also needs 0.5- α votes from rationally inattentive voters. Therefore, allowing m to denote the number of outlets the challenger 'manipulates'

$$\frac{(1-\alpha)m}{n} \geq 0.5 - \alpha$$

$$m/n \geq \frac{0.5-\alpha}{1-\alpha} \rightarrow \boldsymbol{\Gamma^* = \frac{0.5-\alpha}{1-\alpha}}$$

Given this action from the incumbent and any realisation of B α such that $B \geq n\varepsilon \; and \; \alpha < 0.5$ the best response of the challenger is therefore to approach the outlets with the $\lceil \Gamma^* n \rceil$ lowest offers and offer

$$\tau_{C,i} = \boldsymbol{\tau_{I,i}} + 1$$

Since the politicians share the same budget, such an offer is affordable for the challenger and it ensures his electoral win. This contradicts the original hypothesis.

3. $\alpha \geq 0.5$

If the challenger persuades at least one outlet to report honestly, he gains a share of α votes from attentive voters and rationally inattentive votes proportional to the share of honest outlets. Therefore, to win the election, the challenger needs one outlet to report honestly. Total media suppression by the incumbent is obviously a suboptimal strategy in equilibrium since both politicians share the same budget and the incumbent needs to silence all n outlets while the challenger needs to focus on only one.

QED Lemma 2 ■

Media signals

Lemma 2 indicates that a good politician always prevents a bad politician from achieving total media suppression. Given the minimum number of 'manipulated' outlets required to win, there is a disparity in the politicians' disposable income after the minimum media suppression cost given opponent inaction. This will affect the equilibrium outcome when both politicians are exposed as bad. For $\Gamma = \frac{0.5}{1-\alpha}$ and $\Gamma^* = \frac{0.5-\alpha}{1-\alpha}$, the relevant condition will be given by:

$$\lceil \Gamma n \rceil - \lceil \Gamma^* n \rceil \geq 1$$

If the condition above holds, denote:

$$k = (\lceil \Gamma n \rceil - \lceil \Gamma^* n \rceil), k \in \mathbb{N} \; and \; k \geq 1$$

such that:

$$\boldsymbol{B - \lceil \Gamma^* n \rceil - (B - \lceil \Gamma n \rceil) = k}$$

Additionally, the politician budgets can be decomposed as:

$$\boldsymbol{B - \lceil \Gamma n \rceil = l, l \in \mathbb{N} \; and}$$

$$\boldsymbol{B - \lceil \Gamma^* n \rceil = (k + l), k, l \in \mathbb{N}}$$

[Eq. 7]

Case 1: $\boldsymbol{g_I g_C}$

This can occur if both politicians are good or if either politician is bad and no signal is received by the media. In the absence of a signal, the media outlets are unable to report anything and therefore media framing is impossible. Since no signal is received, the best response of all outlets is to report honestly, even if they receive a transfer from either politician. Both sets of voters observe $\Phi_I \Phi_C$ and their posterior beliefs the politicians are good are both higher than γ. As shown in Lemma 1, they both prefer the incumbent. For both politicians, strictly positive transfers do not change the result of the election and reduce utility and hence their equilibrium actions will be to set $\tau_{I,i} = \tau_{C,i} = 0$. From the perspective of media outlets, their reporting strategy can only be honest even in the presence of positive transfers from the politicians, and therefore would accept any positive transfer, if offered. The incumbent wins and obtains utility $2B$, the challenger 0 and all media outlets report honestly. These equilibrium strategies and beliefs are consistent for all realisations of B, n δ, and α.

Case 2: $\boldsymbol{g_I b_C}$

This can occur if the challenger is bad and discovered with probability γ (1-γ)q or if both politicians are bad and only the challenger is discovered, with probability $(1-\gamma)^2 q(1-q)$. With probability γ (1-γ) (1-q) the incumbent is good, while the challenger is bad and undiscovered, in which case equilibrium actions are as outlined in $\boldsymbol{g_I g_C}$. From the voters' perspective, this is the posterior belief induced if they observe $\Phi_I b_C$ in the media or from their preferred outlet.

The bad challenger may attempt to silence the media, in which case more than half the voters do not observe the bad signal about the challenger. However, upon seeing no signal about either politician, both types of voters still prefer to vote for the incumbent. An informed voter would revise his posterior belief to 1, whereas an uninformed voter would still prefer the incumbent. Since media persuasion is both costly and ineffective, it is a strictly suboptimal strategy. Therefore $\boldsymbol{\tau_{C,i}} = 0, \forall i$ and the best response of the incumbent is to also set $\boldsymbol{\tau_{I,i}} = 0, \forall$i. The media outlets then report honestly $\Phi_I b_C$. Should they receive an offer larger at least as large as 1 and $\tau_{I,i} < \boldsymbol{\tau_{C,i}} + 1$, the media outlets would decide to report $\Phi_I \Phi_C$. The incumbent wins and obtains utility 2B, while the challenger obtains 0. A special case where the challenger can suppress the media completely exists, but it is a strictly suboptimal strategy to do so. These equilibrium strategies and beliefs are consistent for all realisations of B, n, δ and $\alpha < 0.5$

Case 3: $\boldsymbol{g_C b_I}$

This can occur if the incumbent is bad and discovered with probability γ (1-γ)q or if both politicians are bad and only the incumbent is discovered, with probability $(1-\gamma)^2 q(1-q)$.With probability γ (1-γ)(1-q) the challenger is good, while the incumbent is bad and undiscovered, in which case equilibrium actions are as outlined in $\boldsymbol{g_I g_C}$. From the voters' perspective, this is the posterior belief induced if they observe $b_I \Phi_C$ in the media or from their preferred outlet.

We know from the properties of Γ that in this scenario, the bad politician needs to manipulate at least half of the media outlets in order to avoid certain defeat in the election.

In this case, the incumbent has B- $\lceil \Gamma n \rceil$ revenue left over from attempting to manipulate $\lceil \Gamma n \rceil$ outlets at the minimum acceptable 'transfer'. Because of the properties of Γ, $n - \lceil \Gamma n \rceil \leq \lceil \Gamma n \rceil$, i.e. the number of outlets the challenger needs to report honestly is smaller than the number of outlets the incumbent needs to 'manipulate'.

By construction, an offer of $\boldsymbol{\tau_{I,i}} + 1$ is acceptable from the challenger and preferable to an offer of $\boldsymbol{\tau_{I,i}}$from the incumbent. It is a best response for any outlet approached by the good politician with this offer to accept it. Consequently, any outlet approached with $\tau_{i,I} \geq 1$ by the incumbent would accept only if the challenger offers less than $\boldsymbol{\tau_{I,i}} + 1$. Given the fact that the challenger moves second and has to convince fewer outlets to report honestly, a challenger strategy that ensures his victory exists for every vector of payments $\boldsymbol{\tau_I}$. Therefore, the equilibrium action for the incumbent is to set $\boldsymbol{\tau_{I,i}} = \mathbf{0}$ for all outlets and for all *B*. Consequently, the equilibrium strategy for the challenger is to set $\boldsymbol{\tau_{C,i}} = \mathbf{0}$ if the incumbent has offered $\tau_{i,I} \geq 1$ to less than $\lceil \Gamma n \rceil$ outlets. If $\tau_{I,i} \geq 1$ for $x \geq \lceil \Gamma n \rceil$ outlets, the optimal action of the challenger is to ensure that only $\lceil \Gamma n \rceil - 1$ outlets remain 'manipulated'. This can be achieved by picking the *x*+1 outlets with the smallest $\tau_{I,i} \geq 1$, letting τ_{I,i^*} denote their set, and:

$$\boldsymbol{\tau_{C,i}} = \begin{cases} 1 \; for \; i \in \tau_{I,i^*} \\ 0 \text{ otherwise} \end{cases}$$

For any outlet that was not approached by any politician, it is a best response to report honestly, $b_I \Phi_C$, and to get the higher audience revenues, which benefits the challenger as the resulting posterior beliefs induce the voters to prefer the challenger. For any offer from the incumbent, $\tau_{I,i} \geq 1$, they would accept it and misreport if and only if $\tau_{C,i} < 1$, i.e. report $\Phi_I \Phi_C$. The resulting politician equilibrium is that neither politician attempts any media suppression. The incumbent would not deviate from this strategy since any spending can be effectively negated at a lower cost by the challenger, ensuring the incumbent's utility is less than *B*. Given the strategy of the incumbent, the challenger will not deviate for the same reason. In this equilibrium, all media outlets report honestly, and the incumbent is revealed as bad by all outlets. Based on Lemma 1, all voters prefer the challenger who then wins the election. These equilibrium strategies and beliefs are consistent for all realisations of *B*, *n*, δ and $\alpha < 0.5$.

Case 4: $\boldsymbol{b_I b_C}$

With probability $(1-q)^2$, neither politician is discovered as bad and the analysis collapses to that observed in case $\boldsymbol{g_I g_C}$. With probability 1-*q* the challenger is not discovered as bad and the analysis collapses to that observed in case $\boldsymbol{b_I g_C}$, and with probability 1-*q* the incumbent is not discovered and the analysis collapses to that observed in case $\boldsymbol{g_I b_C}$.

With probability q^2, both politicians are discovered as bad. Firstly, if $\alpha = 0$, the number of outlets each politician needs to capture depends on whether the number of outlets is odd or even. If n is even, the incumbent needs to capture half the outlets in order to win, whereas the

challenger needs $\frac{n}{2}+1$. Consequently, if the number of outlets is odd, both politicians need $\frac{n+1}{2}$.

Firstly, if *n* is even, and if *x* denotes the number of extra outlets the incumbent attempts to persuade, the condition for the incumbent to be able to win is given by:

$$\frac{B}{\frac{n}{2}+x} \geq B - \frac{xB}{\frac{n}{2}+x} - 1(n - \frac{n}{2} - x)$$

Re-arranging we obtain:

$$\frac{2B}{n+2x} \geq B - \frac{2Bx}{n+2x} - \frac{n}{2} + x$$

$$\frac{n}{2} - x \geq \frac{B(n-2)}{n+2x}$$

By construction, $x \in \{0, 1, 2, \dots, \frac{n}{2}\}$. Both sides of the inequality are strictly increasing in *x*. When *x*=0, the incumbent wins if:

$$1 \geq \frac{2B(n-2)}{n}$$

Which only holds if $2 \geq n$. When x=$\frac{n}{2}$ the incumbent wins if:

$$0 \geq \frac{B(n-2)}{2n}$$

Which only holds if $2 \geq n$.

Therefore, if *n* is even, $\alpha = 0$ and *n*>2, the incumbent is never able to win and consequently his optimal strategy is to not engage in any media manipulation, i.e. $\tau_{Ii} = 0$. The optimal strategy for the challenger is to set $\tau_{C,i} = 1$ $\lceil \Gamma n \rceil$ outlets and zero to the rest. Similarly, if *n* is odd, both politicians need to persuade $\frac{n+1}{2}$ outlets. Therefore, the incumbent can win if:

$$\frac{B}{\frac{n+1}{2}+x} \geq B - \frac{xB}{\frac{n+1}{2}+x} - 1(n - \frac{n+1}{2} - x)$$

$$\frac{2B}{n+2x+1} \geq \frac{B(n+1)}{n+2x} - \frac{n+1+2x}{2}$$

$$2B(1-n) \geq [4x^2 - (n-1)^2]$$

$$0 \geq 4x^2 + (n-1)(2B - n + 1)$$

By construction, $x \in \{0, 1, 2, \dots, \frac{n-1}{2}\}$. Therefore, the incumbent can only win if:

$$B < \frac{n-1}{2}$$

which is equivalent to either politician having insufficient funds to 'silence' enough media outlets. Therefore, if *n* is odd, if sufficient media suppression is 'affordable', the incumbent cannot win. The incumbent's optimal strategy is to set $\tau_{Ii} = 0$, whereas for the challenger it is to set $\tau_{C,i} = 1$ $\lceil \Gamma n \rceil$ outlets and zero to the rest.

If the share of informed voters, α, is strictly positive as outlined above, the incumbent only needs to capture $\lceil \Gamma^* n \rceil$ outlets, while the challenger needs to capture $\lceil \Gamma n \rceil$. To make the subsequent analysis interesting, I assume that *B*> $\lceil \Gamma n \rceil$.

First notice that:

$$\lceil \Gamma \mathrm{n} \rceil + \lceil \Gamma^* \mathrm{n} \rceil \in \{n, n+1\}$$

The argument from the $\alpha = 0$ case extends if α is sufficiently small. The threshold level of the share of informed voters if $\lceil \Gamma n \rceil + \lceil \Gamma^* n \rceil = n + 1$ is the unique solution to the following:

$$\frac{B}{\lceil \Gamma^* n \rceil} = B - (n - \lceil \Gamma^* n \rceil) - \frac{B}{\lceil \Gamma^* n \rceil}$$

Re-arranging, the threshold α is the solution to:

$$B = \frac{\lceil \Gamma^* n \rceil (\lceil \Gamma n \rceil + 1)}{\lceil \Gamma^* n \rceil - 2}$$

Similarly, if $\lceil \Gamma n \rceil + \lceil \Gamma^* n \rceil = n$, the threshold α is the solution to:

$$B = \frac{\lceil \Gamma^* n \rceil \lceil \Gamma n \rceil}{\lceil \Gamma^* n \rceil - 2}$$

If either set of conditions hold, the optimal strategies match the ones from the case when the share of informed voters is zero.

If $\alpha > 0$, any offer from the incumbent of the type: offer 1 to $x \geq \lceil \Gamma n \rceil$ outlets is strictly suboptimal. Since the challenger moves second, he can gain enough outlets by offering them $\tau_{I,i} + 1$ if $\tau_{I,i} > 0$ and 1 to $\lceil \Gamma n \rceil - \lceil \Gamma^* n \rceil$ for the remaining outlets. This offer is affordable for the challenger, for whom it is an optimal strategy.

First, there is a disparity in politicians' disposable income after the minimum media suppression cost given opponent inaction. For $\Gamma = \frac{0.5}{1-\alpha}$ and $\Gamma^* = \frac{0.5-\alpha}{1-\alpha}$:

$$\boldsymbol{B - \lceil \Gamma n \rceil = l, l \in \mathbb{N}\ ad}$$

$$\boldsymbol{B - \lceil \Gamma^* n \rceil = (k + l), k, l \in \mathbb{N}}$$

The condition for which the incumbent can win is given by:

$$1 + \frac{(l+k)}{x} > 1 + \frac{l}{1 + x - \lceil \Gamma^* n \rceil}$$

Where $x \geq \lceil \Gamma^* n \rceil$ is the number of outlets approached by the incumbent.

Re-arranging

$$(l+k)(1 + x - \lceil \Gamma^* n \rceil) > x$$

$$\frac{(l+k)(1 - \lceil \Gamma^* n \rceil)}{k} < x$$

, which always holds, since $(1 - \lceil \Gamma^* n \rceil) \leq 0$ and $l, k, x \in \mathbb{N}$. Therefore, any offer of the type

$\tau_{I,i} = 1 + \frac{(l+k)1}{x}$ to $x \geq \lceil \Gamma^* n \rceil$ outlets

will ensure electoral victory for the incumbent.

The *x* that maximises the incumbent's welfare is given by:

$$\min_{x \geq \lceil \Gamma^* n \rceil, x \in \mathbb{N}} x(1 + \frac{(l+k)}{x})$$

Which is solved by $x = \lceil \Gamma^* n \rceil$. Given this strategy for the incumbent, the best response of the challenger is $\tau_{i,C} = 0, \forall i$, since any positive media transfer would decrease his utility without resulting in electoral victory.

Lemma 3:

This equilibrium strategy is feasible for the incumbent if and only if the minimal acceptable vector of bribes for the challenger is affordable.

Proof:

Suppose the minimal acceptable vector of bribes is affordable for the challenger, i.e.:

$$\lceil \Gamma n \rceil \leq B$$

And the optimal vector of payments for the incumbent described above is not affordable:

$$\lceil \Gamma^* n \rceil \left(1 + \frac{(l+k)}{\lceil \Gamma^* n \rceil}\right) > B$$

$$\lceil \Gamma^* n \rceil \ + (l+k) > B$$

$$(l+k) > B - \lceil \Gamma^* n \rceil$$

By definition:

$$B - \lceil \Gamma^* n \rceil = (k+l)$$

Thus:

$$(l+k) > (k+l)$$

Contradiction

$$\boldsymbol{QED\ Lmma\ 3} \blacksquare$$

Therefore, the incumbent utility maximising action with media suppression is to offer the following vector of payments:

$$\tau_{i,I} = \begin{cases} 1 + \dfrac{(l+k)}{\lceil \Gamma^* n \rceil} & to \lceil \Gamma^* n \rceil \ outlets \\ 0 \ to \ the \ rest \end{cases}$$

if $B \geq \lceil \Gamma n \rceil$

If $B \in [\lceil \Gamma^* n \rceil, \lceil \Gamma n \rceil)$, the incumbent can still afford enough media suppression to allow his victory, but the challenger is unable to afford meaningful media intervention. Therefore, the equilibrium strategy of the incumbent becomes:

$$\tau_{i,I} = \begin{cases} 1 \ to \lceil \Gamma^* n \rceil \text{ outlets} \\ 0 \ to \ the \ rest \end{cases}$$

Finally, if $B < \lceil \Gamma^* n \rceil$, neither politician can afford meaningful media suppression and their equilibrium strategies are $\tau_{i,I} = \tau_{i,C} = 0$

For any vector of payments from the incumbent other than the ones outlined above, there exists a corresponding vector for the challenger which allows him to win. Let τ_I denote a vector of incumbent payments below the optimal one described above.

Let x denote the number of outlets for which $\tau_{I,i}$=0 and order the remaining outlets ascendingly. Then if:

1) $x \geq \lceil \Gamma n \rceil$ **or**

2) $x < \lceil \Gamma n \rceil$ **and**

$$B - x \geq \sum_{j=x+1}^{\lceil \Gamma n \rceil} \tau_{I,j} + (\lceil \Gamma n \rceil - x)$$

$$\tau_{C,i} = \begin{cases} 1 \; to \min(x, \lceil \Gamma n \rceil) \;\; to \; outlets \; s.t \; \tau_{I,i} = 0 \\ \tau_{I,i} + 1 \; to \text{ the first } \max(\lceil \Gamma n \rceil - x, 0) \;\; ordered \; outlets \\ 0 \; otherwise \end{cases}$$

Otherwise $\boldsymbol{\tau_{C,i}} = 0$

From the perspective of the media outlets, their optimal strategy is to report honestly, $b_I b_C$, if they receive no offer higher than 1 or if $\tau_j \geq 1$ and $\tau_{-j} \geq \tau_j + 1$. Otherwise, they accept the higher offer, and they report $\Phi_I b_C$ or $b_I \Phi_C$ depending on the origin of the offer. According to Lemma 1 voters prefer the incumbent if they observe $b_I b_C$ or the candidate not revealed as bad otherwise. Therefore, the incumbent wins the election.

Second equilibrium

A second equilibrium occurs if $\alpha \geq 0.5$.

Histories/Signal received by media	$\Phi_I\Phi_C$	$\Phi_I b_C$	$b_I\Phi_C$	$b_I b_C$	Beliefs
Prior Probability	$\gamma^2 + 2(1-\gamma)\gamma q + (1-\gamma)^2(1-q)^2$	$\gamma(1-\gamma)q + (1-\gamma)^2 q(1-q)$	$\gamma(1-\gamma)q + (1-\gamma)^2 q(1-q)$	$(1-\gamma)^2 q^2$	
Agents and strategies					
Incumbent	$\boldsymbol{\tau_{I,i}} = 0, \forall i,$ $\forall B, \forall\, E(\tau_C)$	$\boldsymbol{\tau_{I,i}} = 0, \forall i,$ $\forall B, \forall\, E(\tau_C)$	$\boldsymbol{\tau_{I,i}} = 0, \forall i$ $\forall B, and\ \forall\, E(\tau_C)\ such\ that\ \forall \boldsymbol{\tau_{C,i}} \geq 1 + \boldsymbol{\tau_{I,i}}$ Otherwise, a) if $B \geq \lceil \Gamma n \rceil$ $\boldsymbol{\tau_{I,i}} = \ to\ all\ outlets$ b) if $B < \lceil \Gamma n \rceil$ then $\tau_{I,i} = 0, \forall i$	$\tau_{I,i} = B - (n-2)1$ to 1 outlet, zero otherwise $\forall\, E(\tau_C)$ if B≥n If $B<n$, $\tau_{I,i} = 0$, $\forall i \forall\, E(\tau_C)$	Informed voters: Believe incumbent is good if they observe $\Phi_I\Phi_C$ or $\Phi_I b_C$ in the media. Believe challenger is good if observe $\Phi_I\Phi_C$ or $b_I\Phi_C$ in the media. Believe both politicians are
Challenger	$\boldsymbol{\tau_{C,i}} = 0, \forall i,$ $\forall B, \forall\, \tau_I$	$\boldsymbol{\tau_{C,i}} = 0, \forall i,$ $\forall B, \forall\, \tau_I$	If $\tau_{I,i} = 0$ for $n-1$ outlets, then $\boldsymbol{\tau_{C,i}} = 0, \forall i\ \forall B.$ If $\tau_{I,i} \geq 1$ for $all\ n$ outlets, pick the outlet with the smallest $\tau_{I,i}$ and offer $\tau_{C,i} = 1 + \boldsymbol{\tau_{I,i}}$	$\tau_{C,i} = 0, \forall i$ if $B < n$ If $B \geq \lceil \Gamma n \rceil 1$, let x denote the number of outlets for which $\tau_{I,i}$=0 and order the remaining outlets ascendingly. Then if x=n **or**	

				x<n **and** $B - x \geq \sum_{j=x+1}^{n} \tau_{I,j} + (n - x)$ $\tau_{C,i}$ $= \begin{cases} 1 \text{ to all } s.t\ \tau_{I,i} = 0 \\ \tau_{I,i} + 1 \text{ to the rest of the outlets} \end{cases}$ Otherwise $\boldsymbol{\tau_{C,i}} = 0$	bad if they observe $b_I b_C$ in the media. Uninformed voters:
Result	Incumbent wins	Incumbent wins	Challenger wins	Incumbent wins	Believe incumbent is good if observe $\Phi_I \Phi_C$ or $\Phi_I b_C$ from their outlet. Believe challenger is good if they observe $\Phi_I \Phi_C$ or $b_I \Phi_C$ from their outlet.
Media Agents	Report honestly,$\Phi_I \Phi_C$, for $\forall\ \tau_I\ and \forall\ \tau_C$. If $\tau_{i,I} \neq 0$ and or $\tau_{C,i} \neq 0$, they would accept the higher offer and still report honestly	Report honestly $\Phi_I b_C$ if $\tau_{C,i} < 1$, $\forall\ \tau_I$ Or if $\tau_{C,i} \geq 1$ and $\tau_I \geq \tau_{C,i} + 1$ Report $\Phi_I \Phi_C$ if $\tau_{C,i} \geq 1$ and $\tau_I < \tau_{C,i} + 1$	Report honestly $b_I \Phi_C$ unless $\begin{cases} \tau_{I,i} \geq 1 \textbf{ and} \\ \tau_{C,i} < \tau_{I,i} + 1 \end{cases}$ then report $\Phi_I \Phi_C$	Report honestly $b_I b_C$ if either set of conditions holds $\tau_{I,i} < 1 \textbf{ and } \tau_{C,i} < 1 \boldsymbol{OR}$ $\tau_{I,j} > 1 \boldsymbol{and}\ \tau_{I,-j} > \tau_{I,j} + 1$ For $j \in \{I, C\}$ Report $\Phi_I b_C$ if $\tau_{I,i} \geq 1\ and\ \tau_{I,i} \geq \tau_{C,i}$ Report $b_I \Phi_C$ if $\tau_{C,i} \geq 1\ and\ \tau_{I,i} < \tau_{C,i}$	

					Believe both politicians are bad if they observe $b_I b_C$ from their outlet
Informed Voters	Vote incumbent unless observe $b_I \Phi_C$ in the media	Vote incumbent unless observe $b_I \Phi_C$ in the media	Vote challenger if they observe $b_I \Phi_C$ in the media. Vote incumbent if they observe $\Phi_I \Phi_C$ in the media	Vote challenger if they observe $b_I \Phi_C$ in the media. Vote incumbent if they observe $b_I b_C$ or $\Phi_I b_C$ in the media	
Uninformed Voters	Vote incumbent unless observe $b_I \Phi_C$ from their outlet	Vote incumbent unless observe $B_I \Phi_C$ from their outlet	Vote challenger if they observe $b_I \Phi_C$ from their outlet. Vote incumbent if they observe $\Phi_I \Phi_C$ from their outlet	Vote challenger if they observe $b_I \Phi_C$ from their outlet. Vote incumbent if they observe $b_I b_C$ or $\Phi_I b_C$ from their outlet	
Equilibrium	$\boldsymbol{\tau_{I,i}} = 0, \forall i, \quad \boldsymbol{\tau_{C,i}} = 0, \forall i$ Media outlets report honestly $\Phi_I \Phi_C$ Both types of voters vote incumbent	$\boldsymbol{\tau_{I,i}} = 0, \forall i,$ $\boldsymbol{\tau_{C,i}} = 0, \forall\, i$ Media outlets report honestly $\Phi_I b_C$ Both types of voters vote incumbent	$\boldsymbol{\tau_{I,i}} = 0, \forall i, \boldsymbol{\tau_{C,i}} = 0, \forall\, i$ Media outlets report honestly $b_I \Phi_C$. Both types of voters vote challenger	$\boldsymbol{\tau_{I,i}} = 0, \forall i, \boldsymbol{\tau_{C,i}} = 0, \forall\, i$ if $B<n$ Otherwise $\tau_{I,i} = B - (n-2)1$ to 1 outlet, $\tau_{I,i} = 0$ to the rest $\boldsymbol{\tau_{C,i}} = 0, \forall\, i$ Media outlets report honestly $b_I b_C$. Both types of voters vote incumbent	

Case 1: $\boldsymbol{g_I g_C}$

This can occur if both politicians are good or if either politician is bad and no signal is received by the media. In the absence of a signal, the media outlets are unable to report anything and therefore media framing is impossible. Therefore the best response of all outlets is to report honestly, even if they receive a transfer from either politician. Both sets of voters observe $\Phi_I \Phi_C$ and their posterior beliefs the politicians are good are both higher than γ. As shown in Lemma 1, they both prefer the incumbent. For both politicians, strictly positive transfers do not change the result of the election and reduce utility, and hence their equilibrium actions will be to set $\tau_{I,i} = \tau_{C,i} = 0$. From the perspective of media outlets, their reporting strategy can only be honest, even in the presence of positive transfers from the politicians, and therefore would accept any positive transfer, if offered. The incumbent wins and obtains utility *2B*, the challenger 0 and all media outlets report honestly. These equilibrium strategies and beliefs are consistent for all realisations of *B*, *n* δ and α.

Case 2: $\boldsymbol{g_I b_C}$

This can occur if the challenger is bad and discovered with probability γ (1-γ)q or if both politicians are bad and only the challenger is discovered, with probability $(1-\gamma)^2 q(1-q)$. With probability γ (1-γ)(1-q) the incumbent is good, while the challenger is bad and undiscovered, in which case equilibrium actions are as outlined in $\boldsymbol{g_I g_C}$. From the voters' perspective, this is the posterior belief induced if they observe $\Phi_I b_C$ in the media or from their preferred outlet. The bad challenger may attempt to silence all media outlets, in which case no voters, informed or otherwise, observe the bad signal about the challenger. However, upon seeing no signal about either politician, following Lemma 1, both types of voters prefer the incumbent. Since media persuasion is both costly and ineffective, it is a strictly suboptimal strategy. Therefore $\boldsymbol{\tau_{C,i} = 0, \forall i}$ and the best response of the incumbent is to also set $\boldsymbol{\tau_{I,i} = 0, \forall i}$. Should they receive an offer larger than 1 and $\boldsymbol{\tau_{I,i} < \tau_{C,i}}$, the media outlets would decide to report $\Phi_I \Phi_C$. The incumbent wins and obtains utility 2*B*, while the challenger obtains 0. A special case where the challenger can suppress the media completely exists, but it is a strictly suboptimal strategy to do so. These equilibrium strategies and beliefs are consistent for all realisations of B, n, δ and $\alpha > 0.5$.

Case 3: $\boldsymbol{g_C b_I}$

This can occur if the incumbent is bad and discovered with probability γ (1-γ)q or if both politicians are bad and only the incumbent is discovered, with probability $(1-\gamma)^2 q(1-$

q). With probability γ (1-γ)(1-q) the challenger is good, while the incumbent is bad and undiscovered, in which case equilibrium actions are as outlined in $\boldsymbol{g_I g_C}$. From the voters' perspective, this is the posterior belief induced if they observe $b_I \Phi_C$ in the media or from their preferred outlet.

We know that since at least half of the voters are informed in this scenario, the bad politician needs to manipulate all media outlets in order to avoid certain defeat in the election. Consequently, the challenger only needs to ensure at least 1 outlet remains 'free' and reports that the incumbent is bad.

In this case, the incumbent has B- $(n-1)$ revenue leftover from attempting to manipulate $(n-1)$ outlets, which is the maximum budget available to persuade the final outlet. To ensure that one outlet strictly prefers to report honestly and he wins the election, the challenger can offer B- $(n-1)$ to one outlet, as the incumbent is unable to afford suppressing the final outlet. It is therefore a suboptimal strategy for the incumbent to attempt media persuasion since it is strictly utility reducing and ineffective. The resulting equilibrium is for the incumbent to offer $\tau_{I,i} = 0, \forall \mathrm{I}$, and the corresponding best response from the challenger is $\tau_{C,i} = 0, \forall i$. The incumbent would not deviate from this strategy since any spending will never lead to electoral victory, ensuring the incumbent's utility is less than B. The challenger will not deviate given the action of the incumbent since he already wins the election. Both politicians obtain utility B in the first period. The media outlets report honestly. These equilibrium strategies and beliefs are consistent for all realisations of B, n, δ and $\alpha > 0.5$.

If the realisation of B is such that the incumbent is unable to afford any meaningful media interference, the equilibrium strategies remain unchanged. Therefore, the equilibrium strategy for the incumbent and the challenger is $\tau_{I,i} = \tau_{C,i} = 0 \; \forall i$ and any realisation of B.

Case 4: $\boldsymbol{b_I b_C}$

With probability $(1-q)^2$, neither politician is discovered as bad and the analysis collapses to that observed in case $\boldsymbol{g_I g_C}$. With probability 1-q the challenger is not discovered as bad and the analysis collapses to that observed in case $\boldsymbol{b_I g_C}$ and with probability 1-q the incumbent is not discovered and the analysis collapses to that observed in case $\boldsymbol{g_I b_C}$. With probability q^2 both politicians are discovered as bad. The challenger needs to capture all the outlets in the market to ensure his victory whereas the incumbent only needs to prevent one outlet from being captured.

As outlined above, the challenger needs to persuade all outlets to misreport his state for him to win the election. Consequently, the incumbent needs to persuade only one outlet to report the state of the challenger honestly. In this case, the challenger would have B- $(n-1)$ revenue leftover from attempting to manipulate $(n-1)$ outlets, which is the maximum budget available to persuade the final outlet. The incumbent can then achieve electoral victory in two ways: full disclosure or media suppression. Under the latter, the incumbent would 'silence' enough outlets such that sufficient rationally inattentive voters do not observe that his type is bad.

Under full disclosure, the incumbent can achieve electoral victory by ensuring enough outlets know the challenger is bad, rather than suppressing news about his own type. Should either type of voter observe that both politicians are bad, they would both prefer the incumbent. If both signals are released across the media outlets, the incumbent would obtain α votes from informed voters and a share of rationally inattentive voters proportional to the number of outlets that report honestly. If the incumbent were to offer $\boldsymbol{\tau_{I,i} \geq \tau_{C,i} + 1}$ for a given challenger offer $\boldsymbol{\tau_{C,i}}$, he could induce the outlet to report honestly the status of both politicians even if $\boldsymbol{\tau_{C,i}}$ is acceptable, i. e. $\boldsymbol{\tau_{C,i} > 1}$. As in $\boldsymbol{g_C b_I}$, the outlet does not risk losing second period revenue since the resulting reporting strategy is sincere. As a result, it is a best response for any outlet to accept such an offer for a given $\boldsymbol{\tau_{C,i}}$ and report honestly $\boldsymbol{b_I b_C}$. Under full disclosure, the equilibrium strategy for the challenger is to offer $\boldsymbol{\tau_{C,i} = 0, \forall i}$ which means that neither politician attempts any media suppression. If feasible, neither politician would attempt signal suppression and all media outlets would report honestly.

The incumbent is then able to ensure one outlet reports $\Phi_I b_C$ by offering any outlet B- $(n-1)$. Since B- $(n-1) < B$, this offer is affordable and results in the incumbent winning the election. Under full disclosure, the incumbent can ensure that enough voters know both politicians are bad, i.e. $b_I b_C$ by offering one outlet $\boldsymbol{\tau_{I,i}} = B - (n-1)$, which is also affordable, as:

$$B - (n-1) < B$$

Full disclosure is an optimal strategy for the incumbent. He will not deviate and spend lower since it will allow the challenger to achieve total media suppression. The incumbent will not spend more since he gains strictly lower utility if he does so. Given the actions of the incumbent, the best response of the challenger is to set $\tau_{C,i} = 0, \forall i$. He will not deviate since he will be unable to achieve total market suppression. The incumbent obtains revenue equal

to $2B - (n-1)$, while the challenger obtains 0. The equilibrium in the media market is that all outlets report honestly, $b_I b_C$.

These equilibrium strategies and beliefs are consistent for all realisations of $B \geq n$, n, δ and $\alpha > 0.5$.

In pure strategy Perfect Bayesian Equilibrium media and political competition and electoral standards ensure that the true state of politicians is always revealed to a subset of the electorate. This implies that a good politician always wins the election whenever he is present. The assumption that bad signals cannot be fabricated is crucial to this result. When both politicians are revealed as bad, the assumptions about incumbency advantage result in the incumbent being able to win the election, while still allowing a subset of the electorate to be correctly informed. Interestingly, in this equilibrium, full disclosure is optimal both in the media market and from the winning politician's perspective, if and only if at least half the voters are informed. This arises due to the presence of political competition with constant budgets when the 'bad' politician needs to capture the entirety of the media whereas his opponent only has to ensure one outlet remains honest. If the share of informed voters is less than half, even when media market capture is rational full media capture is suboptimal as victory can be achieved with fewer funds and fewer captured outlets. This reveals that politician competition is the main driver behind the results obtained in the main section of the paper rather than the revenue structure of the outlets.

Proposition 2:
Under the baseline assumptions, total media suppression never occurs as an equilibrium outcome.

Proof:

Assume such an equilibrium exists and that no media outlets report $\Phi_j b_{-j}$ *or* $b_j b_{-j}$ where $j \in \{I, C\}$. Four cases can be discerned. Note that whenever $\Phi_j b_{-j}$ is overserved, rational Bayesian updating leads voters to prefer politician *j* as he is good with probability strictly larger than zero, whereas *-j* is bad with certainty.

Case 1: $\boldsymbol{\alpha < 0.5}$ and Pr $(\boldsymbol{\theta_I = g | \Phi_I \Phi_C}) \geq$ Pr $(\boldsymbol{\theta_C = g | \Phi_I \Phi_C})$

In this case, media silence benefits the incumbent, who is preferred by all voters when media outlets report $\Phi_I \Phi_C$. However, when the signal outlets received is $b_I \Phi_C$, the challenger has an incentive to provide outlets with inducements such that they report honestly, $b_I \Phi_C$. In this

baseline specification, outlets do not obtain any revenue from their readership and therefore any positive transfer from a politician is enough to influence their report. Their optimal strategy is then to report honestly if neither $\tau_{Ii} > 0$ nor $\tau_{Ci} > 0$ holds or accept the higher political bribe and report accordingly otherwise. Let τ_I denote a vector of payments from the incumbent such that total media suppression occurs. The following conditions then hold:

$$\sum_{i=1}^{n} \tau_{Ii} \leq B \ \ \& \forall \ \tau_{Ii} \geq 1$$

The α informed voters purchase all outlets that publish a signal, whereas the 1-α uninformed voters are split equally across all n outlets. The politicians determine the optimal number of outlets to 'manipulate' based on the behaviour of both types of voters. For the incumbent to win in $b_I \Phi_C$, he requires at least half the votes exclusively from rationally inattentive voters and the number of outlets he needs to 'manipulate', m, is given by:

$$\frac{(1-\alpha)m}{n} \geq 0.5$$

$$m/n \geq \frac{\mathbf{0.5}}{\mathbf{1-\alpha}} \rightarrow \boldsymbol{\Gamma} = \frac{\mathbf{0.5}}{\mathbf{1-\alpha}}$$

Therefore, the incumbent needs at least $\lceil \Gamma n \rceil$ outlets to report $\Phi_I \Phi_C$, where $\lceil \Gamma n \rceil$ denotes the smallest integer larger or equal to Γn. Consequently, the challenger needs n-$\lceil \Gamma n \rceil + 1$ outlets to report honestly in order to win. Given the vector τ_I of incumbent payments defined above, it is then possible to construct a challenger vector of payments τ_C in the following way. Without loss of generality, order the τ_{Ii} in ascending order, and offer:

$$\boldsymbol{\tau_{C,i}} = \begin{cases} \tau_{Ii} + 1 \ to \ the \ first \ n - \lceil \Gamma n \rceil + 1 \text{ outlets} \\ 0 \text{ otherwise} \end{cases}$$

Since:

$$\frac{d\Gamma(\alpha)}{d\alpha} > 0 \text{ and } \Gamma(0) = 0.5 \rightarrow n - \lceil \Gamma n \rceil + 1 \ \leq \lceil \Gamma n \rceil \text{ and}$$

$$\sum_{i=1}^{n} \tau_{Ii} \leq B$$

This entails that:

$$\sum_{i=1}^{n} \tau_{Ci} \leq B \;\&\; \tau_{Ci} > \tau_{Ii} \text{ for } n - \lceil \Gamma n \rceil + 1 \text{ outlets}$$

This vector τ_C is optimal for the challenger as it ensures his electoral victory at minimal cost. The existence of τ_C therefore contradicts that total media suppression can occur as an equilibrium outcome when $\alpha < 0.5$ and Pr $(\theta_I = g|\Phi_I\Phi_C) \geq$ Pr $(\theta_C = g|\Phi_I\Phi_C)$.

Case 2: $\boldsymbol{\alpha \geq 0.5}$ and Pr $\boldsymbol{(\theta_I = g|\Phi_I\Phi_C) \geq}$ Pr $\boldsymbol{(\theta_C = g|\Phi_I\Phi_C)}$

As in the previous case, media silence benefits the incumbent, who is preferred by all voters when media outlets report $\Phi_I\Phi_C$. However, when the signal outlets received is $b_I\Phi_C$, the challenger has an incentive to provide outlets with inducements such that they report honestly, $b_I\Phi_C$. In this case, when the signal media outlets received is $\Phi_j b_{-j}$, politician *-j* needs to 'manipulate' all outlets in order to win. Should one or more outlets report honestly, more than half the electorate will prefer to vote for *j* as his type is good with positive probability, but *-j* is bad with certainty. In this baseline specification, outlets do not obtain any revenue from their readership and therefore any positive transfer from a politician is enough to influence their report. Their optimal strategy is then to report honestly if neither $\tau_{Ii} > 0$ nor $\tau_{Ci} > 0$ holds or accept the higher political bribe and report accordingly otherwise.

In $\Phi_I b_C$, total media suppression is not an optimal strategy for the challenger as he will still lose the election and media interference is costly. In $b_I\Phi_C$, let τ_I denote a vector of payments from the incumbent such that total media suppression occurs. The following conditions then hold:

$$\sum_{i=1}^{n} \tau_{Ii} \leq B \;\; \& \forall \; \tau_{Ii} \geq 1$$

Total media suppression is achieved in this case, but it is not an optimal strategy for the incumbent, as the challenger can win by selecting the outlet with the smallest $\tau_{I,i}$ and offering $\tau_{C,i} = \tau_{I,i} + 1$. Both $\Phi_j b_{-j}$ outcomes contradict the existence of a total media suppression equilibrium in the baseline specification.

Case 3: $\boldsymbol{\alpha < 0.5}$ and Pr$\boldsymbol{(\theta_I = g|\Phi_I\Phi_C) <}$ Pr $\boldsymbol{(\theta_C = g|\Phi_I\Phi_C)}$

In this case, media silence benefits the challenger, who is preferred by all voters when media outlets report $\Phi_I\Phi_C$. However, when the signal outlets received is $\Phi_I b_C$, the incumbent has an incentive to provide outlets with inducements such that they report honestly, $\Phi_I b_C$. In this

baseline specification, outlets do not obtain any revenue from their readership and therefore any positive transfer from a politician is enough to influence their report. Their optimal strategy is then to report honestly if neither $\tau_{Ii} > 0$ nor $\tau_{Ci} > 0$ holds or accept the higher political bribe and report accordingly otherwise.

The α informed voters purchase all outlets that publish a signal, whereas the 1-α uninformed voters are split equally across all *n* outlets. The politicians determine the optimal number of outlets to 'manipulate' based on the behaviour of both types of voters. For the challenger to win in $\Phi_I b_C$, he requires at least half the votes exclusively from rationally inattentive voters and the number of outlets he needs to 'manipulate', *m*, is given by:

$$\frac{(1-\alpha)m}{n} \geq 0.5$$

$$m/n \geq \frac{\mathbf{0.5}}{-\boldsymbol{\alpha}} \rightarrow \boldsymbol{\Gamma} = \frac{\mathbf{0.5}}{\mathbf{1}-\boldsymbol{\alpha}}$$

Let τ_C denote a vector of payments from the challenger such that total media suppression occurs. The following conditions then hold:

$$\sum_{i=1}^{n} \tau_{Ci} \leq B \;\; \& \forall\, \tau_{Ci} \geq \tau_{Ii} \& \forall\, \tau_{Ci} > 0$$

Since he moves second and is preferred by voters who observe $\Phi_I \Phi_C$, there exists a vector of payments from the challenger which ensures the challenger wins. Since this holds for any vector of payments from the incumbent τ_I, it is an optimal strategy for the incumbent to set $\tau_{I,i} = 0, \forall i$. This in turn renders total media suppression a strictly suboptimal strategy for the challenger, who can win the election by only 'manipulating' $\lceil \Gamma n \rceil$ at minimum price, which contradicts the initial hypothesis.

Case 4: $\boldsymbol{\alpha \geq 0.5}$ and $\mathbf{Pr}(\boldsymbol{\theta_I = g | \Phi_I \Phi_C}) < \mathbf{Pr}\,(\boldsymbol{\theta_C = g | \Phi_I \Phi_C})$

In this case, media silence benefits the challenger, who is preferred by all voters when media outlets report $\Phi_I \Phi_C$. However, when the signal outlets received is $\Phi_I b_C$, the incumbent has an incentive to provide outlets with inducements such that they report honestly, $\Phi_I b_C$. In this baseline specification, outlets do not obtain any revenue from their readership and therefore any positive transfer from a politician is enough to influence their report. Their optimal strategy is then to report honestly if neither $\tau_{Ii} > 0$ nor $\tau_{Ci} > 0$ holds or accept the higher political bribe and report accordingly otherwise. In this case, when the signal media outlets received is $\Phi_j b_{-j}$,

politician *-j* needs to 'manipulate' all outlets in order to win. Should one or more outlets report honestly, more than half the electorate will prefer to vote for *j* as his type is good with positive probability, but *-j* is bad with certainty.

In $b_I\Phi_C$, total media suppression is not an optimal strategy for the incumbent as he will still lose the election and media interference is costly. For the challenger to win in $\Phi_I b_C$, the challenger needs to manipulate all outlets. Let τ_C denote a vector of payments from the challenger such that total media suppression occurs. The following conditions then hold:

$$\sum_{i=1}^{n} \tau_{Ci} \leq B \ \& \forall \tau_{Ci} \geq \tau_{Ii} \& \forall \tau_{Ci} > 0$$

Total media suppression is achieved in this case, but it is not an optimal strategy for the challenger, as the incumbent can win by selecting any outlet and offering $\tau_{I,i} = B - n + 1$. Both $\Phi_j b_{-j}$ outcomes contradict the existence of a total media suppression equilibrium in the baseline specification.

It has been shown that total media suppression cannot arise as an equilibrium outcome irrespective of the realisation of α in all possible Bayesian updating scenarios. Several assumptions were critical to reach this result: the fact that the challenger moves second; the incumbent preference tie breaking assumption; and that the budgets are equal. It can be shown that these assumptions are also sufficient in all other possible specifications.

QED Proposition 2■

Section 3.5: Extensions to the benchmark model

Section 3.5.1: Analysis with unequal budgets

In this section I relax the constant budget assumption to analyse whether a similar PBE arises. This extension follows a similar method to the benchmark specification, with the model being solved backwards and while the equilibrium is also being stated in a series of lemmas. Note that if the budget of the incumbent is much larger than that of the challenger, the model approximates that of Trombetta and Rossignoli (2020). The resulting equilibrium has similar features to the benchmark specification if the budget difference is not too large. The incumbent wins if the signal observed by the media is $\Phi_I\Phi_C$ or $\Phi_I b_C$, irrespective of the difference in budget. This arises from the fact that Bayesian updating still indicates the incumbent is preferred by voters if they observe $\Phi_I\Phi_C$ from their media outlet(s). However, if the budget asymmetry is large enough, a bad incumbent can win the election should the signal observed

by media outlets be $b_I \Phi_C$. This is not surprising and mirrors the results obtained by Besley and Prat (2006) and Trombetta and Rossignoli (2020). Obviously, this requires the incumbent to have more funds and that voters do not know which politician has more funds.

If the latter condition does not hold and the voters know which politician has more funds, the resulting equilibrium is quantatively similar to the benchmark specification. If one politician is discovered as bad while the other is not, we obtain the same equilibrium as in the benchmark case. The special case where both politicians have the same budget was considered in the previous section. If the bad politician has lower funds, media capture is trivially decreasing their utility since any attempt can be countered by the good politician. Therefore, media manipulation by the bad politician will not lead to their election victory while decreasing their funds, hence decreasing their utility. If the bad politician has more funds, the Bayesian updating lemma, the voters prefer the politician with fewer funds if they observe their media outlet reporting $\Phi_I \Phi_C$.

If the signal received by media outlets is $b_I b_C$, the resulting equilibrium strategy is dependent on the realisations of both the politician's budgets and the share of informed voters. Again, if the budget of the incumbent is considerably larger than that of the challenger, the model approximates those of Besley and Prat (2006) and Trombetta and Rossignoli (2020). If the budget of the challenger is sufficiently larger than that of the incumbent, he can win by 'enticing' enough outlets to report $b_I \Phi_C$.

Section 3.5.2: Analysis with constant budgets, audience and credibility

In this extension, I relax the assumption of zero audience revenue for the outlets and introduce the credibility extension. The model assumptions for politicians and voters stay the same, however two features are added in the media section. Firstly, keeping in line with Trombetta and Rossignoli (2020), the outlets now can get audience revenues in addition to potential transfers from politicians. Their audience revenue is dependent on their readership, which in turn is determined by the actions of all media outlets.

Then the payoff structure of outlets in the first period is the following:

$$\mathrm{P}_{1i} = \begin{cases} \frac{1}{n^{\delta}} \text{ if no signal is received about either politician or no signal is reported} \\ \left(\frac{1-\alpha}{n} + \alpha\right)^{\delta}, \text{if it receives a bad signal and reports truthfully} \\ \left(\frac{1-\alpha}{n}\right)^{\delta} + \tau_i, \text{if it receives a bad signal and a bribe and accepts it} \end{cases}$$

As in Trombetta and Rossignoli (2020), $\delta \geq 0$ is a coefficient aimed at capturing different types of returns to scale in terms of readership. If no signal is received or no signal is reported, both voters act as if they are inattentive. Therefore, all outlets receive an equal share of the audience and the associated revenue is $\frac{1}{n^{\delta}}$. Alternatively, if the outlets receive a signal and at least one other outlet reports it, their audience depends on whether they depend to report the bad signal themselves. Should they report a bad signal, their audience is an equal share of the rationally inattentive voters, $\frac{1-\alpha}{n}$, and the share of informed voters. Should they accept an offer from a politician, they lose the audience share from informed voters, but they receive payment τ_i from the sender politician.

For each outlet that misreported in the first period, they are publicly revealed to all agents as dishonest with probability λ, in which case it loses their second period audience revenue. In this context, a 'dishonest' or 'captured' media outlet denotes an outlet that does not report the signal truthfully in order to benefit one of the politicians. Since it is optimal for the elected politician to keep the entire second period budget to himself, this revelation eliminates second period revenue from the media outlets revealed to be dishonest in the first period. The assumption that the λ is constant across publication types is made to simplify the initial analysis, but it could be relaxed as a possible extension.

Extending the analysis of Besley and Prat (2006), media outlets are interested in the second period revenues, which are purely audience driven. Since there is no election in the second period, all voters act like rationally inattentive voters and split equally across the outlets. For each outlet that misreported in the first period, they are publicly revealed to all agents as dishonest with probability λ, in which case it loses their second period audience revenue. In this context, a 'dishonest' or 'captured' media outlet denotes an outlet that does not report the signal truthfully in order to benefit one of the politicians. Hence, the expected second period payoff for each of the x outlets 'captured' in the first period is given by:

$$E[\mathrm{P}_{2i}] = (1-\lambda)\sum_{i=0}^{x}\frac{1}{(n-i)^{\delta}}\frac{x!\,\lambda^{i}(1-\lambda)^{x-i}}{i!\,(x-i)!}$$

Which simplifies to:

$$P_{2i} = \begin{cases} \sum_{i=0}^{x} \frac{1}{(n-i)^{\delta}} \frac{x!\,\lambda^{i}(1-\lambda)^{x-i}}{i!\,(x-i)!} & \text{if outlet i was honest and x outlets misreported} \\ (1-\lambda) \sum_{i=0}^{x} \frac{1}{(n-i)^{\delta}} \frac{x!\,\lambda^{i}(1-\lambda)^{x-i}}{i!\,(x-i)!} & \text{if outlet i and other } x-1 \text{ outlets misreported} \\ \frac{1}{n^{\delta}} & \text{if no signal was received in the first period or if all outlets reported honestly} \end{cases}$$

The updated timing of the game is:

- Nature determines the types of the politicians θ_i, where $i \in \{I, C\}$ and $\theta_i \in \{g, b\}$. Only the politicians know their type and their opponent's type prior to the signal. The prior probability that a politician is good is γ for both the incumbent and the challenger and is known by all agents. The shares of informed voters α is determined and is known to all agents.
- Outlets receive a composite signal about the state of the politicians, with each bad politician being revealed with probability q. The signal set outlets may receive is $\{\Phi_I \Phi_C, b_I b_{C,} b_I \Phi_C, \Phi_I b_{C,}\}$. The voters do not know that all outlets received the signal, and their prior belief that a politician is bad is γ. If a politician is revealed as bad, all media agents know it, but the voters do not. Additionally, both politicians know which has been exposed as bad.
- Politicians decide successively on two vectors of bribes τ_I and τ_C towards the media outlets and they specify a reporting strategy towards the outlets they are bribing, with the incumbent moving first. These vectors are unknown only to the voters.
- Media outlets decide simultaneously whether to accept the bribe and report accordingly. The voters do not observe this stage and do not know which, if any, outlets have accepted compensation from the politician. If a news outlet is approached by both politicians, its best response is to adopt the strategy which maximises the outlet's revenue. Should both politicians make equal acceptable offers, it will report the truth. Should the expected payoff after receiving donations from either politician match, they prefer the incumbent.
- Voters simultaneously observe the signal from the media according to their type and vote to maximise their expected utility.

Second period begins.

- The election results are announced along with politician types and are known to all agents.

- For each outlet that supressed a bad signal , it can be revealed as untrustworthy with probability λ. All agents are aware if an outlet is exposed as untrustworthy. Second period audience becomes 0 for any discredited outlet.
- Voters simultaneously consume media and the game ends.

This extension follows a similar method to the benchmark specification, with the model being solved by backward induction and while the equilibrium also being stated in a series of lemmas. The incumbent wins if the signal observed by the media is $\Phi_I\Phi_C$,$\Phi_I b_C$ or $b_I b_C$. The challenger wins if the signal observed by the outlets is $b_I\Phi_C$. This arises from the fact that Bayesian updating still indicates the incumbent is preferred by voters if they observe $\Phi_I\Phi_C$ from their media outlet(s). As a result, the resulting PBE displays qualitatively the same features as the one in the benchmark specification, with the addition that there now exists a minimum amount that a bad politician would have to pay to 'capture' outlets.

Although this extension features more elements than the benchmark specification, it is easier to prove since the addition of audience revenues and credibility considerations both increase the minimum 'bribe' a bad politician would have to pay a single outlet for them to report dishonestly, relative to the benchmark specification. Furthermore, the introduction of credibility considerations additionally hinders the 'capture' efforts of the bad politician to capture media outlets. For a given transfer from the bad politician towards outlet i, $\tau_{B,i}$, the amount the good politician needs to spend to counter it is now always strictly smaller than $\tau_{B,i}$. The intuition behind this result is that while the bad politician needs to account for the potential second period revenue loss in his offer towards the outlet, the good politician does not since honest reporting entails outlets retain their second period revenue with certainty. As a result, the outlet would prefer to accept the smaller transfer and report honestly since first period readership includes the informed voters, they receive a political transfer and keep their second period revenue too.

Section 3.5.3: Analysis with credibility, audience and unequal budgets

Again, I relax the constant budget assumption to analyse whether a similar PBE arises. This extension follows a similar method to the benchmark specification, with the model being solved backwards and while the equilibrium is also being stated in a series of lemmas. Note that if the budget of the incumbent is much larger than that of the challenger, the model approximates that of Trombetta and Rossignoli (2020). The resulting equilibrium has similar features to the previous specification if the budget difference is not too large. The incumbent

wins if the signal observed by the media is $\Phi_I\Phi_C$ or $\Phi_I b_C$, irrespective of the difference in budget. This arises from the fact that Bayesian updating still indicates that the incumbent is preferred by voters if they observe $\Phi_I\Phi_C$ from their media outlet(s). However, if the budget asymmetry is large enough, a bad incumbent can win the election should the signal observed by media outlets be $b_I\Phi_C$. This is not surprising and mirrors the results obtained by Besley and Prat (2006) and Trombetta and Rossignoli (2020). Obviously, this requires the incumbent to have more funds and that voters do not know which politician has more funds.

If the latter condition does not hold and the voters know which politician has more funds, the resulting equilibrium is quantitavely similar to the previous specification. If one politician is discovered as bad while the other is not, we obtain the same equilibrium as in the benchmark case. The special case where both politicians have the same budget was considered in the previous section. If the bad politician has lower funds, media capture is trivially decreasing their utility since any attempt can be countered by the good politician. Therefore, media manipulation by the bad politician will not lead to their election victory while decreasing their funds and, hence decreasing their utility. If the bad politician has more funds, the Bayesian updating lemma, the voters prefer the politician with fewer funds if they observe their media outlet reporting $\Phi_I\Phi_C$.

If the signal received by media outlets is $b_I b_C$, the resulting equilibrium strategy is dependent on the realisations of both the politician's budgets and the share of informed voters. Again, if the budget of the incumbent is considerably larger than that of the challenger, the model again approximates those of Besley and Prat (2006) and Trombetta and Rossignoli (2020). If the budget of the challenger is sufficiently larger than that of the incumbent, he can win by 'enticing' enough outlets to report $b_I\Phi_C$. Full equilibrium and proofs are available in Appendix B.

Section 3.6: Results and Discussion

The seminal paper of Besley and Prat (2006) suggests that the main method employed in curtailing the media's role in spreading political information shifted from outright censorships towards offering economic incentives for more favourable coverage. The papers of Trombetta and Rossignoli (2020), Corduneanu-Huci, and Hamilton (2018) bring evidence of this both theoretically and empirically.

The theoretical model analysed in this paper extends the work of Besley and Prat (2006), Trombetta and Rossignoli (2020), Corduneanu-Huci, and Hamilton (2018) by

providing a rationale for the empirically observed heterogeneous levels of interference. However, a key difference is that the results indicate that in pure strategy Perfect Bayesian Equilibria, media and political competition and electoral standards ensure that the true state of politicians is always revealed to a subset of the electorate. However, if the proportion of voters who actively seek political news is high enough, full disclosure occurs in the media market even if both politicians are bad. The present model can be seen as a generalisation of the previous models as it collapses to that of Trombetta and Rossignoli (2020) if the politician's budgets are highly asymmetrical and to that of Besley and Prat (2006) with the additional assumption that the share of informed voters is zero.

These findings are relevant given the recent attention given to the proliferation of possibly distorted media via the internet and its effects on political results and resulting voter welfare. The results presented in this paper suggest that while media competition mediates the propensity of media suppression, political competition and electoral standards help eliminate total media suppression. In line with the results of Corduneanu-Huci and Hamilton (2018), the politicians prefer the most cost effective alternative when interfering in the media market, which in their framework never occurred in the pure strategy Perfect Bayesian Equilibrium. However, their result relied on the assumption that the politically motivated audience is relatively small and serviced by their own section of the media. In my model, the addition of political competition renders total media suppression strictly suboptimal in pure strategy PBE. With constant budgets and relatively small politically interested audiences, a good politician always prevents a bad politician from winning via media interference and hence full disclosure is observed in the media. Furthermore, the results hold even when the share of informed voters exceeds 50%, in which case the pure strategy equilibrium suggests full media disclosure regardless of the type of the politicians. This result relies on the assumption that voters would prefer one politician when they observe both alternatives as bad. The choice of the incumbent as the favourite in this situation is not crucial as full disclosure is still the equilibrium should voters prefer the challenger.

Similar results hold with asymmetric politician budgets if the differences in budgets are not too large. The results indicate that unlike the constant budget case, a bad politician would be able to achieve meaningful media suppression that would allow him to win the election. Similar to the results of Trombetta and Rossignoli (2020), the model predicts heterogeneous levels of media suppression, with total media suppression only occurring if the share of

informed voters is at least half. This result approximates those of Trombetta and Rossignoli (2020) and holds despite reputation concerns in the media market.

The results found in previous literature on the effects of political interference in media markets indicate that the effects on voter welfare are at best neutral whenever media interference is possible, as a bad politician is able to win an election by capturing the media and concealing his type to the electorate. If the challenger politician is good, the effect on voter welfare is negative. Examples of this are the models of Besley and Prat (2006) and Corduneanu-Huci and Hamilton (2018), where increasing the number of media outlets decreases the probability of successful media capture. The results of Trombetta and Rossignoli (2020) show that the cost of capture is non-monotonic and goes to zero as the number of outlets goes to infinity, which lowers voter welfare by making capture more likely. The policy implications arising from these results is that political interference in media markets is unambiguously bad and should be prevented. However, the results of this paper challenge this policy implication, as the presence of a challenger also involved in the media markets results in a good politician always winning the election whenever he is present. A bad challenger never wins against a good incumbent and a bad incumbent always loses to a good challenger, and therefore, political interference in media markets is voter welfare reducing if and only if the interference is a monopsony. The policy implication is that political involvement in media markets can be beneficial if it is transparent and open to all candidates, rather than just the incumbent.

These results become important in the context of internet proliferation of new media outlets and sliding democratic standards observed in autocracies and democracies alike. They give evidence that larger politically involved sections of the electorate together with political competition can help reduce the likelihood of distortions in the media markets and voter misinformation. Furthermore, it provides evidence that political competition can help promote government accountability even in the presence of dishonest politicians and dishonest media.

Chapter 4: Does longer EU membership improve migrant labour market integration?

Section 4.1: Introduction

Labour market discrimination and integration of immigrants are contentious issues in the UK, which has experienced a migrant growth from 6.4% to 13.4% between 1990 and 2017, according to the 2017 edition of the UN's trend in immigrant stock. Previous literature such as Ottaviano and Peri (2012) suggests that migrant and native labour are not perfect substitutes and as a result different outcomes in the labour market are observed for the two groups, whereas Dustmann, Glitz and Vogel (2010) and Clark and Drinkwater (2008) indicate that non-OECD migrants in the UK are relatively higher educated and better remunerated than their native counterparts. It is therefore important to ask to what extent is it possible for migrants to integrate in the UK labour market, and to analyse whether their labour market performance is fair given their characteristics. As migrants accumulate experience and social capital in the receiving country, their labour market outcomes are also expected to improve. Is this observed in the UK?

In this paper I examine differences in labour market outcomes between UK natives and immigrants from 10 countries that joined the European Union in 2004[4] (hereafter also referred to as EU10 countries). In this context, the absence of wage gaps can be interpreted as a measure of the UK's efficiency at attracting and integrating successful immigrants, whereas the presence of a gap can indicate discrimination. The presence of wage gaps between natives and EU migrants can also give an indication of the extent to which the UK is integrated in the EU labour market.

As a member of the European Union until 2020, the UK was bound by legislation surrounding the freedom of movement that outlaws discrimination of workers from member states. Therefore, it becomes interesting to analyse how changes in the EU membership in 2004 for 10 European countries affected the degree of discrimination and integration in the UK labour market of immigrants from these countries. To assess such changes, I carry out an empirical analysis on whether EU migrant-native wage gaps exist and how they evolve after a country joins the EU.

The European Union (EU) is a unique economic and political partnership between 27 European countries, aiming to increase trade, reduce regional disparities and improve the

[4] Czech Republic, Estonia, Cyprus, Latvia, Lithuania, Hungary, Malta, Poland, Slovakia and Slovenia.

labour market prospects of workers across the union. One avenue through which it operates is the free movement of workers, a fundamental principle enshrined in Article 45 of the Treaty on the Functioning of the European Union (introduced in 1951 as part of the Treaty of Paris). One of the key principles behind the European Single Market is that of the free movement of labour, which can theoretically act as an adjustment mechanism in the face of asymmetric shocks. According to Article 45, Paragraph 2: freedom of labour movement 'shall entail the abolition of any discrimination based on nationality between workers of Member States as regards employment, remuneration and other conditions of work and employment'.

Whether the normative prescription of Article 45 holds in practice is also expected to affect the net present value of migration and consequently the decision to migrate to or remain in the UK from the perspective of potential EU migrants. This also has implications on the larger discussion on intra-EU brain drain, the dynamics of internal EU migration, its long-term effects on the development of sending countries and the sustainability of freedom of movement.

The normative prescription of Article 45 indicates that any pre-existing discrimination in the British market for citizens of countries that have not yet joined the EU should dissipate and/or disappear after accession. In order to test this hypothesis, I propose focusing on the 2004 EU enlargement episode and analyse three periods, namely 1994-2003, 2004-2008 and 2009-2013. The choice of the British labour market is motivated by EU legislation, which allows member states to impose restrictions on workers from new members for up to 7 years after accession. In the context of the accession of the countries which joined the EU in 2004 (see Table 4.1, henceforth referred to EU10 or A10 countries), restrictions were not implemented in Ireland, Sweden and the United Kingdom. This policy choice allowed A10 nationals to enjoy the full benefits of freedom of movement and the protections of Article 45 immediately in these three countries, creating a discontinuity in their legal status which can be used to analyse the benefits of EU membership. In the case of the 2007 and 2013 enlargements, labour restrictions were imposed by all pre-existing member states except Sweden.

Given these considerations, what is the extent of labour market convergence of EU migrants in the UK? Does longer EU membership improve convergence? More generally, does the behaviour of migrants in the labour market change post accession? Do they work second jobs disproportionately to natives? Is labour market convergence different for newer and more established migrants? Are their labour market opportunities converging to those of natives?

Given these questions, this paper makes the following contributions. Firstly, this study adds to the literature of the integration of migrant workers in the British labour market by using the Blinder-Oaxaca decomposition to estimate wage differentials and ascertain to what extent they are motivated by differences in characteristics. It is therefore important to analyse the explained part of the decomposition to understand to what extent the differentials are attributable to different characteristics distributions across groups. Furthermore, this paper analyses whetherEU membership improves the integration prospects of migrant men, both newly arrived and those present in the UK for at least five years. . Secondly, study proposes using the Logit-Oaxaca decomposition proposed by Fairlie (2005) to split the differences in the probability of working a second job into explained and unexplained parts. Thirdly, it adopts an extension proposed by Firpo et al. (2007) to decompose the wage differentials at different quantiles on the wage distribution to analyse which subgroups are better integrated in the British labour market.

In order to correctly interpret the unexplained component of the decompositions as discrimination, every factor which could be reasonably thought to influence individual productivity and wages needs to be taken into consideration. These include differences in education, tenure, occupation type and its characteristics, among others. Even if the unexplained component remains statistically significant, it can be attributed to discrimination only if all relevant characteristics have been included and none present any measurement error. If the remaining unexplained component is statistically insignificant, it can be inferred that the observed differentials are not due to discrimination, but to observed characteristics. In the literature, the unexplained component is usually thought to include discrimination, but it may be biased either way due to unobserved productivity differences or job-market discrimination in the form of job self-selection. These regressions could also provide evidence on whether people with foreign qualifications are disadvantaged in the British labour market, the degree of integration of post-2004 EU migrants, and whether unexplained differences are prominent at the top or bottom of the distribution.

I use the British Labour force Survey (LFS) between 1994-2013 to compare the labour market performance and behaviour of EU10 migrants before and two periods after enlargement to analyse whether two avenues of convergence between natives and migrants occur. This study finds that the wage gap between natives and all EU10 migrant men was 2%, before 2004, which is relatively small and statistically insignificant. This gap however increases immediately after

accession to 46% in 2004-2008, which is statistically significant and explained by the relatively large influx of relatively lower skilled and less experienced migrant men. Differences in characteristics account for 47% of the observed wage differential in this period. In the third period, this decreases to 42%, but remains statistically significant, while 52% of this difference is explained by heterogeneous characteristics between the groups. This suggests that overall, longer EU membership improved the integration prospects of EU10 migrants, but these effects were not homogenous across the migrant group. Generally speaking, migrants who were present in the UK for at least 5 years are better integrated in the British labour market. The results suggest that the large and statistically significant wage differential observed between natives and all migrants in the second period is overwhelmingly composed of wage gaps between natives and migrant men who were present in the UK for less than 5 years, which is of 50%, half of which is explained by differences in characteristics. The equivalent gap between natives and migrants who were in the UK for at least five years is significantly lower and statistically insignificant at 4%. In the third period, the observed wage gap becomes positive and statistically significant for both migrant groups relative to the native population, staying roughly constant at 49% for migrants present for less than 5 years. Differences in characteristics account for 61% of this gap. On the other hand, the gap for migrants present for more than 5 years is 36%, which indicates that migrants present in the UK longer are better integrated in all periods.

The quantile decomposition reveals that prior to 2004, EU10 migrants experienced sticky floors, with a 59% native wage advantage observed between natives and migrants at the bottom of the income distribution, with a statistically significant 53% gap arising from the unexplained component. EU accession is found to have alleviated this phenomenon, both at the bottom and middle of the income distribution. In the former case, this is manifested both through a reduction in the total observed wage gap and an increase in the proportion explained by differences in observed characteristics. At the middle of the income distribution, the difference increases significantly relative to the first period, but it is 97% explained by differences in characteristics. The share explained decreases slightly in the third period to 87%. However, EU accession created a strongly statistically significant ‘glass ceiling’ for EU10 migrants, with a quantile gap of 60% observed in both post-accession periods. The unexplained components are large and statistically significant in both periods, accounting for 45% and 60% wage gaps in favour of natives. The Logit-Oaxaca results surprisingly indicate that natives are

more likely to work a second job in all periods, with the difference being statistically significant in the second and third periods.

Section 4.2: Literature Review

The literature on the expected labour market outcomes of immigrants is rich. Theoretically, the first benchmark model is the human capital of Becker (1964), in which human capital is a function of education, qualifications, labour market experience, and training. An investment in human capital increases productivity in the labour market and remuneration is a positive function of human capital. Individuals are assumed to be utility maximizing and choose their human capital accordingly, and the resulting wage differentials are due to differences in human capital and innate ability. In this model, intuitively, people with higher innate ability also choose higher levels of human capital investment. However, capital acquired in the home country is not always equivalent to capital obtained in the receiving country. As indicated in Constant and Zimmerman (2009), in the case of immigrants, human capital investment undertaken in the host country is an avenue through which they can converge on the earnings of similar natives, but full convergence may not happen because of imperfect skill substitutability. Kahanec (2012) suggests that in this model, heterogeneity in human capital stocks affect individual abilities to integrate, resulting in either positive or negative immigrant self-selection.

Borjas (1987) suggests that immigration from poorer countries suffers from negative selection inasmuch as the workers are relatively lower skilled since higher skilled workers are less likely to migrate. This is supported by Longhi & Rokicka (2012), who observe that pre-2004 EU8[5] migrants 'seem to be negatively selected in terms of education'. Comparing education levels is not straightforward as harmonizing qualification levels across countries can be difficult (Nickell & Salaheen (2017), see Table 4.8), which in turn can make immigrant labour less substitutable to native labour. This may be one of the main factors causing immigrant labour in the UK to be employed below their capabilities, and a main factor causing imperfect migrant integration in the British labour market as seen in Nickell and Salaheen (2017), Longhi and Rokicka (2012), Kahanec (2012), and Dustmann, Frattini and Preston (2007). Baas et al. (2010) also find that the low skilled immigrant subset is more susceptible to wage variations, which could lead to widening gaps relative to the rest of the population when a labour market shock occurs. This would result in widening migrant wage gaps particularly

[5] Czech Republic, Estonia, Latvia, Lithuania, Hungary, Poland, Slovakia and Slovenia

versus skilled natives. Since previous immigrants are more affected by subsequent immigration that natives, this creates a channel via which widening gaps may be observed, shown by Kahanec (2012). Drinkwater et al. (2009) argue that EU8 migrants faced reduced travel costs post accession, which renders the increased inflow feasible and in accordance with migration theory via reducing migration cost and increasing the expected benefit of migration.

Longhi & Rokicka (2012) provide descriptive statistics indicating that EU10 migrant outflow towards the UK was positively affected by newly acquired EU membership. Furthermore, they indicate that the skills composition of the EU10 migrant flow deteriorated after 2004 relative to before EU accession. Further evidence is found by Drinkwater et al. (2009), who indicate that migrants were more positively self-selected in terms of education before the 2004 expansion. Kahanec (2012), using a human capital model, shows that factors which influence the decision to migrate do not have a constant effect on all groups of potential migrants.

Further evidence is given by Longhi and Rokicka (2012) observe who 'larger heterogeneity in more recent migrants' with respect to knowledge of the English language, restricting the potential jobs a migrant could accept, possibly leading to increased wage differentials observed. These gaps can also occur as migrant returns to education may be lower than those of natives as shown by Drinkwater et al. (2009) in the case of Polish workers. Dustmann, Fratinni and Preston (2007) bring further evidence by showing that skills are not perfectly substitutable between natives and migrants and as a result, the latter group are likely to 'temporarily downgrade to less skilled occupations than they are qualified for'.

Longhi & Rokicka (2012), using a probit model, provide evidence that EU8 migrants are more likely to occupationally segregate, with migrants more likely to 'work in manufacturing and financial services and less likely to work in construction, transport and the public sector'; 'immigrants from EU8 countries arrived after the 2004 enlargement work in different types of jobs than those arrived before'. Manacorda et al. (2012) and Baas et al. (2010) provide evidence that immigration more likely 'depresses the earnings of previous immigrants more than those of natives, indicating imperfect substitutability between native and migrant labour. Manacorda et al. (2012) further argue that since the immigrant share is still low and relatively better educated, the effect of reducing the returns to education has been lower on natives than on previous immigrants. This is because these migrants are more substitutable

with previous migrants and that the degree of substitutability between immigrant and native labour increases with the length of time spent by the former in the UK.

Section 4.3: Decomposition Methodology

One of the main methods used in the literature aimed at quantifying potential discrimination is the Blinder-Oaxaca decomposition (see Blinder (1973) and Oaxaca (1973)). Due to its simplicity, it can be used to gain valuable insight into the nature of the gaps, and their evolution over time if each wave is treated as an independent cross-section. Built upon Becker's (1964) Human Capital Theory, starting with usual wage regressions for the groups considered of the form:

$$y_i = \beta_{0j} + \beta_{1j} educ_i + \beta_{2j} exper_i + \beta_{3j} exper_i^2 + \mu_{ij} D_i + \sum_{l=1}^{12} \rho_{jl}\, loc_l + \sum_{k=2}^{5} \tau_{kj}\, w_k + \varepsilon_{ij}$$

[Eq 1]

where the i subscript indexes individuals, k indexes the year, l indexes government regions, and j indexes the groups under consideration. Here y_i denotes log hourly wages, $educ_i$ and $exper_i$ denote years of education (defined as age-education-6) and employment duration at current employer respectively, D_i is a vector of personal characteristics other than education and experience, loc_l are location dummies (control=London), w_k is a year binary variable (control=first year in the period). The location and period dummies are added to account for time and period specific effects which may influence earnings. The wage measure used is inflation adjusted using a GDP deflator from the HM treasury website.

The gap is then defined at the mean, μ, as:

$$(\overline{y_1} - \overline{y_0}) = (\overline{X_1} - \overline{X_0})\, \boldsymbol{\beta_1}(\mu) + (\overline{X_0}\, \beta_1(\mu) - \overline{X_0}\, \boldsymbol{\beta_0}(\mu))$$

[Eq 2]

where the first term of the right-hand side represents the part attributable to differences in mean characteristic differences, generally perceived as representing valid wage differentials arising from differences in characteristics, and the second part which is the unexplained effect. In this context, subscript 1 denotes the group of natives whereas subscript 0 denotes migrants. It is tempting to attribute the entirety of the unexplained component to discrimination, but a better interpretation is that the unexplained component includes all the relevant and omitted factors, amongst which is discrimination. Furthermore, I have:

$$(\overline{X_1} - \overline{X_0})\,\boldsymbol{\beta_1}(\mu) = \sum_{k=1}^{K} (\overline{x_{1,k}} - \overline{x_{0,k}})\,\boldsymbol{\beta_{1,k}}(\mu)$$

[Eq 3]

which allows the estimation of contribution for each variable separately. While the Blinder-Oaxaca decomposition has the advantage of being intuitive and easily implementable, its disadvantage is compounding the effect of any unaccounted factor together with discrimination in the unexplained component. Additionally, this decomposition can only analyse mean differences, it relies on out of sample predictions when the range for the variables X is different across groups, which can lead to misleading results (see Barsky et al., 2002, Longhi, Nicoletti and Platt 2013, Nandi and Nicoletti 2014). Despite these shortcomings, the method still remains widely used because of its intuitiveness and capacity to analyse the contribution of each characteristic individually.

Section 4.3.1 Logit Decomposition

Following Fairlie (2005) for the non-linear decomposition analysis, given a non-linear equation $\bar{y} = F(X, \beta)$, the gap can be defined as:

$$(\overline{y_1} - \overline{y_0}) = \left[\sum_{i=1}^{N^1} \frac{F(X_i^1 \boldsymbol{\beta_0})}{N^1} - \sum_{i=1}^{N^0} \frac{F(X_i^0 \boldsymbol{\beta_0})}{N^0}\right] + \left[\sum_{i=1}^{N^1} \frac{F(X_i^1 \boldsymbol{\beta_1})}{N^1} - \sum_{i=1}^{N^1} \frac{F(X_i^1 \boldsymbol{\beta_0})}{N^1}\right]$$

[Eq 4]

In the equation above N^1 and N^0 denote the sizes of groups 1 and 0 respectively. As in the standard Blinder-Oaxaca decomposition, the first term denotes the part attributable to group differences in the distributions of X, while the second term is the unexplained component, attributable to differences in the group processes in determining levels of y. The standard Blinder-Oaxaca decomposition is a special case of this decomposition. Here, the alternative expression is used since $\overline{y_j}$ need not equal $F(\bar{X}\hat{\beta})$.

Following Fairlie (2005), I define $\overline{y_j}$ as the average probability of the binary variable of interest for group j and let F denote the cumulative distribution function from the logistic distribution. Estimating the first term in the equation above, the total explained component, is achieved by calculating the two sets of predicted probabilities and taking the difference between the average values of the two. As opposed to the standard Blinder-Oaxaca decomposition, identifying the contribution of various variable groups to the observed gap in probability is not straightforward. If the group sizes are equal and there exists a "natural one-

to-one matching" (Fairlie (2005)), coefficient estimates from a logit regression run on a pooled sample, $\hat{\beta}^*$, can be used to determine independent contribution of X_1 to the observed gap:

$$\frac{1}{N^0}\sum_{i=1}^{N^0} F(\hat{\alpha}^* + X_i^1\widehat{\boldsymbol{\beta}}^*_{\mathbf{1}} + X_i^1\widehat{\boldsymbol{\beta}}^*_{\mathbf{0}}) - F(\hat{\alpha}^* + X_i^0\widehat{\boldsymbol{\beta}}^*_{\mathbf{1}} + X_i^1\widehat{\boldsymbol{\beta}}^*_{\mathbf{0}})$$

[Eq 5]

The contribution of each variable to the gap is equal to the change in the average predicted probability from replacing the group 0 distribution with the group 1 for that variable while holding constant the contributions of the rest of the variables. This also has the property that the sum of individual contributions equals the total contribution of all variables evaluated in a full sample regression. However, unlike in the linear case, the independent contributions of the groups depend on the value of the other variable, implying that which variable set is chosen as X_1 matters. This is known as path dependence.

Following Fairlie (2005) standard errors are approximated using the delta method. Letting F denote the logistic probability density function and rewriting equation 5 as:

$$\widehat{D_1} = \frac{1}{N^0}\sum_{i=1}^{N^0} F(X_i^{11}\widehat{\boldsymbol{\beta}}^*) - F(X_i^{01}\widehat{\boldsymbol{\beta}}^*)$$

[Eq 6]

where:

$$F(X_i^{jk}\hat{\beta}^*) = F(\hat{\alpha}^* + X_i^j\widehat{\boldsymbol{\beta}}^*_{\mathbf{1}} + X_i^k\widehat{\boldsymbol{\beta}}^*_{\mathbf{0}})$$

The variance of $\widehat{D_1}$ can be approximated as:

$$Var(\widehat{D_1}) = \frac{\partial\widehat{D_1}}{\partial\hat{\beta}^*}' Var(\hat{\beta}^*)\frac{\partial\widehat{D_1}}{\partial\hat{\beta}^*}$$

[Eq 7]

where:

$$\frac{\partial \widehat{D_1}}{\partial \hat{\beta}^*} = \frac{1}{N^0}\sum_{i=1}^{N^0} F(X_i^{11}\hat{\beta}^*)X_i^{11} - F(X_i^{01}\hat{\beta}^*)X_i^{01}$$

[Eq 8]

In this application, however, the size of the migrant subgroup is smaller than that of natives in all three periods. This issue is addressed by first using pooled coefficient estimates to compute predicted probabilities $\widehat{y_i}$ for each native or migrant observation in the sample. Then a random subsample of natives of size equal to that of the group of migrants and each observation in both groups are separately ranked by the predicted probabilities of being in the top/bottom of their respective distributions. The observations are then matched by their respective relative rankings. Fairlie (2005) however notes that this method yields fairly similar results to matching observations across the two groups randomly. Naturally, the estimated results depend on the choice of natives in the subsample, and optimally the results should approximate those obtained by matching the entire native sample to the migrant one. This is circumvented by drawing a large number of random subsamples, calculating separate decomposition estimates and using the mean estimates to approximate the entire native population results. In the results section below, I use 100 random subsamples in each regression to calculate the means.

This decomposition technique has two drawbacks. Firstly, the unexplained portion of the gap is inherently difficult to interpret, which renders subsequent analysis reliant only on the explained component. The second is the path dependence.

Section 4.3.2: Quantile Decomposition

The paper of Firpo et al. (2007) proposes an extension which permits the decomposition to be extended at any percentiles of the wage distribution. The method entails usage of a Re-Centred Influence Function (RIF), given by:

- RIF(y,q_τ)= $q_\tau + \frac{[\tau - d_\tau]}{f_Y(q_\tau)}$

where the denominator denotes the density distribution function of the log hourly pay variable, $q_\tau$denotes the quantile at which the regression is run while d_τ is a binary variable recording 1 if y$\leq q_\tau$ and zero otherwise. This transformation ensures that the outcome variable, here log of hourly pay, at the mean equals q_τ. In this paper the RIF (y,q_τ) is estimated by superseding

$q_\tau$with the sample estimate of q_τ and using a non-parametric kernel estimation to compute the density distribution.[6] The RIF also satisfies the following properties:

i) The mean equals the τ-quantile: E_Y[RIF(y,q_τ)]= q_τ

ii) The expectation mean given the vector X also equals q_τ: $E_X[E_Y$[RIF(y,q_τ)]|X] = q_τ

According to Firpo et al. (2009), Longhi, Nicoletti and Platt (2013) and Nandi & Nicoletti (2014), if a linear relationship is assumed between the RIF and X, E_Y[RIF (y,q_τ)] can be estimated by:

- RIF(y,q_τ)= $\beta_j(q_\tau)X_j$+u_j, where j is the group indicator (here 1 if British, 0 otherwise), X is a n*k vector of explanatory variables, and $\beta_j(q_\tau)$ is the vector of estimated coefficients at quantile q_τ

The authors prove that, given i) and ii), that the Blinder-Oaxaca equivalent at quantile $q_\tau$is:

- $q_{1,\tau} - q_{0,\tau} = E_Y$[RIF ($y_1$,$q_\tau$)]|$X_1$]- E_Y[RIF (y_0,q_τ)]|X_0]= $\beta_1(q_\tau)\overline{X_1}$-$\beta_0(q_\tau)\overline{X_0}$

Re-arranging the above equation, I arrive at an equivalent of the original Oaxaca decomposition, which the authors call the 'generalized Oaxaca':

- $q_{1,\tau} - q_{0,\tau} = (\overline{X_1} - \overline{X_0})\,\beta_1(q_\tau)$+ $\overline{X_0}(\beta_1(q_\tau) - \beta_0(q_\tau))$

Because of property i), the generalized Oaxaca decomposition also includes the standard Oaxaca decomposition as a special case. It can also be shown that the generalized Oaxaca permits observing the individual influence of each variable in the decomposition:

- $q_{1,\tau} - q_{0,\tau}$=$\sum_{k=1}^{K}(\overline{x_{1,k}} - \overline{x_{0,k}})\,\beta_{1,k}(q_\tau)$ + $\overline{X_0}(\beta_1(q_\tau) - \beta_0(q_\tau))$

Nandi and Nicoletti (2014) points out that this decomposition still has to rely on a linearity assumption and out of sample predictions, but it has the benefit of allowing estimations at quantiles other than the mean, providing a more complete analysis. Furthermore, Longhi, Nicoletti and Platt (2013) indicate that the inferences are 'robust to arbitrary forms of heteroscedasticity', which increases the efficiency of the estimations.

[6] The RIF is estimated using an STATA file rifreg.ado found at: http://faculty.arts.ubc.ca/nfortin/datahead.html

Section 4.4: Data

For this study, I use successive Labour Force Survey (LFS) quarterly surveys of households, comprising a large amount of information on demographic indicators, labour market, and job characteristics for individuals older than 16 who have a private address in the UK. However, according to Longhi and Rokicka (2012) and Saleheen and Shadforth (2006), as the LFS is not focused on immigration, immigrants may be less likely to have a private address in the UK and be less willing to partake in the survey, and therefore be under-represented in the dataset. I define a person interviewed in the dataset born in one of the countries which joined the European Union in 2004 as a A10 national/migrant. For the present study, individuals who are neither a British native nor a national of one of the A10 member states are ignored. I separate the pooled data in three datasets, one before accession (1994-2003) and two after (2004-2008 and 2009-2013). The latter period coincides with the second round of member states reducing immigration restrictions for workers from the 2004 enlargement countries, which should ease the migratory pressure on the UK, Ireland and Sweden. Additionally, the late 2000s recession had settled in the UK, the unemployment rate was increasing in late 2008-early 2009.The third period is thus characterized by a decrease in pull factors in the UK as well as push factors diminishing in the A10 countries, particularly Poland.

The LFS has a rotating structure, with individuals being interviewed in five consecutive quarters, and samples released quarterly. Individuals are asked wage questions in the first and fifth waves of each quarter. To ensure that for each dataset each individual has a unique observation, I use the first wave responses for all years except 1994, 2004 and 2009, for which I also use the fifth wave. The addition of the fifth wave of observations in 1994 2004 and 2009 does not lead to individuals being represented twice in the sample since they correspond to new entrants in 1993, 2003 and 2008 respectively. I restrict the dataset to the subsets of individuals aged 18-65, employed and not in education. I omit outliers in hourly wages by dropping from the sample individuals whose hourly pay exceeds £100. I also restrict the sample by keeping only the men with observations for those with available information on nationality, qualifications and wages data. Any observations with nationalities other than British or one of the A10 set are dropped. Tables 4.3 to 4.9 provide summary statistics for each dataset.

Tables 4.3 to 4.5 outline summary statistics for both the native and migrant populations in the three periods considered in this study. It can be observed that the stock of A10 citizens is relatively small in the period before the expansion, increasing after 2004 and stabilizing in

the third period. Since this is a stock variable, it does not offer any indication whether the change is due to increased inflows, outflows or both. The changes in the immigrant stock are in accordance with the theory outlined above, increasing after the restrictions were lifted in 2004, an increase in the pull factor, and stabilizing after the second period, due to reduced expected gain from migrating to UK, caused by the recession, better macroeconomic conditions in the largest sender and other EU countries lifted the working restrictions for A10 workers, increasing pull factors elsewhere.

Table 4.3 shows that in the first period (1994-2003), the mean age for natives is 39.67, while the mean age for migrants of 39.88, and they are not statistically significantly different. However, the migrant population who arrived in the UK less than five years prior is significantly younger with a mean of 29 years. Conversely, migrants who resided in the UK for more than five years have a mean age of 41.5 years, which is also statistically significant. Table 4.4 shows that in the second period, the mean age of the immigrant population is significantly lower, with an average age of 30.6 years which may indicate an inflow of relatively younger workers in this period. I also find a lower immigrant age in the third period, 2009-2013 (Table 4.5). The differential in age between the native and overall migrant men are strongly statistically significant in both the second and third periods. When separating the migrant population based on time since arriving in the UK, a similar pattern persists in both post-accession periods, where more recent migrants are on average younger than migrants living in the UK for more than five years. In contrast with the first period, both categories of migrants are significantly younger than the native population.

The summary statistics for log hourly pay in the first period indicates no significant difference between the pay received by migrants and natives, when comparing natives with both the whole migrant population or with either migrant subgroup, which suggests that both populations were similarly rewarded in the labour market and good integration of the migrants. The native population in the second period is higher paid relative to the first period, and significantly better paid than the overall migrant population and the migrant population who was present in the UK for less than five years. The migrant population who arrived in the UK before their country's accession to the EU in 2004 are not statistically differently paid relative to natives, indicating this subgroup was relatively better integrated in the British labour market. However, this pattern is not observed in the third period, where the native population is significantly higher paid than the migrant one both when considering recent migrants and migrants who reside in the UK for more than 5 years. However, this is not sufficient to indicate

the presence of discrimination as relevant characteristics may also be significantly different between the two populations. This is analysed in the next section.

The values for the length of continuous employment also suggests an inflow of new migrant men in the second period, a drop from 76 to 19.4 months for the mean values of A10 migrants in the first and second periods respectively. The average length of employment for immigrants increases by 89% in the third period to 3 years, which may be indicative of an improvement in their labour market conditions relative to that of natives. The native mean figure increases slightly but remains roughly constant across the periods. In the case of the migrant subgroups, by construction in all periods, the mean value for the migrants who arrived in the UK more than 5 years prior is significantly higher than that of those who did not. Data in Tables 4.3 to 4.5 also reports comparative statistics on experience in the British labour market. In this context, I define labour market experience as age-education-6 for natives and the difference between the year they arrived in the UK and the year in which they are interviewed. The mean labour market experience for natives in the first period is 22.4 years, while migrants on average have more years of experience in the labour market with 24.87 years. The mean years of British labour market experience increases for natives across the three periods, but falls dramatically for migrants, indicating a large migration inflow in the second and third periods. The mean for the whole migrant population drops to 2.97 and 5.5 years in the second and third periods respectively and as indicated in the literature, may be a vector through which wage differentials between natives and migrants can occur. In both post-accession periods, this holds even for established migrants who were present in the UK for at least 5 years, indicating that migrant workers are less substitutable to natives, in line with the result of Manacorda et al. (2012). Years of experience in the British labour market offer an indication of how integrated a migrant is in the British labour market, due to improved knowledge of language and other local characteristics, improved connections, a longer period in which they could engage in (on the job) job search or longer tenure.

Summary statistics suggest that natives have a relatively higher chance of being employed in the public sector, being a manager and to have been offered training in the three months prior to interview. The figure for natives in the public sector remains roughly constant at 20%, while the immigrant counterpart starts at 23% and stabilises at 4% in the second and third periods. In both post accession periods, the figure for established migrants is relatively higher. The share of natives offered training is stable in all periods at around 26%. On the other hand, share of migrants offered training in the first period is not statistically significantly

different to that of natives, which holds for the migrant population as a whole and both subgroups. However, after accession, the overall figure for migrants is significantly lower in both periods compared to natives, but more established migrants are offered training more.

It can be noticed from Table 4.6 that the distribution of A10 migrants changed over the three periods, particularly in London where it decreased from 36.6% in the first period to 18.83% in the third. Post accession migrants display a more even distribution across the country, especially in the Midlands and North of England. Drinkwater et al. (2009) and Longhi & Rokicka (2012) also find that EU10 migrants are "more likely to be self-employed but earn less than their counterparts who arrived before". If region fixed effects are significant in determining one's wage, the additions of location dummies in the regressions should lead to more precise estimates. The descriptive statistics found in Tables 4.3 to 4.6 suggest that the 2004 expansion had a significant effect on both the number of A10 immigrants and their distribution across the UK.

Table 4.7 illustrates the distribution of migrants in industry sectors. Prior to the 2004 accession, the largest share of immigrant labour was under the managers and administrators and elementary occupations categories (both 17.3%) followed by professional occupations and skilled trades (19.18% and 13.7% respectively) and plant and machine occupations (12%). The statistics suggest that the distribution of immigrants into job categories significantly changed after the 2004 expansion. The share of A10 workers in the managerial and professional occupation categories, likely associated with higher remuneration, drops significantly to 3% and 5.2% respectively. Associate and technical occupation and the clerical categories also suffer similar decreases, although of a lesser magnitude to 3.27 and 1.41% respectively, while sales occupations and skilled trades decrease slightly over the three periods. The share of A10 workers in the plant and elementary occupation increases significantly in the second period, which may indicate job self-selection or a larger share of low skilled labour, for whom the net expected benefit of migration has increased, which can indicate an increase in the mean gap. The share of immigrant workers in managerial positions and professional occupations stay roughly constant in the third period while the share of migrants in clerical occupations rises slightly in the third period to 2.63%. The majority share still lies in elementary and plant occupations. Although the share of immigrants in elementary occupations increases, the summary statistics suggests a slight improvement in the job allocation of the immigrant stock.

The substantial drop in the share of managers in the last two periods can indicate a negative self-selection of migrants into lower paid jobs. The results of Dustman, Fratinni and Preston (2007) indicate that newer migrants may be temporarily underemployed given their human capital, and therefore the distribution of migrants into lower job categories post expansion cannot be, prima facie, as evidence of negative immigrant self-selection in terms of education. The results of Longhi and Rokicka (2012) and Rosso (2013) suggest, however, some negative self-selection of immigrants, at least in the period immediately after accession. One potential issue is endogeneity between wages and occupation, which this decomposition method does not account for. As a result, the results should not be used to deduce any casual relationships.

Table 4.8 illustrates the qualification levels of natives and immigrants in the three periods. Longhi and Rokicka (2012) provide evidence of negative self-selection in the A10 stocks, seen in the prevalence of the "other qualifications" category, which is presumably either not easily comparable to a British equivalent or unrecognized in the British Labour Market. Thus, the other qualification can be interpreted as a proxy for other migrant specific characteristics which may result in a disadvantage in the British labour market. Furthermore, while it may indicate a relatively lower qualification level for the native population, this may not hold in the case of the migrants. The other qualification category represents more than half of the immigrant stock in the first two periods (49.2% and 57.36% respectively) but drops significantly in the third period to 45%. With respect to qualifications corresponding to a degree or equivalent including postgraduate qualifications, the immigrant stock is initially comparable to the native population but suffering a significant shock after the 2004 expansion, providing evidence of negative immigrant self-selection with respects to education.

The share of migrants with qualifications equivalent to a degree or equivalent increases in the third period, becoming relatively comparable to the first period population, but the immigrant stock has now become less represented in this category relative to the British population. A similar trend can be observed for immigrants reporting qualifications in the Higher Education category, but they are relatively less represented in this group in the second period. With regards to secondary education, the share of migrants having qualifications equivalent to an A level varies slightly across the periods, from around 10% in the first and third periods to 7.47% in the second, while the share of those having completed only GCSEs remains roughly constant at around 4.5%. In both categories, they are under-represented relative to natives. Interestingly, the only categories where the immigrant share is larger than

its native counterpart are the 'other' and no qualifications, whose trends are inverse for the native and migrant subsets respectively. The share of natives in the last two categories is progressively declining in both, the share of the immigrant stock in the latter is increasing in the second (from 15.77% to 17.9%), before stabilizing slightly in the third. Based on these considerations, it can be inferred that the human capital of the immigrant stock has been negatively impacted by the 2004 expansion relative to the native, but stabilized in the third period

Table 4.9 attempts to provide a preliminary view of the differences in the return to human capital between the two populations in the sample over the three periods. Overall, the values suggest that the native subsets have a higher mean wage in each qualification, group and time cell except Higher Education and Other qualification individuals in the first period. In the first period, there are only slight mean differentials between native/immigrant men in all categories except for higher education and no qualification brackets.

In the second period, the wages of immigrants with degree education, higher education and A levels cells fall relative to the first, while the wages of other immigrant with GCSE or equivalent and no qualifications rise relative to the first period. Mean wages in all British cells rise slightly, becoming larger than their A10 equivalents. This pattern also holds in the third period, even though the mean wages of all the British cells decrease relative to the second period. The third period also presents an increase in all immigrant cells relative to the second period, which may lead to a smaller mean wage differential compared to the second period. The overall decrease in immigrant mean wages is in accordance with the negative self-selection of migrants after 2004 observed by Longhi and Rokicka (2012) and Rosso (2013) and their observed tendency to be underemployed (also highlighted in Dustmann et al. (2007)). To what extent these patterns explain differentials in wages and hours worked is analysed in the next section.

Section 4.5: Decomposition Analysis Results

In the decomposition analysis I consider the wage gap between UK natives (group 1) and EU10 immigrants (group 0), therefore a positive sign indicates an advantage for the UK natives. The wage gaps are in terms of log wages; therefore, gaps approximate relative, not absolute, changes in wage. I also consider the gap in the probability to work a second job between UK natives (group 1) and EU10 immigrants (group 0). I decompose these gaps into the explained part, differentials arising from differences in explanatory variables, and the

residual unexplained part. If the latter is insignificant, I can conclude that there is no discernible discrimination against A10 migrants. All regressions are run separately for the periods 1994-2003, 2004-2008, 2009-2013. The explanatory variables are grouped together in six categories: personal characteristics (including age, age squared, and dummies on marital status, dependent children, and disability binary variable); qualifications (years of education); and occupation (9 industry division dummies). The last three categories are time and region binary variables, and job characteristics, which includes experience, experience squared, dummies on firm sizes, managerial roles, employment in the public sector and having received job-related training in the last three months.

Section 4.5.1: Blinder-Oaxaca Decomposition of the Wage Gap

Tables 4.10-4.12 display the results of the basic Blinder-Oaxaca decomposition of the difference in the men's log wave between natives and EU10 immigrants separately for the three periods (1994-2003, 2004-2008, 2009-2013). The first table compares natives and all migrants, whereas Tables 4.11 and 4.12 compare natives with newer and more established migrants separately for the three periods.

When comparing wage differentials at the mean for the entire migrant population the differences in the logarithm of earnings increased in the second and third period relative to the first, from 0.02 to 0.46 and 0.42 respectively, the latter two statistically significant at 5% level. In the first period, 1994-2003, there the wage gap is small and not statistically significant different from zero at 5% and 10% levels and neither is the part of the gap explained by differences in characteristics between the two groups nor the residual gap are statistically significant. This does not indicate that migrants are relatively underpaid in the first period, the unexplained component is relatively large. The explained component is statistically insignificant, but negative, indicating an expected wage advantage for migrants in the first period. This result is mainly driven by relatively higher levels of education (-0.041) and location effects (-0.059), both of which are statistically significant. The latter result follows from the statistics in Table 4.6, indicating that migrants are relatively more concentrated in London and the South East of England. These effects are offset by natives having relatively better job characteristics, accounting for a positive native wage advantage of 5%. Furthermore, natives have a significant 41% unexplained advantage given relative qualifications in the first period, despite migrants being relatively better educated. This coefficient accounts for the majority of the unexplained component in the decomposition between natives and all migrants in the first period.

In the second period, the difference in the mean log wage increases considerably to 0.46 and becomes strongly statistically significant, which may be due to the negative self-selection of the immigrant population pointed out in Longhi and Rokicka (2012). Differences in characteristics statistically significantly explain 47.8% of this gap, whereas the remainder is unaccounted for and may include discrimination. Like in the first period, differences in qualifications indicate that migrants have a small wage advantage of 3%, while natives display a large unexplained advantage given their relative educations. The latter coefficient is comparable to its first period equivalent and strongly statistically significant. This result might indicate that native and migrant labour may not be perfectly substitutable. As opposed to the first period, industry has become strongly statistically significant and accounts for 6.5% of the differential in log wages between natives and all migrants. It indicates a 3.1 percentage point advantage for natives, suggesting negative migrant selection across occupations as outlined in Nickell and Salaheen (2017), Longhi and Rokicka (2012) and Dustmann, Frattini and Preston (2007). Given their relative distribution of persona characteristics, the decomposition explains another 10-percentage point advantage for natives, however, its respective unexplained component is significantly larger at 0.6. Differences in job characteristics indicate a positive and statistically significant native wage advantage of 15%, a significant increase from its first period counterpart. On the other hand, differences in job characteristics do not indicate any significant unexplained component in regard to the difference in log wages between natives and EU10 migrants.

The unexplained component attributed to personal characteristics is the largest contributor to the overall unexplained component in the second period. As opposed to its first period counterpart, the unexplained component of the decomposition containing time and region effects both become positive but not statistically significant. Given their distribution across occupations, there is a statistically significant unexplained wage advantage for natives of 22%. This result is consistent with previous literature which pointed out the relative skill diminishment and underemployment of EU10 migrants after the 2004 accession.

In the third period, the wage differential stays relatively constant relative to the second period at 42%, but the proportion explained by differences in characteristics increases slightly to 52%. The pattern explained components in the variable groups is similar to their second period equivalent, with differences in personal characteristics, industry and job characteristics predicting positive gaps of 6.2%, 6.3% and 12% respectively. One notable exception occurs in the positive sign in the unexplained component for education, which reduces 71% between the

two periods, which indicates better integration of migrants in the labour market in the third period. Similarly, the unexplained wage advantage of natives given occupational distribution becomes statistically insignificant in the third period, further indicating improved integration of EU10 migrants relative to the second period. The decomposition again indicates an explained wage advantage for migrants given their educational endowment of 1.3%, a decrease from the second period. Time and region fixed effects both report small but significant positive wage differentials for migrants, at 0.6 and 0.4% respectively. Overall, in the third period I observed relatively constant migrant wage handicaps compared to the second period as well as unexplained wage deficits. The mean Blinder-Oaxaca decomposition suggest that the observed wage differentials are largely explained by variations in qualifications, job characteristics and occupation and the unobserved differentials are significant and relatively small in magnitude post accession. However, these unexplained differentials should not be immediately interpreted as discrimination, since the effects of unobservable and unmeasurable characteristics relevant in wage determination are also present here. Overall, the results indicate that before the accession there was no statistically significant wage gap between the native and migrant male workers. After accession, a positive and statistically significant gap is observed at the mean, but the evidence suggests that EU10 migrants become better integrated in the UK labour market in the third period. Interestingly, differences in qualifications explain a small wage advantage for migrants in all period, while the unexplained component attributed to qualifications drops significantly in the third period. The unexplained component attributable to personal characteristics reduces between the second and third periods, bringing further evidence in favour of better EU10 migrant integration in the UK labour market.

Table 4.11 displays the results of the standard Oaxaca decomposition when comparing natives with recently arrived migrants who were present in the UK for less than five years. In the first period, the gap becomes positive, but remains statistically insignificant, indicating a 10% advantage for natives, out of which 33% is explained by differences in characteristics. As opposed to the whole population decomposition, newer migrants have relatively less valuable personal characteristics relative to natives, for whom the decomposition predicts a 12% wage advantage. Relative to the results in Table 4.10, the migrant wage advantage explained by education is relatively larger, at 15%, while the equivalent unexplained component indicates a statistically insignificant positive wage gap for natives. As expected, differences in job characteristics explain a 16-percentage point wage advantage for natives, which arises because the amount of experience migrants have in this decomposition is limited by construction. As in

the results in the first two columns of Table 4.10, migrants display a small wage advantage of 4% due to their concentration in London and the South of England. However, the unexplained component coefficient for location suggests this migrant subgroup is less geographically concentrated. It is important to note that while not statistically significant, the observed wage in the first period is significantly larger than the one observed between natives and all migrants, and significantly smaller than the ones observed in Table 4.10 in the second and third periods. This suggests that migrants who were in the UK for more than five years were significantly better integrated in the UK labour market than more recent migrants, even before their country acceded to the EU.

The equivalent wage gaps for the second and third periods are larger than those observed in Table 4.10 but follow the same pattern. In both periods, the observed log wage gap remains roughly constant at 0.5, out of which 50% and 60% are explained by differences in characteristics in the second and third periods respectively. On the other hand, the unexplained component is positive in all three periods, but statistically significant in the second and third periods. However, the unexplained component decreases in magnitude in the third period, which follows the pattern observed in the overall decomposition, indicating worsened labour market integration in the second period and improved labour market integration in the third period. Similar to the results observed in the first period, the explanatory pattern of personal characteristics remains constant at 12% in the second period, and at 10% in the third period. However, in the periods post accession, the unexplained component attributed to this category increases and becomes statistically significant, at 0.51 and 0.43 in the second and third periods respectively. Similar to the results in Table 4.10, differences in education indicate a positive wage differential of 3.7% for migrants in the second period and 1.2% in the third. As observed in the analysis in Table 4.10, the unexplained component attributed to this category reduces significantly between the second and third period, indicating relatively improved market integration of migrant skills. Occupational heterogeneity is significant in explaining positive native wage gaps in both post-accession periods, explaining gaps of 16% and 14% respectively. However, in the third period, there is an 8.2 percentage point unexplained wage deficit for migrants. Even though I observe a large and statistically significant wage gap between natives and migrant men from EU10 countries, the same result as in Table 4.10 is observed, with a larger proportion of the gap being attributable to observable characteristics. Again, it is noteworthy that while differences in qualifications indicated an explained statistically significant wage advantage for migrants in all three periods, but significantly reduced post

accession. This is in accordance with Longhi and Rokicka (2012) and Drinkwater et al. (2009) since the stock of new migrants in the second and third periods are relatively less educated. Furthermore, the results indicate that more recent migrants are not as well integrated in the labour market as the overall stock of migrants, suggesting that more established migrants are better integrated and face smaller wage gaps in all three periods.

Table 4.12 displays the decomposition results when comparing natives with established migrants who are present in the UK for at least five years. In the first period, the observed wage differential between natives and the subset of migrant men is negative, statistically insignificant and insignificant in magnitude. Notice this is not consistent with the first period wage differentials observed in Tables 4.10 and 4.11. The observed wage differential indicates a 0.3% advantage for migrants, while differences in characteristics explain a gap of roughly 3% in favour of migrants.

As opposed to the newer migrant decomposition, established migrants have relatively more valuable personal characteristics relative to natives, for whom the decomposition predicts a 2.4% wage advantage. In the second and third periods, differences in personal characteristics now account for roughly 2.9 percentage point wage advantages for natives. However, the unexplained component attributed to personal characteristics decreases relative to the decompositions in Tables 4.10 and 4.11 but remains statistically significant in the third period. While the unexplained component attributed to personal characteristics is statistically significant and positive, it is smaller than the equivalent for migrants present in the UK for less than five years. This can be indicative of more established migrants being better integrated in the labour market in the third period. Unlike the results in Tables 4.10 and 4.11, the migrant wage advantage explained by education becomes statistically insignificant while natives display a large unexplained advantage given their relative educations in the first period. This unexplained native advantage given heterogeneity in education persists in the second period, but changes sign and remains statistically significant in the third period. The explained component attributable to education is negative and strongly statistically significant in the second and third period, indicating migrant wage advantages of 6% and 1% respectively. This indicates that more established, migrants display better returns to education in the British labour market in the third period and are more easily integrated as their experience increases. This result suggests that while some gaps persist even for more established migrants in the third period, there is no evidence of discrimination in terms of education for this group of migrants.

However, it is important to notice that a smaller percentage of the observed wage differential in the third period is explained by differences in observable characteristics in the third period in this decomposition relative to that in Table 4.11. As expected, differences in job characteristics explain a significantly lower native wage advantage, of 3%, 6% and 9% in the three periods respectively. However, in the third period, natives display an unexplained wage advantage resulting from job characteristic, potentially indicating discrimination. As in the results in the first two columns of Table 4.10, migrants display a small wage advantage in all periods due to their concentration in London and the South of England. However, the unexplained component coefficient for location controls suggest this migrant subgroup is again more geographically concentrated in London and the south of England.

Overall, in the third period I observed a significant increase in the observed wage differential between natives and migrants who were in the UK for at least five years, which is statistically significant. This is smaller in magnitude than the one observed in the decomposition in Table 4.11, indicating that more established migrant men are better integrated in the labour market. However, the percentage of the observed gap that is explained by observed characteristics is smaller for more established migrants whereas the unexplained component is roughly equal between the two decompositions, which may indicate an upper limit on the labour market integration an EU10 migrant might achieve. The positive and significant unexplained coefficients for industry and job characteristics may indicate occupational sorting amongst lower paying positions, consistent with restricted job opportunities and the (temporary) underemployment suggested by Nickell and Salaheen (2017), Longhi, Nicoletti and Platt (2013) Longhi and Rokicka (2012), Kahanec (2012), Dustmann, Frattini and Preston (2007). Quantile regressions are carried out in the next subsection to test whether there are sticky floors and glass ceilings, as well as analysing whether the source of the unexplained differences lies at the bottom of the distribution or towards the top.

Section 4.5.2 Quantile Blinder-Oaxaca Decomposition Results

All decomposition results listed in this section are inclusive of time and region effects. Table 4.13 displays the results for wage decomposition over the three periods. In the first period, the decomposition suggests that natives enjoy a 10% wage advantage relative to migrants at the mean, which is statistically significant. Differences in characteristics explain, however, a wage differential of 21.7%, which suggests that migrants enjoy an unattributable wage benefit of 11%. The former is partially explained by region effects (9.1% advantage,

significant at 1% level), and the relative propensity of migrants to settle in London and the south of England in the first period. Similar to the regressions in Table 4.10, job characteristics explain a significant 13.1% wage penalty relative to natives, while qualifications explain a further gap of 9.8% (significant at 1% level). Interestingly, at the bottom 10th quantile, a negative statistically significant migrant wage gap of 60% is observed, of which a significant 10% is explained by differences in explanatory variables (mainly by qualifications and job characteristics). Region fixed effects indicate a 4.1% wage advantage for migrants. The significantly larger quantile gap and resulting unexplained wage penalty of 54% brings weak evidence of sticky floors for migrants in the first period. At the 90th quantile, the overall gap indicates a 10.9% wage advantage for natives, out of which 14.7% is attributable to differences in characteristics. All coefficients at this quantile are significant except for occupation. Region fixed effect which explain a migrant wage advantage of 9.8%. The significant wage advantages may indicate the relative propensity of migrants to settle in London and the south of England in the first period. The results at this quantile bring evidence against the idea of glass ceilings. However, the quantile results in this period are not as strong due to the relatively small migrant sample size.

In the second period, the 50th quantile migrant gap is smaller than at the mean (39.3%) and statistically significant at 99% confidence interval, while the explained components account for the majority of the gap at 38% (also strongly significant). Occupational distribution accounts for a 1.1% native wage advantage, while job characteristics and qualifications explain gaps of 18.8% and 17% respectively; these indicate negative migrant self-selection in terms of qualifications and job distribution. The gap is significantly smaller at the 10th quantile (13.9 %) and strongly significant, although larger than in the first period. However, differences in characteristics explain a strongly significant gap of 7.3%, indicating a positive, significant and unexplained native wage advantage 6.6%. Job characteristics explain most of the explained gap each with 9.4% significant wage penalty, consistent with the relatively lower quality inflow after 2004 and temporary underemployment observed by Longhi and Rokicka (2012). At the 90th quantile, the migrants face a relative and significant wage penalty of 60.3%, of which only 32.6% (or 54%, and significant) is attributable to observable differences, which results in an unexplained and significant wage penalty of 27.6%, which includes unobserved characteristics, including discrimination. However, quantifying discrimination is not within the scope of this paper. The largest contributors to the gap are differences in job characteristics (9.1%), personal traits (12.7%) and qualifications (14.2%), all of which are statistically significant. Like the first

period, region effects are significant at all quantiles, despite the change in location distribution (see Table 4.6). Meanwhile, time effects become significant in all quantiles due to the sensitivity of the migrant stock to macroeconomic conditions (Blanchflower and Shadworth (2009)). These results disprove the existence of sticky floors in the second period but provide evidence of glass ceilings for migrants.

In the third period, a significant (at 99% confidence interval) migrant wage penalty of 3.8% is observed at the 50th quantile. Similar to the second period, differences in explanatory variables explain most of the migrant gap (3.3%), resulting in a significant and unexplained migrant wage disadvantage of 5.1%. Occupation distribution explains 49% of the total explained gap (1.6%), while job characteristics and qualifications also contribute 16% and 12.6% each towards explaining the observed gap. Personal traits indicate a migrant wage gap of 3.1%, which indicates a reduction in the negative migrant self-selection observed after 2004. At the 10th quantile, migrants suffer a statistically significant wage penalty of 12.3% (as opposed to 13.9% in the second period), while differences in characteristics predict a penalty of 10.3%, and thus an unexplained wage advantage of 2% (both are significant at 99% confidence interval). Similar to the second period, a substantial amount of the explained component is due to the coefficient of occupation (2.3%) and job characteristics (7%). Personal traits indicate a smaller migrant wage penalty relative to the second period (1.7%), indicating a relative improvement in the set of migrant characteristics. The 90th quantile decomposition indicates an even negative and significant wage differential for migrants (relative to the second period), while the explained component is relatively smaller at 24.2%. Personal characteristics explains the largest share of the gap at 9.6%, and qualifications (8.6%). The wage penalty explained by job characteristics is now smaller at 5.7%, which may indicate a lower share of new migrants in this period. The unexplained component is also relatively larger at 36.5%, giving evidence of a glass ceiling for the migrant population. Interestingly, the pattern of migrant relative wage advantages explained by time effects observed in the second period no longer applies. Instead, they predict negligible and significant wage penalties of 0.6%, 0.5% and 0.6% in the 10th, 50th and 90th quantiles respectively.

When analysing the pattern of integration in the British labour market across periods, two findings emerge. Firstly, the integration of migrants at the bottom of the income distribution improves significantly immediately after accession (period 1 to period 2), and between the second and third period. The proportion of the observed gap explained by differences in observable characteristics increases from 10% in the first period to 83% in the

third period, indicating that low income EU10 migrants are considerably better able to integrate as a result of their country's accession in 2004. The unexplained component remains positive and statistically significant for this migrant subset across all three periods, but it reduced greatly from a 53% wage disadvantage in the first period, to a 2% wage disadvantage in the third. This suggests that not only did migrants not display 'sticky floors' after their country joined the EU, their joining did in fact greatly enhance their labour market integration. At the 50th quantile, the observed wage differential between natives and migrants increased significantly from 10% in the first period to roughly 39% in the second and third periods. Whereas in the first period, the decomposition indicated that the observed wage gap should be even higher at 21%, in the second and third period, the majority of the observed gap is explained by differences in observed characteristics. In the second period, the proportion explained by the decomposition is 97%, whereas in the third period 87% is explained. The unexplained component is positive in both post-accession periods and increases from 1% in the second period to 5% and statistically significant in the third. While this suggests slightly worse integration five years after the accession, migrants in the middle of the income distribution are considerably better integrated in the British labour market relative to before 2004.

Secondly, for migrants at the top of the income distribution, we observe a glass ceiling after the 2004 EU enlargement. In the first period, the observed differential of 10% is positive but not statistically significant. However, the observed differential in the second and third period increases to 60% which is strongly statistically significant in both periods. The evolution of the explained component follows a similar pattern. In the period before the accession, the decomposition results suggest an expected wage differential of 14% in favour of natives. In the second period it increases to indicate an expected wage differential of 32%, while the third period equivalent results decreased to indicate a 24% explained wage advantage for natives. This result compounds the diminished integration observed from the 60% wage advantage for natives as a smaller percentage of the observed differential is explained by differences in observed characteristics. The unexplained component evolves in a similar manner, from negative and statistically insignificant (3% migrant wage advantage) before the accession, increasing to 27% and 36% in the second and third periods respectively. These results are significant, indicating that the proportion of the observed wage differential not attributable to observed characteristics increases from 45% in the second period to 60% in the third. These three results indicate that the integration of migrant men from the EU10 countries deteriorated significantly after their country joined the EU and that a 'glass ceiling effect' is observed. These

results indicate that UK policy intervention is required to address this effect and improve EU10 migrant integration at the top of the income distribution.

Section 4.5.3 Decomposition of the Gap in the Probability to Work a Second Job

Table 4.14 displays the results of the Logit-Oaxaca decomposition of the gap in the probability of working a second job between UK natives and EU10 immigrant for the pre-accession period 1994-2003 and the two post-accession periods 2004-2008 and 2009-2013. The decomposition results are expressed in terms of the difference in probability of working a second job.

As mentioned previously this methodology uses pooled coefficient estimates to compute predicted probabilities$\widehat{y_i}$ for each native or migrant observation in the sample and then matches them by their respective relative rankings. Since the sample of natives is larger, the estimates obtained depend on the choice of natives sampled. To alleviate this, I draw 100 random subsamples, calculate separate decomposition estimates and use the mean estimates to approximate the entire native population results. The contribution estimates reported below are the mean values obtained from replicating the decomposition 100 times. One potential drawback of this methodology is that coefficient order can potentially affect the results. To alleviate the concern of path dependence, all models are run with a randomised variable order and therefore the reported results are approximations across all possible orders.

In the first period, the Blinder-Oaxaca decomposition reveals that in the first period the gap in the probability of working unpaid overtime is 0.0037% where a positive sign indicates natives were more likely to work a second job. Alternatively, natives in the first period were 12.2% more likely to work a second job. However, heterogeneity in endowments across the two groups predicts that migrants should be more likely to work a second job at 0.046%, although this result is not statistically significant. In the second and third periods, the observed gap in the probability of working a second job increases to 1.78% and 1.5% respectively (natives were 34% and 36% more likely to work a second job), with the percentage explained by heterogeneity in characteristics at 62% and 71.3% respectively. In the second period, the differential increases may be due to the negative self-selection of the immigrant population pointed out in Longhi and Rokicka (2012) and the migrant downgrading suggested by Rosso (2013).

In the first period, differences in personal and job characteristics indicate that migrants are expected to be more likely to work a second job, predicting probability differences of

0.025% and 0.047% respectively. In the second and third periods, personal characteristics predict natives are more likely to work a second job at roughly 0.1% in both periods. Differences in job characteristics become positive and insignificant in the third period, but still predict a small increase in the probability that a migrant is working a second job (0.08%). Differences in education indicate that natives are more likely to work a second job in all periods, by 0.13% in the first period and increasing to 0.55% and 0.36% in the latter two. This is consistent with the migrant stock post accession becoming relatively less skilled or less substitutable to native labour compared to the migrant population in the first period. Differences in occupation initially are statistically insignificant in the first period. In the second and third periods, they indicate that given occupational distribution, natives are more likely to work in sectors which are more likely to be supplemented by a second job. The increase in probability in the latter two periods is of 0.46% and 0.53% respectively. The persistent hours gaps suggested by the occupation category in the latter two periods provides evidence against occupational sorting of migrants amongst lower paying positions which may require to be supplemented by a second income. This can also be consistent with restricted job opportunities and the (temporary) underemployment suggested by Nickell and Salaheen (2017), Longhi, Nicoletti and Platt (2013) Longhi and Rokicka (2012), Kahanec (2012), Dustmann, Frattini and Preston (2007). The results in this section suggest that the probability of working a second job is correlated with lower education, as it is a mean to supplement one's income, but is correlated with longer hours and lower pay.

This can be interpreted as homogenization of the two groups post expansion, but the persistent unexplainable difference indicates that migrant work substitutability remains imperfect. However, the unexplained differentials in all periods should not be immediately interpreted as discrimination since the effects of unobservable and unmeasurable characteristics relevant in wage determination are also present here.

Section 4.6 Conclusion

This paper quantified and analysed the wage differentials between native and immigrant men from countries which joined the European Union in 2004, as well as analysing whether more established migrants are better integrated in the British labour market, both before and after 2004. To discern if EU membership had any effect on the relative outcomes of the migrants, I compare the hourly wage of natives and migrants in the period immediately preceding accession and two post-accession periods, 2004-2008 and 2009-2013. This study provides evidence of remuneration gaps between natives and A10 immigrants in the post-

accession periods, but not in the pre-accession period. While the observed wage differential is statistically insignificant in the first period, I find a difference in hourly wage between natives and migrants of about 46% in the second period and of about 42% in the third period. The increase in the gap in the post-accession periods can be seen as a result of a relatively lower skilled immigrant inflow after 2004, as also indicated in Longhi and Rokicka (2012) and by the increase in the percentage of EU10 migrants with Other Qualifications (Table 4.8) .

Secondly, the paper attempted to ascertain to what extent these differentials are arising because of differences in the distribution of characteristics across the two populations. Using the Blinder-Oaxaca decomposition method, the paper provides evidence that the differences in education, job and personal characteristics explain about 50% of the afore-mentioned gaps in earnings in the post accession periods.

The paper then uses the Re-Centred Influence Function extension proposed by Firpo et al. (2007) to extend the analysis into a decomposition at the quantiles of the wage distribution and provide a more detailed examination of the relative labour market outcomes of immigrants in the three periods. The results suggest that joining the European Union improved the labour market outcomes of A10 nationals at the bottom of the wage distribution at the cost of creating a glass ceiling for the migrants at the top.

The quantile Oaxaca decomposition results indicate that migrants at the bottom of the wage distribution (10^{th} percentile) faced a large and unexplained wage gap in the first period of 53%. This unexplained wage gap at the bottom of the distribution is reduced to only 6% and 2% in the post accession periods. On the contrary, in the post-accession periods, there seems to be a large unexplained wage gap at the top of the wage distribution suggesting that policy interventions are needed to reduce this glass ceiling effect for EU10 immigrants. The results in the middle of the distribution indicate that while the observed wage gap increased in the second period and decreased slightly in the third, better labour market integration for EU10 migrants is observed relative to the period before the 2004 enlargement.

This paper also looks at the difference between natives and immigrants in the probability of having a second job and uses the Logit-Oaxaca decomposition to analyse if this difference is explained by differential characteristics. The results indicate that natives are more likely to work a second job in all three periods, with the predicted probability gap of 3.8% in the pre-accession period and of 1.8% and 1.5% in the two post-accession periods. Contrary to expectations, migrants do not seem to work a second job to increase their low earnings. These

differences are however small in magnitude and indicate some convergence in the labour market outcomes of natives and migrants.

Tables referenced in the fourth chapter

Table 4.1: EU Membership Enlargements			
Date	Existing members	New members	Number of member states
01.01.1957	None	Belgium, France, Germany, Italy, Luxembourg, The Netherlands	6
01.01.1973	The above	Denmark, Ireland, United Kingdom	9
01.01.1981	The above	Greece	10
01.01.1986	The above	Spain, Portugal	12
01.01.1995	The above	Austria, Finland, Sweden	15
01.05.2004	The above	Czech Republic, Estonia, Cyprus, Latvia, Lithuania, Hungary, Malta, Poland, Slovakia, Slovenia	25
01.01.2007	The above	Romania, Bulgaria	27
01.07.2013	The above	Croatia	28

Table 4.2: Macroeconomic Characteristics of A10 Countries 2004				
Country	GDP per capita (Constant 2000 US dollars)	Unemployment rate		Population (millions)
		Youth	Overall	
Czech Republic	11 947.09	19.9	8.3	10.1971
Cyprus	22 781.79	8.7	4.6	1.0158
Estonia	8 849.32	26.5	10.1	1.3626
Hungary	10 102.5	14.4	6.1	10.1071
Latvia	6 080.79	19.3	11.7	2.2631
Lithuania	6 709.74	21.2	10.9	3.3771
Malta	14 064.02	18.3	7.2	0.4012
Poland	6 639.39	40.1	19.1	38.1822
Slovakia	10 491.81	32.8	18.4	5.3723
Slovenia	17 316.17	14	6.3	1.997
UK (Reference)	39 822.94	11.4	4.7	59.987

Table 4.3: Summary Statistics by Period and Subgroup, First Period (1994-2003)								
	Natives		All migrants		Migrants present for up to five years		Migrants present for more than 5 years	
	Mean (Standard Deviation)	T-test p-value H_0: $mean(0) - mean(1) = 0$	Mean (Standard Deviation)	T-test p-value H_0: $mean(0) - mean(1) = 0$	Mean (Standard Deviation)	T-test p-value H_0: $mean(0) - mean(1) = 0$	Mean (Standard Deviation)	T-test p-value H_0: $mean(0) - mean(1) = 0$
Period 1 (1994-2003)								
AGE	39.67 (11.55)	1	39.88 (11.16)	0.18	29.04 (7.92)	0.00	41.53 (10.62)	0.01
Log Hourly Pay	2.11 (0.59)	1	2.09 (0.63)	0.73	2.00 (0.58)	0.20	2.12 (0.62)	0.75

Length of Continuous Employment (Months)	108.40 (109.74)	1	76.46 (84.39)	0.00	19.56 (19.26)	0.00	86.00 (86.90)	0.00
Years in labour market	22.42 (12.27)	1	24.87 (15.79)	0.00	3.08 (1.47)	0.04	28.42 (14.02)	0.00
Years of education	12.25 (2.60)	1	12.96 (4.13)	0.00	14.75 (3.80)	0.00	12.69 (4.14)	0.00
Less than secondary education dummy	0.60 (0.49)	1	0.36 (0.48)	0.00	0.13 (0.34)	0.00	0.40 (0.49)	0.00
Secondary education dummy	0.24 (0.42)	1	0.29 (0.45)	0.02	0.37 (0.45)	0.02	0.28 (0.45)	0.08

Tertiary education dummy	0.16 (0.37)	1	0.34 (0.47)	0.00	0.50 (0.50)	0.00	0.31 (0.46)	0.00
Public Sector	0.21 (0.41)	1	0.23 (0.42)	0.46	0.13 (0.34)	0.17	0.24 (0.43)	0.24
Manager dummy	0.35 (0.48)	1	0.27 (0.45)	0.00	0.19 (0.40)	0.01	0.29 (0.45)	0.03
Training opportunity	0.27 (0.44)	1	0.23 (0.42)	0.13	0.25 (0.25)	0.72	0.23 (0.25)	0.18
Divorced	0.06 (0.25)	1	0.07 (0.26	0.61	0.00 (0.00)	0.05	0.08 (0.28)	0.23
Has dependent children	0.42 (0.55)	1	0.43 (0.58)	0.59	0.33 (0.51)	0.22	0.45 (0.59)	0.31
Disabled	0.13 (0.34)	1	0.11 (0.31	0.38	0.00 (0.00)	0.00	0.12 (0.32)	0.06

Firm size larger than 250	0.87 (0.34)	1	0.88 (0.32)	0.47	0.67 (0.47)	0.00	0.92 (0.26)	0.01
Number of observations in group	235677		302		52		250	

Table 4.4: Summary Statistics by Period and Subgroup, Second Period								
	Natives		All migrants		Migrants present for less than five years		Migrants present for more than 5 years	
	Mean (Standard Deviation)	T-test p-value H_0: $mean(0) - mean(1) = 0$	Mean (Standard Deviation)	T-test p-value H_0: $mean(0) - mean(1) = 0$	Mean (Standard Deviation)	T-test p-value H_0: $mean(0) - mean(1) = 0$	Mean (Standard Deviation)	T-test p-value H_0: $mean(0) - mean(1) = 0$
Period 2 (2004-2008)								
AGE	41.64 (11.81)	1	30.62 (8.45)	0.00	29.85 (10.67)	0.00	37.21 (11.52)	0.00
Log Hourly Pay	2.40 (0.57)	1	1.94 (0.45)	0.00	1.90 (0.42)	0.00	2.36 (0.55)	0.43

Length of Continuous Employment (Months)	109.26 (111.82)	1	19.43 (29.70)	0.00	14.72 (12.51)	0.00	60.89 (70.19)	0.00
Years in labour market	23.81 (13.17)	1	2.97 (5.34)	0.00	1.86 (1.25)	0.00	12.93 (11.99)	0.00
Years of education	12.67 (2.93)	1	13.47 (4.42)	0.00	13.46 (4.37)	0.00	14.35 (4.13)	0.00
Less than secondary education dummy	0.53 (0.50)	1	0.12 (0.32)	0.00	0.10 (0.31)	0.00	0.23 (0.42)	0.00
Secondary education dummy	0.26 (0.44)	1	0.52 (0.50)	0.00	0.26 (0.44)	0.00	0.31 (0.50)	0.18

Tertiary education dummy	0.21 (0.41)	1	0.35 (0.48)	0.00	0.35 (0.48)	0.00	0.46 (0.47)	0.00
Public Sector	0.22 (0.41)	1	0.04 (0.18)	0.00	0.02 (0.30)	0.00	0.12 (0.33)	0.01
Manager dummy	0.46 (0.50)	1	0.14 (0.35)	0.00	0.12 (0.32)	0.00	0.34 (0.48)	0.00
Training opportunity	0.28 (0.45)	1	0.18 (0.39)	0.00	0.18 (0.39)	0.00	0.21 (0.41)	0.05
Divorced	0.08 (0.27)	1	0.04 (0.18)	0.00	0.03 (0.17)	0.00	0.06 (0.24)	0.51
Has dependent children	0.41 (0.49)	1	0.24 (0.43)	0.00	0.23 (0.42)	0.00	0.42 (0.49)	0.89
Disabled	0.14 (0.35)	1	0.02 (0.13)	0.00	0.01 (0.12)	0.00	0.05 (0.23)	0.00

Firm size larger than 250	0.58 (0.49)	1	0.60 (0.49)	0.15	0.6 (0.49)	0.17	0.60 (0.49)	0.58
Number of observations in group	106185		1132		1033		99	

Table 4.5: Summary Statistics by Period and Subgroup, Third period								
	Natives		All migrants		Migrants present for less than five years		Migrants present for more than 5 years	
	Mean (Standard Deviation)	T-test p-value H_0: $mean(0) - mean(1) = 0$	Mean (Standard Deviation)	T-test p-value H_0: $mean(0) - mean(1) = 0$	Mean (Standard Deviation)	T-test p-value H_0: $mean(0) - mean(1) = 0$	Mean (Standard Deviation)	T-test p-value H_0: $mean(0) - mean(1) = 0$
Period 3 (2009-2013)								
AGE	42.55 (11.95)	1	33.14 (8.46)	0.00	31.29 (7.9)	0.00	35.06 (8.54)	0.00
Log Hourly Pay	2.52 (0.58)	1	2.1 (0.43)	0.00	2.03 (0.42)	0.00	2.16 (0.45)	0.00

Length of Continuous Employment (Months)	115.22 (112.99)	1	36.78 (29.86)	0.00	26.44 (19.35)	0.00	47.30 (32.79)	0.00
Years in labour market	24.43 (13.27)	1	5.49 (4.40)	0.00	3.30 (1.46)	0.00	7.52 (4.74)	0.00
Years of education	13.10 (4.61)	1	13.94 (5.63)	0.00	13.87 (6.36)	0.00	14.01 (4.85)	0.00
Less than secondary education dummy	0.47 (0.50)	1	0.11 (0.31)	0.00	0.10 (0. 30)	0.00	0. 10 (0. 30)	0.00

Secondary education dummy	0.28 (0.45)	1	0.52 (0.50)	0.00	0.55 (0.50)	0.04	0. 50 (0.50)	0.00
Tertiary education dummy	0.28 (0.45)	1	0.36 (0.48)	0.00	0.34 (0.47)	0.00	0.39 (0.49)	0.00
Public Sector	0.22 (0.41)	1	0.04 (0.20)	0.00	0.03 (0.16)	0.00	0.05 (0.33)	0.01
Manager dummy	0.45 (0.50)	1	0.17 (0.38)	0.00	0.14 (0.34)	0.00	0.22 (0.41)	0.00
Training opportunity	0.26 (0.44)	1	0.2 (0.4)	0.00	0.18 (0.39)	0.00	0.19 (0.39)	0.05
Divorced	0.08 (0.27)	1	0.05 (0.21)	0.00	0.04 (0.20)	0.00	0.06 (0.23)	0.00

Has dependent children	0.41 (0.49)	1	0.41 (0.49)	0.75	0.36 (0.48)	0.00	0.47 (0.50)	0.00
Disabled	0.15 (0.35)	1	0.05 (0.22)	0.00	0.04 (0.35)	0.00	0.06 (0.23)	0.00
Firm size larger than 250	0.57 (0.5)	1	0.64 (0.48)	0.00	0.63 (0.48)	0.00	0.63 (0.48)	0.00
Number of observations in group	96349		2085		1186		899	

Table 4.6: Distribution of A10 Migrants Across the UK (Percentages)			
Aggregate Region	Period 1 (1994-2003)	Period 2 (2004-2008)	Period 3 (2009-2013)
1. Scotland	3.42	6.71	9.41
2. Northern Ireland	1.02	5.74	5.40
3. North England	13.01	23.14	21.27
4. Midlands	6.16	19.94	18.40
5. Wales	4.1	2.91	2.67
6. South of England	35.61	27.38	30.78
7. London	36.64	14.13	12.04

Table 4.7: Distribution of Migrants by Occupation Group (Percentages)						
Major Occupation Group	Period 1 (1994-2003)		Period 2 (2004-2008)		Period 3 (2009-2013)	
	EU10	Natives	EU10	Natives	EU10	Natives
1. Managers and Administrators	20.21	19.35	3	20.57	3.3	16.6
2.Professional Occupations	19.18	12.33	5.21	14.64	5.21	18.7
3. Associate Professional and Technical Occupations	7.88	10.73	3.27	14.14	3.78	15.77
4. Clerical, Secretarial Occupations	5.82	7.2	1.41	5.4	2.63	5.48
5.Skilled Trades Occupations	13.7	17.63	17.76	15.13	17.78	14.3
6. Personal Services	10.27	5.5	2.47	2.53	1.63	3.05
7. Sales Occupations	6.85	4.08	2.12	3.62	2.39	4.51
8. Plant and Machine Occupations	11.99	14.94	27.3	12.93	28.39	11.37
9. Elementary Occupations	20.21	8.19	37.46	11	34.8	10.19

Table 4.8: Highest Qualification for Both Subgroups (Percentages)						
Highest qualification	Period 1 (2000-2003)		Period 2 (2004-2008)		Period 3 (2009-2013)	
	UK	EU10	UK	EU10	UK	EU10
1. Degree or Equivalent	17.23	14.19	21.2	10.13	25.79	17.81
2. Higher Education	9.77	3.15	9.77	2.7	10.25	5.57
3. GCE A Level or Equivalent	24.07	10.1	24.4	7.47	23.17	10.04
4. GCSE Grades A-C or Equivalent	21.42	4.41	22.5	2.29	22.42	4.66
5. Other Qualifications	12.51	49.21	9.68	57.36	8.52	45.03
6. No Qualifications	13.09	15.77	11.32	17.96	7.76	14.95

Table 4.9: Mean Log (Wages) by Qualification Cell and Group												
Highest Qualification	Period 1 (1994-2003)				Period 2 (2004-2008)				Period 3 (2009-2013)			
	UK		EU10		UK		EU10		UK		EU10	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
1. Degree or Equivalent	2.57	0.56	2.37	0.78	2.79	0.56	2.31	0.60	2.88	0.57	2.41	0.55
2. Higher Education	2.31	0.51	2.46	0.42	2.57	0.49	2.25	0.56	2.66	0.51	2.28	0.43
3. GCE A Level or Equivalent	2.07	0.52	2.04	0.54	2.34	0.49	1.96	0.40	2.44	0.51	2.11	0.33
4. GCSE Grades A-C or Equivalent	1.96	0.56	1.94	0.46	2.22	0.52	2.06	0.40	2.32	0.51	2.13	0.36
5. Other Qualifications	1.83	0.49	2.00	0.62	2.12	0.46	1.92	0.41	2.22	0.46	2.03	0.40
6. No Qualifications	1.75	0.50	1.74	0.39	2.02	0.45	1.80	0.40	2.11	0.45	2.14	0.30

Table 4.10: Oaxaca decomposition of the difference in mean log wage. All EU10 immigrants.						
Comparison Groups	First Period, all migrants (1994-2003)		Second Period, all migrants (2004-2008)		Third Period, all migrants (2009-2013)	
Overall Difference between UK natives and EU10 immigrants	0.020 (0.03)		0.46*** (0.01)		0.42*** (0.01)	
	Explained component	Unexplained component	Explained component	Unexplained component	Explained component	Unexplained component
	-0.02 (0.024)	0.047 (0.031)	0.22*** (0.010)	0.24*** (0.013)	0.22*** (0.00)	0.19*** (0.00)
Contributions						
Intercept		-0.11 (0.44)		-1.11*** (0.19)		-0.44*** (0.13)
Personal Characteristics	-0.00	0.18	0.10***	0.60***	0.06***	0.47***

	(0.00)	(0.41)	(0.00)	(0.17)	(0.00)	(0.12)
Qualifications	-0.04** (0.01)	0.41*** (0.12)	-0.03*** (0.00)	0.42*** (0.041)	-0.01*** (0.00)	0.12*** (0.02)
Industry Sector	0.00 (0.00)	-0.30 (0.18)	0.03*** (0.00)	0.22*** (0.051)	0.06*** (0.00)	-0.02 (0.03)
Job Characteristics	0.05*** (0.00)	-0.17 (0.12)	0.15*** (0.0043)	-0.00 (0.02)	0.12*** (0.00)	0.07*** (0.02)
Location	-0.05*** (0.00)	-0.075 (0.04)	-0.00 (0.00)	0.047 (0.03)	-0.00*** (0.00)	-0.017 (0.01)
Time	0.02*** (0.00)	0.11 (0.074)	-0.026*** (0.00)	0.054 (0.05)	-0.00*** (0.00)	-0.06 (0.57)
	Natives regression	Migrant regression	Natives regression	Migrant regression	Migrants arrived between 2004 and 2008	Migrants arrived before 2004
Observations	124.472	290	106.304	1132	96544	2092

Adjusted R^2 Natives	0.41		0.38		0.35	
Adjusted R^2 Migrants		0.38		0.30		0.27

Standard errors in parentheses; Notes: *** Significant at 1% confidence interval **Significant at 5% confidence interval *Significant at 10% confidence interval

Personal Characteristics: age, age squared, years in the UK, dummies on dependent children; Qualifications: Years of education; Reference category: No qualifications; Industry Sector: 9 industry sector dummies. Job Characteristics: current employment duration, current employment duration squared, dummies on firm sizes, managerial roles, employment in the public sector and having received job-related training in the last three months' time: Yearly dummies; Location: Region dummies; Control Location=London; Control wave=1994, 2004 and 2008

Table 4.11: Oaxaca decomposition of the difference in mean log wage. EU10 immigrants resident in the UK for 5 years or less.						
Comparison Groups	Natives and Migrants present for 5 years or less (1994-2003)		Natives and Migrants present for 5 years or less years (2004-2008)		Natives and Migrants present for 5 years or less (2009-2013)	
Overall Difference between UK natives and EU10 immigrants	0.10 (0.09)		0.50*** (0.01)		0.49*** (0.012)	
	Explained component	Unexplained component	Explained component	Unexplained component	Explained component	Unexplained component
	0.033 (0.066)	0.072 (0.070)	0.25*** (0.01)	0.25*** (0.01)	0.30*** (0.00)	0.19*** (0.01)
Contributions						
Intercept		-1.02* (0.51)		-1.09*** (0.21)		-0.46** (0.17)

Personal Characteristics	0.12*** (0.02)	0.19 (0.79)	0.12*** (0.00)	0.51** (0.19)	0.10*** (0.00)	0.43** (0.15)
Qualifications	-0.15*** (0.03)	0.32 (0.34)	-0.03*** (0.00)	0.42*** (0.04)	-0.01*** (0.00)	0.12*** (0.03)
Industry Sector	-0.00 (0.01)	-1.56*** (0.29)	0.03*** (0.00)	0.23*** (0.05)	0.06*** (0.00)	0.01 (0.04)
Job Characteristics	0.16*** (0.01)	-0.11 (0.13)	0.16*** (0.00)	0.04 (0.02)	0.14*** (0.00)	0.08*** (0.02)
Location	-0.04*** (0.00)	-0.14* (0.07)	-0.00 (0.00)	0.01 (0.02)	-0.00** (0.00)	0.01 (0.02)
Time	-0.05*** (0.01)	0.27* (0.13)	-0.02*** (0.00)	0.10 (0.06)	0.00*** (0.00)	-0.01 (0.01)
	Natives regression	Migrant regression	Natives regression	Migrant regression	Natives regression	Migrant regression
Observations	124.472	52	106304	1033	96544	1186

Adjusted R^2 Natives	0.398		0.38		0.34	
Adjusted R^2 Migrants		0.70		0.19		0.23

Standard errors in parentheses; Notes: *** Significant at 1% confidence interval **Significant at 5% confidence interval *Significant at 10% confidence interval

Personal Characteristics: age, age squared, years in the UK, dummies on dependent children; Qualifications: Years of education; Reference category: No qualifications; Industry Sector: 9 industry sector dummies. Job Characteristics: current employment duration, current employment duration squared, dummies on firm sizes, managerial roles, employment in the public sector and having received job-related training in the last three months' time: Yearly dummies; Location: Region dummies; Control Location=London; Control wave=1994, 2004 and 2008

Table 4.12: Oaxaca decomposition of the difference in mean log wage. EU10 immigrants resident in the UK for more than 5 years.						
Comparison Groups	Natives and Migrants present for more than 5 years (1994-2003)		Natives and Migrants present for more than 5 years (2004-2008)		Natives and Migrants present for more than 5 years (2009-2013)	
Overall Difference between UK natives and EU10 immigrants	-0.00 (0.04)		0.04 (0.04)		0.36*** (0.01)	
	Explained component	Unexplained component	Explained component	Unexplained component	Explained component	Unexplained component
	-0.03 (0.02)	0.03 (0.03)	0.01 (0.03)	0.03 (0.03)	0.16*** (0.00)	0.20*** (0.01)
Contributions						
Intercept		-0.23 (0.59)		0.32 (0.52)		-0.26 (0.59)
Personal Characteristics	-0.02**	0.25	0.03**	0.03	0.02***	0.12**

	(0.00)	(0.56)	(0.01)	(0.49)	(0.00)	(0.0)
Qualifications	-0.02 (0.01)	0.40*** (0.12)	-0.06*** (0.02)	0.45*** (0.13)	-0.01*** (0.00)	-0.06*** (0.01)
Industry Sector	0.00 (0.00)	-0.01 (0.16)	0.01 (0.00)	-0.54*** (0.15)	0.06*** (0.00)	0.09** (0.03)
Job Characteristics	0.03** (0.01)	-0.26 (0.15)	0.06*** (0.01)	-0.17* (0.07)	0.09*** (0.00)	0.05* (0.02)
Location	-0.05*** (0.00)	-0.09* (0.04)	-0.01*** (0.00)	0.07 (0.04)	-0.00*** (0.00)	0.02 (0.04)
Time	0.03*** (0.00)	-0.02 (0.01)	-0.01*** (0.00)	-0.14 (0.08)	-0.00*** (0.00)	-0.33 (0.20)
	Natives regression	Migrant regression	Natives regression	Migrant regression	Natives regression	Migrant regression
Observations	124.472	250	69.241	121	96.544	1218
Adjusted R^2 Natives	0.39		0.40		0.34	

Adjusted R^2 Migrants		0.36		0.20		0.22

Standard errors in parentheses; Notes: *** Significant at 1% confidence interval **Significant at 5% confidence interval *Significant at 10% confidence interval

Personal Characteristics: age, age squared, years in the UK, dummies on dependent children; Qualifications: Years of education; Reference category: No qualifications; Industry Sector: 9 industry sector dummies. Job Characteristics: current employment duration, current employment duration squared, dummies on firm sizes, managerial roles, employment in the public sector and having received job-related training in the last three months' time: Yearly dummies; Location: Region dummies; Control Location=London; Control wave=1994, 2004 and 2008

Table 4.13: Oaxaca decomposition of the difference in log wage at the 10th, 50th and 90th deciles. All EU10 immigrants.										
Period and Percentile		Quantile Gap (Log)	Explained	Unexplained	Explained by					
					Personal	Qualifications	Occupation	Job char.	Region	Time
Period 1	10	0.59*** (0.10)	0.06*** (0.01)	0.53*** (0.09)	0.03*** (0.00)	-0.02*** (0.00)	0.02*** (0.00)	0.07*** (0.00)	-0.04*** (0.00)	-0.01* (0.00)
	50	0.10** (0.05)	0.21*** (0.02)	-0.11** (0.04)	0.06*** (0.01)	0.09*** (0.01)	0.03*** (0.00)	0.13*** (0.01)	-0.08*** (0.00)	-0.01* (0.00)
	90	0.10 (0.07)	0.14*** (0.03)	-0.03 (0.07)	0.09*** (0.01)	0.06** (0.02)	0.01 (0.00)	0.09*** (0.01)	-0.09*** (0.00)	-0.01* (0.01)
Period 2	10	0.14*** (0.01)	0.07*** (0.00)	0.06*** (0.01)	0.02*** (0.00)	-0.01*** (0.00)	0.00* (0.00)	0.09*** (0.00)	-0.00*** (0.00)	-0.02*** (0.00)
	50	0.39*** (0.00)	0.38*** (0.00)	0.01 (0.01)	0.04**** (0.00)	0.17*** (0.00)	0.01*** (0.00)	0.18*** (0.00)	-0.02*** (0.00)	-0.05*** (0.00)
	90	0.60*** (0.02)	0.32*** (0.00)	0.27*** (0.02)	0.12**** (0.00)	0.14*** (0.00)	0.01** (0.00)	0.09*** (0.00)	-0.02*** (0.00)	-0.02*** (0.00)
Period 3	10	0.12*** (0.00)	0.10*** (0.00)	0.02** (0.00)	0.01*** (0.00)	-0.00*** (0.00)	0.02*** (0.00)	0.07*** (0.00)	-0.00*** (0.00)	-0.00*** (0.00)

	50	0.38*** (0.00)	0.33*** (0.00)	0.05*** (0.00)	0.03*** (0.00)	0.12*** (0.00)	0.02*** (0.00)	0.16*** (0.00)	-0.01*** (0.00)	-0.00*** (0.00)
	90	0.60*** (0.01)	0.24*** (0.00)	0.36*** (0.01)	0.09*** (0.00)	0.08*** (0.00)	0.01*** (0.00)	0.05*** (0.00)	-0.01*** (0.00)	-0.00*** (0.00)

Standard errors in parentheses;

Legend: *** Significant at 1% confidence interval **Significant at 5% confidence interval *Significant at 10% confidence interval

Control Location=London; Control waves=1994, 2004 and 2009 respectively

Table 4.14: Decomposition of the difference in the probability of working a second job between UK natives and EU10 immigrants.			
	First Period: 1994-2003	Second Period:2004-2008	Third Period:2009-2013
Comparison group	Natives and all EU10 immigrants	Natives and all EU10 immigrants	Natives and all EU10 immigrants
Average probability of working a second job	4.32%	3.82%	3.97%
	Decomposition	Decomposition	Decomposition
Overall Difference	.00377	0.0178**	0.015***
Total Explained	-.00046	0.0111***	0.0107***
Variables			
Personal Characteristics	-0.00025*** (0.000065)	0.0010** (0.00036)	0.00097** (0.00035)
Qualifications	0.0013*** (0.00019)	0.0055*** (0.00052)	0.0036*** (0.00044)
Industry Sector	-0.00040 (0.00022)	0.0046*** (0.00028)	0.0053*** (0.00034)
Job Characteristics	-0.00047***	-0.00088*	0.00052

	(0.00011)	(0.00040)	(0.00038)
Location	0.0014*** (0.00028)	0.00051** (0.00017)	0.00022 (0.00014)
Time	-0.0020*** (0.00028)	0.00031* (0.00015)	0.00012* (0.000061)

Standard errors in parentheses;

Notes: *** Significant at 1% confidence interval **Significant at 5% confidence interval *Significant at 10% confidence interval

Personal Characteristics: age, age squared, years in the UK, dummies on marital status, dependent children, gender, and disability

Qualifications: Years of education; Reference category: No qualifications

Industry Sector: 9 industry sector dummies.

Job Characteristics: current employment duration, current employment duration squared, dummies on firm sizes, managerial roles, employment in the public sector and having received job-related training in the last three months, basic usual hours

Time: Yearly dummies; Location: Region dummies; Control Location=London; Control waves=1994, 2004 and 2009 respectively

Chapter 5: Conclusion

The fundamental result of the first paper is that economic indicators were the main predictors of Leave vote percentages, whereas total weighted media coverage had a relatively small and overall statistically insignificant effect on aggregate. The first set of results is in line with the literature of Becker et al (2017) and Fetzer (2018). My results found that constituencies which experienced a decrease in public spending and EU funding and those which experienced increases in unemployment in the year preceding the referendum were estimated to vote more in favour of exiting the EU. These effects are statistically significant and significant in magnitude. The results of Fetzer (2018) with regards to the effects of austerity on the propensity to vote Leave are also found in this paper as one percentage point decrease in public spending in the year before the referendum has strong effect in all specifications, larger in magnitude than 2.3 percentage points. A one percentage point increase in the unemployment rate at the 2011 census is estimated to have significantly increased the propensity to vote to Leave by approximately four percentage points in the benchmark specifications. Furthermore, a constituency level of one percentage point decrease in unemployment in the year before the referendum is estimated to have decreased the Leave vote by roughly 2.2 to 2.4 percentage points in all specifications.

The use of the weighted media variables proposed in this paper mitigates the drawbacks of the relevant article count variable (defined in Equation 1) used in previous literature such as Lamla and Lein (2008) and van der Wiel (2009), leading to more accurate estimates and thus eliminating bias resulting from measurement error.

The benchmark weighted media results indicate that newspaper coverage of migrant crime had an initial small, positive, and statistically insignificant effect on the constituency level percentage Leave votes regardless of the version of coverage used. Further analysis reveals heterogeneities of effects both from relative population qualifications and from coverage type. The estimates suggest that negative coverage had a small, statistically significant and positive effect in constituencies in the bottom three quartiles of the qualification distribution. This provides evidence in favour of information bias since overall media coverage had a statistically significant effect which persists after accounting for migration and economic indicators. However, this effect dissipates in the top quartile of the qualification distribution for all coverage types and all periods. Furthermore, the estimated results of the dummies

indicating the relative qualification levels of the electorates are all negative and strongly statistically significant in all specifications. This suggests that the percentage of Leave votes is significantly lower in constituencies with higher proportions of highly educated people.

Analysis on publication type and stance reveals a heterogeneous pattern of effects. Coverage originating in broadsheets which favoured Remain was not found to have significantly altered the referendum vote. On the other hand, tabloids which supported Remain were successful at dissuading voters from voting Leave, which is significant in all specifications. Coverage originating in broadsheets which supported Leave is estimated to have increased the Leave vote percentage in the last six months before the referendum for all coverage types and in both periods in the case of anti-EU coverage. Anti-EU coverage is also found to be positive and statistically significant if originating in Leave-supporting tabloids, but not if found in the other two definitions. This finding is similar to that of Chiang and Knight (2011) in the sense that credibility affects the effect of media coverage. These results are also in line with the studies of Williams and Dickinson (1993), Banerjee et al. (2013), and Larreguy et al. (2014), among others, where being exposed to information can alter both the perception and voting actions of individuals in line with the message being relayed. Furthermore, this study finds evidence that not only are more qualified electorates less prone to media persuasion, but also that voters are aware of the reputation of a newspaper. This result is strengthened by Larreguy et al. (2014) and Chiang and Knight (2011), who also find evidence of voter sophistication along the lines of reputation and expected bias. A further analysis of this manner in the context of the 2014 Scottish independence referendum would be an interesting avenue for further research, with a specific focus on pro-independence coverage

Looking at the effects of migrant shares, only the migrant stock from the post-2004 accession countries is estimated to have had a positive statistically significant effect on the propensity to vote Leave, which is consistent across specifications. On the other hand, the estimated effects of the pre-2004 accession EU migrants and non-EU migrants is still strongly statistically significant and negative, indicating the propensity to vote Leave is lower in areas that are more cosmopolitan and possible migrant self-selection. On the other hand, recent net migration is not a significant predictor of the percentage of Leave votes. Overall, the effects of economic coefficients are largely identical between the regressions involving all types of negative coverage considered in this paper. Although there is some evidence of significant media effects in this study, the results on economic factors are significant predictors of the

constituency level Leave votes in all specifications. The results on migrant shares are similar to those obtained by Becker et al. (2017).

The first paper has two main limitations. The first is the endogeneity issue outlined above, since the resulting OLS estimates would be biased upwards. However, given that the results are already relatively small, this reinforces the finding that the overall media effect is insignificant in magnitude. The Leave voting intention variable obtained from the May 2015 British Election Study survey is added to account for potential differences in latent Euroscepticism and to alleviate concerns about endogeneity.

The second paper generalises the seminal paper of Besley and Prat (2006) and those of Trombetta and Rossignoli (2020) and Corduneanu-Huci, and Hamilton (2018). However, a key difference is that my theoretical results indicate that in pure strategy Perfect Bayesian Equilibria, media and political competition and electoral standards ensure that the true state of politicians is always revealed to a subset of the electorate. However, if the proportion of voters who actively seek political news is high enough, full disclosure occurs in the media market even if both politicians are bad. The present model can be seen as a generalisation of the previous models as it collapses to that of Trombetta and Rossignoli (2020) if the politician's budgets are highly asymmetrical and to that of Besley and Prat (2006) with the additional assumption that the share of informed voters is zero.

These findings are relevant given the recent attention given to the proliferation of possibly distorted media via the internet and its effects on political results and resulting voter welfare. The results presented in this paper suggests that while media competition mediates the propensity of media suppression, political competition and electoral standards help eliminate total media suppression. In line with the results of Corduneanu-Huci and Hamilton (2018), the politicians prefer the most cost-effective alternative when interfering in the media market, which in their framework never occurred in the pure strategy Perfect Bayesian Equilibrium. However, their result relied on the assumption that the politically motivated audience is relatively small and serviced by their own section of the media. In my model, the addition of political competition renders total media suppression strictly suboptimal in pure strategy PBE. With constant budgets and relatively small politically interested audiences, a good politician always prevents a bad politician from winning via media interference and hence full disclosure is observed in the media. Furthermore, the results hold even when the share of informed voters exceeds 50%, in which case the pure strategy equilibrium suggests full media disclosure

regardless of the type of the politicians. This result relies on the assumption that voters would prefer one politician when they observe both alternatives as bad. The choice of the incumbent as the favourite in this situation is not crucial as full disclosure is still the equilibrium should voters prefer the challenger.

Similar results hold with asymmetric politician budgets if the differences in budgets are not too large. The results indicate that unlike the constant budget case, a bad politician would be able to achieve meaningful media suppression that would allow him to win the election. Similar to the results of Trombetta and Rossignoli (2020), the model predicts heterogeneous levels of media suppression, with total media suppression only occurring if the share of informed voters is at least half. This result approximates those of Trombetta and Rossignoli (2020) and holds despite reputation concerns in the media market.

These results become important in the context of internet proliferation of new media outlets and sliding democratic standards observed in autocracies and democracies alike. They give evidence that larger politically involved sections of the electorate together with political competition can help reduce the likelihood of distortions in the media markets and voter misinformation. Furthermore, it provides evidence that political competition can help promote government accountability even in the presence of dishonest politicians and dishonest media. The results of this paper challenge this policy implication that media interference is bad, as the presence of a challenger also involved in the media markets results in a good politician always winning the election whenever he is present. A bad challenger never wins against a good incumbent and a bad incumbent always loses to a good challenger, and therefore, political interference in media markets is voter welfare reducing if and only if the interference is a monopsony. The policy implication is that political involvement in media markets can be beneficial if it transparent and open to all candidates, rather than just the incumbent. Given the results obtained in the first period, an interesting avenue for future research would be to analyse theoretically the trade-off between de facto media independence and the probability of being captured by politicians.

The final paper quantified and analysed the wage differentials between native and immigrant men from countries which joined the European Union in 2004, as well as analysing whether more established migrants are better integrated in the British labour market, both before and after 2004. To discern if EU membership had any effect on the relative outcomes of the migrants, I compare the hourly wage of natives and migrants in the period immediately

preceding accession and two post-accession periods, 2004-2008 and 2009-2013. This study provides evidence of remuneration gaps between natives and EU10 immigrants in the post-accession periods, but not in the pre-accession period. While the observed wage differential is statistically insignificant in the first period, I find a difference in hourly wage between natives and migrants of about 46% in the second period and of about 42% in the third period. The increase in the gap in the post-accession periods can be seen as a result of a relatively lower skilled immigrant inflow after 2004, as also indicated in Longhi and Rokicka (2012) and by the increase in the percentage of EU10 migrants with Other Qualifications (Table 4.8).

Secondly, the paper attempted to ascertain to what extent these differentials are arising because of differences in the distribution of characteristics across the two populations. Using the Blinder-Oaxaca decomposition method, the paper provides evidence that the differences in education, job and personal characteristics explain about 50% of the afore-mentioned gaps in earnings in the post accession periods.

The quantile Oaxaca decomposition results indicate that migrants at the bottom of the wage distribution (10^{th} percentile) faced a large and unexplained wage gap in the first period of 53%. This unexplained wage gap at the bottom of the distribution is reduced to only 6% and 2% in the post accession periods. On the contrary, in the post-accession periods, there seems to be a large unexplained wage gap at the top of the wage distribution suggesting that policy interventions are needed to reduce this glass ceiling effect for EU10 immigrants. The results in the middle of the distribution indicate that while the observed wage gap increased in the second period and decreased slightly in the third, better labour market integration for EU10 migrants is observed relative to the period before the 2004 enlargement.

This paper also looks at the difference between natives and immigrants in the probability of having a second job and uses the Logit-Oaxaca decomposition to analyse if this difference is explained by differential characteristics. The results indicate that natives are more likely to work a second job in all three periods, with the predicted probability gap of 3.8% in the pre-accession period and of 1.8% and 1.5% in the two post-accession periods. Contrary to expectations, migrants do not seem to work a second job to increase their low earnings. These differences are however small in magnitude and indicate some convergence in the labour market outcomes of natives and migrants.

Appendix A: Appendices to Chapter 2

Appendix A1: Keywords used to build the main specification variables

The keywords used to construct the main set of media coverage variables are:

- For the first version of articles on crime involving migrants
    - ‘Migration’ OR ‘migrant’ OR ‘migrants’ OR ‘immigrant’ OR ‘immigrants’ OR ‘foreigner’ OR ‘foreigners’ OR ‘refugee’ OR ‘refugees ‘in the main text AND
    - ‘crime’ OR ‘criminal’ OR ‘violent’ OR ‘felon’ OR ‘felony’ OR ‘charged’ OR ‘arson’ OR ‘arsonist’ OR ‘arrested’ OR ‘crack-down’ OR ‘crook’ OR ‘smuggling’ OR ‘smuggler’ OR ‘murder’ OR ‘murderer’ in the headline
- For articles displaying migration pressure on public resources
    - ‘Migration’ OR ‘migrant’ OR ‘migrants’ OR ‘immigrant’ OR ‘immigrants’ OR ‘foreigner’ OR ‘foreigners’ OR ‘refugee’ OR ‘refugees‘ in the main text
    - ‘Burden’ OR ‘burdened’ OR ‘add to NHS bill’ OR ‘cost the’ OR ‘cripple’ OR ‘strain’ OR ‘strain on’ OR ‘health tourism’ or ‘bleed the NHS’ OR ‘out of control’ OR ‘NHS charged’ OR ‘Filled’ OR ‘overrun’ OR ‘over capacity’ OR ‘full’ OR ‘strain’ OR ‘strained’ OR ‘flooded with’ OR ‘overcrowded’ or ‘overcrowding’ OR ‘disruption’ OR ‘culture dilution’ OR ‘cultural dilution’ OR ‘overwhelmed’ OR ‘at breaking point’ OR ‘cost the taxpayer’ OR ‘bourne by the taxpayer’ OR ‘Claiming’ OR ‘claiming credit’ OR ‘getting’ OR ‘accruing’ OR ‘claim’ OR ‘get’ or ‘accrue’ or ‘obtain’ OR ‘cheap labour’ OR ‘undercutting’ OR ‘race to the bottom’ OR ‘force wages down’ or ‘forces wages down’ or ‘holding wages down’ OR ‘keeps wages down’ OR ‘keeping wages down’ in the headline
    - ‘National Health Service’ OR ‘JSA’ OR ‘welfare system’ OR ‘NHS’ OR ‘NHS care’ OR ‘health services’ OR ‘hospital’ OR ‘school’ OR ‘schools’ OR ‘schooling’ OR ‘college’ OR ‘sixth form’ OR ‘6th form’ OR ‘Child Benefit’ OR ‘Unemployment Benefit’ OR ‘council tax benefit’ OR ‘disability Benefit’ OR ‘Child Benefits’ OR ‘council tax Benefits’ OR ‘disability Benefits’ OR ‘welfare’ OR ‘job seekers allowance’ OR ‘income support’ OR ‘universal credit employment’ OR ‘support allowance’ OR ‘social care’ OR ‘housing benefit’ in the main text
    - United Kingdom as index
- For the aggregate anti-EU sentiment variable

- ‘crime’ OR ‘criminal’ OR ‘violent’ OR ‘felon’ OR ‘felony’ OR ‘charged’ OR ‘arson’ OR ‘arsonist’ OR ‘arrested’ OR ‘crack-down’ OR ‘crook’ OR ‘smuggling’ OR ‘smuggler’ OR ‘murder’ OR ‘murderer’ in the text
- ‘EU’ or ‘EUSSR’ OR ‘Migration’ OR ‘migrant’ OR ‘migrants’ OR ‘immigrant’ OR ‘immigrants’ OR ‘foreigner’ OR ‘foreigners’ OR ‘refugee’ OR ‘refugees’
- ‘Burden’ OR ‘burdened’ OR ‘add to NHS bill’ OR ‘cost the’ OR ‘cripple’ OR ‘strain’ OR ‘strain on’ OR ‘health tourism’ or ‘bleed the NHS’ OR ‘out of control’ OR ‘NHS charged’ OR ‘Filled’ OR ‘overrun’ OR ‘over capacity’ OR ‘full’ OR ‘strain’ OR ‘strained’ OR ‘flooded with’ OR ‘overcrowded’ OR ‘overcrowding’ OR ‘disruption’ OR ‘culture dilution’ OR ‘cultural dilution’ OR ‘overwhelmed’ OR ‘at breaking point’ OR ‘cost the taxpayer’ OR ‘bourne by the taxpayer’ OR ‘Claiming’ OR ‘claiming credit’ OR ‘getting’ OR ‘accruing’ OR ‘claim’ OR ‘get’ or ‘accrue’ OR ‘obtain’ OR ‘cheap labour’ OR ‘undercutting’ OR ‘race to the bottom’ OR ‘force wages down’ OR ‘forces wages down’ or ‘holding wages down’ OR ‘keeps wages down’ OR ‘keeping wages down’ OR ‘National Health Service’ OR ‘JSA’ OR ‘welfare system’ OR ‘NHS’ OR ‘NHS care’ OR ‘health services’ OR ‘hospital’ OR ‘school’ OR ‘schools’ OR ‘schooling’ OR ‘college’ OR ‘sixth form’ OR ‘6th form’ OR ‘Child Benefit’ OR ‘Unemployment Benefit’ OR ‘council tax benefit’ OR ‘disability Benefit’ OR ‘Child Benefits’ OR ‘council tax Benefits’ OR ‘disability Benefits’ OR ‘welfare’ OR ‘job seekers allowance’ OR ‘income support’ OR ‘universal credit employment’ OR ‘support allowance’ OR ‘social care’ OR ‘housing benefit’
- ‘corrupt’ OR ‘autocratic’ OR ‘evil’ OR ‘devil’ OR ‘unelected’ or ‘unaccountable’ OR ‘technocrat’ OR ‘mafia’ OR ‘dying’ OR ‘incompetent’ OR ‘unnecessary red tape’ OR ‘useless’ OR ‘red tape’ OR ‘dictator’ OR ‘dictatorial’

The time period used in the construction of the variables is 23-June-2015 to 22-June-2016 for the one-year versions, and 23-December-2015 to 22-June-2016 for the six month versions respectively.

Appendix A2: Analysis with unweighted variables

Table 2.10 below lists the estimated effects of the exposure to migrant crimes variable, with varying controls. The results of the analysis surprisingly indicate negative and statistically significant effects in the first five specifications. The first specification includes just age bands as controls, one factor commonly thought to have influenced the voting outcome. The result indicates a reduction of 0.15 percentage points per unit increase in article count, or a 4.92 percentage points decrease per standard deviation increase in the number of articles involving migrant crime. Unsurprisingly, the age bands are jointly significant in this specification. The second column adds constituency controls, which increases the estimated coefficient to 0.148 percentage point decrease, or a 4.76 percentage point decrease per standard deviation. However, the constituency controls are jointly significant at 95% confidence level, but not at 99%. Interestingly, this specification is the only one where the null is not rejected in the Breusch-Pagan heteroskedasticity test, unlike all subsequent sets of results, where the first two specifications do not reject the null. Adding labour market conditions at the time of the census as controls significantly increases the predictive power of the model to 61%, increasing the estimated coefficient for the effect of exposure to migrant crime to denote a 4.25 percentage point reduction per standard deviation increase, which is also statistically significant. The addition of migrant shares and GVA measures induces the same increase in the estimated coefficient to 3.84 and 3.76 percent decrease in the estimated constituency level Leave vote per standard deviation increase in migrant crime articles.

Interestingly, adding changes in public spending, unemployment, and EU funding as controls changes the sign to denote a 0.33 percentage point increase in the estimated constituency Leave vote. The goodness of fit also increases relatively more compared to the addition of newer controls in previous specifications. To further test this, the media variable on labour market articles is added as an independent variable, which again increases the coefficient for migrant crime articles to a 2.14 percentage point increase per standard deviation. This result is statistically significant and has the expected sign. The coefficient for the labour market articles is also statistically significant, indicating a 3.37 percentage point decrease in the estimated Leave vote per 1 standard deviation increase. Adding further media variables as controls is not feasible due to multicollinearity concerns. All specifications have been tested for the multicollinearity and do not display it. Additionally, joint significance tests have been carried out for each pre-existing and new control variable group in all specifications above.

The null hypothesis is rejected at 99% significance level in all cases except for constituency controls in the second specification, for which it can be rejected at 95% significance level.

Economic indicators were found to be statistically significant as well. Constituency level Gross Value Added (GVA) per head were estimated to decrease the expected vote to Leave by 0.79 per standard deviation increase in wage decline article regressions to 0.825 percentage points per standard deviation in labour market articles regression. Interestingly, relative increases in this measure were found to decrease the expected Leave vote, but they are statistically insignificant except in the migration level regression, where a percentage point increase led to percentage points decreases between 0.017 and 0.052 in expected registered Leave votes. Constituency level unemployment changes were also found highly statistically significant in all specifications, with estimated effects ranging from a 3.05 to 2.67 percentage points decrease in the estimated Leave vote per percentage point increase in unemployment in the year preceding the referendum. The coefficients on the effects of changes in EU funding indicate that a percentage point decrease in regional EU funding increases the expected Leave vote share between 0.022 and 0.053 percentage points. The final set of coefficients indicate a correlation between a decrease in regional public spending and Leave vote, with a 1% decrease in public spending leading to 3.04 percentage points increase in constituency level Leave vote. The estimated effect is weaker in the case of other media variables, and it becomes insignificant in the context of migrant level article regressions. The results indicate that changes in economic conditions are strongly statistically significant determinants of the expected constituency level Leave vote.

Table 2.10: Analysis results with unweighted media coverage			
Dependent Variable: Vote Leave percentage			
VARIABLES	Leaveperc	Leaveperc	Leaveperc
Migrant Crime 1 year, unweighted	-5.602***	-4.921***	-4.329***
	(0.325)	(0.299)	(0.680)
Constituency Readership	1.898	-0.554	-1.525
	(1.232)	(1.154)	(0.991)
Age 30 to 44	-1.274***	0.182	0.118
	(0.309)	(0.329)	(0.229)
Age 45 to 59	1.305***	0.180	-0.988***
	(0.373)	(0.410)	(0.266)
Age over 60	-0.330*	0.542***	1.120***
	(0.178)	(0.182)	(0.131)
Population Density	-0.130***	0.0454	-0.0346
	(0.0289)	(0.0404)	(0.0310)
Male share	0.638	1.216**	1.536***
	(0.557)	(0.548)	(0.374)
Household Married	-0.166	0.00501	0.792***
	(0.115)	(0.128)	(0.0981)
Share Other Pre 2004 EU		-4.421***	-1.360**
		(0.749)	(0.688)
Share Post 2004 EU		1.050***	1.240***
		(0.340)	(0.296)
Born Other		-0.259***	-0.535***
		(0.0936)	(0.0776)
Net migr. 1 year		0.00164	0.0267
		(0.0615)	(0.0442)
Unemployed census 2011			5.297***
			(0.328)
Change in Unemployment			-2.988***
			(0.929)

Decrease in EU funding			0.0238***
			(0.00867)
Change in Public Spending			-0.500
			(0.419)
GVA per head			0.0473**
			(0.0200)
Change in wage median.			0.0904**
			(0.0446)
Constant	34.63	-19.97	-73.18***
	(31.60)	(30.81)	(21.24)
Observations	632	632	632
R-squared	0.445	0.589	0.750

Table 2.11: Analysis results with unweighted media coverage Dependent Variable: Vote Leave percentage

VARIABLES	Leaveperc	Leaveperc	Leaveperc
Resource Constraint 1 year, unweighted	-5.905***	-5.430***	-5.134***
	(0.376)	(0.326)	(0.527)
Constituency Readership	2.838**	0.410	-0.382
	(1.171)	(1.074)	(0.915)
Age 30 to 44	-1.534***	0.0146	-0.0732
	(0.301)	(0.326)	(0.219)
Age 45 to 59	1.665***	0.462	-0.715***
	(0.385)	(0.409)	(0.249)
Age over 60	-0.446**	0.428**	1.026***
	(0.178)	(0.180)	(0.124)
Population Density	-0.117***	0.0672*	-0.0176
	(0.0279)	(0.0398)	(0.0291)
Male share	0.634	1.204**	1.519***
	(0.552)	(0.555)	(0.342)
Household Married	-0.237**	-0.0613	0.731***
	(0.114)	(0.127)	(0.0922)
Share Other Pre 2004 EU		-4.589***	-1.442**
		(0.771)	(0.705)
Share Post 2004 EU		0.982***	1.183***
		(0.340)	(0.293)
Born Other		-0.292***	-0.576***
		(0.0963)	(0.0781)
Net migr. 1 year		-0.0343	-0.00560
		(0.0611)	(0.0422)
Unemployed census 2011			5.357***
			(0.324)
Change in Unemployment			-2.122**
			(0.935)

Decrease in EU funding			0.0327***
			(0.00716)
Change in Public Spending			-0.415
			(0.299)
GVA per head			0.0515**
			(0.0217)
Change in wage median.			0.0706
			(0.0443)
Constant	36.85	-17.26	-70.66***
	(31.83)	(31.55)	(19.99)
Observations	632	632	632
R-squared	0.448	0.607	0.774

Appendix A3: Results of alternative specifications of the media variables

The second and the third version of the migrant crime variables are given by the following definitions:

- For the second version of articles on crime involving migrants
    - (‘migrant’ OR ‘immigrant’ OR ‘foreigner’ OR ‘refugee ‘OR ‘migrant crisis’) in the main text AND
    - ‘guilty of’ OR ‘convicted’ OR ‘condemned’ OR ‘sent to prison’ OR imprisoned’ in the main text AND
    - ‘delinquent’ OR ‘delinquency’ OR ‘culprit’ OR ‘guilty’ OR ‘crime’ OR ‘criminal’ OR ‘violent’ OR ‘violence’ OR ‘felon’ OR ‘felony’ OR ‘charged’ OR ‘arson’ OR ‘arsonist’ OR ‘arrested’ OR ‘crack-down’ OR ‘crook’ OR ‘smuggling’ OR ‘smuggler’ OR ‘murder’ OR ‘murderer’ OR ‘rape’ OR ‘rapist’ OR ‘riot’ OR ‘theft’ OR ‘thief’ OR ‘pickpocket’ OR ‘stolen’ OR ‘steal’ in the headline
- For the third version of articles on crime involving migrants
    - ‘migrant crime’ OR ‘migrant criminal’ OR ‘migrant criminals’ OR ‘violent migrant’ OR ‘migrant involved in violence’ OR ‘migrant felon’ OR ‘migrant felony’ OR ‘felony committed by migrant’ OR ‘migrant crook’ OR ‘crook migrant’ OR ‘migrant con’ OR ‘con migrant’ OR ‘migrant culprit’ OR ‘migrant delinquent’ OR ‘migrant delinquency’ OR ‘migrant charged’ OR ‘migrant arsonist’ OR ‘migrant arson’ OR ‘arrested migrant’ OR ‘migrant arrested’ OR ‘suspected migrant’ OR ‘suspect migrant’ OR ‘crack-down on migrant’ OR ‘migrant smuggler’ OR ‘migrant smuggled’ OR ‘smuggling migrant’ OR ‘migrant smuggling’ OR ‘migrant murderer’ OR ‘migrant murderers’ ‘migrant killed’ or ‘migrant kills’ or ‘killed by migrant’ OR ‘murdered by migrant’ OR ‘killer migrant’ OR ‘migrant murders’ OR ‘migrant murdered’ OR ‘migrant rapes’ OR ‘raped by migrant’ OR ‘migrant rapist’ OR ‘migrant rapists’ OR ‘migrant gang’ OR ‘gang of migrants’ OR ‘migrant riot’ OR ‘migrant rioters’ OR ‘migrant theft’ OR ‘migrant thief’ OR ‘migrant thieves’ OR ‘stolen by migrant’ OR ‘migrant steals’ in the headline OR
    - ‘immigrant crime’ OR ‘immigrant criminal’ OR ‘immigrant criminals’ OR ‘violent immigrant’ OR ‘immigrant involved in violence’ OR ‘immigrant felon’ OR ‘immigrant felony’ OR ‘felony committed by immigrant’ OR ‘immigrant

crook' OR 'crook immigrant' OR 'immigrant con' OR 'con immigrant' OR 'immigrant culprit' OR 'immigrant delinquent' OR 'immigrant delinquency' OR 'immigrant charged' OR 'immigrant arsonist' OR 'immigrant arson' OR 'arrested immigrant' OR 'immigrant arrested' OR 'suspected immigrant' OR 'suspect immigrant' OR 'crack-down on immigrant' OR 'immigrant smuggler' OR 'immigrant smuggled' OR 'smuggling immigrant' OR 'immigrant smuggling' OR 'immigrant murderer' OR 'immigrant murderers' OR 'immigrant killed' OR 'immigrant kills' OR 'killed by immigrant' OR 'murdered by immigrant' OR 'killer immigrant' OR 'immigrant murders' OR 'immigrant murdered' OR 'immigrant rapes' OR 'raped by immigrant' OR 'immigrant rapist' OR 'immigrant rapists' OR 'immigrant gang' OR 'gang of immigrants' OR 'immigrant riot' OR 'immigrant rioters' OR 'immigrant theft' OR 'immigrant thief' OR 'immigrant thieves' OR 'stolen by immigrant' OR 'immigrant steals' in the headline OR

- 'foreigner crime' OR 'foreigner criminal' OR 'foreigner criminals' OR 'violent foreigner' OR 'foreigner involved in violence' OR 'foreigner felon' OR 'foreigner felony' OR 'felony committed by foreigner' OR 'foreigner crook' OR 'crook foreigner' OR 'foreigner con' OR 'con foreigner' OR 'foreigner culprit' OR 'foreigner delinquent' OR 'foreigner delinquency' OR 'foreigner charged' OR 'foreigner arsonist' OR 'foreigner arson' OR 'arrested foreigner' OR 'foreigner arrested' OR 'suspected foreigner' OR 'suspect foreigner' OR 'crack-down on foreigner' OR 'foreigner smuggler' OR 'foreigner smuggled' OR 'smuggling foreigner' OR 'foreigner smuggling' OR 'foreigner murderer' OR 'foreigner murderers' OR 'foreigner killed' OR 'foreigner kills' OR 'killed by foreigner' OR 'murdered by foreigner' OR 'killer foreigner' OR 'foreigner murders' OR 'foreigner murdered' OR 'foreigner rapes' OR 'raped by foreigner' OR 'foreigner rapist' OR 'foreigner rapists' OR 'foreigner gang' OR 'gang of foreigners' OR 'foreigner riot' OR 'foreigner rioters' OR 'foreigner theft' OR 'foreigner thief' OR 'foreigner thieves' OR 'stolen by foreigner' OR 'foreigner steals' in the headline OR
- 'refugee crime' OR 'refugee criminal' OR 'refugee criminals' OR 'violent refugee' OR 'refugee involved in violence' OR 'refugee felon' OR 'refugee felony' OR 'felony committed by refugee' OR 'refugee crook' OR 'crook refugee' OR 'refugee con' OR 'con refugee' OR 'refugee culprit' OR 'refugee

delinquent' OR 'refugee delinquency' OR 'refugee charged' OR 'refugee arsonist' OR 'refugee arson' OR 'arrested refugee' OR 'refugee arrested' OR 'suspected refugee' OR 'suspect refugee' OR 'crack-down on refugee' OR 'refugee smuggler' OR 'refugee smuggled' OR 'smuggling refugee' OR 'refugee smuggling' OR 'refugee murderer' OR 'refugee murderers' OR 'refugee killed' OR 'refugee kills' OR 'killed by refugee' OR 'murdered by refugee' OR 'killer refugee' OR 'refugee murders' OR 'refugee murdered' OR 'refugee rapes' OR 'raped by refugee' OR 'refugee rapist' OR 'refugee rapists' OR 'refugee gang' OR 'gang of refugees' OR 'refugee riot' OR 'refugee rioters' OR 'refugee theft' OR 'refugee thief' OR 'refugee thieves' OR 'stolen by refugee' OR 'refugee steals

Table 2.12: Summary Statistics for the Unweighted Media variables					
Variable	Observations	Mean	Std. Dev.	Min	Max
Vote Leave percent	632	52.11	11.4	18.48	74.96
Un-weighted media variables					
Migrant crime articles	632	713.97	32.17	668	782
Migrant crime articles, 2nd version	632	149.55	7.33	130	172
Migrant crime articles, 3rd version	632	77.9	7.87	74	103
Weighted media variables					
Migrant crime articles, 1st version	632	442.08	71.79	190.66	745.46
Migrant crime articles, 2nd version	632	96.54	11.32	32.62	145.38
Migrant crime articles, 3rd version	632	78.25	10.29	29.28	116.99

Table 2.13: Estimation of the effect of first alternative unweighted newspapers coverage of migrant induced crime in the year before the EU referendum:

Dependent Variable: Vote Leave percentage

VARIABLES	(1)	(2)	(3)	(4)	(5)
Unweighted migrant crime version 2	-4.672***	-4.834***	-2.988***	-2.933***	-2.968***
	(0.277)	(0.276)	(0.521)	(0.521)	(0.517)
Constituency Readership	-2.191***	-2.238***	-2.147***	-2.117***	-2.289***
	(0.815)	(0.804)	(0.794)	(0.793)	(0.789)
Age 30 to 44	0.175	0.123	0.136	0.149	0.169
	(0.216)	(0.213)	(0.210)	(0.210)	(0.210)
Age 45 to 59	-1.301***	-1.288***	-1.048***	-1.066***	-1.143***
	(0.276)	(0.272)	(0.275)	(0.275)	(0.274)
Age over 60	1.180***	1.198***	1.117***	1.117***	1.124***
	(0.149)	(0.147)	(0.146)	(0.146)	(0.145)
Population Density	-0.0280	-0.0288	-0.0392	-0.0426*	-0.0419*
	(0.0247)	(0.0243)	(0.0241)	(0.0242)	(0.0243)
Male share	1.769***	1.743***	1.501***	1.490***	1.345***
	(0.374)	(0.369)	(0.369)	(0.368)	(0.371)
Household Married	0.847***	0.813***	0.776***	0.772***	0.791***
	(0.0926)	(0.0917)	(0.0910)	(0.0908)	(0.0905)
Unemployed census 2011	5.430***	5.340***	5.198***	5.174***	5.161***
	(0.295)	(0.292)	(0.290)	(0.290)	(0.289)
Share Other Pre 2004 EU	-1.184***	-1.144***	-0.937***	-0.920***	-1.336***
	(0.339)	(0.335)	(0.334)	(0.334)	(0.363)
Share Post 2004 EU	1.049***	1.058***	1.145***	1.127***	1.218***
	(0.207)	(0.205)	(0.203)	(0.203)	(0.205)
Born Other	-0.554***	-0.552***	-0.514***	-0.514***	-0.515***
	(0.0679)	(0.0670)	(0.0668)	(0.0666)	(0.0674)
Change in Unemployment	-1.448	-2.173**	-2.703***	-2.560***	-2.478***
	(0.881)	(0.887)	(0.884)	(0.886)	(0.880)
Decrease in EU funding		0.0347***	0.0454***	0.0460***	0.0439***
		(0.00823)	(0.00852)	(0.00851)	(0.00849)

Change in Public Spending			-1.355***	-1.396***	-1.347***
			(0.326)	(0.326)	(0.325)
Change in GVA					-0.0512
					(0.0452)
GVA per head					4.81e-05***
					(1.60e-05)
Change in wage median.				0.0871*	0.0848*
				(0.0477)	(0.0474)
Constant	-82.68***	-78.94***	-67.34***	-66.24***	-58.67***
	(21.46)	(21.20)	(21.11)	(21.08)	(21.24)
Observations	632	632	632	632	632
R-squared	0.726	0.734	0.741	0.743	0.747

Table 2.14: Estimation of the effect of 2nd alternative unweighted newspapers coverage of migrant crime in the year before the referendum

Dependent Variable: Vote Leave percentage

VARIABLES	(1)	(2)	(3)	(4)	(5)
Unweighted migrant crime version 3	-5.068***	-5.060***	-4.284***	-4.231***	-4.250***
	(0.274)	(0.274)	(0.671)	(0.670)	(0.668)
Constituency Readership	-0.737	-0.780	-0.978	-0.961	-1.119
	(0.803)	(0.803)	(0.818)	(0.816)	(0.813)
Age 30 to 44	0.146	0.131	0.136	0.149	0.158
	(0.209)	(0.209)	(0.209)	(0.209)	(0.209)
Age 45 to 59	-1.039***	-1.053***	-1.003***	-1.022***	-1.101***
	(0.270)	(0.270)	(0.273)	(0.273)	(0.272)
Age over 60	1.098***	1.111***	1.096***	1.097***	1.099***
	(0.145)	(0.145)	(0.145)	(0.145)	(0.144)
Population Density	-0.0298	-0.0289	-0.0324	-0.0360	-0.0337
	(0.0238)	(0.0238)	(0.0240)	(0.0240)	(0.0241)
Male share	1.490***	1.493***	1.442***	1.433***	1.273***
	(0.364)	(0.364)	(0.366)	(0.365)	(0.368)
Household Married	0.829***	0.823***	0.809***	0.805***	0.821***
	(0.0897)	(0.0897)	(0.0903)	(0.0902)	(0.0899)
Unemployed census 2011	5.426***	5.395***	5.331***	5.306***	5.285***
	(0.286)	(0.287)	(0.291)	(0.291)	(0.289)
Share Other Pre 2004 EU	-1.062***	-1.055***	-0.990***	-0.973***	-1.405***
	(0.329)	(0.329)	(0.333)	(0.332)	(0.362)
Share Post 2004 EU	1.072***	1.077***	1.109***	1.090***	1.188***
	(0.201)	(0.201)	(0.202)	(0.202)	(0.204)
Born Other	-0.523***	-0.523***	-0.513***	-0.513***	-0.520***
	(0.0659)	(0.0659)	(0.0663)	(0.0661)	(0.0670)
Change in Unemployment	-2.811***	-3.061***	-3.135***	-2.976***	-2.898***
	(0.852)	(0.869)	(0.871)	(0.873)	(0.868)
Decrease in EU funding		0.0113	0.0190*	0.0199**	0.0180*
		(0.00800)	(0.0100)	(0.0100)	(0.00996)

Change in Public Spending			-0.526	-0.568	-0.536
			(0.415)	(0.415)	(0.413)
Change in GVA					-0.0307
					(0.0451)
GVA per head					4.88e-05***
					(1.59e-05)
Change in wage median.				0.0903*	0.0889*
				(0.0474)	(0.0471)
Constant	-72.52***	-71.90***	-68.72***	-67.60***	-59.11***
	(20.86)	(20.85)	(20.99)	(20.95)	(21.11)
Observations	632	632	632	632	632
R-squared	0.743	0.744	0.744	0.746	0.750

Table 2.15: Summary statistics: Media variables and constituency readership by newspaper							
	Migrant Crime 1 year	Migrant crime 6 months	Resource constraint 1 year	Resource constraint 6 months	Aggregate coverage 1 year	Aggregate coverage 6 months	Readership
The Guardian	181.62 (120.20)	96.09 (63.60)	90.15 (59.66)	54.48 (36.06)	104.67 (69.28)	55.47 (36.71)	0.27 (0.03)
Daily Telegraph	78.16 (46.97)	48.48 (29.13)	29.68 (17.83)	19.78 (11.89)	22.78 (13.63)	13.87 (8.28)	0.78 (0.08)
Sunday Telegraph	18.19 (10.78)	9.79 (5.92)	27.61 (11.99)	7.57 (4.50)	9.09 (5.40)	3.78 (2.25)	0.60 (0.06)
The Times	37.69 (23.52)	31.08 (18.95)	18.76 (11.01)	15.11 (8.97)	24.73 (14.92)	3.67 (2.04)	0.71 (0.08)
Sunday Times	19.24 (10.95)	0.00 (0.00)	12.63 (7.47)	7.89 (4.67)	36.32 (21.49)	20.53 (12.14)	1.27 (0.14)
Financial Times	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.34 (0.03)
The Observer	8.55 (5.47)	6.22 (3.98)	5.44 (3.48)	3.88 (2.48)	10.49 (6.72)	4.66 (2.98)	0.32 (0.03)
Independent	5.19 (1.71)	2.99 (0.98)	1.26 (0.41)	0.47 (0.15)	1.65 (0.54)	1.18 (0.38)	0.08 (0.01)
i-independent	28.67 (11.63)	16.72 (6.78)	5.97 (2.42)	2.98 (1.21)	4.77 (1.93)	3.58 (1.45)	0.46 (0.05)
Metro	54.71 (73.37)	33.30 (44.66)	9.26 (3.85)	6.13 (2.62)	3.04 (1.36)	0.00 (0.00)	2.13 (0.24)
Daily Mail	173.14 (63.93)	97.99 (40.98)	111.32 (45.60)	65.17 (29.24)	97.32 (41.72)	71.93 (31.29)	2.44 (0.27)
Daily Express	21.41 (2.41)	15.38 (1.73)	23.41 (2.66)	12.03 (1.36)	8.75 (1.02)	4.71 (0.55)	0.66 (0.07)
Sunday Express	4.12 (0.46)	1.17 (0.13)	3.53 (0.40)	1.76 (0.20)	2.35 (0.26)	1.17 (0.13)	0.58 (0.06)
The Sun	114.50 (50.86)	74.94 (33.40)	118.38 (54.70)	94.29 (43.71)	73.34 (34.00)	48.89 (22.66)	2.77 (0.31)
Daily Mirror	13.71 (6.07)	9.15 (4.03)	10.64 (4.76)	7.60 (3.40)	7.60 (3.40)	6.08 (2.72)	1.21 (1.138)
The People	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.43 (0.04)	0.00 (0.00)	0.43 (0.04)
Daily Star	13.02 (1.48)	3.26 (0.36)	0.82 (0.09)	0.00 (0.00)	0.01 (0.03)	0.01 (0.03)	0.81 (0.09)

Appendix B: Appendices to Chapter 3

In this appendix, I analyse the benchmark specification of the model, as outlined in section 3.2. Extending the analysis of Besley and Prat (2006), media outlets are interested in the second period revenues, which are purely audience driven. The revenue function in the second period has the same form as in the first, with the addendum that all voters act rationally inattentive in the absence of an election. With probability λ, an outlet can be revealed to have been dishonest, in which case it loses all audience revenues. Once an outlet is discredited, their status becomes public knowledge.

Proposition 3

In the Perfect Bayesian Equilibrium, the incumbent wins in cases $\Phi_I\Phi_C$, $\Phi_I b_C$, $b_I b_C$, while the incumbent wins if case $b_I\Phi_C$. Consequently, a good politician always wins the election if he is present and a bad incumbent defeats a bad challenger. Outlets report honestly if at in all cases. The minimum share of outlets a bad challenger needs to 'influence' to win the election is $\Gamma = \frac{\mathbf{0.5}}{\mathbf{1-\alpha}}$, unless both politicians are revealed as bad. If both politicians are bad, the number of outlets the incumbent needs to capture $\Gamma^* = \frac{\mathbf{0.5-\alpha}}{\mathbf{1-\alpha}} \leq \Gamma$.

Focusing first on the voter side, the following lemma outlines the equilibrium Bayesian updating if voters use weakly undominated strategies.. Note that inattentive voters only notice the bad signal if their publication is publishing it, while attentive voters if it is published by at least one outlet.

Corollary 1.

The turnover of politicians is non-decreasing in q, n, τ, λ and α

Proof of Proposition 3

Depending on the realisation of α, two cases can be distinguished, which alter the strategies and outcomes in case $b_I b_C$. The perfect Bayesian equilibrium is composed of two parts occurring successively: a bargaining component between the politicians and media outlets and the electoral game. The bargaining equilibrium determines whether the media is an accurate purveyor of information. Denote the event that an outlet reports a received bad signal as 'captured media' or 'captured outlet' and if all outlets are captured 'total media capture'

Firstly, if $\alpha < 0.5$ the following strategies and beliefs constitute a Perfect Bayesian Equilibrium.

Histories/ Signal received by media	$\Phi_I\Phi_C$	$\Phi_I b_C$	$b_I\Phi_C$	$b_I b_C$	Beliefs
Prior Probability	$\gamma^2 + 2(1-\gamma)\gamma q + (1-\gamma)^2(1-q)^2$	$\gamma(1-\gamma)q + (1-\gamma)^2 q(1-q)$	$\gamma(1-\gamma)q + (1-\gamma)^2 q(1-q)$	$(1-\gamma)^2 q^2$	
Agents and strategies					Informed voters: Believe incumbent is good if they observe $\Phi_I\Phi_C$ or $\Phi_I b_C$ in the media. Believe challenger is good if observe $\Phi_I\Phi_C$ or $b_I\Phi_C$ in the media. Believe both politicians are bad if they observe $b_I b_C$ in the media.
Incumbent strategy	$\boldsymbol{\tau_{I,i}} = 0, \forall i,$ $\forall B, \forall\, E(\tau_C)$	$\boldsymbol{\tau_{I,i}} = 0, \forall i,$ $\forall B\ \forall\, E(\tau_C)$	$\boldsymbol{\tau_{I,i}} = 0, \forall i$ $\forall B, and\ \forall\, E(\tau_C)\ such\ that\ \boldsymbol{\tau_{C,i}} < \tau^{**}(\boldsymbol{\tau_{I,i}})$ for $n + 1 - \lceil \Gamma n \rceil$ outlets Otherwise, a) if $B \geq \lceil \Gamma n \rceil \tau^*(\lceil \Gamma n \rceil)$ $\boldsymbol{\tau_{I,i}} = \begin{cases} \tau^*(\lceil \Gamma n \rceil)\ to\ \lceil \Gamma n \rceil \text{ outlets} \\ 0\ otherwise \end{cases}$ b) if $B < \lceil \Gamma n \rceil \tau^*\ \tau_{I,i} = 0, \forall i$	If $\lceil \Gamma n \rceil - \lceil \Gamma^* n \rceil = 0$ and $B\frac{\lceil \Gamma n \rceil - 1}{\lceil \Gamma n \rceil} > \tau^*(n)\lceil \Gamma n \rceil$ $\tau_{i,I} = \begin{cases} \tau^*(n)\ to \lceil \Gamma n \rceil\ outlets \\ 0\ to\ the\ rest \end{cases}$ If $\lceil \Gamma n \rceil - \lceil \Gamma^* n \rceil = 0$ and $B\frac{\lceil \Gamma n \rceil - 1}{\lceil \Gamma n \rceil} \leq \tau^*(n)\lceil \Gamma n \rceil$ $\boldsymbol{\tau_{I,i}} = 0, \forall i$	

				If $\lceil \Gamma n \rceil - \lceil \Gamma^* n \rceil \geq 1$ $\boldsymbol{\tau_{I,i}}$ $= \begin{cases} \tau^*(n) + \frac{(l+k)\tau^*(n)}{\lceil \Gamma^* n \rceil} \; to \; \lceil \Gamma^* n \rceil \; out \\ 0 \; to \; the \; rest \end{cases}$ if $B \geq \lceil \Gamma n \rceil \tau^*(n)$ If $B \in [\lceil \Gamma^* n \rceil \tau^*(n), \lceil \Gamma n \rceil \tau^*(n))$ $\tau_{i,I} = \begin{cases} \tau^*(n) \; to \lceil \Gamma^* n \rceil \; outlets \\ 0 \; to \; the \; rest \end{cases}$ If $B \in [\lceil \Gamma^* n \rceil \tau^*(\lceil \Gamma^* n \rceil), \lceil \Gamma^* n \rceil \tau^*(n))$ $\tau_{i,I} = \begin{cases} \tau^*(\lceil \Gamma^* n \rceil) \; to \lceil \Gamma^* n \rceil \; outlets \\ 0 \; to \; the \; rest \end{cases}$ Otherwise $\boldsymbol{\tau_{I,i}} = 0, \forall i$	Uninformed voters: Believe incumbent is good if observe $\Phi_I \Phi_C$ or $\Phi_I b_C$ from their outlet. Believe challenger is good if they observe $\Phi_I \Phi_C$ or $b_I \Phi_C$ from their outlet. Believe both politicians are bad if they observe $b_I b_C$ from their outlet
Challenger strategy	$\boldsymbol{\tau_{C,i}} = 0, \forall i,$ $\forall B, \forall \tau_I$	$\boldsymbol{\tau_{C,i}} =$ $0, \forall i \forall B, \forall \tau_I$	If $\tau_{I,i} < \tau^*(\lceil \Gamma n \rceil)$ for $\lceil \Gamma n \rceil$ outlets, then $\boldsymbol{\tau_{C,i}} = 0, \forall i \; \forall B$. If $\tau_{I,i} \geq \tau^*(\lceil \Gamma n \rceil)$ for $x \geq \lceil \Gamma n \rceil$ outlets, pick the x+1 outlets with the smallest $\tau_{I,i} \geq \tau^*(\lceil \Gamma n \rceil)$, τ_{I,i^*}, and	If $\lceil \Gamma n \rceil - \lceil \Gamma^* n \rceil = 0$ and $B \frac{\lceil \Gamma n \rceil - 1}{\lceil \Gamma n \rceil} > \tau^*(n) \lceil \Gamma n \rceil$ $\boldsymbol{\tau_{C,i}} = 0 \; , \forall i$	

			$\boldsymbol{\tau_{C,i}} = \begin{cases} \tau^{**}(\boldsymbol{\tau_{I,i}}) \; for \; i \in \tau_{I,i^*} \\ 0 \; otherwise \end{cases}$	If $\lceil \Gamma n \rceil - \lceil \Gamma^* n \rceil = 0$ and $B \frac{\lceil \Gamma n \rceil - 1}{\lceil \Gamma n \rceil} \leq \tau^*(n) \lceil \Gamma n \rceil$ Order the outlets in ascending order of offers received from the incumbent and offer: $\tau_{C,i} = \begin{cases} \tau^*(n) \; to \; the \; first \; n - \lceil \Gamma n \rceil \; ou \\ \frac{B}{\lceil \Gamma n \rceil} + 1 \; to \; 1 \; of \; the \; remaining \end{cases}$ If $\lceil \Gamma n \rceil - \lceil \Gamma^* n \rceil \geq 0$ $\boldsymbol{\tau_{C,i}} = 0 \, , \forall i$ if $B < \lceil \Gamma n \rceil \tau^*(\lceil \Gamma n \rceil)$ If $B \geq \lceil \Gamma n \rceil \tau^*(\lceil \Gamma n \rceil)$, let x denote the number of outlets for which $\tau_{I,i}$=0 and order the remaining outlets ascendingly. Then if 1)$x \geq \lceil \Gamma n \rceil$ or 2) $x < \lceil \Gamma n \rceil$ and	

				$B - x\tau^*(\lceil \Gamma n \rceil) \geq \sum_{j=x+1}^{\lceil \Gamma n \rceil} \tau_{I,j} + (\lceil \Gamma n \rceil - x)$ $\tau_{C,i}$ $= \begin{cases} \tau^*(\min(x, \lceil \Gamma n \rceil)) \; to \; \min(x, \lceil \Gamma n \rceil) \\ \tau_{I,i} + 1 \; to \text{ the first } \max(\lceil \Gamma n \\ 0 \; otherwis \end{cases}$ Otherwise $\boldsymbol{\tau_{C,i}} = 0$	
Electoral result	Incumbent wins	Incumbent wins	Challenger wins	Incumbent wins	
Media Agents	Report honestly, $\Phi_I \Phi_C$, for $\forall \tau_I \; and \; \forall \tau_C$. If $\tau_{i,I} \neq 0$ and or $\tau_{C,i} \neq 0$, they would accept the higher offer and	Report honestly $\Phi_I b_C$ if $\tau_{C,i} < \tau^*(x), \forall \tau_I$ Or if $\tau_{C,i} \geq \tau^*(x)$ and $\tau_I \geq \tau^{**}(\boldsymbol{\tau_{C,i}})$	Report honestly $b_I \Phi_C$ unless $\begin{cases} \tau_{I,i} \geq \tau^*(x) \; \textbf{and} \\ \tau_{C,i} < \tau^{**}(\boldsymbol{\tau_{I,i}}) \end{cases}$ then report $\Phi_I \Phi_C$	Report honestly $b_I b_C$ if either set of conditions holds $\tau_{I,i} < \tau^*(x) \textbf{and } \tau_{C,i} < \tau^*(x) \; \boldsymbol{OR}$ $\tau_{I,j} > \tau^*(x) \; \boldsymbol{and} \; \tau_{I,-j} > \tau^{**}(\tau^*(x))$ For $j \in \{I, C\}$	

	still report honestly	Report $\Phi_I\Phi_C$ if $\tau_{C,i} \geq \tau^*(x)$ and $\tau_I < \tau^{**}(\boldsymbol{\tau_{C,i}(x)})$		Report $\Phi_I b_C$ if $\tau_{I,i} \geq \tau^*(x) and\ \tau_{I,i} \geq \tau_{C,i}$ Report $b_I\Phi_C$ if $\tau_{C,i} \geq \tau^*(x) and\ \tau_{I,i} < \tau_{C,i}$	
Informed Voters	Vote incumbent unless observe $b_I\Phi_C$ in the media	Vote incumbent unless observe $b_I\Phi_C$ in the media	Vote challenger if they observe $b_I\Phi_C$ in the media. Vote incumbent if they observe $\Phi_I\Phi_C$ in the media	Vote challenger if they observe $b_I\Phi_C$ in the media. Vote incumbent if they observe b_Ib_C or Φ_Ib_C in the media	
Uninformed Voters	Vote incumbent unless observe $b_I\Phi_C$ from their outlet	Vote incumbent unless observe $b_I\Phi_C$ from their outlet	Vote challenger if they observe $b_I\Phi_C$ from their outlet. Vote incumbent if they observe $\Phi_I\Phi_C$ from their outlet	Vote challenger if they observe $b_I\Phi_C$ from their outlet. Vote incumbent if they observe b_Ib_C or Φ_Ib_C from their outlet	

Lemma 4

In any Perfect Bayesian equilibrium with weakly undominated strategies, voters will prefer to vote for the incumbent unless he is revealed as bad and the challenger is not.

Proof:

For the purposes of this proof, a distinction between how informed and uninformed voters update their beliefs needs to be made.

From the perspective of any voter:

$$\Pr(\theta_I = g | s_i = b_I b_C) = 0 < \gamma$$

$$\Pr(\theta_I = g | s_i = b_I \Phi_C) = 0 < \gamma$$

$$\Pr(\theta_C = g | s_i = b_I b_C) = 0 < \gamma$$

$$\Pr(\theta_C = g | s_i = \Phi_I \mathrm{b}_C) = 0 < \gamma$$

which results from the structure of the game for both politicians and for both voter types and all histories. Therefore, upon observing a bad signal, a voter knows the politician is bad and that he would get utility of 0 if said politician is in office. It is therefore an optimal decision for the voter to not vote for a politician revealed as bad. If both politicians are revealed as bad, the voters prefer the incumbent due to the incumbency advantage.

Voters are aware that bad politicians could attempt to 'silence' outlets, but they update their beliefs differently. An attentive voter acts like one in Besley and Prat (2006) and can be misled only by total media suppression.

Applying Bayes' rule for both the incumbent and the challenger from the perspective of an informed voter based on the politician strategies outlined above:

$$\Pr(\theta_I = g | s_i = \Phi_I \Phi_C) = \frac{\Pr(\theta_I = g)\Pr(s_i = \Phi_I \Phi_C | \theta_I = g)}{\Pr(s_i = \Phi_I \Phi_C)}$$

$$= \frac{\gamma[\gamma + (1-\gamma)(1-q)]}{\gamma^2 + (1-\gamma)^2(1-q)^2 + 2\gamma(1-\gamma)(1-q)}$$

Similarly, about the challenger:

$$\Pr(\theta_C = g | s_i = \Phi_I \Phi_C) = \frac{\Pr(\theta_I = g)\Pr(s_i = \Phi_I \Phi_C | \theta_I = g)}{\Pr(s_i = \Phi_I \Phi_C)}$$

$$= \frac{\gamma[\gamma + (1-\gamma)(1-q)]}{\gamma^2 + (1-\gamma)^2(1-q)^2 + 2\gamma(1-\gamma)(1-q)}$$

This is because an informed voter observes a pair of null signals in the following three situations only: both politicians are good, both politicians are good and undiscovered, or one politician is good and the other is bad and undiscovered. Therefore, given politician j is good, a null signal can be observed only if a bad signal is not observed about politician -j.

Now consider Bayesian updating for an informed voter when they observe signal $b_j\Phi_{-j}$

$$\Pr\left(\theta_j = g \middle| s_i = b_j\Phi_{-j}\right) = 0 < \gamma$$

$$\Pr\left(\theta_j = g \middle| s_i = \Phi_j \mathrm{b}_{-j}\right) = \frac{\Pr\left(\theta_j = g\right)\Pr\left(s_i = \Phi_j \mathrm{b}_{-j} \middle| \theta_j = g\right)}{\Pr\left(s_i = \Phi_j \mathrm{b}_{-j}\right)}$$

$$= \frac{\gamma q(1-\gamma)}{\gamma q(1-\gamma) + q(1-\gamma)^2(1-q)} > 0$$

This happens that upon observing signal $\Phi_j \mathrm{b}_{-j}$, informed voters know with certainty that politician -j is bad, while there is a strictly positive probability that politician j is good. Consequently, they would prefer politician –j as $\Pr\left(\theta_{-j} = g \middle| s_i = \Phi_j \mathrm{b}_{-j}\right) = 0$.

Because of the incumbency advantage assumption, informed voters prefer to vote for the incumbent should they observe $\mathrm{b}_I \mathrm{b}_C$. Thus, the only scenario in which an informed voter prefers the challenger to the incumbent is if the expected utility from electing the challenger exceeds that of electing the incumbent, i.e. after observing $b_i\Phi_c$

I now analyse how an uninformed voter updates their belief. As above, after observing a bad signal, the posterior belief that the politician is bad becomes one. The difference with respect to the informed voters is that in case $b_I b_C$ some rationally inattentive voters can be subscribed to an outlet which misreports one of the politicians. The winning politician depends on the realisation of B and α, which in turn determine the fraction of outlets suppressed in equilibrium.

Let n denote the total number of outlets. Therefore, in both histories, the probability that an inattentive voter reads a captured outlet is $\frac{\lceil \Gamma n \rceil}{n}$ or $\frac{\lceil \Gamma^* n \rceil}{n}$, where m_i and m_c denote the number of outlets silenced by the incumbent and challenger respectively. For the incumbent:

$$\Pr(\theta_I = g | s_i = \Phi_I\Phi_C) = \frac{\Pr(\theta_I = g)\Pr(s_i = \Phi_I\Phi_C | \theta_I = g)}{\Pr(s_i = \Phi_I\Phi_C)}$$

$$= \frac{\gamma[\gamma + (1-\gamma)(1-q)]}{\gamma^2 + (1-\gamma)^2(1-q)^2 + 2\gamma(1-\gamma)(1-q)}$$

Similarly, for the challenger

$$\Pr(\theta_C = g | s_i = \Phi_I\Phi_C) = \frac{\Pr(\theta_I = g)\Pr(s_i = \Phi_I\Phi_C | \theta_I = g)}{\Pr(s_i = \Phi_I\Phi_C)}$$

$$= \frac{\gamma[\gamma + (1-\gamma)(1-q)]}{\gamma^2 + (1-\gamma)^2(1-q)^2 + 2\gamma(1-\gamma)(1-q)}$$

As with informed voters, we can observe that $\Pr(\theta_I = good | s_{i,I} = \Phi) = \Pr(\theta_C = good | s_{i,I} = \Phi)$, and due incumbency advantage assumption both informed and rationally inattentive voters prefer to vote for the incumbent after observing a pair of null signals.

Now consider Bayesian updating for a rationally inattentive voter when they observe signal $b_I\Phi_C$. Since outlets cannot fabricate a bad signal, it follows that

$$\Pr(\theta_I = g | s_i = \mathrm{b}_I\Phi_C) = 0$$

In the case of the challenger

$$\Pr(\theta_C = g | s_i = \mathrm{b}_I\Phi_C) = \frac{\Pr(\theta_C = g)\Pr(s_i = \mathrm{b}_I\Phi_C | \theta_C = g)}{\Pr(s_i = \mathrm{b}_I\Phi_C)}$$

$$= \frac{\gamma(1-\gamma)q}{\gamma(1-\gamma)q + (1-\gamma)^2q(1-q) + \omega(1-\gamma)^2q^2\frac{\lceil \Gamma^* \mathrm{n} \rceil}{n}} > 0$$

Where ω denotes the probability that $\lceil \Gamma \mathrm{n} \rceil - \lceil \Gamma^* \mathrm{n} \rceil \geq 1$. This expression is strictly larger than 0 since $\gamma > 0$ and $q > 0$. Therefore, an inattentive voter will prefer to vote for the challenger upon observing $\mathrm{b}_I\Phi_C$. A similar argument occurs for the $\Phi_I\mathrm{b}_C$ signal. Since outlets cannot fabricate a bad signal, it follows that $\Pr(\theta_C = g | s_i = \Phi_I\mathrm{b}_C) = 0$. In the case of the incumbent, we have:

$$\Pr(\theta_I = g | s_i = \Phi_I\mathrm{b}_C) = \frac{\Pr(\theta_I = g)\Pr(s_i = \Phi_I\mathrm{b}_C | \theta_I = g)}{\Pr(s_i = \Phi_I\mathrm{b}_C)}$$

$$= \frac{\gamma(1-\gamma)q}{\gamma(1-\gamma)q + (1-\gamma)^2q(1-q)} > 0$$

Because of the incumbency advantage assumption, rationally inattentive voters prefer to vote for the incumbent should they observe $b_I b_C$. Thus, the only scenario in which a rationally inattentive voter prefers the challenger to the incumbent is if the expected utility from electing the challenger exceeds that of electing the incumbent, i.e. after observing $b_i \Phi_c$

QED Lemma 4■Lemma 4 shows that a voter prefers to vote for a politician not revealed as bad, with a preference for the incumbent if both their respective signals match. This belief system is compliant with Bayes' rule and a best response given the strategies of the politicians.

If both politicians are good, neither has an incentive to offer a positive bribe to the media outlets as they will receive no signal with probability one and thus eliminating and scope for media manipulation. Subsequent analysis will focus on the latter three scenarios with at least one bad politician. Given Lemma 7, a bad politician would have an interest in partially or totally preventing the media from signalling his type since otherwise he loses the election with certainty. In a similar fashion, a good politician whose opponent is bad has an incentive in offering $\tau > 0$ to enough outlets to ensure his opponent does not win. Furthermore, for either politician, it is an optimal strategy to not offer any bribe in the second period, rendering audience revenue the only source of income for outlets in the second period, which is expected to increase the minimum transfer an outlet would accept. In the absence of media interference, due to incumbency advantage, the incumbent is expected to win in $\boldsymbol{g_I g_C}$ and if he is good while a bad challenger remains undiscovered.

The α informed voters purchase all outlets that publish a signal, whereas the 1-α uninformed voters are split equally across all n outlets. The politicians determine the optimal number of outlets to 'manipulate' based on the behaviour of both types of voters. First consider the case in which the challenger is revealed as bad and the incumbent is not, i.e. $g_i b_c$. In this case, even if successful at media persuasion, Lemma 7 indicates the challenger would still lose the election as both types of voters prefer the incumbent in the absence of a signal.

In the case $\boldsymbol{b_i g_c}$, the incumbent benefits from media manipulation as it potentially enables him to win the election extract utility B in the second period. Based on Lemma 7, the incumbent gains the votes of informed readers, α, if and only if he achieves total media suppression. However, it can be shown this cannot occur in equilibrium.

If only one politician is discovered as bad, he requires at least half the votes exclusively from rationally inattentive voters and the number of outlets he needs to 'manipulate', m, is given by:

$$\frac{(1-\alpha)m}{n} \geq 0.5$$

$$m/n \geq \frac{\mathbf{0.5}}{\mathbf{1-\alpha}} \rightarrow \mathbf{\Gamma} = \frac{\mathbf{0.5}}{\mathbf{1-\alpha}}$$

Silencing additional outlets may not change the election result, but it is costly in terms of bribes and hence it is optimal to silence at least ⌈Γn⌉ outlets. In this context, ⌈Γn⌉ denotes the smallest integer larger or equal to Γn. If the share of informed voters is at least half, the incumbent needs to achieve total media suppression if he is to win.

Should both politicians be discovered as bad, as shown in Lemma 7, informed voters prefer the incumbent and therefore needs to 'manipulate' fewer outlets than a bad challenger. As before, the bad challenger requires to obtain at least half the votes exclusively from rationally inattentive voters and requires to 'influence' ⌈Γn⌉ outlets. For the incumbent, he needs to ensure that the number of voters that find the challenger is bad is at least half. If m denotes the number of outlets the incumbent manipulates:

$$\alpha + \frac{(1-\alpha)m}{n} \geq 0.5$$

$$m/n \leq \frac{0.5-\alpha}{1-\alpha} \rightarrow \mathbf{\Gamma}^* = \frac{\mathbf{0.5-\alpha}}{\mathbf{1-\alpha}} < \mathbf{\Gamma}$$

Additional definitions:

Denote by τ' a bribe. Bribe τ' is said to be:

1. 'acceptable' if an outlet is willing to accept it in exchange for favourable coverage and not deviate.
2. 'affordable' if it allows the politician to 'manipulate' at least ⌈Γn⌉ or ⌈Γ*n⌉ outlets for the challenger if both politicians are revealed as bad.

First notice two properties of Γ:

$$\frac{d\Gamma(\alpha)}{d\alpha} > 0 \; and \; \Gamma(0) = 0.5$$

which indicate that if attempted, the number of outlets a bad politician would have to supress is always at least half.

In all scenarios, a bribe τ is acceptable to outlet i if:

$$\tau \geq P(\sigma_i = b|\sigma_{-i}) - P(\sigma_i = \varphi|\sigma_{-i})$$

i.e. it is higher than the additional revenue a newsagent would earn if he decided to report the true signal compared to the revenue earned with false reporting, given the actions of the other media agents. A news outlet faces a risk λ of being exposed as untrustworthy in the second period if it chooses to misreport either signal. In addition to the analyses of Trombetta and Rossignoli (2020) and Besley and Prat (2006), a bribe would have to account for the possibility that second period revenues may become zero. Since outlets are interested in maximising the sum of revenue in both periods π, an 'acceptable' bribe must hedge the potential loss. As observed in the equations above, the acceptable bribe is strictly larger than the difference in potential first period revenues. Since in the second period, revenues depend only on being revealed as dishonest and not voter distribution, they are accounted for with the second and fourth terms of τ^*. The minimal acceptable bribe now depends on the number of outlets that accept a bribe from the bad politician.

$$\tau(x) \geq P(\sigma_i = b|\sigma_{-i}) - P(\sigma_i = \varphi|\sigma_{-i})$$

In this context, if a bad signal is revealed, the outlet gains $\left(\frac{1-\alpha}{n} + \alpha\right)^{\delta}$ in the first period if it reports honestly. Assuming x outlets reported dishonestly outlet i has an expected second period payoff of

$$E[\mathrm{P}_{2i}(x)] = \sum_{i=0}^{x} \frac{1}{(n-i)^{\delta}} \frac{x!\,\lambda^i(1-\lambda)^{x-i}}{i!\,(x-i)!}$$

If outlet i misreported without receiving a bribe, it obtains $\left(\frac{1-\alpha}{n}\right)^{\delta}$ in the first period, while it's expected payoff in the second period if x outlets misreported is

$$E[\mathrm{P}_{2i}(x)] = (1-\lambda) \sum_{i=0}^{x} \frac{1}{(n-i)^{\delta}} \frac{x!\,\lambda^i(1-\lambda)^{x-i}}{i!\,(x-i)!}$$

Replacing we obtain

$$\tau(x) > P(\sigma_i = b|\sigma_{-i}) - P(\sigma_i = \varphi|\sigma_{-i})$$

$$\tau(x) > \left(\frac{1-\alpha}{n} + \alpha\right)^{\delta} + \sum_{i=0}^{x} \frac{1}{(n-i)^{\delta}} \frac{x!\,\lambda^i(1-\lambda)^{x-i}}{i!\,(x-i)!} - \left(\frac{1-\alpha}{n}\right)^{\delta} - (1$$

$$- \lambda) \sum_{i=0}^{x} \frac{1}{(n-i)^{\delta}} \frac{x!\,\lambda^i(1-\lambda)^{x-i}}{i!\,(x-i)!}$$

$$\tau^*(x) > \left[\left(\frac{1-\alpha}{n} + \alpha\right)^{\delta} - \left(\frac{1-\alpha}{n}\right)^{\delta}\right] + \lambda \sum_{i=0}^{x} \frac{1}{(n-i)^{\delta}} \frac{x!\,\lambda^i(1-\lambda)^{x-i}}{i!\,(x-i)!}$$

Therefore $\tau^*_{J,i}(x) = \left[\left(\frac{1-\alpha}{n} + \alpha\right)^{\delta} - \left(\frac{1-\alpha}{n}\right)^{\delta}\right] + \lambda \sum_{i=0}^{x} \frac{1}{(n-i)^{\delta}} \frac{x!\lambda^i(1-\lambda)^{x-i}}{i!(x-i)!} + 1$ denotes the minimal transfer outlet i would accept from bad politician J in exchange for favourable coverage if x outlets misreport and politician –J is inactive.

Notice, however, that the good politician can influence outlets at a lower cost. Firstly, his interest is in the media outlets reporting sincerely the status of the bad politician. Should a media outlet accept a $\tau > 0$ from the good politician, their reporting strategy remains sincere and they do not risk losing second period revenue despite being influenced by a politician. Secondly, because the reporting strategy remains sincere, the good politician's 'bribe' does not have to take into account the loss in audience revenue an outlet would sustain if influenced by the other politician. Therefore, the minimum acceptable bribe from the good politician, τ^{**}, is defined as:

$$\tau^{**} = \mathrm{P}_{\tau(x)*} - \mathrm{P}_{honest}$$

$$\tau^{**}(\tau_{bad}) = \max\left(0, \left(\frac{1-\alpha}{n}\right)^{\delta} + \tau_{bad} - \left[\left(\frac{1-\alpha}{n} + \alpha\right)^{\delta} - \lambda \sum_{i=0}^{x} \frac{1}{(n-i)^{\delta}} \frac{x!\,\lambda^i(1-\lambda)^{x-i}}{i!\,(x-i)!}\right]\right)$$

$$< \tau_{bad}$$

where $\mathrm{P}_{\tau(x)*}$ and P_{honest} denote the expected payoff for the outlet if it accepts $\tau^*(x)$ from the bad politician and if it reports honestly respectively.

If approached by more than one politician, it is optimal given their information set for the outlet to accept the offer which maximises their expected payoff. This implies that any offer from a bad politician is rejected if lower than τ^* in the case of the bad politician. Should

accepting two offers have the same expected return, they would prefer the one which allows them to remain honest.

Lemma 5

If the budget is constant across politicians and a bad signal is received about only one politician, then total media suppression will not occur in equilibrium.

Proof of Lemma 5:

Any media interference from the challenger is ineffective since both types of voters prefer the incumbent if no signal is observed. It remains to show that total media suppression will not occur in case $\boldsymbol{b_i g_c}$.

Suppose ∃ a vector of payment from the incumbent to the media outlets τ_I such that achieves total media suppression. Therefore, it must be the case that it is a winning strategy for the incumbent (but not necessarily optimal) and:

$$\sum_{i=1,..,n} \tau_{i,I} \leq B \; and \; \tau_{Ii} \geq \tau^*(n) > 0 \; \forall i$$

Three cases are distinguished depending on the realisations of B and α

1. $B \leq n\tau^*(n)$

Under this realisation of B, total media suppression is impossible as the incumbent does not have enough funds to achieve it even if the challenger is inactive

1. $B \geq n\tau^*(n) \; and \; \alpha < 0.5$

If the challenger persuades at least one outlet to report honestly, he gains a share of α votes from attentive voters and rationally inattentive votes proportional to the share of honest outlets. In order for the challenger to win the election, he also needs 0.5- α votes from rationally inattentive voters. Therefore, allowing m to denote the number of outlets the challenger 'manipulates'

$$\frac{(1-\alpha)m}{n} \geq 0.5 - \alpha$$

$$m/n \geq \frac{0.5-\alpha}{1-\alpha} \rightarrow \boldsymbol{\Gamma^* = \frac{0.5-\alpha}{1-\alpha}}$$

Silencing additional outlets does not change the election result, but it is costly in terms of bribes and hence it is optimal to silence $\lceil m/n \rceil$ outlets. In this context, $\lceil m/n \rceil$ denotes the smallest integer larger or equal to m/n. Without loss of generality, order the τ_{Ii} in ascending order.

Given this action from the incumbent and any realisation of B α such that $B \geq n\tau^*(n)\ and\ \alpha < 0.5$ the best response of the challenger is therefore to approach the outlets with the $\lceil \Gamma^* n \rceil$ lowest offers and offer

$$\tau_{C,i} = \tau^{**}(\boldsymbol{\tau_{I,i}})$$

Since the politicians share the same budget, such an offer is affordable for the challenger and it ensures his electoral win. This contradicts the original hypothesis.

2. $\alpha \geq 0.5$

If the challenger persuades at least one outlet to report honestly, he gains a share of α votes from attentive voters and rationally inattentive votes proportional to the share of honest outlets. Therefore, to win the election, the challenger needs one outlet to report honestly. Total media suppression by the incumbent is obviously a suboptimal strategy in equilibrium since both politicians share the same budget and the incumbent needs to silence all n outlets while the challenger only needs to focus on one.

***QED Lemma* 5■**

Lemma 5 indicates that a good politician always prevents a bad politician from achieving total media suppression. Given the minimum number of 'manipulated' outlets required to win, there is a disparity in politician disposable income after the minimum media suppression cost given opponent inaction. This will affect the equilibrium outcome when nature exposes both politicians as bad. For $\Gamma = \frac{0.5}{1-\alpha}\ and\ \Gamma^* = \frac{0.5-\alpha}{1-\alpha}$, the relevant condition will be given by:

$$\lceil \Gamma \mathrm{n} \rceil - \lceil \Gamma^* \mathrm{n} \rceil \geq 1$$

If the condition above holds, denote:

$$k = (\lceil \Gamma \mathrm{n} \rceil - \lceil \Gamma^* \mathrm{n} \rceil), k \in \mathbb{N}\ and\ k \geq 1$$

such that

$$\boldsymbol{B - [\Gamma^{*}n]\tau^{*}(n) - (B - [\Gamma n]\tau^{*}(n)) = k\tau^{*}(n)}$$

Additionally, the politician budgets can be decomposed as:

$$\boldsymbol{B - [\Gamma n]\tau^{*}(n) = l\tau^{*}(n) + \varepsilon, \varepsilon < \tau^{*}(n)\ and\ l \in \mathbb{N}\ and}$$

$$\boldsymbol{B - [\Gamma^{*}n]\tau^{*}(n) = (k + l)\tau^{*}(n) + \varepsilon, \varepsilon < \tau^{*}(n)\ and\ k, l \in \mathbb{N}}$$

Case 1: $\boldsymbol{g_I g_C}$

This can occur if both politicians are good or if either politician is bad and no signal is received by the media. In the absence of a signal, the media outlets are unable to report anything and therefore media framing is impossible. Since no signal is received, the best response of all outlets is to report honestly, even if they receive a transfer from either politician. Both sets of voters observe $\Phi_I\Phi_C$ and their posterior beliefs the politicians are good are both higher than γ. As shown in Lemma 7, they both prefer the incumbent. For both politicians, strictly positive transfers do not change the result of the election and reduce utility and hence their equilibrium actions will be to set $\tau_{I,i} = \tau_{C,i} = 0$. From the perspective of media outlets, their reporting strategy can only be honest even in the presence of positive transfers from the politicians, and therefore would accept any positive transfer, if offered. The incumbent wins and obtains utility 2B, the challenger 0 and all media outlets report honestly. These equilibrium strategies and beliefs are consistent for all realisations of B, n δ and α.

Case 2: $\boldsymbol{g_I b_C}$

This can occur if the challenger is bad and discovered with probability γ (1-γ)q or if both politicians are bad and only the challenger is discovered, with probability $(1-\gamma)^2 q(1-q)$. With probability γ (1-γ)(1-q) the incumbent is good, while the challenger is bad and undiscovered, in which case equilibrium actions are as outlined in $\boldsymbol{g_I g_C}$. From the voters' perspective, this is the posterior belief induced if they observe $\Phi_I b_C$ in the media or from their preferred outlet.

The bad challenger may attempt to silence the media, in which case more than half the voters do not observe the bad signal about the challenger. However, upon seeing no signal about either politician, both types of voters still prefer to vote for the incumbent. An informed voter would revise his posterior belief to 1, whereas an uninformed voter would still prefer the incumbent. Since media persuasion is both costly and ineffective, it is a strictly suboptimal strategy. Therefore $\boldsymbol{\tau_{C,i} = 0, \forall i}$ and the best response of the incumbent is to also set $\boldsymbol{\tau_{I,i} = 0, \forall}$i.

The media outlets then report honestly $\Phi_I b_C$. Should they receive an offer larger or at least as large as $\tau^*(x)\ and\ \tau_{I,i} < \tau^{**}(\boldsymbol{\tau_{I,i}})$, the media outlets would decide to report $\Phi_I \Phi_C$. The incumbent wins and obtains utility 2B, while the challenger obtains 0. A special case where the challenger can suppress the media completely exists, but it is a strictly suboptimal strategy to do so. These equilibrium strategies and beliefs are consistent for all realisations of B, n, δ and $\alpha < 0.5$.

Case 3: $\boldsymbol{g_C b_I}$

This can occur if the incumbent is bad and discovered with probability γ (1-γ)q or if both politicians are bad and only the incumbent is discovered, with probability $(1-\gamma)^2 q(1-q)$. With probability γ(1-γ)(1-q) the challenger is good, while the incumbent is bad and undiscovered, in which case equilibrium actions are as outlined in $\boldsymbol{g_I g_C}$. From the voters' perspective, this is the posterior belief induced if they observe $b_I \Phi_C$ in the media or from their preferred outlet.

As above, the condition for market suppression is given by:

$$B \geq [\pi(no\ bribe|\sigma_{-i}) - \pi(bribe|\sigma_{-i})]\lceil \Gamma n \rceil$$

$$B \geq \left[\left[\left(\frac{1-\alpha}{n} + \alpha\right)^{\delta} - \left(\frac{1-\alpha}{n}\right)^{\delta}\right] + \lambda \sum_{i=0}^{x} \frac{1}{(n-i)^{\delta}} \frac{x!\,\lambda^i (1-\lambda)^{x-i}}{i!\,(x-i)!} + 1\right] \lceil \Gamma n \rceil$$

while the minimal acceptable bribe from the incumbent is given by:

$$\tau^*(x) = \left[\left(\tfrac{1-\alpha}{n} + \alpha\right)^{\delta} - \left(\tfrac{1-\alpha}{n}\right)^{\delta}\right] + \lambda \textstyle\sum_{i=0}^{x} \frac{1}{(n-i)^{\delta}} \frac{x!\lambda^i(1-\lambda)^{x-i}}{i!(x-i)!} + 1.$$

We know from the properties of Γ that in this scenario, the bad politician needs to manipulate at least half of the media outlets in order to avoid certain defeat in the election.

In this case, the incumbent has B- $\lceil \Gamma n \rceil \tau^*(\lceil \Gamma n \rceil)$ revenue leftover from attempting to manipulate $\lceil \Gamma n \rceil$ outlets at the minimum acceptable 'transfer'. Because of the properties of Γ, $n - \lceil \Gamma n \rceil \leq \lceil \Gamma n \rceil$, i.e.the number of outlets the challenger needs to report honestly is smaller than the number of outlets the incumbent needs to 'manipulate'.

By construction, an offer of $\tau^{**}(\boldsymbol{\tau_{I,i}})$ is acceptable from the challenger and preferable to an offer of $\boldsymbol{\tau_{I,i}}$ from the incumbent. The outlet does not risk losing second period revenue if it accepts $\tau^{**}(\boldsymbol{\tau_{I,i}})$ from the good politician since the reporting strategy in this case is sincere.

Therefore, it is a best response for any outlet approached by the good politician with this offer to accept it. Consequently, any outlet approached with $\tau_{i,I} \geq \tau^*(x)$ by the incumbent would accept only if the challenger offers less than $\tau^{**}(\boldsymbol{\tau_{I,i}})$. Because $n - \lceil \Gamma n \rceil \leq \lceil \Gamma n \rceil \; and \; \tau^{**}(\tau^*) < \tau^*(x)$, offering $\tau^{**}(\tau^*(x))$ to all outlets is affordable for the challenger, which indicates that the challenger can ensure all outlets report honestly given any vector of payments from the incumbent. Therefore, the equilibrium action for the incumbent is to set $\boldsymbol{\tau_{I,i}} = \mathbf{0}$ for all outlets and for all B. Consequently, the equilibrium strategy for the challenger is to set $\boldsymbol{\tau_{C,i}} = \mathbf{0}$ if the incumbent has offered $\tau_{i,I} \geq \tau^*(x)$ to less than $\lceil \Gamma n \rceil$ outlets. If $\tau_{I,i} \geq \tau^*(x)$ for $x \geq \lceil \Gamma n \rceil$ outlets, the optimal action of the challenger is to ensure that only $\lceil \Gamma n \rceil - 1$ outlets remain 'manipulated'. This can be achieved by picking the x+1 outlets with the smallest $\tau_{I,i} \geq \tau^*(x)$, letting τ_{I,i^*} denote their set, and:

$$\boldsymbol{\tau_{C,i}} = \begin{cases} \tau^{**}(\boldsymbol{\tau_{I,i}}) \; for \; i \in \tau_{I,i^*} \\ 0 \; otherwise \end{cases}$$

For any outlet that was not approached by any politician, it is a best response to report honestly, $b_I \Phi_C$, and get the higher audience revenues, which benefits the challenger as the resulting posterior beliefs induce the voters to prefer the challenger. For any offer from the incumbent, $\tau_{I,i} \geq \tau^*(x)$, they would accept it and misreport if and only if $\tau_{C,i} < \tau^{**}(\boldsymbol{\tau_{I,i}})$, i.e. report $\Phi_I \Phi_C$.

Therefore, since he moves second and $\tau^{**}(\tau^*(x)) < \tau^*(x) \; \forall x$, for any offer made by the incumbent, the challenger can afford to pay a vector of $\tau_{C,i}$ which ensures total media transparency. The resulting politician equilibrium is that neither politician attempts any media suppression. The incumbent would not deviate from this strategy since any spending can be effectively negated at a lower cost by the challenger, ensuring the incumbent's utility is less than B. Given the strategy of the incumbent, the challenger will not deviate for the same reason. In this equilibrium, all media outlets report honestly, and the incumbent is revealed as bad by all outlets. Based on Lemma 7, all voters prefer the challenger who then wins the election. These equilibrium strategies and beliefs are consistent for all realisations of B, n, δ and $\alpha < 0.5$.

Case 4: $\boldsymbol{b_I b_C}$

With probability $(1-q)^2$ neither politician is discovered as bad and the analysis collapses to that observed in case $\boldsymbol{g_I g_C}$. With probability 1-q the challenger is not discovered

as bad and the analysis collapses to that observed in case $\boldsymbol{b_I g_C}$ and with probability 1-q the incumbent is not discovered and the analysis collapses to that observed in case $\boldsymbol{g_I b_C}$

With probability q^2 both politicians are discovered as bad. As outlined above, the incumbent only needs to capture $\lceil \Gamma^* n \rceil$ outlets, while the challenger needs to capture$\lceil \Gamma n \rceil$. To make the subsequent analysis interesting, I assume that B> $\lceil \Gamma n \rceil \tau^*(\lceil \Gamma n \rceil)$.

First notice that:

$$\lceil \Gamma n \rceil + \lceil \Gamma^* n \rceil \in \{n, n+1\}$$

Any offer from the incumbent of the type: offer $\tau^*(x)$ to $x \geq \lceil \Gamma^* n \rceil$ outlets is strictly suboptimal. Since the challenger moves second, he can gain enough outlets by offering them $\tau^*(x) + 1$ and $\tau^*(x)$ to $\lceil \Gamma n \rceil - \lceil \Gamma^* n \rceil$ for the remaining outlets. This offer is affordable for the challenger, for whom it is an optimal strategy.

A first equilibrium occurs if $\lceil \Gamma n \rceil - \lceil \Gamma^* n \rceil = 0$. In this case, for x denoting the number of additional outlets the incumbent attempts to silence in excess of the $\lceil \Gamma n \rceil$ required, the condition for the incumbent being able to win is:

$$\frac{B}{\lceil \Gamma n \rceil} > B - (n - \lceil \Gamma n \rceil)\tau^*(n)$$

Re-arranging, the incumbent wins if

$$B\frac{\lceil \Gamma n \rceil - 1}{\lceil \Gamma n \rceil} > \tau^*(n)\lceil \Gamma n \rceil$$

If the condition above holds, the incumbent utility maximising action with media suppression is to offer the following vector of payments:

$$\tau_{i,I} = \begin{cases} \tau^*(n) \; to \lceil \Gamma n \rceil \; outlets \\ 0 \; to \; the \; rest \end{cases}$$

While for the challenger, it is optimal to offer $\tau_{i,C} = 0$ for all outlets. Consequently, if

$$B\frac{\lceil \Gamma n \rceil - 1}{\lceil \Gamma n \rceil} \leq \tau^*(n)\lceil \Gamma n \rceil$$

any attempt at media interference from the incumbent is strictly suboptimal as the challenger as the challenger can win by capturing the n-$\lceil \Gamma n \rceil$ outlets not approached by the incumbent at minimum cost while having enough funds to persuade the final outlet. Formally, for the

incumbent it is optimal to offer $\tau_{i,I} = 0$ for all outlets. For the challenger, without loss of generality, order the outlets in ascending order of offers received from the incumbent and offer:

$$\tau_{i,C} = \begin{cases} \tau^*(n) \; to \; the \; first \; n - \lceil \Gamma \mathrm{n} \rceil \; outlets \\ \dfrac{B}{\lceil \Gamma \mathrm{n} \rceil} + 1 \; to \; the \; rest \end{cases}$$

Secondly, if $\lceil \Gamma \mathrm{n} \rceil - \lceil \Gamma^* \mathrm{n} \rceil \geq 1$ there is a disparity in politician disposable income after the minimum media suppression cost given opponent inaction. For $\Gamma = \frac{0.5}{1-\alpha} \; and \; \Gamma^* = \frac{0.5-\alpha}{1-\alpha}$ denote

$k = (\lceil \Gamma \mathrm{n} \rceil - \lceil \Gamma^* \mathrm{n} \rceil), k \in \mathbb{N} \; and \; k \geq 1$ such that

$$\boldsymbol{B - \lceil \Gamma^* n \rceil \tau^*(n) - (B - \lceil \Gamma n \rceil \tau^*(n)) = k \tau^*(n)}$$

Additionally, the politician budgets can be decomposed as:

$$\boldsymbol{B - \lceil \Gamma n \rceil \tau^*(n) = l \tau^*(n) + \mu, \mu < \tau^*(n) \; and \; l \in \mathbb{N} \; and}$$

$$\boldsymbol{B - \lceil \Gamma^* n \rceil \tau^*(n) = (k + l) \tau^*(n) + \mu, \mu < \tau^*(n) \; and \; k, l \in \mathbb{N}}$$

The condition for which the incumbent can win is given by:

$$\tau^*(n) + \frac{(l+k)\tau^*(n)}{x} > \tau^*(n) + \frac{l\tau^*(n)}{1 + x - \lceil \Gamma^* \mathrm{n} \rceil}$$

Where $x \geq \lceil \Gamma^* \mathrm{n} \rceil$ is the number of outlets approached by the incumbent.

Re-arranging

$$(l+k)\tau^*(n)(1 + x - \lceil \Gamma^* \mathrm{n} \rceil) > l\tau^*(n)x$$

$$\frac{(l+k)(1 - \lceil \Gamma^* \mathrm{n} \rceil)}{k} < x$$

which always holds, since $(1 - \lceil \Gamma^* \mathrm{n} \rceil) \leq 0$ and $l, k, x \in \mathbb{N}$. Therefore, any offer of the type $\tau_{I,i} = \tau^*(n) + \frac{(l+k)\tau^*(n)}{x}$ to $x \geq \lceil \Gamma^* \mathrm{n} \rceil \; outlets$ will ensure electoral victory for the incumbent.

The x that maximises the incumbent's welfare is given by:

$$\min_{x \geq \lceil \Gamma^* \mathrm{n} \rceil, x \in \mathbb{N}} x(\tau^*(n) + \frac{(l+k)\tau^*(n)}{x})$$

Which is solved by $x = \lceil \Gamma^* \mathrm{n} \rceil$. Given this strategy for the incumbent, the best response of the challenger is $\tau_{i,C} = 0, \forall i$, since any positive media transfer would decrease his utility without resulting in electoral victory.

Lemma 6:

This equilibrium strategy is feasible for the incumbent if and only if the minimal acceptable vector of bribes for the challenger is affordable.

Proof of Lemma 6:

Suppose the minimal acceptable vector of bribes is affordable for the challenger, i.e.:

$$\tau^*(n)\lceil \Gamma \mathrm{n} \rceil \leq B$$

And the optimal vector of payments for the incumbent described above is not affordable:

$$\lceil \Gamma^* \mathrm{n} \rceil \left(\tau^*(n) + \frac{(l+k)\tau^*(n)}{\lceil \Gamma^* \mathrm{n} \rceil}\right) \geq B$$

$$\lceil \Gamma^* \mathrm{n} \rceil \, \tau^*(n) + (l+k)\tau^*(n) \geq B$$

$$(l+k)\tau^*(n) \geq B - \lceil \Gamma^* \mathrm{n} \rceil \, \tau^*(n)$$

By definition:

$$B - \lceil \Gamma^* \mathrm{n} \rceil \tau^*(n) = (k+l)\tau^*(n) + \mu$$

Thus

$$(l+k)\tau^*(n) \geq (k+l)\tau^*(n) + \mu$$

Contradiction

QED Lemma 6 ■

Therefore, the incumbent utility maximising action with media suppression is to offer the following vector of payments:

$$\tau_{i,I} = \begin{cases} \tau^*(n) + \dfrac{(l+k)\tau^*(n)}{\lceil \Gamma^* \mathrm{n} \rceil} \; to \lceil \Gamma^* \mathrm{n} \rceil \; outlets \\ 0 \; to \; the \; rest \end{cases}$$

if $B \geq \lceil \Gamma \mathrm{n} \rceil \tau^*(n)$.

Contrary to the previous histories, the incumbent can achieve electoral victory by ensuring enough outlets know the challenger is bad, rather than suppressing news about his own type. Should either type of voter observe that both politicians are bad, they would both prefer the incumbent. If both signals are released across the media outlets, the incumbent would obtain α votes from informed voters and a share of rationally inattentive voters proportional to the number of outlets that report honestly. If the incumbent were to offer $\boldsymbol{\tau_{I,i}} \geq \tau^{**}(\boldsymbol{\tau_{C,i}})$ for a given challenger offer $\boldsymbol{\tau_{C,i}}$, he could induce the outlet to report honestly the status of both politicians even if $\boldsymbol{\tau_{C,i}}$ is acceptable, i. e. $\boldsymbol{\tau_{C,i}} > \tau^{*}(x), x \in \mathbb{N}$. As in $\boldsymbol{g_C b_I}$, the outlet does not risk losing second period revenue since the resulting reporting strategy is sincere. As a result, it is a best response for any outlet to accept such an offer for a given $\boldsymbol{\tau_{C,i}}$ and report honestly $\boldsymbol{b_I b_C}$. Under full disclosure, the equilibrium strategy for the challenger is to offer $\boldsymbol{\tau_{C,i}} = \mathbf{0}, \forall \boldsymbol{i}$ which means that neither politician attempts any media suppression. If feasible, both politicians refrain from signal suppression and all media outlets report honestly.

Lemma 6.1

If both politicians are exposed as bad, the share of informed voters is less than half and their budgets are equal, full disclosure is always a suboptimal strategy.

Proof of Lemma 6.1:

In order for the incumbent to win the election with full disclosure, he needs 0.5- α votes from rationally inattentive voters, which is equivalent to a share $\lceil \Gamma^{*} \mathrm{n} \rceil$ of outlets. Since $\lceil \Gamma \mathrm{n} \rceil + \lceil \Gamma^{*} \mathrm{n} \rceil \in \{n, n+1\}$, two cases emerge. If $\lceil \Gamma \mathrm{n} \rceil + \lceil \Gamma^{*} \mathrm{n} \rceil = n + 1$, influencing $\lceil \Gamma^{*} \mathrm{n} \rceil$ outlets ensures the challenger has one fewer outlet than required. If $\lceil \Gamma \mathrm{n} \rceil + \lceil \Gamma^{*} \mathrm{n} \rceil = n$, the incumbent would still win by 'influencing' $\lceil \Gamma^{*} \mathrm{n} \rceil$ outlets as the election will result in a tied vote. However, the optimal strategy of the incumbent will differ slightly in these cases.

If $\lceil \Gamma \mathrm{n} \rceil + \lceil \Gamma^{*} \mathrm{n} \rceil = n + 1$, then the challenge has $B - (\lceil \Gamma \mathrm{n} \rceil - 1)\tau^{*}(\lceil \Gamma \mathrm{n} \rceil)$ funds remaining to persuade the final outlet to misreport his state. The best response for the incumbent is to offer any $\lceil \Gamma^{*} \mathrm{n} \rceil$ outlets the sum:

$$\boldsymbol{\tau_{I,i}} = \tau^{**}(B - (\lceil \Gamma \mathrm{n} \rceil - 1)\tau^{*}(\lceil \Gamma \mathrm{n} \rceil))$$

which simplifies to:

$$\boldsymbol{\tau_{I,i}} = B - (\lceil \Gamma \mathrm{n} \rceil - 2)\tau^{*}(\lceil \Gamma \mathrm{n} \rceil) + 1$$

If $\lceil \Gamma n \rceil + \lceil \Gamma^* n \rceil = n$, then the challenge has $B - (\lceil \Gamma n \rceil - 1)\tau^*(\lceil \Gamma n \rceil)$ funds remaining to persuade the final outlet to misreport his state. The best response for the incumbent is to offer any $\lceil \Gamma^* n \rceil$ outlets the sum:

$$\boldsymbol{\tau_{I,i}} = \tau^{**}(B - \lceil \Gamma n \rceil \tau^*(\lceil \Gamma n \rceil))$$

which simplifies to:

$$\boldsymbol{\tau_{I,i}} = B - (\lceil \Gamma n \rceil - 1)\tau^*(\lceil \Gamma n \rceil) + 1$$

Therefore, the incumbent utility maximising action with full disclosure is to offer the following vector of payments:

$$\tau_{i,I} = \begin{cases} B - (\lceil \Gamma n \rceil - 2)\tau^*(\lceil \Gamma n \rceil) + 1 \; to \lceil \Gamma^* n \rceil \; outlets \; if \; \lceil \Gamma n \rceil + \lceil \Gamma^* n \rceil = n + 1 \\ B - (\lceil \Gamma n \rceil - 1)\tau^*(\lceil \Gamma n \rceil) + 1 \; to \lceil \Gamma^* n \rceil \; outlets \; if \; \lceil \Gamma n \rceil + \lceil \Gamma^* n \rceil = n \\ 0 \; to \; the \; rest \end{cases}$$

However, in order for this strategy to be feasible, it is required that:

$$B - \lceil \Gamma^* n \rceil (B - (\lceil \Gamma n \rceil - 1)\tau^*(\lceil \Gamma n \rceil) + 1) > 0$$

Or equivalently:

$$\lceil \Gamma^* n \rceil (\lceil \Gamma n \rceil - 1)\tau^*(\lceil \Gamma n \rceil) > B(\lceil \Gamma^* n \rceil - 1) + \lceil \Gamma^* n \rceil$$

$$\lceil \Gamma^* n \rceil (\lceil \Gamma n \rceil - 1)\tau^*(\lceil \Gamma n \rceil) > B(\lceil \Gamma^* n \rceil - 1) + B - B + \lceil \Gamma^* n \rceil$$

$$\lceil \Gamma^* n \rceil (\lceil \Gamma n \rceil - 1)\tau^*(\lceil \Gamma n \rceil) > B(\lceil \Gamma^* n \rceil) - B + \lceil \Gamma^* n \rceil$$

$$(\lceil \Gamma n \rceil - 1)\tau^*(\lceil \Gamma n \rceil) > B - \frac{B}{\lceil \Gamma^* n \rceil} + 1$$

$$\tau^*(\lceil \Gamma n \rceil) > \frac{B}{\lceil \Gamma^* n \rceil} - 1$$

To arrive at the equilibrium strategy for the incumbent, compare incumbent utility under full disclosure and media suppression. In the latter case, his utility maximising action is:

$$\tau_{i,I} = \begin{cases} \tau^*(n) + \dfrac{(l+k)\tau^*(n)}{\lceil \Gamma^* n \rceil} \; to \lceil \Gamma^* n \rceil \; outlets \\ 0 \; to \; the \; rest \end{cases}$$

And therefore, his utility in the first period is:

$$U_I = B - \lceil \Gamma^* n \rceil (\tau^*(n) + \frac{(l+k)\tau^*(n)}{\lceil \Gamma^* n \rceil})$$

$$U_I = \tau^*(n)\lceil \Gamma^* \mathrm{n} \rceil + (l+k)\tau^*(n)$$

$$U_I = \mu, \mu < \tau^*(\mathrm{n})$$

Under full disclosure, the incumbent's utility in the first period is given by:

$$U_I = \begin{cases} B - \lceil \Gamma^* \mathrm{n} \rceil (B - (\lceil \Gamma \mathrm{n} \rceil - 2)\tau^*(\lceil \Gamma \mathrm{n} \rceil) + 1\ ) \ if\ \lceil \Gamma \mathrm{n} \rceil + \lceil \Gamma^* \mathrm{n} \rceil = n+1 \\ B - \lceil \Gamma^* \mathrm{n} \rceil (B - (\lceil \Gamma \mathrm{n} \rceil - 1)\tau^*(\lceil \Gamma \mathrm{n} \rceil) + 1\ ) \ if\ \lceil \Gamma \mathrm{n} \rceil + \lceil \Gamma^* \mathrm{n} \rceil = n \end{cases}$$

Beginning with the incumbent utility under full disclosure if $\lceil \Gamma \mathrm{n} \rceil + \lceil \Gamma^* \mathrm{n} \rceil = n$, we show that full disclosure is always a suboptimal strategy.

$$B - \lceil \Gamma^* \mathrm{n} \rceil (B - (\lceil \Gamma \mathrm{n} \rceil - 1)\tau^*(\lceil \Gamma \mathrm{n} \rceil) + 1\ ) < \mu$$

$$B(1 - \lceil \Gamma^* \mathrm{n} \rceil) + \lceil \Gamma^* \mathrm{n} \rceil (\lceil \Gamma \mathrm{n} \rceil - 1)\tau^*(\lceil \Gamma \mathrm{n} \rceil) - \lceil \Gamma^* \mathrm{n} \rceil < \mu$$

$$\lceil \Gamma^* \mathrm{n} \rceil (\lceil \Gamma \mathrm{n} \rceil - 1)\tau^*(\lceil \Gamma \mathrm{n} \rceil) < \mu + \lceil \Gamma^* \mathrm{n} \rceil + B(\lceil \Gamma^* \mathrm{n} \rceil - 1)$$

$$(\lceil \Gamma \mathrm{n} \rceil - 1)\tau^*(\lceil \Gamma \mathrm{n} \rceil) < \frac{\mu}{\lceil \Gamma^* \mathrm{n} \rceil} + 1 + B - \frac{B}{\lceil \Gamma^* \mathrm{n} \rceil}$$

$$(\lceil \Gamma \mathrm{n} \rceil - 1)\tau^*(\lceil \Gamma \mathrm{n} \rceil) - B + \frac{B}{\lceil \Gamma^* \mathrm{n} \rceil} - 1 < \frac{\mu}{\lceil \Gamma^* \mathrm{n} \rceil}$$

Affordability implies $(\lceil \Gamma \mathrm{n} \rceil - 1)\tau^*(\lceil \Gamma \mathrm{n} \rceil) - B < 0$ and full disclosure feasibility implies:

$$\tau^*(\lceil \Gamma \mathrm{n} \rceil) > \frac{B}{\lceil \Gamma^* \mathrm{n} \rceil} - 1 \rightarrow$$

$$(\lceil \Gamma \mathrm{n} \rceil - 1)\tau^*(\lceil \Gamma \mathrm{n} \rceil) - B + \frac{B}{\lceil \Gamma^* \mathrm{n} \rceil} - 1 < (\lceil \Gamma \mathrm{n} \rceil)\tau^*(\lceil \Gamma \mathrm{n} \rceil) - B \leq 0$$

On the right-hand side, both μ and $\lceil \Gamma^* \mathrm{n} \rceil$ are positive by construction.

Using similar calculations for the case when $\lceil \Gamma \mathrm{n} \rceil + \lceil \Gamma^* \mathrm{n} \rceil = n + 1$, full disclosure if a suboptimal strategy if

$$(\lceil \Gamma \mathrm{n} \rceil - 2)\tau^*(\lceil \Gamma \mathrm{n} \rceil) - B + \frac{B}{\lceil \Gamma^* \mathrm{n} \rceil} - 1 < \frac{\mu}{\lceil \Gamma^* \mathrm{n} \rceil}$$

Which also always holds for the same reasons. This concludes the proof.

***QED Lemma* 6.1■**

Therefore, the equilibrium strategy for the incumbent is

$$\tau_{i,I} = \begin{cases} \tau^*(n) + \dfrac{(l+k)\tau^*(n)}{\lceil \Gamma^* \mathrm{n} \rceil} \ to \lceil \Gamma^* \mathrm{n} \rceil \ outlets \\ 0 \ to \ the \ rest \end{cases}$$

if $B \geq \lceil \Gamma \mathrm{n} \rceil \tau^*(n)$.

If $B \in [\lceil \Gamma^* \mathrm{n} \rceil \tau^*(n), \lceil \Gamma \mathrm{n} \rceil \tau^*(n))$, the incumbent can still afford enough media suppression to allow his victory, but the challenger is unable to afford meaningful media intervention. Therefore, the equilibrium strategy of the incumbent becomes:

$$\tau_{i,I} = \begin{cases} \tau^*(n) \ to \lceil \Gamma^* \mathrm{n} \rceil \ outlets \\ 0 \ to \ the \ rest \end{cases}$$

If $B \in [\lceil \Gamma^* \mathrm{n} \rceil \tau^*(\lceil \Gamma^* \mathrm{n} \rceil), \lceil \Gamma^* \mathrm{n} \rceil \tau^*(n))$, the incumbent is still able to afford meaningful media suppression and his equilibrium strategy becomes:

$$\tau_{i,I} = \begin{cases} \tau^*(\lceil \Gamma^* \mathrm{n} \rceil) \ to \lceil \Gamma^* \mathrm{n} \rceil \ outlets \\ 0 \ to \ the \ rest \end{cases}$$

Finally, if $B < \lceil \Gamma^* \mathrm{n} \rceil \tau^*(\lceil \Gamma^* \mathrm{n} \rceil)$, neither politician can afford meaningful media suppression and their equilibrium strategies are $\tau_{i,I} = \tau_{i,C} = 0$

For any vector of payments from the incumbent other than the ones outlined above, there exists a corresponding vector for the challenger which allows him to win. Let τ_I denote a vector of incumbent payments below the optimal one described above.

Let x denote the number of outlets for which $\tau_{I,i}$=0 and order the remaining outlets ascendingly. Then if:

1) x≥⌈Γn⌉ **or**

2) x<⌈Γn⌉ **and**

$$B - x\tau^*(n) \geq \sum_{j=x+1}^{\lceil \Gamma \mathrm{n} \rceil} \tau_{I,j} + (\lceil \Gamma \mathrm{n} \rceil - \mathrm{x})$$

$$\tau_{C,i} = \begin{cases} \tau^*(n) \ to \min(x, \lceil \Gamma n \rceil) \ \ to \ outlets \ s.t \ \tau_{I,i} = 0 \\ \tau_{I,i} + 1 \ to \text{ the first } \max(\lceil \Gamma n \rceil - x, 0) \ \ ordered \ outlets \\ 0 \ otherwise \end{cases}$$

Otherwise $\boldsymbol{\tau_{C,i}} = 0$

If the minimal acceptable vector of bribes for the challenger is not affordable, then the best response of the incumbent is to offer $\tau_{I,i} = 0$ to all outlets. This is the minimum cost at which the incumbent can guarantee he wins as the challenger lacks the funds to implement a media intervention strategy that would allow him to win. Since the challenger is not able to afford any media market intervention which would allow him to win, he will not deviate.

From the perspective of the media outlets, their optimal strategy is to report honestly, $\mathrm{b}_I\mathrm{b}_C$, if they receive no offer higher than $\tau^*(n)$ or if $\tau_j \geq \tau^*(n)$ and $\tau_{-j} \geq \tau^{**}(\tau_j)$. Otherwise, they accept the higher offer, and they report $\Phi_I\mathrm{b}_C$ or $\mathrm{b}_I\Phi_C$ depending on the origin of the offer. According to Lemma 7, voters prefer the incumbent if they observe $\mathrm{b}_I\mathrm{b}_C$, therefore, the incumbent wins the election.

Second Equilibrium

A second equilibrium occurs if $\alpha \geq 0.5$. In this case, more than half of the voters are informed voters, in which case the game becomes a two-politician extension of Besley and Prat (2006). The crucial difference with respect to the first two cases is that a bad politician needs to suppress all the outlets in the market. If at least one outlet remains 'free' the informed voters find out about the bad politician and thus ensuring his loss unless he is the incumbent and the challenger is also revealed as bad.

Histories/Signal received by media	$\Phi_I\Phi_C$	$\Phi_I b_C$	$b_I\Phi_C$	$b_I b_C$	Beliefs
Prior Probability	$\gamma^2 + 2(1-\gamma)\gamma q + (1-\gamma)^2(1-q)^2$	$\gamma(1-\gamma)q + (1-\gamma)^2 q(1-q)$	$\gamma(1-\gamma)q + (1-\gamma)^2 q(1-q)$	$(1-\gamma)^2 q^2$	
Agents and strategies					
Incumbent	$\boldsymbol{\tau_{I,i}} = 0, \forall i,$ $\forall B, \forall\, E(\tau_C)$	$\boldsymbol{\tau_{I,i}} = 0, \forall i,$ $\forall B, \forall\, E(\tau_C)$	$\boldsymbol{\tau_{I,i}} = 0, \forall i$ $\forall B, and\ \ \forall\, E(\tau_C)\ such\ that\ \forall \boldsymbol{\tau_{C,i}} \geq \tau^{**}(\boldsymbol{\tau_{I,i}})$ Otherwise, a) if $B \geq \lceil \Gamma n \rceil \tau^*(n)$ $\boldsymbol{\tau_{I,i}} = \tau^*(n)\ to\ all\ outlets$ b) if $B < \lceil \Gamma n \rceil \tau^*(n)$,then $\tau_{I,i} = 0, \forall i$	$\tau_{I,i} = B - (n-2)\tau^*(n)$ to 1 outlet, zero otherwise $\forall\, E(\tau_C)$ if $B \geq n\tau^*(n)$ If $B < n\tau^*(n)$, $\tau_{I,i} = 0$, $\forall i \forall\, E(\tau_C)$	Informed voters: Believe incumbent is good if they observe $\Phi_I\Phi_C$ or $\Phi_I b_C$ in the media. Believe challenger is
Challenger	$\boldsymbol{\tau_{C,i}} = 0, \forall i,$ $\forall B, \forall\, \tau_I$	$\boldsymbol{\tau_{C,i}} = 0, \forall i,$ $\forall B, \forall\, \tau_I$	If $\tau_{I,i} < \tau^*(n)$ for $n-1$ outlets, then $\boldsymbol{\tau_{C,i}} = 0, \forall i\ \forall B.$	$\tau_{C,i} = 0, \forall i$ if $B < n\tau^*(n)$	

			If $\tau_{I,i} \geq \tau^*(n)$ for $all\ n$ outlets, pick the outlet with the smallest $\tau_{I,i}$ and offer $\tau_{C,i} = \tau^{**}(\boldsymbol{\tau_{I,i}})$	If $B \geq n\tau^*(n)$, let x denote the number of outlets for which $\tau_{I,i}$=0 and order the remaining outlets ascendingly. Then if 1) x=n **or** 2) x<n **and** $B - x\tau^*(n) \geq \sum_{j=x+1}^{n} \tau_{I,j} + (\mathrm{n} - \mathrm{x})$ $\tau_{C,i}$ $= \begin{cases} \tau^*(n)\ to\ all\ s.t\ \tau_{I,i} = 0 \\ \tau_{I,i} + 1\ to\ the\ rest\ of\ the\ outlets \end{cases}$ Otherwise $\boldsymbol{\tau_{C,i}} = 0$	good if observe $\Phi_I\Phi_C$ or $\mathrm{b}_I\Phi_C$ in the media. Believe both politicians are bad if they observe $\mathrm{b}_I\mathrm{b}_C$ in the media.
Result	Incumbent wins	Incumbent wins	Challenger wins	Incumbent wins	Uninformed voters:
Media Agents	Report honestly,$\Phi_I\Phi_C$, for $\forall\ \tau_I\ and\ \forall\ \tau_C$. If	Report honestly $\Phi_I\mathrm{b}_C$ if $\tau_{C,i} < \tau^*(x), \forall\ \tau_I$	Report honestly $\mathrm{b}_I\Phi_C$ unless $\begin{cases} \tau_{I,i} \geq \tau^*(x)\ \textbf{and} \\ \tau_{C,i} < \tau^{**}(\boldsymbol{\tau_{I,i}}) \end{cases}$	Report honestly $\mathrm{b}_I\mathrm{b}_C$ if either set of conditions holds	Believe incumbent is good if observe $\Phi_I\Phi_C$

	$\tau_{i,I} \neq 0$ and or $\tau_{C,i} \neq 0$, they would accept the higher offer and still report honestly	Or if $\tau_{C,i} \geq \tau^*(x)$ and $\tau_I \geq \tau^{**}(\boldsymbol{\tau_{C,i}})$ Report $\Phi_I\Phi_C$ if $\tau_{C,i} \geq \tau^*(x)$ and $\tau_I < \tau^{**}(\boldsymbol{\tau_{C,i}(x)})$	then report $\Phi_I\Phi_C$	$\tau_{I,i} < \tau^*(x)$**and** $\tau_{C,i} < \tau^*(x)$ ***OR*** $\tau_{I,j} > \tau^*(x)$ ***and*** $\tau_{I,-j} > \tau^{**}(\tau^*(x))$ For $j \in \{I, C\}$ Report $\Phi_I b_C$ if $\tau_{I,i} \geq \tau^*(x) and\ \tau_{I,i} \geq \tau_{C,i}$ Report $b_I\Phi_C$ if $\tau_{C,i} \geq \tau^*(x) and\ \tau_{I,i} < \tau_{C,i}$	or $\Phi_I b_C$ from their outlet. Believe challenger is good if they observe $\Phi_I\Phi_C$ or $b_I\Phi_C$ from their outlet. Believe both politicians are bad if they observe $b_I b_C$ from their outlet
Informed Voters	Vote incumbent unless observe $b_I\Phi_C$ in the media	Vote incumbent unless observe $b_I\Phi_C$ in the media	Vote challenger if they observe $b_I\Phi_C$ in the media. Vote incumbent if they observe $\Phi_I\Phi_C$ in the media	Vote challenger if they observe $b_I\Phi_C$ in the media. Vote incumbent if they observe $b_I b_C$ or $\Phi_I b_C$ in the media	
Uninformed Voters	Vote incumbent unless observe $b_I\Phi_C$ from their outlet	Vote incumbent unless observe $B_I\Phi_C$ from their outlet	Vote challenger if they observe $b_I\Phi_C$ from their outlet. Vote incumbent if they observe $\Phi_I\Phi_C$ from their outlet	Vote challenger if they observe $b_I\Phi_C$ from their outlet. Vote incumbent if they observe $b_I b_C$ or $\Phi_I b_C$ from their outlet	

Case 1: $\boldsymbol{g_I g_C}$

This can occur if both politicians are good or if either politician is bad and no signal is received by the media. In the absence of a signal, the media outlets are unable to report anything and therefore media framing is impossible and the best response of all outlets is to report honestly, even if they receive a transfer from either politician. Both sets of voters observe $\Phi_I \Phi_C$ and their posterior beliefs the politicians are good are both higher than γ. As shown in Lemma 7, they both prefer the incumbent. For both politicians, strictly positive transfers do not change the result of the election and reduce utility and hence their equilibrium actions will be to set $\tau_{I,i} = \tau_{C,i} = 0$. From the perspective of media outlets, their reporting strategy can only be honest even in the presence of positive transfers from the politicians, and therefore would accept any positive transfer, if offered. The incumbent wins and obtains utility 2B, the challenger 0 and all media outlets report honestly. These equilibrium strategies and beliefs are consistent for all realisations of B, n δ and α.

Case 2: $\boldsymbol{g_I b_C}$

This can occur if the challenger is bad and discovered with probability γ (1-γ)q or if both politicians are bad and only the challenger is discovered, with probability $(1-\gamma)^2 q(1-q)$. With probability γ (1-γ)(1-q) the incumbent is good, while the challenger is bad and undiscovered, in which case equilibrium actions are as outlined in $\boldsymbol{g_I g_C}$. From the voters' perspective, this is the posterior belief induced if they observe $\Phi_I \mathrm{b}_C$ in the media or from their preferred outlet. The bad challenger may attempt to silence all media outlets, in which case no voters, informed or otherwise, observe the bad signal about the challenger. However, upon seeing no signal about neither politician, following Lemma 7, both types of voters prefer the incumbent. Since media persuasion is both costly and ineffective, it is a strictly suboptimal strategy. Therefore $\boldsymbol{\tau_{C,i}} = 0, \forall i$ and the best response of the incumbent is to also set $\boldsymbol{\tau_{I,i}} = 0, \forall$i. Should they receive an offer larger or at least as large as $\tau^*(n)$ *and* $\tau_I < \tau^{**}(\boldsymbol{\tau_{C,i}})$, the media outlets would decide to report $\Phi_I \Phi_C$. The incumbent wins and obtains utility 2B, while the challenger obtains 0. A special case where the challenger can suppress the media completely exists, but it is a strictly suboptimal strategy to do so. These equilibrium strategies and beliefs are consistent for all realisations of B, n, δ and $\alpha > 0.5$.

Case 3: $\boldsymbol{g_C b_I}$

This can occur if the incumbent is bad and discovered with probability γ (1-γ)q or if both politicians are bad and only the incumbent is discovered, with probability $(1-\gamma)^2 q(1-q)$.With probability γ (1-γ)(1-q) the challenger is good, while the incumbent is bad and undiscovered, in which case equilibrium actions are as outlined in $\boldsymbol{g_I g_C}$. From the voters' perspective, this is the posterior belief induced if they observe $\mathrm{b}_I \Phi_C$ in the media or from their preferred outlet.

As above, the condition for market suppression is given by:

$$B \geq [\pi(no\ bribe|\sigma_{-i}) - \pi(bribe|\sigma_{-i})]n$$

$$B \geq \left[\left[\left(\frac{1-\alpha}{n}+\alpha\right)^{\delta} - \left(\frac{1-\alpha}{n}\right)^{\delta}\right] + \lambda \sum_{i=0}^{x} \frac{1}{(n-i)^{\delta}} \frac{x!\,\lambda^i (1-\lambda)^{x-i}}{i!\,(x-i)!} + \varepsilon\right] n$$

while the minimal acceptable bribe from the incumbent is given by:

$$\tau^*(x) = \left[\left(\tfrac{1-\alpha}{n}+\alpha\right)^{\delta} - \left(\tfrac{1-\alpha}{n}\right)^{\delta}\right] + \lambda \textstyle\sum_{i=0}^{x} \frac{1}{(n-i)^{\delta}} \frac{x!\lambda^i(1-\lambda)^{x-i}}{i!(x-i)!} + \varepsilon.$$

We know that since at least half of the voters are informed in this scenario, the bad politician needs to manipulate all media outlets in order to avoid certain defeat in the election. Consequently, the challenger only needs to ensure at least 1 outlet remains 'free' and reports that the incumbent is bad.

In this case, the incumbent has B- $(n-1)\tau^*(n)$ revenue leftover from attempting to manipulate $(n-1)$ outlets, which is the maximum budget available to persuade the final outlet. Because the challenger is good, as before, he is able to persuade that outlet to report honestly at lower cost. By construction, an offer of $\tau^{**}(\tau^*(n))$is acceptable from challenger and preferable to $\tau^*(n)$ from the incumbent. Since $\tau^{**}(\tau^*(n)) < \tau^*(n)$, an offer of B- $(n-1)\tau^{**}$ from the challenger is also affordable for the politician. To ensure that one outlet strictly prefers to report honestly and he wins the election, the challenger can offer B- $(n-1)\tau^{**}(\tau^*)$ to one outlet as the incumbent is unable to afford suppressing the final outlet. It is therefore a suboptimal strategy for the incumbent to attempt media persuasion since it is strictly utility reducing and ineffective. The resulting equilibrium is for the incumbent to offer $\tau_{I,i} = 0, \forall \mathrm{I}$, and the corresponding best response from the challenger is $\tau_{C,i} = 0, \forall i$. The incumbent would not deviate from this strategy since any spending will never lead to electoral victory, ensuring

the incumbent's utility is less than B. The challenger will not deviate given the action of the incumbent since he already wins the election. Both politicians obtain utility B in the first period. The media outlets report honestly. These equilibrium strategies and beliefs are consistent for all realisations of B such that l≥ n, n, δ and $\alpha > 0.5$.

If the realisation of B is such that the incumbent is unable to afford any meaningful media interference, the equilibrium strategies remain unchanged. Therefore, the equilibrium strategy for the incumbent and the challenger is $\tau_{I,i} = \tau_{C,i} = 0 \; \forall i$ and any realisation of B.

Case 4: $\boldsymbol{b_I b_C}$

With probability $(1-q)^2$ neither politician is discovered as bad and the analysis collapses to that observed in case $\boldsymbol{g_I g_C}$. With probability 1-q the challenger is not discovered as bad and the analysis collapses to that observed in case $\boldsymbol{b_I g_C}$ and with probability 1-q the incumbent is not discovered and the analysis collapses to that observed in case $\boldsymbol{g_I b_C}$.With probability q^2 both politicians are discovered as bad. The challenger needs to capture all the outlets in the market to ensure his victory whereas the incumbent only needs to prevent one outlet from being captured.

As above, the condition for market suppression is given by:

$$B \geq [\pi(no\ bribe|\sigma_{-i}) - \pi(bribe|\sigma_{-i})]n$$

$$B \geq \left[\left[\left(\frac{1-\alpha}{n} + \alpha\right)^{\delta} - \left(\frac{1-\alpha}{n}\right)^{\delta}\right] + \lambda \sum_{i=0}^{x} \frac{1}{(n-i)^{\delta}} \frac{x!\,\lambda^i (1-\lambda)^{x-i}}{i!\,(x-i)!} + \varepsilon\right] n$$

while the minimal acceptable bribe if x outlets are 'suppressed' from either politician is given by:

$$\tau^*(x) = \left[\left(\frac{1-\alpha}{n} + \alpha\right)^{\delta} - \left(\frac{1-\alpha}{n}\right)^{\delta}\right] + \lambda \sum_{i=0}^{x} \frac{1}{(n-i)^{\delta}} \frac{x!\,\lambda^i (1-\lambda)^{x-i}}{i!\,(x-i)!} + \varepsilon$$

As outlined above, the challenger needs to persuade all outlets to misreport his state for him to win the election. Consequently, the incumbent only needs to persuade one outlet to report the state of the challenger honestly. In this case, the challenger would have B- $(n-1)\tau^*(n)$ revenue leftover from attempting to manipulate $(n-1)$ outlets, which is the maximum budget available to persuade the final outlet. The incumbent can then achieve electoral victory in two ways: full disclosure or media suppression. Under the latter, the

incumbent would 'silence' enough outlets such that enough rationally inattentive voters do not observe that his type is bad.

Under full disclosure, the incumbent can achieve electoral victory by ensuring enough outlets know the challenger is bad, rather than suppressing news about his own type. Should either type of voter observe that both politicians are bad, they would both prefer the incumbent. If both signals are released across the media outlets, the incumbent would obtain α votes from informed voters and a share of rationally inattentive voters proportional to the number of outlets that report honestly. If the incumbent were to offer $\boldsymbol{\tau_{I,i}} \geq \tau^{**}(\boldsymbol{\tau_{C,i}})$ for a given challenger offer $\boldsymbol{\tau_{C,i}}$, he could induce the outlet to report honestly the status of both politicians even if $\boldsymbol{\tau_{C,i}}$ is acceptable, i. e. $\boldsymbol{\tau_{C,i}} > \tau^{*}(n)$. As in $\boldsymbol{g_C b_I}$, the outlet does not risk losing second period revenue since the resulting reporting strategy is sincere. As a result, it is a best response for any outlet to accept such an offer for a given $\boldsymbol{\tau_{C,i}}$ and report honestly $\boldsymbol{b_I b_C}$. Under full disclosure, the equilibrium strategy for the challenger is to offer $\boldsymbol{\tau_{C,i}} = \mathbf{0}, \forall \boldsymbol{i}$ which means that neither politician attempts any media suppression. If feasible neither politician would attempt signal suppression and all media outlets report honestly.

The incumbent is then able to ensure one outlet reports $\Phi_I \mathrm{b}_C$ by offering any outlet B- $(n-1)\tau^{*}(n)$. Since B- $(n-1)\tau^{*}(n) < B$, this offer is affordable and results in the incumbent winning the election. Under full disclosure, the incumbent can ensure that enough voters know both politicians are bad, i.e. $\mathrm{b}_I \mathrm{b}_C$ by offering one outlet $\boldsymbol{\tau_{I,i}} = \tau^{**}(\mathrm{B} - (n-1)\tau^{*}(n))$, which is also affordable, as:

$$\tau^{**}(\mathrm{B} - (n-1)\tau^{*}(n)) = \mathrm{B} - (n-2)\tau^{*}(n) + 1 < \mathrm{B} - (n-1)\tau^{*}(n) < B$$

Full disclosure is an optimal strategy for the incumbent. He will not deviate and spend lower since it will allow the challenger to achieve total media suppression. The incumbent will not spend more since he gains strictly lower utility if he does so. Given the actions of the incumbent, the best response of the challenger is to set $\tau_{C,i} = 0, \forall i$. He will not deviate since he will be unable to achieve total market suppression. The incumbent obtains revenue equal to $2B - \tau^{**}(\mathrm{B} - (n-1)\tau^{*}(n))$, while the challenger obtains 0. The equilibrium in the media market is that all outlets report honestly, $\mathrm{b}_I \mathrm{b}_C$.

These equilibrium strategies and beliefs are consistent for all realisations of B≥nτ^{*}, n, δ and $\alpha > 0.5$.

As in Trombetta and Rossignoli (2020) the equilibrium beliefs and strategies depend on the realisations of α, which can lead to heterogeneous levels of capture required to win the election, which is consistent with those observed empirically. However, a key difference is that the results indicate that in pure strategy Perfect Bayesian equilibria media and political competition and electoral standards ensure that the true state of politicians is always revealed to a subset of the electorate. This implies that if at least one of the politicians is not revealed as bad, he will win the election as the choice voters face is either between politician they know with certainty to be bad and one which is good with some strictly positive probability or two politicians of the latter type. This further implies that a good politician always wins the election whenever he is present. The assumption that bad signals cannot be fabricated is crucial to this result. When both politicians are revealed as bad, the assumptions about incumbency advantage result in the incumbent being able to win the election, while still allowing a subset of the electorate to be correctly informed.

Interestingly, in this equilibrium, full disclosure is optimal both in the media market and from the winning politician's perspective, if and only if the share of informed voters is at least half. This arises due to the presence of political competition with constant budgets when the 'bad' politician needs to capture the entirety of the media whereas his opponent only has to ensure one outlet remains honest. If the share of informed voters is less than half, even when media market capture is rational full media capture is suboptimal as victory can be achieved with fewer funds and fewer captured outlets.

The equilibrium in the media market suggests that irrespective of the share of informed voters, whenever at least one politician is not revealed as bad, we observe all outlets reporting honestly. The same strategies are observed when both politicians are bad, and the share of informed voters is at least half. This result is in line with previous literature where supply side bias is diminished by increased media market competition. The results on voter welfare suggest that media and political competition together increase voter welfare.

QED Proposition 4■

Abbreviations

A10	Countries which joined the European Union in 2004
BES	British Election Study
EU	European Union
EU10	Countries which joined the European Union in 2004
EU15	Existing EU members before 2004
GCE	General Certificate of Education
GCSE	General Certificate of Secondary Education
GE	General Election
GVA	Gross Value Added
LFS	Labour Force Survey
NRS	National Readership Survey
NUTS	Nomenclature of Territorial Units for Statistics
OECD	Organisation for Economic Cooperation and Development
OLS	Ordinary Least Squares
ONS	Office of National Statistics
PBE	Perfect Bayesian Equilibrium
QED	Quod Erat Demonstrandum
RIF	Re-centred Influence Function
RQF	Regulated Qualifications Framework
UK	United Kingdom
US	United States

Bibliography

Anand, B., Di Tella, R. and Galetovic, A. (2007). Information or opinion? Media bias as product differentiation. *Journal of Economics & Management Strategy*, 16(3), 635-682.

Anderson, S.P. and McLaren, J. (2012). Media mergers and media bias with rational consumers. *Journal of the European Economic Association*, 10(4), 831-859.

Bandyopadhyay, S., Chatterjee, K. and Roy, J. (2015). *Manufacturing extremism: political consequences of profit-seeking media*. Department of Economics, University of Birmingham.

Banerjee, A., Kumar, S., Pande, R. and Su, F. (2011). Do informed voters make better choices? Experimental evidence from urban India. *Unpublished manuscript*, Harvard University.

Baron, D.P. (2006). Persistent media bias. *Journal of Public Economics*, 90(1-2), 1-36.

Barsky, R., Bound, J., Charles, K. K. and Lupton, J. P. (2002). Accounting for the black–white wealth gap: a nonparametric approach. *Journal of the American Statistical Association* 97(459), 663-673.

Becker G. (1964). *Human Capital*, 2nd edn. New York: Columbia University Press.

Becker, S.O., Fetzer, T. and Novy, D. (2017). Who voted for Brexit? A comprehensive district-level analysis. *Economic Policy*, 32(92), 601-650.

Bednarek, M. (2005). 'He's nice but Tim': Contrast in British newspaper discourse. In *Proceedings from the Corpus Linguistics Conference Series* [Accessed 23 September 2014]. Available at: http://bit. ly/1wvvGf4 Google Scholar.

Berg, G. and Zia, B. (2013). Harnessing emotional connections to improve financial decisions: Evaluating the impact of financial education in mainstream media. *World Bank Policy Research Working Paper*, (6407).

Bernhardt, D., Krasa, S. and Polborn, M. (2008). Political polarization and the electoral effects of media bias. *Journal of Public Economics*, 92(5-6), 1092-1104.

Besley, T. and Prat, A. (2006). Handcuffs for the grabbing hand? Media capture and government accountability. *American Economic Review*, 96(3), 720-736.

Bjorvatn K., Cappelen, A., Helgesson S., L., Sorensen, E., and Tungodden, B. (2015). Teaching through television: Experimental evidence of entrepreneurship education in Tanzania. NHH Dept. of Economics Discussion Paper (3).

Blackaby, D, Leslie, D, Murphy, P, and O'Leary, N. (1997). The ethnic wage gap and employment differentials in the 1990s: evidence for Britain. *Economics Letters* 58, 97-103.

Blinder, A.S. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human Resources*, 436-455.

Blinder, A.S. and Krueger, A.B. (2004). What does the public know about economic policy, and how does it know it? (No. w10787). *National Bureau of Economic Research.*

Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *The Quarterly Journal of Economics*, 118(4), 1335-1374.

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

Bun, M.J. and Harrison, T.D. (2018). OLS and IV estimation of regression models including endogenous interaction terms. *Econometric Reviews*, 1-14.

Bursztyn, L. and Cantoni, D. (2014). A Tear in the Iron Curtain: The Impact of Western Television on Consumption Behavior. *NBER Working Paper*, (w20403).

Cagé, J. (2014). Media competition, information provision and political participation. *Unpublished manuscript*, Harvard University.

Chakraborty, A. and Ghosh, P. (2016). Character endorsements and electoral competition. *American Economic Journal: Microeconomics*, 8(2), 277-310.

Chan, J. and Suen, W. (2008). A spatial theory of news consumption and electoral competition. *The Review of Economic Studies*, 75(3), 699-728.

Chiang, C.F. and Knight, B. (2011). Media bias and influence: Evidence from newspaper endorsements. *The Review of Economic Studies*, 78(3), 795-820.

Clark, K. and Drinkwater, S. (2008). The labour-market performance of recent migrants. *Oxford Review of Economic Policy*, 24(3), 495-516.

Cohen, B. C. (1963). *The press and foreign policy*. Princeton, N.J.: Princeton University Press.

Colantone, I. and Stanig, P. (2016). Global Competition and Brexit (No. 1644). BAFFI CAREFIN, Centre for Applied Research on International Markets Banking Finance and Regulation, Universita'Bocconi, Milano, Italy.

Constant, A, and Zimmerman, K. (2009). Migration, ethnicity and economic integration. *IZA Discussion Paper* 4620, 1-38.

Corduneanu-Huci, C. and Hamilton, A.J. (2018). Selective control: the political economy of censorship (No. 8556). *The World Bank Policy Research Working Paper*.

DellaVigna, S., Enikolopov, R., Mironova, V., Petrova, M. and Zhuravskaya, E. (2011). Cross-border media and nationalism: Evidence from Serbian radio in Croatia (No. w16989). *National Bureau of Economic Research.*

DellaVigna, S. and Kaplan, E. (2006). The Fox News effect: Media bias and voting (No. w12169). *National Bureau of Economic Research*.

Dobson, J. (2008). Labour Mobility and Migration within the EU following the 2004 Central and East European Enlargement. *Employee Relations* 31(02), 121-138.

Doms, M.E. and Morin, N.J. (2004). Consumer sentiment, the economy, and the news media. *FRB of San Francisco Working Paper*, (2004-09).

Drinkwater, S, Eade, J, and Garapich, M. (2009). Poles Apart? EU enlargement and the labour market outcomes of immigrants in the United Kingdom. *International Migration* 47(01), 161-190.

Drufuca, S.M. (2014). Information, media and elections: incentives for media capture. *Working Paper.*

Duggan, J. and Martinelli, C. (2011). A spatial theory of media slant and voter choice. *The Review of Economic Studies*, 78(2), 640-666.

Durante, R. and Knight, B. (2012). Partisan control, media bias, and viewer responses: Evidence from Berlusconi's Italy. *Journal of the European Economic Association*, 10(3), 451-481.

Durante, R., Pinotti, P. and Tesei, A. (2015). The Political Legacy of Entertainment TV (No. 10738). *CEPR Discussion Papers.*

Dustmann, C., and Glitz. A. (2011). Migration and education. In Hanushek, E. A., Machin, S., and Woessmann, L. (eds.). *Handbook of the Economics of Education, Volume 4*. Amsterdam: Elsevier, North-Holland, pp. 327–439.

Dustmann, C., Glitz, A. and Vogel, T. (2010). Employment, wages, and the economic cycle: Differences between immigrants and natives. *European Economic Review*, 54(1), 1-17.

Dustmann, C., Hatton, T., and Preston, I. (2005). The Labour Market Effects of Immigration. *The Economic Journal*, 115(507), F297-F299.

Dutta, P., Ravallion, M., Murgai, R., and van de Walle, D. (2013). Testing Information Constraints on India's Largest Antipoverty Program. *World Bank, Policy Research Working Paper* (No. 6598).

Eisensee, T. and Strömberg, D. (2007). News droughts, news floods, and US disaster relief. *The Quarterly Journal of Economics*, 122(2), 693-728.

Enikolopov, R., Petrova, M. and Zhuravskaya, E. (2011). Media and political persuasion: Evidence from Russia. *The American Economic Review*, 101(7), 3253-3285.

Enikolopov, R., Petrova, M. and Sonin, K. (2018). Social media and corruption. *American Economic Journal: Applied Economics*, 10(1), 150-74.

Facchini, G, Margalit, Y, Nakata, H. (2016). *Countering Public Opposition to Immigration: The Impact of Information Campaigns*. London: Centre for Economic Policy Research.

Fairlie, R.W. (2005). An extension of the Blinder-Oaxaca decomposition technique to logit and probit models. *Journal of Economic and Social Measurement*, 30(4), 305-316.

Ferraz, C. and Finan, F. (2008). Exposing corrupt politicians: the effects of Brazil's publicly released audits on electoral outcomes. *The Quarterly Journal of Economics*, 123(2), 703-745.

Fetzer, Thiemo. (2018) Did austerity cause Brexit? Working Paper. Coventry: University of Warwick. Department of Economics. Warwick economics research papers series (WERPS) (1170). (Unpublished).

Firmstone, Julie. (2016). Newspapers' editorial opinions during the referendum campaign. EU Referendum Analysis 2016: Media, Voters and the Campaign. Bournemouth: The Centre for the Study of Journalism, Culture and Community.

Firpo, S., Fortin, N. and Lemieux, T. (June 2007). Decomposing wage distributions using recentered influence function regressions. *University of British Columbia*.

Firpo, S., Fortin, N. and Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica*, *77*(3), 953-973.

Fortin, N, Lemieux, T, and Firpo, S. (2010) Decomposition Methods in Economics. *NBER Working Paper* 16045, 1-114.

Garber, S. and Bower, A.G. (1999). Newspaper coverage of automotive product liability verdicts. *Law and Society Review*, 93-122.

Gehlbach, S. and Sonin, K. (2014). Government control of the media. *Journal of Public Economics*, 118, 163-171.

Gentzkow, M. and Shapiro, J.M. (2006). Media bias and reputation. *Journal of political Economy,* 114(2), 280-316.

Gerber, A. S., Karlan, D. and Bergan, D. (2009). Does the media matter? A field experiment measuring the effect of newspapers on voting behavior and political opinions. *American Economic Journal: Applied Economics*, 1(2), 35-52.

Grigorieff, A., Roth, C. and Ubfal, D. (2016). Does information change attitudes towards immigrants? Evidence from survey experiments. IZA DP No. 10419.

Hahn, J., Ham, J. C., and Moon, H. R. (February 2011). The Hausman test and weak instruments. *Journal of Econometrics*, *Elsevier*. 160(2), 289-299.

Hausman, J., Stock, J. H. and Yogo, M. (2005). Asymptotic properties of the Hahn–Hausman test for weak-instruments. *Economics Letters*, 89(3), 333-342.

Heitmüller, A. (2005). A Note on Decompositions in Fixed Effects Models in the Presence of Time-Invariant Characteristics. *IZA Discussion Paper* 1886, 1-8.

Hobolt, S. B. (2016). The Brexit vote: a divided nation, a divided continent. *Journal of European Public Policy*, 23(9), 1259-1277.

Hunt, P. (2012). From the bottom to the top: a more complete picture of the immigrant-native wage gap in Britain. *IZA Journal of Migration*, 1(1), 9.

Jacobsen, G. D. (2011). The Al Gore effect: an inconvenient truth and voluntary carbon offsets. *Journal of Environmental Economics and Management*, 61(1), 67-78.

Jann, B. (2008). A Stata implementation of the Blinder-Oaxaca decomposition. *Stata Journal*, *8*(4), 453-479.

Kahanec, M. (2012). Labour Mobility in an Enlarged European Union. *IZA Discussion Paper* 6485, 1-48.

Kahanec, M, and Zimmerman, K. (2008). International Migration, Ethnicity, and Economic Inequality. *IZA Discussion Paper* 3450, 1-57.

Kasper, M., Kogler, C. and Kirchler, E. (2015). Tax policy and the news: An empirical analysis of taxpayers' perceptions of tax-related media coverage and its impact on tax compliance. *Journal of Behavioral and Experimental Economics*, 54, 58-63.

Kearney, M. S. and Levine, P. B. (2015). Early childhood education by MOOC: Lessons from Sesame Street (No. w21229). *National Bureau of Economic Research*.

Kosse, A. (2013). Do newspaper articles on card fraud affect debit card usage? *Journal of Banking & Finance*, 37(12), 5382-5391.

Lamla, M. J. and Lein, S. M. (2008). The role of media for consumers' inflation expectation formation. *KOF Swiss Economic Institute*, Working Paper, (201).

Larreguy, H. A., Marshall, J. and Snyder Jr, J. M. (2014). Revealing malfeasance: How local media facilitates electoral sanctioning of mayors in Mexico (No. w20697). *National Bureau of Economic Research.*

Lemos, S. (2017). Mind the gap: A detailed picture of the immigrant-native earnings gap in the UK using longitudinal data between 1978 and 2006. *Regional Science and Urban Economics*, 63, 57-75.

Longhi, S, Nicoletti, C, and Platt, L. (2013). Explained and Unexplained Wage Gaps across the Main Ethno-Religious Groups in Great Britain. *Oxford Economic Papers*. 65, 471–493.

Longhi, S, and Rokicka, M. (2012). European Immigrants in the UK Before and After the 2004 Enlargement: Is there a Change in Immigrant Self-Selection? *Norface Migration Discussion Paper* 30, 1-29.

Manacorda, M., Manning, A. and Wadsworth, J. (2012). The impact of immigration on the structure of wages: theory and evidence from Britain. *Journal of the European Economic Association*, 10(1), 120-151.

McAuliffe, M., Weeks, W. and Koser, K. (2015). Media and migration: Comparative analysis of print and online media reporting on migrants and migration in selected origin and destination countries.in McAuliffe. M and Koser, K (Eds.), *A long way to go: Irregular migration patterns, processes, drivers and decision making.* (p. 277-315) ANU Press, Canberra.

McGuigan, M., McNally, S. and Wyness, G. (2012). Student Awareness of Costs and Benefits of Educational Decisions: Effects of an Information Campaign. CEE DP 139. *Centre for the Economics of Education*.

Moore, M. and Ramsay, G. (2017). *UK media Coverage of the 2016 EU Referendum campaign*. Centre for the Study of Media, Communication and Power, King's College London.

Morton, R.B. and Tyran, J.R. (2011). Let the experts decide? Asymmetric information, abstention, and coordination in standing committees. *Games and Economic Behaviour*, 72(2), 485-509.

Mullainathan, S. and Shleifer, A. (2005). The market for news. *American Economic Review*, 95(4), 1031-1053.

Murphy, R. (2014). Trade Unions in the Age of Litigation.

Nagelhout, G. E., van den Putte, B. de Vries, H., Crone, M., Fong, G. T. and Willemsen, M. C. (2011). The influence of newspaper coverage and a media campaign on smokers' support for smoke-free bars and restaurants and on second hand smoke harm awareness: findings from the International Tobacco Control (ITC) Netherlands Survey. *Tobacco control*, tc-2010.

Nandi, A. and Nicoletti, C. (2014). Explaining personality pay gaps in the UK. *Applied Economics* 46(26), 3131-3150.

Nickell, S. and Saleheen, J. (2017). The impact of EU and Non-EU immigration on British wages. *IZA Journal of Development and Migration* 7(1), 1–28.

Niederdeppe, J., Fowler, E. F., Goldstein, K. and Pribble, J. (2010). Does local television news coverage cultivate fatalistic beliefs about cancer prevention? *Journal of Communication*, 60(2), 230-253.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review,* 693-709.

Oliveros, S. and Várdy, F. (2015). Demand for slant: how abstention shapes voters' choice of news media. *The Economic Journal*, 125(587), 1327-1368.

Olken, B. A. (2009). Do television and radio destroy social capital? Evidence from Indonesian villages. *American Economic Journal: Applied Economics*, 1(4), 1-33.

Ottaviano, G. I. and Peri, G. (2012). Rethinking the effect of immigration on wages. *Journal of the European Economic Association*, 10(1), 152-197.

Prat, A. and Strömberg, D. (2013). The Political Economy of Mass Media. In D. Acemoglu, M. Arellano, & E. Dekel (Eds.) *Advances in Economics and Econometrics: Volume 2, Applied Economics: Tenth World Congress* (Vol. 50, p. 135-187). Cambridge University Press.

Rabin, M. and Schrag, J. L. (1999). First impressions matter: A model of confirmatory bias. *The Quarterly Journal of Economics*, 114(1), 37-82.

Rosso, A. (2013). Skill premia and immigrant-native wage gaps. LLAKES Research Paper, 45.

Rudiger, J. (2013). Cross-checking the media. *Working Paper European University Institute Florence.*

Saleheen, J. and Shadforth, C. (2006). The economic characteristics of immigrants and their impact on supply. *Bank of England Quarterly Bulletin, Winter*.

Skupnik, C. (2014). EU enlargement and the race to the bottom of welfare states. *IZA Journal of Migration*, 3(1), 1-2.

Strömberg, D. (2015). Media and Politics. *Annual Review of Economics*, 7(1), 173-205.

Trombetta, F. and Rossignoli, D., 2020. The price of silence: Media competition, capture, and electoral accountability. *European Journal of Political Economy, p.101939.*

Tversky, A. and Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5(2), 207-232.

van der Wiel, K. (2009). Have You Heard the News? How Real-Life Expectations React to Publicity. *IZA DP* 4064.

Williams, P. and Dickinson, J. (1993). Fear of crime: Read all about it? The relationship between newspaper crime reporting and fear of crime. *British Journal of Criminology*, 33(1), 33-56.

Yanagizawa-Drott, D. (2014). Propaganda and conflict: Evidence from the Rwandan genocide. *The Quarterly Journal of Economics*, 129(4), 1947-1994.

Zavodny, M. (2006). Does watching television rot your mind? Estimates of the effect on test scores. *Economics of Education Review*, 25(5), 565-573.