ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics
Master Thesis [International Economics]

Title thesis: EU Immigration and the English Labour Market: A Shift-Share Instrument Analysis
Name student: Louis Emyr Williams
Student ID number: 612569

Supervisor: Prof. Laura Hering
Second assessor: Prof. Olivier Marie

Date final version: 11/08/2022

EU Immigration and the English Labour Market: A Shift-Share Instrument analysis

Louis Emyr Williams

August 11, 2022

Abstract

In this study, shift-share instruments are exploited in order to counteract the endogeneity bias induced by EU immigrants selecting into locations with higher wages. Shift-share instruments exploit national inflows of EU migrants by interacting them with the previous settlement location of EU migrants. Using regional and time-specific effects and two so-called *Past Settlement Instruments*, the short (2 year) and long (5 year) term impacts of EU migration on wages at the municipality level in England are examined. The study provides evidence for the long-term impact of migration having a small, positive impact on wages. In the short-term, the impact of migration on wages is found to be negative, albeit lacking statistical significance. The multiple instrumentation approach allows for the conflating short and long run impacts of migration to be "untangled" and estimated uniquely. In a tumultuous time in the English labour market - after the UK's exit from the European Union (EU) in 2020 - this paper offers a cautionary view into the short and long-run impacts of EU migration on wages.

Contents

1 Introduction **4**

2 Literature Review **7**

2.1 Endogeneity Bias and "Shift-Share" Instruments 8

2.2 Untangling the Short and Long term Effects of Immigration 10

2.3 Migration and the English Labour market 13

2.4 Contribution to Existing Literature 14

3 Data **15**

3.1 Source and Content 15

3.2 Sample selection 16

3.3 Descriptive Statistics 16

4 Methodology **18**

4.1 Method & Reasoning 18

4.2 Variables 20

4.2.1 Dependent Variable 20

4.2.2 Independent Variables 20

4.2.3 Instrumental Variables 21

5 Results & Discussion **24**

5.1 Main Results 24

5.2 Robustness Checks 26

5.3 Study Limitations 28

6 Conclusion & Discussion **30**

7 Appendix **35**

7.1 Appendix A 35

7.2 Appendix B 35

7.3 Appendix C 36

7.4 Appendix D 37

7.5 Appendix E 38

1 Introduction

Perhaps the most salient reason for the United Kingdom voting to leave the European Union (EU) was immigration. The UK was one of the 12 original signatories of the Maastricht Treaty in 1992 - the beginning of the EU in its current form - allowing for the free movement of persons between the 12 member states. However, it was not until the A8 countries joined the EU in 2004 that immigration became a prominent issue in British politics. The UK, along with its neighbour Ireland and Sweden, were initially the only EU members to allow unrestricted migration from these countries (Lemos and Portes, 2008; Vargas-Silva, 2014). Due to the ease of entry and the wage differential between the UK and the A8 countries, it became a prime location for working age EU migrants from 2004 onwards. This study will focus on the impact of EU migration on wages in England, leading to the following research questions being formulated.

1. Does an increase in the rate of migration lead to an increase of the median wage within a municipality in the short-term?

2. Does an increase in the rate of migration lead to an increase of the median wage within a municipality in the long-term?

Figure 1 conveys the EU-born working age population (16-64) within England between 2004-2020. From the lowest point in 2004 to the peak in 2017, the working age EU population increased by a factor of over 2.5. The dramatic increase in EU immigrants moving to the UK/England in recent years has motivated academics to examine the impact on the labour market. Furthermore, with the exit of the UK from the EU (Brexit), a further avenue of analysis has been opened.

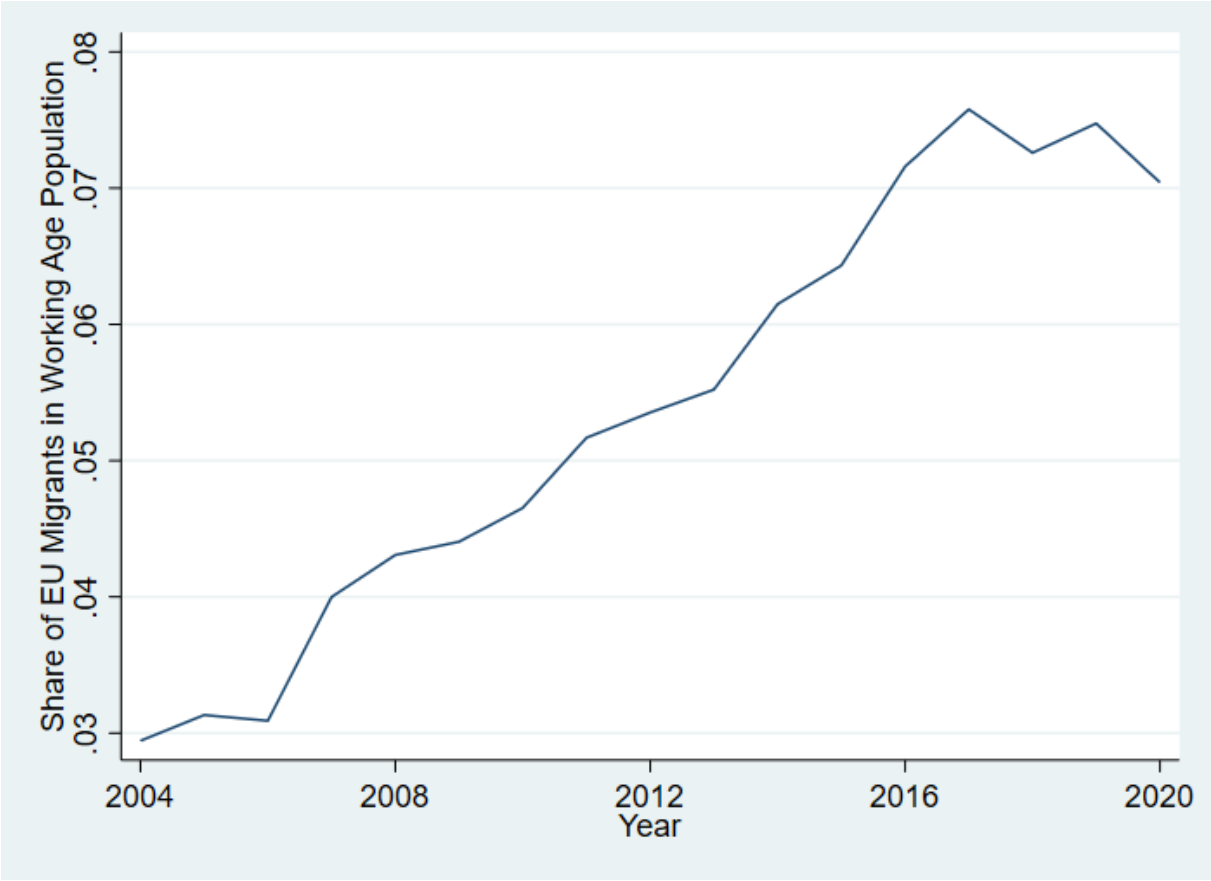

Figure 1: Share of EU migrants in total working age population in England

Evaluating the effects of migration on the labour market has critical implications for policymakers. In the case of the UK, the free movement of persons within the bloc was deeply unpopular for a large proportion of the electorate. As a result, policymakers must appropriately weigh up the positives and negatives of immigration. In the short-run, negative impacts can arise (for native workers) due to an increased labour supply. With migrant workers willing to work at wages that are higher compared to their home country, but lower than the wage in the destination country, wages can stagnate or fall. On the other hand, migrants can have positive medium and long-term effects for the host country. For example, migrants from the A8 countries made positive contributions to the UK's fiscal budget in four of the first five years after entering the EU, even whilst the UK was running a budget deficit in each of those years. Though many EU migrants enter low-skilled (low wage) jobs, they have high labour market participation and employment rates, resulting in a net positive fiscal contribution (Vargas-Silva, 2014). Moreover, migrants further contribute to aggregate demand through additional consumption. Migrants require goods and services, requiring the creation of additional jobs to meet the extra demand. However, an often-neglected impact of migration on the host country is the supplementary demand for public goods. Public transport, for example, will need to be maintained on a more frequent basis, incurring extra costs. Nonetheless, this study

will focus on the impacts of migration on the English labour market.

Questions related to the understanding of migration and its impacts are not confined to policymakers - academics have been confronted by the issue of causal inference in migration literature for decades. In the instance of evaluating the effect of migration on wages, it can be argued they affect each other simultaneously, meaning OLS estimates will be affected by endogeneity bias (Wooldridge, 2002). For the sake of the argument, an increase in the rate of migration can decrease wages, due to a larger labour supply. On the other hand, migrants potentially select into areas that have the greatest wage compared to their home country.

To counteract this issue, previous literature has exploited "shift-share" instruments (otherwise known as "Bartik Instruments") to isolate the effect of immigration on labour markets. First coined by Bartik (1991), local level industry shares were interacted with national level industry growth rates to measure common exposure to labour supply shocks (Goldsmith-Pinkham et al., 2020). Blanchard and Katz (1992) introduced the so-called "Bartik instruments" into migration literature shortly after, examining the effect of employment shocks (such as an influx of immigrants to a specific region) on labour market outcomes in US states between 1950-1990.

Recent literature has used a dynamic spatial panel approach (Fingleton and Szumilo, 2019; Jaeger et al., 2018), allowing for the impacts of migration on labour market outcomes to be analysed at a more localised level. By combining this with fixed effects estimation, which accounts for time-invariant municipality effects, a comprehensive analysis of the effect of immigration on labour market outcomes in England can be conducted. Jaeger et al. (2018), takes the paper by Card (2009) a step further by implementing multiple past settlement instruments to decompose the short and long-run effects of immigration on wages. This step forward in the literature will be utilised in the forthcoming analysis.

Hereby, this paper aims to complement existing literature on the impact of EU migration on wages in England. For this purpose, a cross-area analysis is conducted on a total of 89 areas. For simplification purposes, these areas will hereafter be referred to as municipalities. An explanation on the selection of the municipalities will be made in the Data section (section 3). Following the introduction, section 2 will present the theoretical framework in the form of a literature review, in which the theoretical evidence supporting the subject matter will be discussed. The data used for the empirical analysis will be discussed in section 3, alongside a table of descriptive statistics. The methodology chosen will be conveyed in section 4, with the associated results presented and discussed in section 5. Finally, the concluding statements will be made section 6.

2 Literature Review

Immigration is defined as the action of a person moving to another area, whether that is locally, regionally or on an international basis. In the context of this study, immigration refers to the entry of EU citizens into England. Positive net migration (immigration minus emigration) has been the main driver of population growth in England since the 1990s (ONS, 2021). With the A8 countries becoming members of the EU in 2004, and the UK allowing unrestricted entry of migrants from these countries the same year, the contribution of migration to England's population growth has only increased. Hereby, this has resulted in a vast array of literature examining the impacts of positive net migration on the UK. Dustmann et al. (2013) examined the impacts of EU migrants on the fiscal deficit in the UK, finding a small, positive net contribution. Other studies, such as those released by the Migration Observatory, focus on the native population's attitudes to immigration. They found that in recent years since the referendum on the UK's membership to the EU in 2016, attitudes to immigration from the had softened (Migration Observatory, 2020). However, this study will examine the impact of immigration on labour market outcomes, with a focus on the impact of immigration on wages at the municipality level.

Studies including those by De Silva et al. (2010) use natural experiments in order to find the impacts of migration on wages. After Hurricane Katrina in 2005, many evacuees arrived in the Houston Metropolitan Area, with the population increasing by 3.1% between November 2005 and August 2006 (McIntosh, 2008). Most of the evacuees were low-skilled and less educated relative to the rest of the Houston population. De Silva et al. (2010) utilise Differences-in-Difference-in-Differences (DDD) methodology, comparing the trend in wages of workers in Dallas, which did not have a labour supply shock comparable to Houston's in the same time period. In Houston, the authors find that the relative wages of workers in the low-skilled bracket compared to the high-skilled wage bracket dropped by 0.7% in comparison to Dallas. By using Dallas as a comparison city, it allows for broader regional differences to be controlled for, as well as the impacts of immigration on the demand side. Immigrants also demand goods, which can help to stimulate the economy and increase wages in the medium/long run. Ideally, the following study would also utilise a natural experiment. However, as the focus is on the English labour market and there has not been a labour supply shock specific to one municipality, the study is unable to do this. Thereafter, solutions to this will be discussed in the subsequent sections of the literature review.

Initially, this section will focus on the central issue of endogeneity bias and how literature has countered this with shift-share instruments. After, there will be a focus on the study by Jaeger et al. (2018), one of the most influential papers in migration literature in

the past decade. Thereafter, literature examining the English/British labour market will be reviewed, culminating in the creation of this study's hypothesis.

2.1 Endogeneity Bias and "Shift-Share" Instruments

As discussed in the introduction, the most frequent identification issue arises due to endogeneity bias. Endogeneity bias emerges in situations in which the explanatory variable is correlated with the structural error term (Wooldridge, 2002). If the explanatory variable correlates with the error term, the effect of the unobserved factors in the error term will also impact the estimated parameter of the explanatory variable (Antonakis et al., 2010). In the case of examining the effects of immigration, it can be argued that wages and the rate of immigration affect each other simultaneously. For example, a migrant is more likely to move to an area with higher wages, to maximise the wage differential between the wages they receive in their host and home country (Jaeger et al., 2018). Contrastingly, immigrants can also impact the wages within an area through a plethora of channels. In the short-run, immigrants can place downward pressure on wages (in industries where they are substitutes for native workers) as they increase the labour supply. In the longer term, immigrants can help grow productivity, by encouraging specialisation or complementing the existing skill set of native workers. In time, an increase in productivity should lead to wage growth.

Bartik's (1991) book is widely considered an important contribution to a range of topics in Economics, such as public, finance, and macroeconomics. Migration literature is among those which Bartik's work has been essential for (Goldsmith-Pinkham et al., 2020), with Jaeger et al. (2018) stating in their study "few literatures rely so heavily on a single instrument or variants thereof". Within the study, Bartik analyses the impact of business growth on local-level labour market outcomes (including wages, employment and labour force participation). However, both business growth at the local level and labour market outcomes can simultaneously affect each other. For instance, it is viable that positive business growth has a positive effect on wages. On the other hand, if a local area experiences an increase in business growth (perhaps caused by lower corporation tax), wages may also change simultaneously, resulting in endogeneity bias. As a solution, Bartik proposes a "shift-share" instrument, which interacts local industry shares (calculated at an initial base period where t=0) with national-level industry growth rates. Bartik argues that the impact of the "shift-share" instrument on business growth is exogenous, because the variation within the instrument comes from the alteration in national-level industry growth rates. The impact of national-level industry growth rates are then weighted by the local-level industry shares (at basis time $t=0$). One could argue that the exclusion restriction is violated, as Bartik makes the strong assumption that national-level indus-

try growth rates have no impact on local-level wages, unemployment and labour force participation. That being said, Bartik's work had a profound impact on many areas in economics literature, let alone on migration literature.

Blanchard and Katz (1992) provide the first instance of shift-share instruments within labour economics literature. Comparing US states between 1952-1990, the authors examine the impact of labour demand shocks on the level of unemployment and wages, with internal migration between US states assumed to be the adjustment mechanism. Again, the impact of labour demand shocks on these outcomes is potentially endogenous. Nevertheless, they discover that a labour demand shock has positive effect on wages and a negative impact on unemployment. They decide to check the robustness of their findings by using a shift-share instrument. Like Bartik (1991), Blanchard and Katz (1992) utilise employment growth rates on a national scale, weighted with state industry employment shares (from the previous year) to characterise labour demand shocks. Industries are defined at the two-digit level. Blanchard and Katz (1992) argue their instrument is valid, as state-specific labour demand shocks are feasibly uncorrelated with national industry-level employment rates.

David Card has been one of the most significant contributors to migration literature in the past decades since Bartik's groundbreaking work in 1991. Beforehand, Altonji and Card (1991) conducted a comparison between native's labour market outcomes in 120 US cities in 1970 and 1980 using census data. By comparing the difference in the share of immigrants in each time period, the authors can correlate the impact that the change in immigrants has had on natives. In order to account for city-specific effects that may have occurred between 1970 and 1980, the authors use first-differences. As there are only two time periods (1970 and 1980), this is equivalent to using fixed effects, however, the issue of endogeneity remains. To counteract this, Altonji and Card (1991) propose an instrument inspired by the paper by Bartel (1989). She finds that post-1964 immigrants to the US cluster into areas where there already exists a sizeable immigrant population. Altonji and Card (1991) exploit this fact by using past immigrant shares as an instrument for the current share of immigrants. Feasibly, the share of immigrants in an area 10 years previously has no impact on wages in the current period. Furthermore, the share of immigrants 10 years previously has an impact on the current share of immigrants - the exclusion restriction is satisfied. Using the instrumental variable analysis, the authors find that a 1% increase in the fraction of immigrants leads to a 1.2% decrease in the wages for less skilled natives. On the contrary, the study is limited by the fact it encompasses only two time periods, so the findings should be approached with caution. The combination of the papers by Altonji and Card (1991) and Bartik (1991) paved the way for the past settlement shift-share instrument to be applied to migration literature.

Thereafter was a lull in the use of the past settlement instrument in migration literature, until the paper by Card and DiNardo (2000) spurred a flurry of activity. Further literature, such as the papers by Ottaviano and Peri (2005) and Kugler and Yuksel (2008) also exploited the past settlement instrument constructed by Altonji and Card (1991). Building on this, the paper by Card (2009) "Immigration and Inequality" - which created a shift-share instrument using the past settlement of immigrants - contributed heavily to migration literature of the past decade. Card (2009) creates a past settlement instrument exploiting variation in national inflows of migrants within a yearly time period:

$$(\sum_{m} \lambda_{mj} M_m)/P_j = \sum_{m} [N_{mj}/P_j] M_m/N_m \tag{1}$$

Where λ_{mj} is the share of migrants from country m who lived in city j at a previous date, M_m is the number of migrants arriving in the US and P_j is the total population in city j. N_{mj} represents the new arrivals of immigrants into city j. The instrument is constructed by predicting where new immigrants from country m will move to based off the number of migrants from country m that are already settled in city j. A weighted average of national-level inflows from country m is combined with a weighting that depends on the proportion of the total migrants from country m (national-level) that are settled in city j. Card argues that national inflow rates of country m migrants are exogenous to specific city labour market conditions, therefore the predicted inflows into city j will also be exogenous. This can be argued against, if migrants from country m are overwhelmingly based in city j and city j experiences an increase in wages, then this may spark an increase in national inflows from country m, meaning the instrument does not exhibit exogeneity. Although there are inevitably differences in opinion over the validity of the past settlement instrument to identify labour market effects, it remains the most accepted method of identification within migration literature.

2.2 Untangling the Short and Long term Effects of Immigration

With past settlement instruments playing a core role in migration literature in the past two decades, papers such as those by Jaeger et al. (2018) have attempted to improve the identification procedure. Literature by Card (2009), Ottaviano et al. (2013) and Dustmann et al. (2013) all exploited the shift-share instrument procedure coined by Bartik (1991). However, Jaeger et al. (2018) argues that the impact of immigration can impact labour market outcomes in contrasting ways in the short and long-run. If a single instrument is used, the short and long-run effects are conflated into the estimate, making causal interpretation near impossible. The authors exploit US census data on a decadal basis between 1960 and 2011. It is argued that large-scale immigration can reduce wages in a region in the short-run, in the form of a labour supply shock. A labour supply shock can

be interpreted as reducing the bargaining power of native workers, therefore wages fall. In the case of the UK, a vocal proportion of the population argued that low-skilled migrants from the EU were "taking" jobs that natives could do and pushed down real wages Travis (2016). However, in the long-run, Jaeger et al. (2018) argue immigration can have positive impacts on wages, as the labour market adjusts to a new equilibrium. To account for these opposing flows, a multiple instrumentation approach is proposed. Instead of only using current immigrant inflows to construct the instrument, a further instrument is constructed using past immigrant inflows. It has been discovered that during the 1970s that there is a negative impact on wages from positive current inflows of migrants (those who immigrated during the 1970s), with estimates having a greater magnitude of negativity than previous literature. Using the past inflow of immigrants (from the 1960s), they observe a positive impact on wages in the 1970s. The migrants from the past decade have contributed positively to wages in the subsequent decade. To summarise, a positive labour supply shock has a temporarily negative effect on wages. As the economy adjusts to a new general equilibrium, the effect of past migrant inflows on wages becomes positive.

The strength of the identification strategy proposed by Jaeger et al. (2018) depends heavily on assumptions and the data being utilised. Firstly, the data needs to exhibit sufficient variation in national inflows from one period to the next. If not, the instruments will be highly correlated with one another, as shift-share instruments rely on changes to national inflows for the variations within the instrument. This is a potential pitfall of this study, which will be discussed in the limitations section. A further robustness check will be carried out with only one past settlement instrument, to test the impact of the instruments being correlated with one another. In order to utilise a lagged past settlement instrument, it needs to be lagged by the correct number of time periods, otherwise the true long-term effects will not be captured. Therefore, assuming the correct length of time of adjustment is key. Spatial adjustments in the US are estimated to take around a decade (Blanchard and Katz, 1992; **?**). In the case of the analysis by Jaeger et al. (2018), the instrument exploiting past immigrant inflows are lagged by a time period of 1, due to the decadal nature of the data. Cohen-Goldner and Paserman (2011) find in Israel that in the short-run there is a negative impact of high-skilled migrants on wages, but that the effects wear off within 6-7 years. The decade long adjustment period potentially applies only to larger countries such as the US, where there are high fixed costs associated with internal migration. Thereafter, Jaeger et al. (2018) formulate the following first-stage regressions:

$$m_{jt} = \pi_{10} + \alpha_{11}\widetilde{m}_{jt} + \alpha_{12}\widetilde{m}_{jt-1} + u_{jt} \tag{2}$$

$$m_{jt-1} = \pi_{20} + \alpha_{21}\widetilde{m}_{jt} + \alpha_{22}\widetilde{m}_{jt-1} + v_{jt} \tag{3}$$

Where $\widetilde{m}_{jt}$ is the past settlement instrument for current immigrant inflows and $\widetilde{m}_{jt-1}$ is the past settlement instrument for the previous period's immigration inflows. The past settlement is constructed using the methodology from the paper by Card (2009), discussed in section 2.2. The instruments are used to construct estimates for immigrant inflows to area j in in times t and *t-1*. The second stage equation estimated by Jaeger et al. (2018) is thereafter as follows:

$$\Delta lnw_{jt} = \beta_0 + \beta_1 m_{jt} + \beta_2 m_{jt-1} + _{jt} \tag{4}$$

Where β_1 captures the short-term and β_2 captures long-term impacts of immigration inflows on the rate of change in wages.

Although arguably a step forward in migration literature, there are some caveats to the author's data and methodology. As mentioned above, the study relies on US census data, with observations taken every 10 years. Measuring inflows of migration on this infrequent basis can produce limitations. For example, temporary migrants that were present in the labour market between 1971-1979 would not have been measured, despite spending a substantial amount of time in the labour market. Nevertheless, these migrants impacted the wages of regions they were in within the short-term. These omitted migrants would have negatively affected wages in the framework of Jaeger et al. (2018), and would have been correlated positively with the measured immigrant inflows. As a result, the estimate for β_1 is potentially negatively biased, meaning the true value for the estimate should be less negative. Furthermore, the methodology relies on sufficient variation in national inflows. With large-scale migration to the US taking a hold in the 1970s, there was sufficient variation between national inflows between the 1960s and 1970s. However, the estimates lose significance when comparing decades after that, because national inflow immigration rates had stabilised (the instruments become weak due to high serial correlation). In order to implement the methodology for the English labour market, adjustments need to be made to the model, which are specified in section 4 (Methodology).

A potential caveat for this study arises from the annual nature of the dataset. It arises from the reduced form equation, which estimates the impact of current and past immigration inflows on the rate of change in wages. In the context of the paper's by Jaeger et al. (2018) and Card (2009), the decadal nature of the data allows a long enough adjustment period so that the impacts of immigration inflows in the current 10 year period have impacted the median wage. Within this study where the time period is one year, this does not allow enough time for the migration flows to have an impact on the median wage. To overcome this, the inflows will be lagged by 2 years so that migration flows have had time to impact the wages within an area.

2.3 Migration and the English Labour market

Research by Dustmann et al. (2013) is unique in the context of the English labour market, applying the work of Card (2009) to examine the impacts of immigration along the wage distribution. By organising the workforce into different skill groups characterised by their level of education, Dustmann and co-authors find that the impacts of migration affect natives asymmetrically. It builds on the paper by Dustmann et al. (2008), where the authors observe that migrants moving to the UK fill in jobs that they are overqualified for. 46% of new immigrants have an education up to the age of 21+, compared to only 16% of natives. However, 26% of newly arrived immigrants filled in "routine" and "semi-routine" jobs, categorised by Dustmann et al. (2008) as the lowest paid (skilled) occupation categories. In other words, migrants are working in jobs below their observable skill level. As a result, Dustmann et al. (2013) propose that migrants should be categorised by where they place among the wage distribution, due to migrant's tendency to downgrade on arrival. The authors find that new migrants depress wages below the 20th percentile of the wage distribution, but increase wages in the upper part of the wage distribution. In the context the study by Jaeger et al. (2018), migration may lead to a depression of wages in the short-run as the low-skilled labour market plausibly adjusts faster to labour supply shocks. A potential explanation for this is the short-term nature of numerous low-skilled occupations; low-skilled workers are often employed on temporary or zero-hour contracts. This implies that native workers can be released from their jobs in a simple manner, being replaced by migrants who are willing to work for lower wages. This acts as a potential explanation of the short-term negative impacts of migration in a single-good economy, where labour is assumed to have the same skill level.

Dustmann et al. (2013) find a positive effect of migration on wages in the upper percentiles of the wage distribution. Workers in the upper echelon of the wage distribution are more likely to be employed on permanent, open-ended contracts ((Rouvroye et al., 2021)), meaning they have more protection from a labour supply shock and a depression of wages in the short-run. In the long-run (again in the context of the paper by Jaeger et al. (2018)), immigrants can counteract the slow response of the native labour supply to economic conditions. Borjas (2001) argues that immigrants "grease the wheels of the labour market", as they are more responsive to regional differences in economic opportunities. Therefore, migrants help to create efficiency gains in the form of increased wages for immigrants (and a higher mean wage for the total population). The study by Borjas (2001) focuses on the mobility of native and migrant labour within the US, and how this induces wage growth. Arguably, EU migrants in the UK are more mobile than migrants in the US, due to the smaller size of the country and lower costs of internal migration. Gregg et al. (2004) however, find that labour mobility amongst low-skilled workers in

the UK is low, due to housing market constraints. On the other hand, as EU migrants are relatively more skilled compared to native workers, migrants are perhaps more mobile than native workers in low-skilled occupations.

2.4 Contribution to Existing Literature

By utilising other papers within the literature, this study is able to uniquely contribute. Firstly, the study by Bartik (1991) used shift-share instruments to examine the impact of business growth on labour market outcomes. By using national industry growth rates, which were plausibly unrelated to business growth rates at the local level, Bartik argued that he had a solution to the endogeneity problem. The study by Card (2009) finds small, positive effects of immigration on wages, exploiting a past settlement instrument that had been first adopted by Altonji and Card (1991)). Wages and immigration plausibly affect each other simultaneously, making causal inference difficult. By combining the past settlement instrument and the shift-share element of Bartik's (1991) study, it allows a more compelling argument to be made for exogeneity. National inflows of migrants has an impact on the inflows of migrants to a local area, whilst plausibly not affecting the wage rate in that area - satisfying the exclusion restriction. Finally, Jaeger et al. (2018) understood the short and long-run impacts of migration on wages could run in different directions - with negative effects in the short-run and positive in the longer term.

With this in mind, this study will examine the impacts of EU migration on the English labour market since the accession of the A8 countries to the EU in 2004, which allowed for unrestricted migration into the UK. Contrastingly to previous literature, the study will focus on the impacts of migration on wages at the municipality level over 17 years (2004-2020) on a yearly basis. In addition, fixed effects will be used in order to control for year-specific and municipality time-invariant effects. Furthermore, using two past settlement instruments, one constructed using current immigrant flows and another using past immigrant flows, the short and long-run impacts will be analysed. It is argued that the mechanism for the short and long-run impacts of migration on wages holds for the English labour market, thus the following hypotheses are formulated:

1. An increase in EU migrant inflows to a municipality decreases the median weekly wage in the short-run.

2. An increase in EU migrant inflows to a municipality increases the median weekly wage in the short-run.

3 Data

3.1 Source and Content

The study uses public data compiled by the Office for National Statistics (ONS). Specifically, this paper exploits data provided by the ONS at the local level. The database provides statistics on over 300 districts in England, which were then aggregated into larger municipalities. An ideal alternative dataset would be the decadal censuses, the last occurring in 2021. These withhold precise data on social and economic indicators in a local area, making them an ideal data source for this type of analysis. On the other hand, with there only being three censuses (2001, 2011 and 2021) since the EU was formed, the sample size of the data set will be too small.

Firstly, median wage data was compiled from the Annual Survey for Hours and Earnings. 1% of workers in each municipality (excluding self-employed workers) were surveyed on a yearly basis to find the yearly median wage in the area. Their yearly median wages were then divided by 52 to find the median weekly earnings. Median wage is preferred as a measurement of weekly wage over the mean wage, because it is not affected as much by outliers. The mean wage does however offer a suitable robustness check.

An EU migrant is defined as an EU-born UK resident. EU migrant flows were calculated by subtracting the previous periods number of (working age) EU migrants from the number in the current period. Alternatively, there exists data on immigrant inflows for each municipality. However, this counts total migrants, and does not decompose the data into more detailed sub-groups. Therefore, this second-best data collection method must be used. Furthermore, there are limitations to the length of time the data is available for. Ideally, data on migrant inflows would be available pre-2004, before the large-scale migration prompted by the A8 countries acceding to the EU. However, data on migrant inflows runs post-2011 only, placing restrictions on the sample size and undermining the internal and external validity of findings (Faber and Fonseca, 2014). Therefore, the method of subtracting the previous period's (working age) EU migrant population from the current EU migrant population is utilised. The ONS estimates population size to the nearest thousand, so in smaller areas there may exist the issue of a lack of variation. As a robustness check, municipalities with less than 3000 working age EU migrants in the initial time period will be removed. In areas where there are many municipalities in a small area (such as London, which is split into 32 boroughs), they are clustered together. This is to account for workers who may live in one borough, but receive their wages by working in another. Thereafter, London is split into two areas; Inner and Outer London, with the latter consisting of 13 of the 32 boroughs.

3.2 Sample selection

The 89 "municipalities" (as they will hereafter be referred to) are made up of non-metropolitan county councils, unitary authorities (UAs), metropolitan counties and Inner/Outer London. Up until the 1990s, most of England was governed at a local level by county councils in two-tiered system. The lower, more localised tiers are known as district councils. County councils are responsible for education, waste disposal and social services, whereas district council are responsible for waste collection, leisure and recreation. UAs, the first of which appeared in 1994, are responsible for all of these services, creating a one-tiered system of local governance. Furthermore, there are six metropolitan counties, including the West Midlands (which includes Birmingham and surrounding cities) and Greater Manchester. The municipalities were selected because they exhibit enough variation in migration flows to be able to examine the effect of it on wages. None of the municipalities overlap each other.

Some municipalities have had their boundaries altered over time, with Bournemouth and Poole remaining as separate unitary authorities until 2019, when they combined into one, larger UA (named Bournemouth, Christchurch and Poole). Moreover, Cheshire governed as a county council until 2009, thereafter it was split into two separate UAs: Cheshire East, and Cheshire West and Chester. If changes have been made to the municipality during the sample's time frame, the municipality that exists for the longest is selected for the sample. Therefore, Bournemouth, Poole, Cheshire East, and Cheshire West and Chester are included in the sample rather than Cheshire and Bournemouth, Poole and Christchurch.

The impacts of migration on the English labour market is not only absorbing from an empirical perspective (due to the mass EU migration since 2004), but also from the viewpoint of Brexit. England voted to leave the EU with a share of 53.4%, the largest of any of the UK's constituent countries. With immigration a key issue of the Leave campaign, the sample years (2004-2020) were selected to investigate whether the concerns of native workers were justified in terms of negative labour market outcomes. Limitations arise from the ONS having no public availability on weekly median wages before 2004, which would have allowed for further analysis to be undertaken on the impacts of migration from the A8 countries on the English labour market.

3.3 Descriptive Statistics

Table 1 conveys descriptive statistics for the variables being utilised in this study. The main dependent variable (ln(Median Wage)) has 1480 observations. On the grounds of robustness, missing observations for the variables of interest were dropped (*ln(Median*

Wage), *Current Migration Flow* and *Past Migration Flow*) in order to create a balanced panel data set.

Table 1: Descriptive Statistics

Variable	Obs.	Mean	Std. Dev	Min	Max
Year	1513	-	-	2004	2020
Median Wage (GBP)	1480	406.731	64.565	251	638.2
Mean Wage (GBP)	1489	480.038	83.566	308.2	848.2
ln(Median Wage)	1480	5.996	0.156	5.525	6.459
ln(Mean Wage)	1489	6.159	0.169	5.731	6.743
Total (working) Population (000s)	1513	376.7	501.1	19	3496
EU (working) Population (000s)	1501	20.9	50.7	1	487
Non-EU (working) Population (000s)	1504	43.8	132.4	1	1100
UK (working) Population (000s)	1513	312.2	351.0	17	2064
Total EU Population in UK	1513	1837.2	565.6	945	2627
Current Migration Flow	1319	0.0032	0.011	-0.041	0.057
Past Migration Flow	1055	0.0036	0.010	-0.039	0.056
Ln(Current Migration Flow)	1319	-4.720	0.750	-7.703	-2.871
Ln(Past Migration Flow)	1055	-4.780	0.745	-7.703	-2.878
Past Settlement Inst 1	1335	0.0025	0.0029	-0.0008	0.028
Past Settlement Inst 2	979	0.0029	0.0029	-0.0003	0.031

Note: Values that are not calculated due to the categorical nature of the variable are represented by a slash (-). The monetary variables; Median Wage and Mean Wage are measured in Pound Sterling per week. The working age populations (Total, EU, Non-EU and UK) are estimated to the nearest thousand. Current and Past Migration flows are calculated as the growth in the EU migrant population divided by the total population in the prior period. The Past Settlement Instruments are calculated using the methodology of Jaeger et al. (2018) (refer section 2.3). More information on these variables will be given in the Methodology section (section 4).
Source: Office for National Statistics

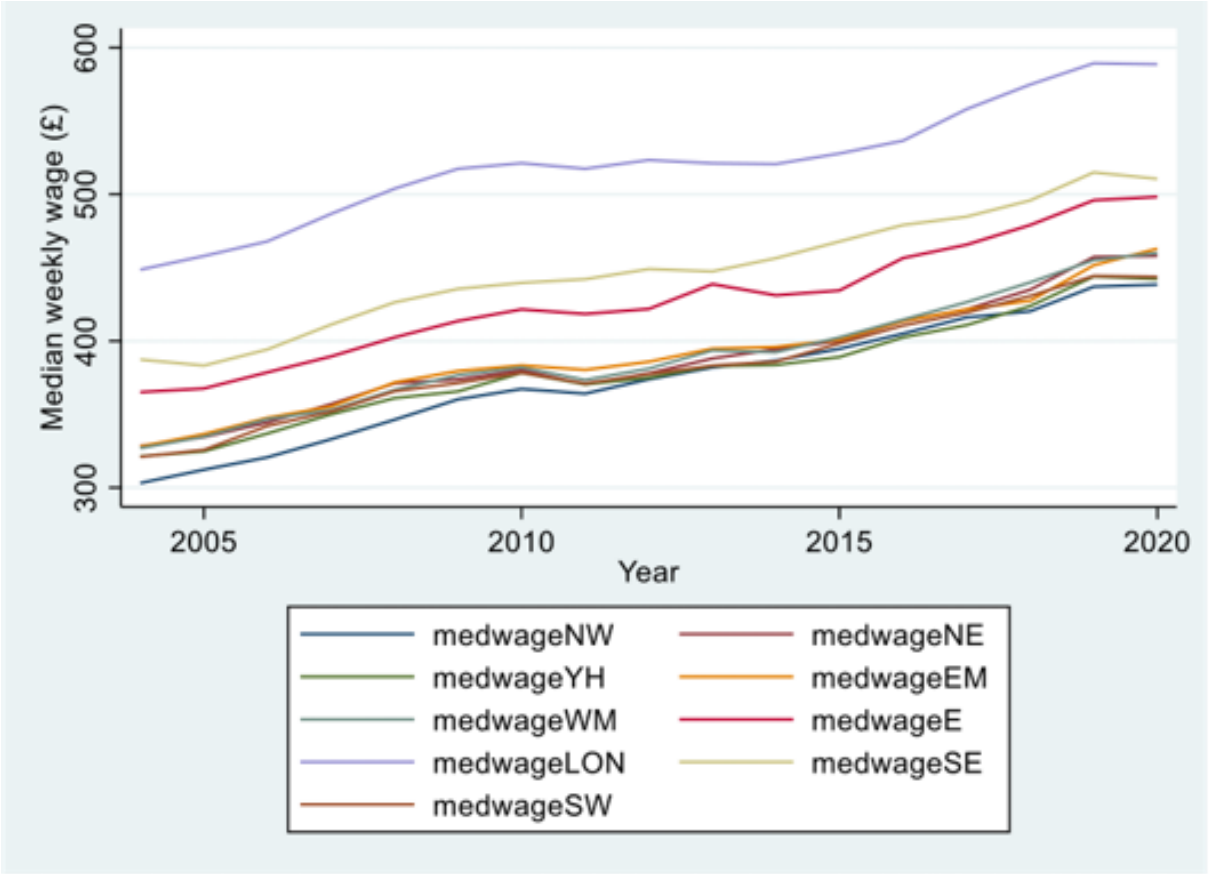

Figure 2: Median Weekly Wage (£) by Region

Figure 2 shows the median weekly wage by region in England between 2004-2020. England is split into 9, larger administrative regions: North West (NW), North East (NE), Yorkshire & the Humber (YH), East Midlands (EM), West Midlands (WM), East of England (E), London (LON), South East (SE), South West (SW). The figure exhibits the disparities in income between the South East (incl. London) and the rest of England.

4 Methodology

4.1 Method & Reasoning

With the purpose of investigating how immigration EU workers effects wages at the municipality level in the English labour market, it was hypothesised that an increase in migrant flows leads to a decrease in the median wage of a municipality in the short-term. Additionally, it was hypothesised that an increase in migration flows leads to an increase in the median wage of a municipality in the long-run (defined as a 5 year period). To investigate these hypotheses, the following two-way fixed effects model was estimated:

$$ln(w_{jt}) = \alpha_j + \beta_1 ln(m_{jt-1}) + \beta_2 ln(m_{jt-5}) + \alpha_t + \epsilon_{jt} \tag{5}$$

The variable *Median Wage (GBP)* is the median weekly wage of municipality j in year t. The natural logarithm of this variable is taken so that the dependent variable for this analysis is *ln(Median Wage)*. This therefore allows the coefficients to be interpreted as elasticities, as the natural logarithm is taken whilst calculating both the independent variables.

For this study, the independent variables are *Current Migration Flows* and *Past Migration Flows*, with the variables, like the dependent variable, being municipality j and year t specific. *Current Migration Flows* represents the impact of EU migration on the labour supply in municipality j in the short-run, with *Past Migration Flows* accounts for the impact of EU migration on the labour supply in the long run. β_1 captures the impact of the previous year's migration flows on median wages, and β_2 captures the impacts of migration flows 5 years previously on median wages.

On top of that, α_t is the year fixed effect used to control for time-specific heterogeneity, that is driven by time trends. For example, it accounts for any impacts of the Great Recession and the impacts of the Brexit referendum vote, both of which could have impacted migration flows and the median weekly wage of a municipality. α_j accounts for the regional fixed effects, accounting for time invariant factors that may impact the median wage and the migration flows into a region. These fixed effect variables give every region their own intercept in the regression, thereafter accounting for region-specific and year-specific effects.

Potentially, there are other time-varying factors that impact the median wage and haven't been accounted for by the region and year fixed effects. However, with a very localised dataset the availability of indicators are difficult to come by without permission, so no control variables are included in the two-way fixed effects model. This results in the potential caveat of Omitted Variable Bias (OVB), but only if the instrumental variables (*Past settlement Inst 1* and *Past Settlement Inst 2*) are correlated with the error term. In statistical terms that is:

1. $Cov(ps_{jt-1}, \epsilon_{jt}) = 0$
2. $Cov(ps_{jt-5}, \epsilon_{jt}) = 0$

However, the instrumental variables are required to be valid and strong for there to be a case for causal inference. These factors will be discussed in the limitations section. Finally, ϵ_{jt} represents the error term, which captures the amount the equation varies.

Firstly, the model will be estimated for the full sample period (2004-2020). Standard errors will be clustered at the regional level. Thereafter, additional robustness checks will be undertaken, to test the strength of the instrumental variables and the results.

4.2 Variables

4.2.1 Dependent Variable

The Dependent variable of choice for this study is the *Median (Weekly) Wage*. The median wage is defined as the weekly gross pay in Pound Sterling (GBP), which is estimated on a yearly basis. These are calculated by randomly selecting 1% of the working age population from Her Majesty's Revenue & Customs' (HRMC) Pay As You Earn (PAYE) database. The natural logarithm of the median wage was taken. By virtue of this, the coefficient of the independent variables (discussed in section 4.2.2) can be interpreted as elasticity's.

Median weekly wage was viewed as the most appropriate choice of dependent variable to analyse the impacts of EU migration on the English Labour Market. Firstly, wages, along with unemployment, are key indicators on the health of the labour market. A prominent argument of the Brexit referendum was that EU migrants were depressing real wages, as they were willing to work for a lower wage than native workers. The following empirical analysis will conduct an examination to whether this was the case between 2004-2020. Secondly, as mentioned in section 3.1, the median wage was preferred over the mean wage. The wage distribution is often positively skewed, with the highest earners causing the mean wage to be higher than the median. Hereby, the median wage was deemed to be the most suitable dependent variable for the forthcoming analysis.

It is important to note that these are estimates of a 1% sample size of each municipality. For the smallest municipalities, this results in very small sample sizes, so estimates of their median wage must be approached with caution. Inaccuracies arising from the measurement of the dependent variable will be discussed in the Limitations section.

4.2.2 Independent Variables

The independent variables for the empirical analysis are *Ln(Current Migration Flows)* and *Ln(Past Migration Flows)*. All estimates of population were calculated to the nearest thousand. This is potentially problematic in the smaller municipalities, some of which have very little growth in EU migrants over time. As a robustness check, municipalities with an initial EU population of below 3000 will be excluded.

The aim of this study is to untangle the effects of migration on wages in the short and long run in the context of the English labour market. Jaeger et al. (2018), find that using a single independent variable conflates the short and long run effects of migration. Therefore, two separate instruments are used to decompose these effects. *Ln(Current Migration Flows)* is used to capture the the impacts of recent migration on the wages in the current period. *Ln(Past Migration Flows)* captures the impact of past migration flows of current wages. *Ln(Past Migration Flows)* is lagged by 5 years, as this study assumes this is the time of adjustment of the English labour market to labour supply shocks. Although a strong assumption, with the US taking a decade or longer to spatially adjust (Greenaway-McGrevy and Hood, 2016), for a smaller economy such as England these adjustments plausibly happen faster. Furthermore, if the *Ln(Past Settlement Flow)* variable is lagged any further, it severely limits the number of observations in the sample. The main independent variables were calculated as follows:

$$Ln(CurrentMigrationFlow) = log(\Delta M_{jt-1} - L_{jt-2}) - log(L_{jt-2}) \tag{6}$$

$$Ln(PastMigrationFlow) = log(\Delta M_{jt-5} - L_{jt-6}) - log(L_{jt-6}) \tag{7}$$

Where ΔM_{jt-1} and ΔM_{jt-5} are the changes in the number of EU migrants in area j between time period's *t-1* and *t-2* & time period's *t-5* and *t-6* respectively. *Current Migration Flow* is lagged by one time period as it is assumed that migrant flows in the current period have no impact on the wages in the current period. L_{jt-2} and L_{jt-6} represent the total population in area j in time period's *t-2* and *t-6* respectively. Effectively, *Current Migration Flow* and *Past Migration Flow* measure the impact of the change in the number of EU migrants on the total labour supply. The more positive the flow of EU migrants into area j at a certain point in time is, the greater the values of *Ln(Current Migration Flow)* and *Ln(Past Migration Flow)*.

4.2.3 Instrumental Variables

As previously mentioned, the key issue of identification within migration literature arises from endogeneity bias between migration flows and labour market outcomes. Migration may affect the wages of an area, albeit in polarising ways in the short and long run. However, the median wage will also impact the choice of municipality a migrant will choose to live in. A migrant will want to maximise their utility (in the form of higher wages), so plausibly they will select a municipality in which their wages are maximised. These simultaneous effects mean a simple OLS not feasible if unbiased and consistent estimates are desired. Instrumental variables are a potential solution to this issue, although criterion must be met in order to have a valid instrument (these are specified below). In order to have causal estimates, the explanatory variable must be exogenous, but it has been established that *Current Migration Flow* and *Past Migration Flow* is not. Therefore, an

instrument that is not related to the error term; ($Cov(Z, \epsilon) = 0$) can be used to isolate the variation in the explanatory variable that is exogenous. Two shift-share instruments are thereafter proposed: *Past Settlement Inst 1* and *Past Settlement Inst 2*. They are defined as follows:

$$\widetilde{ps}_{jt-1} = \sum_j \frac{M_{jt^{2001}}}{M_{t^{2001}}} \frac{\Delta M_{t-1}}{L_{jt-2}} \tag{8}$$

$$\widetilde{ps}_{jt-5} = \sum_j \frac{M_{jt^{2004}}}{M_{t^{2004}}} \frac{\Delta M_{t-5}}{L_{jt-6}} \tag{9}$$

Where $\widetilde{ps}_{jt-1}$ is *Past Settlement Inst 1* and $\widetilde{ps}_{jt-5}$ is *Past Settlement Inst 2*. The first part of the equation, $\frac{M_{jt^{2001}}}{M_t^{2001}}$ and $\frac{M_{jt^{2004}}}{M_t^{2004}}$, represents the share of EU migrants in area j at base period's t^{2001} and t^{2004}. $M_{jt^{2001}}$ and $M_{jt^{2004}}$ being the number of EU migrants in area j in each of the base period's. M_t^{2001} and M_t^{2004} refer to the total number of working age EU migrants in England in each of the base periods. The base period for equations 8 and 9 are selected in order to minimise their correlation with current shifts in the migrant population, reducing the likelihood of the instruments being weak. $\frac{\Delta M_{t-1}}{L_{jt-2}}$ and $\frac{\Delta M_{t-5}}{L_{jt-6}}$ represent the shift in migration flows to area j in the previous time periods. ΔM_{t-1} and ΔM_{t-5} are the national-level migrant flows in periods *t-1* and *t-5*, with L_{jt-2} and L_{jt-6} being the total workforce in area j in time periods *t-2* and *t-6*.

Using the past settlement of EU migrants conceivably satisfies the criterion for a valid instrument. Instruments need to satisfy the following restrictions:

1. **Meaningful First Stage**: The instrument is correlated with the independent variable.

2. **Exclusion Restriction**: The instrument only affects the dependent variable through its effect on the independent variable.

3. **Exogeneity**: The instrument should not be related to any unobservable factors (the error term ϵ_{jt}) that influence the dependent variable.

The Past Settlement instruments plausibly satisfy the meaningful first stage - the instrument is correlated with the independent variable. If there is a greater proportion of EU migrants already living in area j, this potentially means that there will be greater migrant flows to these areas compared to others that have a lower number of EU migrants. A possible reason for this includes family ties, new migrants may want to move close to family to ease the transition of migrating. Additionally, migrating to areas with high numbers of the same nationality can reduce the size of the initial language barrier. Therefore, a positive relationship is expected between the Past Settlement Instruments

and migration flows.

The exclusion restriction is also arguably satisfied. However, it is not possible to test for whether this is the case, so economic intuition is relied on to explain whether this criteria is met. Within the *Past Settlement Instruments*, variation arises from changes in national migration flows of EU workers. National migration flows has an impact on migration flows at the municipality level, thereby satisfying the requirement of a meaningful first stage. In years where there is a greater volume of migration from the EU, it would be expected that there is greater migration flows at a municipality level too. For the exclusion restriction to be fully satisfied, the past settlement instruments must have no impact on the wages in the current period. It is plausible that this is the case, although arguments can be made against this. For example, the number of migrants in area j in the past could impact the wages in the current period, especially if there is a large enough labour supply shock that means the labour market takes a number of years to adjust to.

Finally, the instrument must be exogenous, that is, it should not be related to unobservable factors within the error term. An unobservable factor that influences wages over time and is plausibly related to the past settlement of migrants could be labour productivity within an area. Migrants can have different skill levels compared to migrants that fosters complementarity, leading to an increase in Total Factor Productivity (TFP), thereafter leading to an increase in wages. However, in the case of the UK, there has been net positive migration in the last 15 years, but TFP growth has stagnated, leaving the manner of the relationship unclear. On the other hand, the influx of EU migrants working in lower-skilled jobs may have instead hurt productivity. With a larger labour supply, and the ability to hire different workers if required, firms potentially have a smaller incentive to invest in lower-skilled workers. This acts as a blockage on productivity, potentially explaining the very low productivity growth in the UK/England in the past 15 years compared to its G7 peers (Strauss, 2021). To summarise, the instruments could potentially be exogenous, as the relationship between migration flows and productivity is unclear. Ideally, productivity would be included in the reduced form regression as a control, but data availability at the municipality level limits this.

4.2.3.1 First Stage The first stage regressions are motivated off the paper by Jaeger et al. (2018), albeit with a tweak to the regression in which *Past Migration Flow* is the dependent variable. The first stage regressions take on the following form:

$$m_{jt-1} = \pi_{10} + \gamma_{11}\widetilde{ps}_{jt-1} + \gamma_{12}\widetilde{ps}_{jt-5} + u_{jt} \tag{10}$$

$$m_{jt-5} = \pi_{20} + \gamma_{21}\widetilde{ps}_{jt-1} + \gamma_{22}\widetilde{ps}_{jt-5} + v_{jt} \tag{11}$$

Estimates for *Current Migration Flow* are constructed using both past settlement instruments, because both *Past Settlement Inst 1* and *Past Settlement Inst 2*, both influence current migration flows. The number of EU migrants in England 5 years prior to the current time period can plausibly have an impact on the migration flows of today, as well as the number who were previously settled within the last time period. Therefore, two instruments are used for the first first stage regression. In Jaeger et al. (2018), they use both past settlement instruments to estimate m_{jt-1}, where m_{jt-1} is the equivalent of *Past Settlement Inst 2* in this study (the time period *t-1* is equivalent to 10 years). However, it is not logical to use the past settlement of migrants in the previous period (1 year in this study) to calculate the migration flows from 5 years previously. Thereafter, it was decided that it was not appropriate to utilise *Past Settlement Inst 1* for the second first stage regression.

4.2.3.2 Second Stage Now the endogenous variables have been have been estimated as a function of exogenous variables, they can now be made use of to estimate the impact of wages, without the issue of endogeneity bias. Thereafter, the estimates from the first stages will be implemented into the second stage equations, in order to examine the short and long run impacts of changing migration flows on wages. The second stage equation is also based off the paper by Jaeger et al. (2018), with some key differences:

$$ln(w_{jt}) = \beta_0 + \beta_1 m_{jt-1} + \beta_2 m_{jt-5} + \epsilon_{jt} \tag{12}$$

β_1 represents the short term impacts of migration on wages, with β_2 accounting for the longer term impacts. As both m_{jt-1} and m_{jt-5} were calculated using natural logarithms, the coefficients β_1 and β_2 can both be interpreted as elasticity's. The coefficients will represent the impact of a 1% increase in migration flows in terms of a percentage.

5 Results & Discussion

5.1 Main Results

Initially, a Davidson-MacKinnon test was conducted to determine whether a fixed effects with instrumental variables or an OLS model should be implemented. The p-value of the test showed that the null hypothesis should be rejected, and so a fixed effects with instrumental variables. The results of the Davidson-MacKinnon test can be viewed in Appendix A. In addition, the Cragg-Donald Wald F statistic was conducted to examine whether the second stage equation was weakly identified. The Sanderson-Windmeijer F-test was conducted to test the instruments for over-identification (Sanderson and Windmeijer, 2016). Further tests will be discussed in the robustness section.

Table 2: First and Second Stage IV Regression: Full Sample

	(1) First Stage: Ln(CMF)	(2) First Stage: Ln(PMF)	(3) Second Stage: Ln(Med Wage)
Past Settlement Inst 1	-.2946	.4507**	
	(.2490)	(.2706)	
Past Settlement Inst 2	.5002**	.5135**	
	(.2431)	(.2027)	
Ln(Current Migration Flow)			-.1386
			(0.0956)
Ln(Past Migration Flow)			0.1563**
			(0.0779)
Region Effects	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes
Observations	950	950	950
R-Squared			-3.9696
SW S stat	4.51	8.17	
CD Wald F			2.18
Sargan-Hansen Stat			0.00
F-Stat			5.12

Standard Errors in Parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2 withholds the results for the first and second stage regressions for the full sample. For the first stage regressions (Columns 1 and 2 in table 2), it was assumed that municipality's that hosted a larger population of EU migrants in the past were more likely to attract EU migrants in the future. The argument for this is based off the assumption that migrants will select into moving to specific municipality's to be close to family members. Alternatively, migrants may want to soften the adjustment of living in a new country (new language and different culture), and so they move to a municipality that has many migrants from their country. In column (1) *Past Settlement Inst 1* has a negative, but statistically insignificant impact on *Ln(Current Migration Flows)*. The statistical insignificance is perhaps expected, as migrants that settled 2 years ago plausibly will not yet have an impact on rate of immigration to a municipality. However, *Past Settlement Inst 2* has a positive, statistically significant (at the 5% level) impact on the rate of change in current migration flows. The coefficient of 0.5002 suggests that, ceteris paribus, an increase of 10% in the proportion of EU migrants settled will result in a 5% increase in migrant flows over a 5 year period. This provides evidence for the claim that

EU migrants are more likely to move to a municipality where there there already is a high proportion. Furthermore, a meaningful first stage is a key requisite for arguing whether an instrumental variable is valid.

Column (2) displays the results for the second first stage regression. Both *Past Settlement Inst 1* and *Past Settlement Inst 2* have a positive and statistically significant impact on the rate of change of past migration flows (5 years previously). Again, the positive coefficients provide evidence for the previous settlement location of migrants having a positive impact on migration flows.

The final column (3) displays the results for the second stage. The second stage regresses the rate of change in the past and current migration flows against the change in the weekly median wage. The coefficient of *Ln(Current Migration Flows)* is -.1386, with both the dependent and independent variable are in logarithmic form . Therefore, a 1% increase in *Current Migration Flows* leads to a 0.14% decrease in wages. This finding is consistent with the first hypothesis of this study, however it is statistically insignificant, with a p-value of 0.145. Despite the insignificance of the coefficient, it is not a great distance from significance at the 10% level. Intuitively, the negative sign of this coefficient can be explained by the fact in the short run an increase in labour supply will increase competition for the jobs (the number of vacancies is assumed to be fixed in the short run). This allows employers to keep wages at the same level, or even reduce them, particularly for lower skilled workers.

The coefficient for *Ln(Past Migration Flows)* has a positive, statistically significant (at the 5% level coefficient of 0.1563. Ceteris paribus, an increase of 1% in the net migration flows 5 years ago will lead to a 0.16% increase in the median wage in the current period. Again, this finding is consistent with the hypothesis that in the longer run positive net migration will lead to an increase in median weekly wages. The new migrants will stimulate additional demand in the economy, as they too will consume goods and pay taxes to help fund government spending. The initial results are in line with the hypotheses formulated in section 2.4. However, to confirm the validity of the methodology used and the results, a variety of robustness checks are thereafter undertaken.

5.2 Robustness Checks

Firstly, the sample adjusted so that any municipality's with a working age EU population of less than 3000 were excluded from the sample. In general, these were the smaller of the municipality's within the sample and exhibited less variation in their working age EU population than others in the sample. Therefore, these were excluded in order to examine

if the results hold for municipality's that are impacted to a greater extent. The results from omitting municipality's with a working age population can be viewed in Appendix B. Although the results have a slightly higher p-value, in general it holds, with the coefficients of interest remaining within a narrow band.

As an additional robustness check, the dependent variable was changed so that it was now *Ln(Mean Wage)*, instead of *Ln(Median Wage)*. The relationships observed in table 2 in the main results section remain consistent, with the coefficients similar to one another. The results of the regression in which the dependent variable is *Ln(Mean Wage)* can be observed in Appendix C.

Previous literature, such as studies by Card (2009) and Kugler and Yuksel (2008) used one past settlement instrument in order to examine the impacts migration on wages, unlike this study which uses two. Two instruments are used in order to decompose the short and long run impacts of EU migration into the English labour market. Previously, it has been argued that the short and long run effects conflate when a single past settlement instrument is used. However, a useful comparison can be made to previous literature's findings by using a single past settlement instrument. For this robustness check, the instrument that is more suitable is *Past Settlement Inst 2*. In the Main Results section, *Ln(Current Migration Flows)* was statistically insignificant for the second stage, so it is more appropriate to use *Ln(Past Migration Flows)* as the independent variable and use *Past Settlement Inst 2* as the instrument. Furthermore, it can be argued that using *Past Settlement Inst 2* is more suitable due to the time it takes for wages to adjust after a labour supply shock. This instrument is lagged by 5 years. Observing appendix D, the independent variable *Ln(Past Migration Flow)* has now lost its significance. Although problematic in terms of comparing to other literature, there is a potential explanation for this. By omitting the other independent variable of this study, it becomes apparent that the argument put forward by Jaeger et al. (2018) - that the short and long term effects of migration on wages conflate - is a reason as to why the variable *Ln(Past Migration Flows)* becomes significant when *Current Migration Flows* is added to the model. Testing the joint significance of the two independent variables gives a F-stat of 2.15, meaning we cannot reject the null of the variables being jointly significant, and so the short and long term effects of migration on wages do conflate.

Additionally, the reduced form equation was estimated. The dependent variable *Ln(Median Wage)* was regressed on the two instrumental variables. The reduced form is useful at it solves for the endogenous variables in the model. The results for the reduced form can be viewed in Appendix E. In this case, the coefficient for *Past Settlement Inst 1* is negative, meaning that a greater number of EU migrants being settled in a munic-

ipality 2 years ago will lead to a reduction in wages in the short run. The coefficient flips sign for *Past Settlement 2*, suggesting that the greater the number of EU migrants settled in a municipality 5 years, the greater the wages in the current period. The reduced form coefficients back up the economic theory behind this study - in the short-run, increased immigration leads to a reduction in wages and in the long-run there is a positive relationship with wages.

5.3 Study Limitations

The robustness checks undertaken offer additional support for the findings in the main results, however, they are not as comprehensive as desired. There are a variety of limitations from the data set used and the methodology implemented, including a lack of controls and the aggregated nature of the data collected on EU migrants.

Perhaps one of the key limitations of this study is the lack of controls. Controls were not included due to the difficulty in obtaining data at the municipality level. This point is illustrated by the fact that population data is only estimated to the nearest one thousand. Controls that would have been suitable include the education level of the working age population. The education level is plausibly positively correlated with the wages one would receive. If a worker has spent more time in education they can attain a higher skill level and be more productive, thus demanding higher wages. In the context of this study, it is important to control for this, as EU migrants have often filled in jobs that are below the skill level their education would suggest (Dustmann et al., 2013). Therefore, more concrete data on the education level of native workers and their respective wages would be desirable. This is especially important in order to examine the impacts of EU migration at different levels of the wage spectrum. To counteract this, regional level fixed effects were implemented so that the impacts of the education attainment in each region were accounted for, but this still has limitations in itself. By taking regional fixed effects, the strong assumption that each of the municipality's have the same educational level must be made, when in reality this may not hold.

Another control variable that is also not accounted for due to the lack of availability of controls is productivity. Intuitively, productivity (output per worker) has a positive impact on wages and also a positive impact on migration flows. If a municipality has better economic opportunities (productivity can be viewed as a sign of economic activity), then they are more likely to attract migrants. The positive impact of productivity on both migration flows and wages means that the estimates for both *Ln(Current Migration Flows)* and *Ln(Past Migration Flows)* may be overestimated. Although fixed effects aids with this issue, the problem of the impacts of migration affecting different areas of the

wage spectrum arises again. By aggregating the labour market into one skill level, we cannot untangle the impacts of EU migrants on high or low-skilled labour. In the future, studies examining the impact of migration on the English labour market should exploit census data that becomes available every 10 years. Potentially, specific municipality data could be extrapolated so that controls can be utilised for every year in a large dataset, this would help to nullify the problems caused by OVB.

The first robustness check conducted was to remove any municipality's with a EU migrant working age population of 3000. This was to ensure that the municipality's which exhibited little variation were excluded from the dataset, thereafter the results held. This issue can also be countered by using decadal census data. UK Censuses offer comprehensive, exact data on the population in each municipality, with data on the birth place of migrants and their country of birth. As a solution, the dataset can be combined with national growth rates in population for migrants to create estimates of the exact population in each municipality, instead of to the nearest thousand. This would generate variation in each observation, ending the requirement for smaller municipality's to be removed from the dataset.

In the future, studies can focus on the impacts on specific population groups from the EU, such as Polish migrants, who's population living in the UK topped one million in 2016 and 2017 (decreasing thereafter). By focusing on EU migrants, this study neglects the fact that the characteristics of migrants from the EU greatly vary. Again, this limitation arises from the caveats of using the ONS public dataset, the observations of population (and wages) are estimates.

The central theme of this study is the use of two past settlement instrumental variables in order to remove the exogeneity bias that arises between wages and migration. As has already been established, wages intuitively will increase migration flows, as migrants will observe the greater economic opportunities these municipality's have and thereafter move there. Migration also plausibly influences wages in contrasting directions, as the results convey. However, it can be argued that the past settlement instruments do not satisfy the exogeneity that is required to for an instrument to be valid. The past settlement choice of migrants is plausibly related to unobservable factors. For example, the median wage 5 years ago is credibly related to the choice of whether a migrant decided to settle in a particular municipality 5 years ago. Furthermore, the median wage 5 years ago logically has an impact on the wage in the current period, serving as an argument against the validity of the past settlement instruments used in this study. However in any case, if one uses an instrumental variable, arguments against the validity can be made.

Perhaps the greatest caveat of this study is the weakness of the instruments used. In table 2 the Cragg-Donald Wald F-statistic is given as 2.18, a strong indication of weak instruments (it is widely accepted that an F-statistic below 10 is strong evidence of a weak instrument). Due to the annual nature of the data and the reliance on changes in national inflows of EU migrants for the variation within the instruments, there was little variation within the instruments. However, with the dataset commencing in 2004 on the advent of the A8 countries joining the EU and sparking a rapid increase in immigration to England, it was thought there would be enough variation in national inflows to nullify this problem. There is clearly not enough variation in national inflows of migrants to prevent the instruments being correlated with each other (and therefore weak). In order to prevent this, future studies can exploit data-sets over a longer period of time, and thereafter lag *Past Settlement Inst 2*, by a greater amount than 5 years. In this study, it was decided to lag this instrument by 5 years because England is a small country. It was assumed that England would adjust faster to labour supply shocks than larger countries such as the US. To counter the weak instrument problem, a longer lag period on *Past Settlement Inst 2* was considered, to reduce the correlation between the two instruments. However, this reduced the number of the observations in the sample by an excessive amount. The problem of weak instruments is commonplace with the use of shift-share instruments in migration literature (Jaeger et al., 2018). Future studies can make use of a longer time period, so that if multiple instruments are used, they can be lagged by a sufficient number of time periods that they are not heavily correlated.

6 Conclusion & Discussion

To answer the research questions, it was hypothesised that in the short-run wages within a municipality would be negatively impacted by an increase in migration flows. By means of a two-way fixed effects model used in conjunction with two past settlement instrumental variables, it was observed that a 1% increase in the migration flows leads to a 0.14% decrease in the median weekly wage. However, although the coefficient for the impact of *Ln(Current Migration Flows)* on wages as a p-value of just above 0.1, it is still insignificant. Therefore, the coefficient should be interpreted with caution. Additional robustness checks, such as reducing the sample to include municipality's that exhibited enough variation, also showed a negative relationship between current migration flows and wages. Thereafter, the initial research question can be answered with reasonable confidence. Future studies can allow a greater time period for the lag of the short-run variable, to allow for time for the impacts of a labour supply shock to filter through to wages.

In answering the second research question, it was hypothesised that an increase in mi-

gration flows in the long-run (5 years prior) would lead to an increase in the growth of the median weekly wage. Again using past settlement instrumental variables, it was observed that a 1% increase in migration flows in the long-run leads to an increase of 0.15% in the median wage in the current period. Furthermore, the estimate for the coefficient of *Ln(Past Migration Flows)* was statistically significant at the 5% level. The significance of the estimates also held for the robustness checks, meaning it can be stated with some confidence that in the longer term increased migration from the EU has a positive impact on the median wage in England.

This paper compliments existing literature which has attempted to examine the impacts of migration on wages and other labour market outcomes. This paper demonstrates that migration does have a positive impact on the labour market in the long-run, but not before there is an initial depressing of wages. The findings are consistent with the paper by Jaeger et al. (2018). Their extension to previous literature was to add an extra instrumental variable to untangle the short and long-run effects of immigration, which act in opposite directions. It is argued in the longer term that migrants contribute to higher wages by increasing the aggregate consumption of the economy. To meet this demand, firms potentially invest to increase productivity, thereby increasing wages. The findings of this paper corroborate the findings of Jaeger et al. (2018), and prove that the multiple instrument methodology can be applied to smaller countries such as England.

There are considerable limitations to this study however, with the lack of availability of controls being particularly problematic. The omission of control variables that can impact the current and past migration flows and also the the rate of change in the median wage leads to a possibility of OVB. Depending on the relationship of the omitted variables with the dependent and independent variables, the estimates could under or overestimated. Furthermore, by only focusing on the median wage for each municipality and not different wage brackets such as the lowest 20 percentile, it is unclear where the impacts of EU migration affect on the wage distribution. Intuitively, it is the lowest skilled workers in England that get impacted the most negatively by the additional labour supply, as EU migrants fill the lowest paid/skilled jobs (Dustmann et al., 2013).

Perhaps the main weakness of this paper is the weakness of the instruments used. Due to the lack of variation in the rate of change in national inflows of EU migrants, the instruments are heavily correlated. Over time, England experienced a steady increase in the share of EU migrants, explaining the lack of variation in the instruments. After the Brexit referendum, the number of EU migrants living in England has dropped. Future studies can exploit the impact the Brexit referendum has had on national inflows of migrants from the EU, as this will potentially aid the weak instrument problem.

Despite the shortcomings of some aspects of the methodology and the data, this paper offers a suitable framework that can be adjusted and applied to other country's labour markets, as well as demonstrating the importance of the overarching relationship between migration and wages for policymakers. Either encouraging or placing quotas on migration can be utilised to provide the best outcome for the current population.

The results can be built on by examining the impacts of EU migration on wages at a more specific level - the district level. By using ONS decadal census data, detailed data can be extrapolated to estimate a specific value for each year, instead of to the nearest thousand given by the yearly estimates. In addition, in a longer time frame, *Past Settlement Inst 2* can be lagged by a greater number of years than five. This would also be a potential solution to solving the weak instrument issue, caused by the correlation in national inflows of EU migrants over time.

References

Altonji, J. G. and Card, D. (1991). The effects of immigration on the labor market outcomes of less-skilled natives. In *Immigration, trade, and the labor market*, pages 201–234. University of Chicago Press.

Antonakis, J., Bendahan, S., Jacquart, P., and Lalive, R. (2010). On making causal claims: A review and recommendations. *The leadership quarterly*, 21(6):1086–1120.

Bartel, A. P. (1989). Where do the new US immigrants live? *Journal of Labor Economics*, 7(4):371–391.

Bartik, T. J. (1991). Who benefits from state and local economic development policies?

Blanchard, O. and Katz, L. (1992). Regional Evolutions. *Brookings Papers on Economic Activity*, 23(1):1–76.

Borjas, G. J. (2001). Does immigration grease the wheels of the labor market? *Brookings papers on economic activity*, 2001(1):69–133.

Card, D. (2009). Immigration and inequality. *American Economic Review*, 99(2):1–21.

Card, D. and DiNardo, J. (2000). Do immigrant inflows lead to native outflows? *American Economic Review*, 90(2):360–367.

Cohen-Goldner, S. and Paserman, M. D. (2011). The dynamic impact of immigration on natives' labor market outcomes: Evidence from Israel. *European Economic Review*, 55(8):1027–1045.

De Silva, D. G., McComb, R. P., Moh, Y.-K., Schiller, A. R., and Vargas, A. J. (2010). The effect of migration on wages: evidence from a natural experiment. *American Economic Review*, 100(2):321–26.

Dustmann, C., Frattini, T., and Preston, I. P. (2013). The effect of immigration along the distribution of wages. *Review of Economic Studies*, 80(1):145–173.

Dustmann, C., Glitz, A., and Frattini, T. (2008). The labour market impact of immigration. *Oxford Review of Economic Policy*, 24(3):477–494.

Faber, J. and Fonseca, L. M. (2014). How sample size influences research outcomes. *Dental press journal of orthodontics*, 19:27–29.

Fingleton, B. and Szumilo, N. (2019). Simulating the impact of transport infrastructure investment on wages: a dynamic spatial panel model approach. *Regional Science and Urban Economics*, 75:148–164.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624.

Greenaway-McGrevy, R. and Hood, K. K. (2016). Worker migration or job creation? persistent shocks and regional recoveries. *Journal of Urban Economics*, 96:1–16.

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

Kugler, A. and Yuksel, M. (2008). Effects of low-skilled immigration on us natives: Evidence from Hurricane Mitch. Technical report, National Bureau of Economic Research.

Lemos, S. and Portes, J. (2008). New labour? the impact of migration from central and eastern european countries on the UK labour market.

McIntosh, M. F. (2008). Measuring the labor market impacts of hurricane katrina migration: Evidence from houston, texas. *American Economic Review*, 98(2):54–57.

ONS (2021). Overview of the UK population: January 2021. *ONS*.

Ottaviano, G. and Peri, G. (2005). Rethinking the Gains from Immigration: Theory and Evidence from the US.

Ottaviano, G. I., Peri, G., and Wright, G. C. (2013). Immigration, offshoring, and American jobs. *American Economic Review*, 103(5):1925–59.

Rouvroye, L., van Dalen, H. P., Henkens, K., and Schippers, J. J. (2021). Employers' views on flexible employment contracts for younger workers: Benefits, downsides and societal outlook. *Economic and Industrial Democracy*, page 0143831X211053378.

Sanderson, E. and Windmeijer, F. (2016). A weak instrument f-test in linear iv models with multiple endogenous variables. *Journal of econometrics*, 190(2):212–221.

Strauss, D. (2021). The challenge of unlocking the UK's low productivity. *The Financial Times*.

Travis, A. (2016). Are EU migrants really taking british jobs and pushing down wages? *The Guardian*.

Vargas-Silva, C. (2014). *The fiscal impact of immigration in the UK*. Migration Observatory-University of Oxford.

Wooldridge, J. M. (2002). Econometric analysis of cross section and panel data MIT press. *Cambridge, MA*, 108(2):245–254.

7 Appendix

7.1 Appendix A

Table 3: Davidson-MacKinnon Test (1993)

	Coef.
F-Test Value	109.3
P-value	5.1e-43

7.2 Appendix B

Table 4: IV Regression: Sample Excluding Smallest Municipalities

	(1) First Stage: Ln(CMF)	(2) First Stage: Ln(PMF)	(3) Second Stage: Ln(Med Wage)
Past Settlement Inst 1	-.3340	.4309*	
	(.2640)	(.2214)	
Past Settlement Inst 2	.5607**	.4667**	
	(.2605)	(.2185)	
Ln(Current Migration Flow)			-.1450
			(.0919)
Ln(Past Migration Flow)			0.1580*
			(.0896)
Region Effects	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes
Observations	889	889	889
R-Squared			-4.5935
SW S stat	5.17	7.19	
CD Wald F			2.36
Sargan-Hansen Stat			0.00
F-Stat			4.29

Standard Errors in Parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

7.3 Appendix C

Table 5: IV Regression: *Ln(Mean Wage)* as the Dependent Variable

	(1) First Stage: Ln(CMF)	(2) First Stage: Ln(PMF)	(3) Second Stage: Ln(Mean Wage)
Past Settlement Inst 1	-.3033	.4457*	
	(.2496)	(.2073)	
Past Settlement Inst 2	.5039**	.5121**	
	(.2436)	(.2023)	
Ln(Current Migration Flow)			-.1776
			(.1127)
Ln(Past Migration Flow)			.1752*
			(.9370)
Region Effects	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes
Observations	932	932	932
R-Squared			-5.096
SW S stat	4.62	8.21	
CD Wald F			2.23
Sargan-Hansen Stat			0.00
F-Stat			3.26

Standard Errors in Parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

7.4 Appendix D

Table 6: IV Regression: Single Past Settlement Instrument

	(1) First Stage: Ln(PMF)	(2) Second Stage: Ln(Med Wage)
Past Settlement Inst 2	.4751	
	(.2161)	
Ln(Past Migration Flow)		0.9097
		(4.272)
Region Effects	Yes	Yes
Year Effects	Yes	Yes
Observations	995	995
R-Squared		0.2924
SW S stat	6.26	
CD Wald F		6.132
Sargan-Hansen Stat		0.00
F-Stat		4.29

Standard Errors in Parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

7.5 Appendix E

Table 7: Reduced Form

	(1)
	Reduced Form
	Ln(Median Wage)
Past Settlement Inst 1	1.052
	(1.035)
Past Settlement Inst 2	.6268
	(.4460)
Area Effects	Yes
Year Effects	Yes
Observations	968
R-Squared	0.2211
F-Stat	143.1

Standard Errors in Parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$