MIGRATION ADVISORY COMMITTEE

EEA migration in the UK: Annexes

September 2018

Migration Advisory Committee
2nd Floor
Peel Building
2 Marsham Street
London
SW1P 4DF
https://www.gov.uk/government/organisations/migration-advisory-committee

EEA migration in the UK: Annexes

Migration Advisory Committee

September 2018

Contents

Annex A: MAC Commission

Home Secretary
2 Marsham Street
London
SW1P 4DF
www.gov.uk/home-office

Professor Alan Manning, Chair, Migration Advisory Committee
Sent via email only

27 July 2017

Dear Professor Manning,

I am attaching to this letter a commission for advice from the Migration Advisory Committee (MAC). It covers both the impacts on the United Kingdom labour market of the UK's exit from the European Union and also, since the two issues are clearly closely linked, how the UK's immigration system should be aligned with a modern industrial strategy. I am grateful for the indications that I have been given of the MAC's willingness to take on this work.

I am sure you do not underestimate the significance of the task which we are asking the MAC to undertake and I thought the Committee might find it helpful if I were to set out some relevant background information.

The Government has been clear that it respects the outcome of the Referendum on the UK's membership of the European Union, and will make a success of the UK's exit from the EU. We are working towards the goal of achieving sustainable levels of net migration but we also want to ensure all economic impacts are well understood and prepared for.

The Government also said that after the UK leaves the EU, free movement will end but migration between the UK and the EU will continue. Migration benefits the UK, economically, culturally and socially. Our businesses, agriculture, public services, voluntary organisations and universities rely to a greater or lesser extent on migration for labour, skills and ideas. Britain is a tolerant country, open for business and will stay that way. We will remain a hub for internatio5
nal talent and our departure from the EU must be seen in this context.

But sharply increased levels of net migration since 1997, from both the EU and beyond, have given rise to public concern about pressure on public services and wages. These concerns about the sustainability of unrestricted migration from the

EU featured strongly in the debate surrounding the referendum on the United Kingdom's EU membership on 23 June 2016. The public must have confidence in our ability to control immigration from the EU. Although net migration from the EU has fallen over the last year, we cannot exercise control over the type and volume of EU migration at present, as free movement gives EU citizens extensive rights to reside.

As set out in the Government's Command Paper Legislating for the United Kingdom's withdrawal from the European Union, the Repeal Bill will convert EU-derived law into UK law as it stands at the moment at which we leave the EU. Without further change, that would mean that the free movement of EU citizens to the UK would continue, albeit as part of UK law. We have therefore committed to introduce to Parliament an Immigration Bill to repeal the current EU-derived free movement provisions so as to be able, on our exit from the EU, to bring EU citizens fully within the scope of UK law. This will mean that, in future, we will be able to apply different immigration rules and requirements according to the UK's economic and social needs at the time, and reflecting our future deep and special partnership with the EU, including on any implementation arrangements following the UK's departure.

We do not envisage moving to that future system in a single step when we leave the EU. It will be in the interests of migrants, employers and the UK authorities, to have a predictable, well understood process which moves gradually from the free movement regime to a new set of arrangements.

Our first priority is to safeguard the position of existing EU residents in the UK and UK nationals in the EU. So, the first phase of our immigration proposals was to publish our fair and serious offer on 26 June[1]. This set out our proposals that qualifying EU citizens, arriving and resident before a specified date, would be able to apply for 'settled status' in UK law once they have accumulated five years' continuous residence – meaning that they would be free to reside in any capacity and exercise any lawful activity, and to access public funds and services. Those arriving and resident before the specified date but who had not yet accrued five years' residence would be able to remain until they accumulate those five years' residence. They would all have adequate time to apply for their documentation after our exit as there would be a 'grace period' of up to two years.

EU citizens arriving after the specified date but before exit (if the specified date we agree with the EU is prior to withdrawal) would be allowed to remain in the UK for at least the temporary 'grace period', and, may subsequently become eligible to settle permanently depending on their status and the rules in place at the time.

As part of a smooth and orderly transition as we leave the EU, the second phase of our immigration proposals is based on a temporary implementation period to ensure there is no cliff-edge on the UK's departure for employers or individuals. This includes the 'grace period' during which those EU citizens who arrived before the

[1] Safeguarding the Position of EU Citizens Living the UK and UK Nationals Living in the EU, published on 26 June 2017

specified date will have time to obtain their documentation from the Home Office. During this period there will also be a straightforward system for the registration and documentation of new arrivals (as well as for those who arrived after the specified date but before exit, if appropriate). A registration system that enables EU citizens to demonstrate their right to live and work in the UK is the basic requirement to be able to operate any system of immigration control.

After this implementation period, we will move to the third phase which will be our long-term arrangements covering the migration of EU citizens, designed according to economic and social needs at the time, and reflecting our future deep and special partnership with the EU.

The Government will want to ensure that decisions on the long-term arrangements are based on evidence. The commission that we are now asking the MAC to undertake is very much part of this. I very much hope that in undertaking its work the MAC will want to consult widely and that those affected will take the opportunity to make sure their voices are heard.

Alongside that, the Government will be undertaking its own extensive programme of engagement and evidence gathering with all interested parties including business, industry, trades unions, educational institutions and many others, to ensure we strike a balance on future EU migration arrangements. It is important that those affected contribute to the design of future arrangements and start to consider how they might adapt to a future immigration system.

Only when all of this concluded, and we have the MAC's advice, will we determine what the future long-term immigration rules for EU citizens should be. The Government will be able to set and adjust the successor arrangements to meet the needs of our wider immigration policy, our economic circumstances and the deep and special partnership we seek to agree with the EU, as well as trade agreements with other countries. I would be grateful if the MAC could report by September 2018, though it would be helpful if you felt able to provide interim reports throughout the period that you are working on this commission.

I look forward to receiving the MAC's advice on these important issues and I shall be publishing this letter.

Rt Hon Amber Rudd MP

COMMISSION FOR THE MIGRATION ADVISORY COMMITTEE

The Government has made clear that part of its immigration policy is to continue to reduce net migration, towards sustainable levels and to end free movement as we leave the European Union. Against that background, and to support future policy development, the Government would welcome advice and evidence from the MAC in respect of current patterns of EU and European Economic Area (EEA) migration and the role of migration in the wider economy and society.

<u>EU and EEA Migration</u>

- Drawing on existing sources where appropriate, the MAC should set out current patterns of EU and EEA migration, looking at:
    - sectors,
    - regional distribution,
    - skill levels,
    - duration of assignments,
    - self-employment, entrepreneurs, part time, agency, temporary and seasonal workers; and
    - any other characteristics the MAC considers relevant;

The MAC should consider the evolution of EU and EEA migration since 2000 and possible future trends (absent new immigration controls).

- What are the methods of recruitment used by UK employers to employ EU and EEA migrants and how does this impact on UK workers?

- What are the economic and social costs and benefits, including fiscal impacts to the UK economy and impacts on public services and infrastructure of EU and EEA migration?

- Is it possible to estimate the potential impact of any future reductions in EU and EEA migration (whether occurring naturally or through policy), at a range of levels and how may these be felt differently across the economy and society? This may include a consideration of the impacts on the different parts of the UK, within the context of designing a UK-wide immigration system. How could business adjust if EU and EEA net migration was substantially reduced? What mitigating actions could be taken by employers and government and over what timescale?

<u>Aligning the UK immigration system with a modern industrial strategy</u>

- What is the current impact of immigration, both EU, EEA and non-EEA, on the competitiveness of UK industry, including on productivity, innovation and labour market flexibility?

- What impact does immigration have on skills and training?

- Is there any evidence that the free availability of unskilled labour has contributed to the UK's relatively low rate of investment in some sectors?

- Are there advantages to focussing migrant labour on highly skilled jobs or across the entire skills spectrum?

- Does the shortage occupation list need to be amended to include skills shortages at lower skills levels than NQF6?

Where relevant to the above, we would welcome detail of what lessons can be drawn from the approach taken by other countries.

The MAC is asked to report by September 2018. The MAC may wish to provide interim reports throughout that period.

Annex B: Technical Annex

Defining migrant groups

B.1. In what follows, and in line with previous MAC reports, we use **country-of-birth** to define **migrants as those not born** in the UK. In contrast, some recent ONS publications have argued it is preferable to use nationality. Neither measure is perfect and the use of one over the other depends on the purpose for which the statistics are being used. Individuals have only one country of birth while they might have multiple nationalities (and our data sources typically record only one) and individuals might also be eligible for nationalities they do not currently hold. However, some foreign-born individuals will be eligible for British citizenship from birth in which case they would not be subject to any migration controls and it may be misleading to categorise them as migrants. In the rest of this note, we exclude Irish-born from our EEA-born migrant definition, as it is assumed that the Common Travel Area will continue between the UK and Ireland. And the group we refer to as EU13+ are those countries who were members of EU before 2004 plus EEA members plus Switzerland.

B.2. This leads to migrant groups being classified as follows:

- **UK and Ireland**
- **EU13+**: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden, Iceland, Liechtenstein, Norway and Switzerland.
- **New Member States (NMS)**: Bulgaria, Croatia, Cyprus, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Romania, Slovakia, Slovenia
- **Non-EEA**

Source of data is APS/LFS. The coding of country of birth has changed over time and we use 3 different variables; CRY (changed to CRY01 in 2001), CRYO and CRYOX7. CRYOX7 is used after 2007, and a combination of CRY and CRYO is used until 2006.

Defining skill levels

B.3. Skill levels are determined by occupation and their corresponding Regulated Qualifications Framework (RQF) level. RQF level 1 and 2 are determined to be ‘low skill’, levels 3 and 4 are ‘medium skill’ and level 6+ is ‘high skill’.

B.4. Skill level is determined by occupation (classified using SOC10 coding). Each occupation corresponds to an RQF level and is categorised into the low,

medium and high categories discussed above. The mapping from occupations to RQF levels comes from Immigration Rules Appendix J[2].

SIC and SOC code conversation

B.5. Where SIC and SOC codes need to be converted for consistency, we convert to the newer coding via proportional mapping. SIC92 codes are converted to SIC07 and SOC2000 is converted to SOC2010.

Use of survey data

B.6. Throughout the report we use survey data, such as the Labour Force Survey (LFS), Annual Population Survey (APS), Annual Survey of Household Earnings (ASHE) and the EU LFS, to name a few.

B.7. Inherent to all survey data is uncertainty, either through sampling variability or measurement error. The figures we present throughout the report that use survey data are subject to uncertainty. Readers should be aware that there are a range of values around the statistics and time series we report which the underlying survey data support. As we are interested in broad trends we do not overly emphasise these ranges unless they influence our interpretation.

Combining the LFS and APS

B.8. In a number of figures we present time series calculated from a combination of LFS (1997-2003) and APS (2004-2017) sources. The choice to combine the two datasets is not strictly recommended by the ONS, who advise using a consistent dataset across years. Equally combining LFS quarters is also not recommended due to individuals remaining in the LFS sample for five consecutive quarters. However, as we are interested in disaggregated series we wish to take advantage of the larger sample size of the APS. Equally we are interested in trends prior to the A8 accession (i.e. pre-2004), leaving us with a limited choice between using estimates from individual LFS quarters or from aggregated LFS quarters combined with APS estimates from 2004 onwards. We judge the later suits are needs better, with internal sense checking showing little difference between the two approaches.

[2] https://www.gov.uk/guidance/immigration-rules/immigration-rules-appendix-j-codes-of-practice-for-skilled-work

Annex C: Labour Market Impacts

Key Figures

C.1. *Figure 1.2* – Median hourly pay ("HOURPAY") for UK-born by occupational skill level using our quarter LFS averages between 1997 and 2003 and the APS between 2004 and 2017. Mappings between SOC2010 codes and RQF skill levels are taken from the "Immigration Rules Appendix J: codes of practice for skilled work"[3]. These mappings are in turn based on methodology developed by the MAC which takes account of earnings, qualifications and SOC skill level[4]. Proportional mappings were used to translate SOC 1990 and SOC 200 occupation codes into SOC 2010[5]. Values are adjusted by CPI and then indexed to 2004.

C.2. *Figure 1.4* – Regression analysis using linear probability model, estimated using 2017 APS. Adjusted estimates control for age, sex, region (GOVTOF) and age left full time education (EDAGE).

C.3. *Figure 1.5* – Employed UK-born ("ILODEFR" = 1) as a percentage of total UK-born population for those aged between 16 and 64 by highest qualification obtained ("HIQUAL"), using four quarter LFS averages between 1997 and 2003 and the APS between 2004 and 2017. "Other qualification" and "Don't Know" responses excluded.

Replication & extension of Dustmann, Fabbri and Peston (2005)

Replication

C.4. We start by replicating the main estimates from DFP (2005)[6] by regressing UK-born labour market outcomes (employment/unemployment/participation rates) on the immigrant/UK-born ratio, the average age of immigrants and UK-born workers, the relative UK-born worker skill supplies and year dummies using LFS data for 17 UK regions between 1983 and 2000. As in the original paper we use three and four-period lags of the immigrant/UK-born ratio as instruments. Table C.1 below reports the original and replicated coefficients estimated for the immigrant/UK-born ratio.

[3] https://www.gov.uk/guidance/immigration-rules/immigration-rules-appendix-j-codes-of-practice-for-skilled-work

[4] See Chapter 4 of MAC report "Analysis of the points-based system: Tier 2 at NQF level 6" for more details

[5] SOC mappings taken from https://github.com/dncnbrn/SOCmapping

[6] http://www.ucl.ac.uk/~uctpb21/Cpapers/ecoj_1038.pdf

Table C.1: Original and replicated DFP(2005) estimates				
	DFP specification by UK-born group			
Dependent variable	All	Adv. Education	Inter. Education	Low. Education
(1) Original DFP Estimates (1983-2000) – All immigrants				
Employment rate	-0.070	0.111	-0.179(***)	-0.028
Unemployment rate	0.066	0.001	0.098 (**)	-0.034
Participation rate	-0.035	0.108	-0.108 (**)	-0.063
(2) Replicated Estimates (1983-2000) – All immigrants				
Employment rate	-0.089	x	x	x
Unemployment rate	0.119	x	x	x
Participation rate	-0.021	x	x	x

All rows use Labour Force Survey data for those aged 16-64.
Regressions include time dummies, the average age of immigrants and UK-born workers, and the relative UK-born worker skill supplies.
Low education refers to no formal qualification; intermediate education to O-levels (or equivalent); and advanced education to A-levels or college/ university degrees.
Estimates uses third and fourth lag of immigrant ratio as instruments.
x – not estimated/replicated.
Statistical significance – (***) 1%, (**) 5%, (*) 10%

C.5. Our replicated estimates are very similar to the original estimates, giving us confidence to extend the original analysis.

Extension

C.6. We extend the original estimates by estimating the same model over a longer time period, 1997-2017. In addition we use instruments based on the 1991 share of immigrants rather than lags, as these generate more precise estimates. Finally, we also separately estimate the impact of EU and non-EU immigrants and further disaggregate the impact on the UK-born by looking at the impact on the young (16-24) and older (25-64) workers. The results of this are reported in Table 1.2 in the main report.

Replication & Extension of Dustmann, Frattini and Preston (2013)

Replication

C.7. We start by replicating the estimates from Table 4 of the original DFP (2013)[7] paper. This involves regressing changes to the percentiles of the UK-born wage distribution on changes in the fraction of immigrants to natives, additionally controlling for changes in the average age of immigrant and native workers, ratio of high (or intermediate) to low-educated native workers and time dummies for 17 UK regions between 1997 and 2005. The estimates resulting from using 1991 based instruments are reported in Table C.2 below.

Table C.2: Original and replicated DFP (2013) estimates

Dependent variable	Original Estimates	Replication
Impact of all immigrants on UK-born wage percentiles (1997-2005)		
5th	-0.353	-0.265
10th	-0.217	-0.111
25th	0.237 (***)	0.228 (***)
50th	0.409 (***)	0.386 (***)
75th	0.441 (***)	0.403 (***)
90th	0.299 (***)	0.299 (**)
95th	0.301	0.284

Estimates resulting from using 1991 immigrant shares as instrument.
Statistical significance – (***) 1%, (**) 5%, (*) 10%

C.8. Our replication efforts again match the original estimates very closely.

Extension

C.9. As with our extension to DFP (2005) we extend DFP (2013) by increasing the number of years of data used, increasing the sample by 12 years and covering the period 1997-2017. Again, the results of this extension are reported in Table 1.4 in the main report.

Robustness checks for DFP (2005) and DFP (2013)

[7] http://www.ucl.ac.uk/~uctpb21/Cpapers/Review%20of%20Economic%20Studies-2013-Dustmann-145-73.pdf

C.10. In addition to extending the period covered we also consider how robust the specifications used in the original papers are to changes in the variables controlled for and the estimation techniques used.

The baseline specification

C.11. In DFP (2013) the authors use the regional migrant share in 1991 as an instrument for the migrant share in their model. As the model is estimated in first-differences the growth rate of the outcome variable is uncorrelated with this time-invariant instrument, providing the identifying assumption.

C.12. DFP (2013) uses a version of Arellano-Bond in which the first-stage is allowed to be different in each year. We take this specification, using the time-invariant 1991 instrument and the Arellano-Bond estimation technique, as the baseline specification. We then estimate a further 12 specifications that deviate from this baseline in some way, separately reporting the coefficients for the all immigrant share and the EU immigrant share. These are reported in Tables C.3 and C.4.

C.13. In the subsections that follow we explain the alternative specifications, grouped into the nature of the deviation from the baseline.

Estimation Method and Choice of Instrument

C.14. An alternative to allowing the first stage estimation to be different in each year is to have it constant – this is reported as specification (2). Similarly we can estimate this specification using STATA's "ivregress" command rather than the Arellano-Bond approach – this is reported as specification (3).

C.15. Equally as an alternative to using the 1991 migrant share in a way that does not change over time, as has been criticised by Goldsmith-Pinkham, Sorkin and Swift (2018)[8], we can construct the predicated migrant share in each year by applying the growth rate in the aggregate stock of migrants to the initial share. Using this instrument and the "ivregress" command generates specification (4).

Region Trends

C.16. One advantage of using an instrument that varies over time is that it allows for the introduction of region trends without losing identification when estimating in first-differences. This is reported as specification (5).

Levels vs Differences

[8] http://www.nber.org/papers/w24408

C.17. DFP (2005) and DFP (2013) estimate their models in first-differences, which eliminates the time-invariant region fixed effects. An alternative to this approach is to estimate in levels and to include region fixed effects (and region trends if they are important). Estimates for the DFP model in levels with region fixed effects are reported in specification (6) while specification (7) additional adds-in region trends. The addition of region-specific trends changes many of the results. This is important because it suggests that it may be difficult to distinguish the impact of migration from other factors that trend over this period.

Computation of standard errors

C.18. Standard errors are important because they determine how precise the estimates are and, consequently, how confident we can be in our conclusions. In both DFP studies the standard errors are clustered by region, as is standard practice. If standard errors are not clustered each observation is treated as independent of every other. This is often not plausible. For example, in the current context, observations from the same region may well be correlated if there are unmodeled regional shocks. Clustering was introduced because of fears that reported standard errors were too low and, hence, estimates were too precise.

C.19. In specification (8) we show what happens to the standard errors of the estimates in the differenced model (specification 4) when they are not clustered. Similarly in specification (9) we show how the standard errors change in when not clustered at the region level for specification (6). It is worth noting that standard errors decrease without clustering in the levels specification but increase in the first-difference specification.

The specification of the migrant share variable

C.20. Both DFP papers use the migrant share as the relevant variables, as is standard in the literature. This has been recently criticised by Card and Peri (2016)[9] in their review of Borjas (2014)[10]. In their model it is the size of the labour force that influences labour market outcomes and immigration affects outcomes to the extent that it raises the size of the labour force. However, the size of the labour force is not necessarily well-captured by the immigrant share. If, for example, the native labour force falls, then the migrant share would rise but this would be associated with a fall in the labour force. Card and Peri

[9] Card, David and Peri, Giovanni, (2016), Immigration Economics by George J. Borjas: A Review Essay, *Journal of Economic Literature*, 54, 1333-49.

[10] Borjas, George J. (2014) "*Immigration Economics*" Cambridge, Mass: Harvard University Press.

suggest using the change in the labour force that can be ascribed to immigration as the measure of the impact of migration. Specification (10) shows what happens when this variable is used in specification (4). The Card-Peri critique is in the context of OLS estimates and it is not clear whether IV removes the problem.

C.21. If you take the Card and Peri model seriously then simply including the log of total labour supply as the relevant regressor and instrumenting using the predicted change using migrant share would address their critique of the standard approach. In specification (11) we do this to specification (4).

C.22. There may be reasons why the migrant share has an impact on native outcomes independent of the impact on total employment. This might happen if migrants and natives are not perfect substitutes as in Ottaviano and Peri (2008)[11] and Manacorda, Manning and Wadsworth (2010)[12]. But there is insufficient power to distinguish between the two effects in the data used here.

C.23. Another way in which we may wish to deviate from the use of the migrant share is to move away from using levels and to use the log of the migrant share instead – this is what specification (12) reports. Using this as the variable of interest results in a change in sign. The reason is a region like London has had the largest absolute change in the share of migrants but not the largest proportional change. The log of the migrant share would be closer to the "right" variable to use if you start with a CRS CES production function in which migrants and natives were the two inputs. However, there are some practical reasons to prefer use of the level as a rise of the migrant share from 1% to 2% might be expected to have the same impacts as a rise from 10% to 11% as implied by the level specification but not 10% to 20% as implied by the log specification.

Impact of the minimum wage

C.24. In our final specification (13) we include a variable based on a Kaitz index that is designed to capture the extent to which the minimum wage bites within a region. In regions where the minimum wage is closer to the average wage there is arguably greater potential for growth in labour supply to push the equilibrium wage below the minimum wage possibly generating greater impacts on the labour market along the extensive margin.

C.25. Table C.5 below gives a summary of each of the 13 specifications discussed above. The conclusion is that while the estimates from the original

[11] http://www.nber.org/papers/w14188.pdf

[12] http://personal.lse.ac.uk/manacorm/manacorda_manning_wadsworth.pdf

Table C.3: Summary of regression specifications used for robustness analysis

Specification	Independent variable	Estimation method	Instrument	Levels or first-differences?	Region fixed effects?	Region trends?	S.E's clustered at region?	Minimum wage variable?
Baseline specification								
(1)	Immigrants/Natives	Arellano-Bond	1991 share with time varying 1st-stage	First-differences	No	No	?	No
Estimation Method and Choice of Instrument								
(2)	Immigrants/Natives	Arellano-Bond	1991 share with constant 1st-stage	First-differences	No	No	?	No
(3)	Immigrants/Natives	ivregress	1991 share with constant 1st-stage	First-differences	No	No	Yes	No
(4)	Immigrants/Natives	ivregress	Time-varying predicted share (1991 base)	First-differences	No	No	Yes	No
Region fixed-effects and trends								
(5)	Immigrants/Natives	ivregress	Time-varying predicted share (1991 base)	First-differences	Yes	No	Yes	No
Estimation in levels								
(6)	Immigrants/Natives	ivregress	Time-varying predicted share (1991 base)	Levels	Yes	No	Yes	No
(7)	Immigrants/Natives	ivregress	Time-varying predicted share (1991 base)	Levels	Yes	Yes	Yes	No
Clustering of standard errors								
(8)	Immigrants/Natives	ivregress	Time-varying predicted share (1991 base)	First-differences	No	No	No	No
(9)	Immigrants/Natives	ivregress	Time-varying predicted share (1991 base)	Levels	Yes	No	No	No
Choice of independent variable								
(10)	Immigration contribution to total population change	ivregress	Time-varying predicted share (1991 base)	First-differences	No	No	Yes	No
(11)	Log of total population	ivregress	Time-varying predicted share (1991 base)	First-differences	No	No	Yes	No
(12)	Log of migrant share	ivregress	Time-varying predicted share (1991 base)	First-differences	No	No	Yes	No
Role of the minimum wage								
(13)	Immigrants/Natives	ivregress	Time-varying predicted share (1991 base)	First-differences	No	No	Yes	Yes

specifications stand up remarkably well when adding more data, these estimates are sensitive to whether controls for region-specific trends are included.

Table C.4: Estimates of the impact of EU migrants on UK-born labour market outcomes by model specification and UK-born group

Outcome	Native group	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Employment rate	All	-0.31*** [0.04]	-0.40*** [0.08]	-0.40*** [0.07]	0.15 [0.12]	0.31** [0.14]	-0.44 [0.30]	0.36 [0.28]	0.15 [0.16]	-0.44*** [0.10]	0.25 [0.21]	-0.35** [0.16]	0.16 [0.22]	0.27* [0.14]
	High-skill	-0.21*** [0.04]	-0.17* [0.10]	-0.17* [0.09]	-0.03 [0.18]	0.02 [0.21]	-0.26 [0.21]	0.10 [0.18]	-0.03 [0.18]	-0.26*** [0.08]	-0.05 [0.30]	-0.61** [0.26]	0.26 [0.39]	0.03 [0.20]
	Inter-skill	-0.56*** [0.06]	-0.95*** [0.13]	-0.95*** [0.08]	0.13 [0.20]	0.43* [0.25]	-0.50*** [0.16]	0.35 [0.33]	0.13 [0.32]	-0.50*** [0.09]	0.22 [0.34]	-0.29 [0.22]	0.20 [0.24]	0.45* [0.24]
	Low-skill	-0.34*** [0.08]	-0.22 [0.18]	-0.22 [0.22]	0.88* [0.48]	1.17** [0.54]	-0.52* [0.29]	0.24 [0.52]	0.88* [0.49]	-0.52*** [0.12]	1.48* [0.76]	0.16 [0.48]	-0.06 [0.28]	0.84 [0.53]
	Older	-0.17*** [0.04]	-0.16** [0.08]	-0.16*** [0.05]	0.15** [0.08]	0.25*** [0.09]	-0.31 [0.27]	0.28 [0.27]	0.15 [0.16]	-0.31*** [0.09]	0.26* [0.13]	-0.24** [0.11]	0.10 [0.13]	0.23** [0.10]
	Youth	-0.96*** [0.08]	-1.45*** [0.19]	-1.45*** [0.29]	0.06 [0.42]	0.51 [0.46]	-1.07** [0.43]	0.87*** [0.33]	0.06 [0.42]	-1.07*** [0.18]	0.10 [0.70]	-0.77* [0.41]	0.39 [0.58]	0.33 [0.42]
Unemployment rate	All	0.09*** [0.03]	0.29*** [0.06]	0.29*** [0.08]	-0.03 [0.07]	-0.13* [0.07]	0.25 [0.17]	-0.65*** [0.12]	-0.03 [0.11]	0.25*** [0.06]	-0.05 [0.12]	0.09 [0.08]	-0.03 [0.06]	-0.04 [0.08]
	High-skill	0.09*** [0.03]	0.24*** [0.06]	0.24*** [0.05]	0.02 [0.09]	-0.05 [0.12]	0.17 [0.12]	-0.36*** [0.08]	0.02 [0.11]	0.17*** [0.04]	0.03 [0.16]	0.20 [0.13]	-0.07 [0.13]	0.04 [0.11]
	Inter-skill	0.20*** [0.04]	0.42*** [0.10]	0.42*** [0.10]	0.03 [0.15]	-0.09 [0.19]	0.29** [0.14]	-0.81*** [0.20]	0.03 [0.23]	0.29*** [0.07]	0.06 [0.25]	0.15 [0.15]	-0.09 [0.14]	-0.07 [0.17]
	Low-skill	-0.25*** [0.08]	-0.20 [0.18]	-0.20 [0.15]	-0.09 [0.46]	-0.07 [0.56]	-0.03 [0.18]	-1.03*** [0.16]	-0.09 [0.44]	-0.03 [0.10]	-0.15 [0.77]	-0.21 [0.33]	0.13 [0.23]	0.03 [0.51]
	Older	0.08*** [0.02]	0.19*** [0.05]	0.19*** [0.06]	-0.01 [0.07]	-0.07 [0.08]	0.21 [0.14]	-0.53*** [0.11]	-0.01 [0.11]	0.21*** [0.05]	-0.02 [0.11]	0.06 [0.06]	-0.03 [0.05]	-0.03 [0.09]
	Youth	0.39*** [0.07]	0.90*** [0.16]	0.90*** [0.25]	-0.14 [0.45]	-0.47 [0.51]	0.58* [0.30]	-1.15*** [0.16]	-0.14 [0.38]	0.58*** [0.12]	-0.24 [0.76]	0.35 [0.34]	-0.10 [0.25]	-0.15 [0.44]
Participation rate	All	-0.28*** [0.03]	-0.23*** [0.07]	-0.23*** [0.05]	0.14 [0.10]	0.24* [0.12]	-0.30 [0.19]	-0.12 [0.33]	0.14 [0.14]	-0.30*** [0.07]	0.23 [0.18]	-0.30** [0.13]	0.15 [0.19]	0.26* [0.14]
	High-skill	-0.15*** [0.04]	0.02 [0.08]	0.02 [0.07]	-0.01 [0.17]	-0.02 [0.21]	-0.13 [0.13]	-0.22 [0.22]	-0.01 [0.17]	-0.13** [0.06]	-0.02 [0.29]	-0.46** [0.18]	0.22 [0.30]	0.07 [0.21]
	Inter-skill	-0.47*** [0.05]	-0.71*** [0.11]	-0.71*** [0.11]	0.19 [0.15]	0.43** [0.20]	-0.32*** [0.07]	-0.31 [0.36]	0.19 [0.30]	-0.32*** [0.07]	0.32 [0.26]	-0.20 [0.16]	0.15 [0.16]	0.46** [0.21]
	Low-skill	-0.52*** [0.08]	-0.29* [0.17]	-0.29 [0.25]	0.96* [0.52]	1.27** [0.58]	-0.54*** [0.17]	-0.46 [0.55]	0.96* [0.50]	-0.54*** [0.10]	1.60** [0.80]	0.09 [0.40]	0.00 [0.21]	0.97* [0.52]
	Older	-0.14*** [0.03]	-0.05 [0.07]	-0.05 [0.05]	0.15* [0.08]	0.21** [0.10]	-0.18 [0.18]	-0.11 [0.32]	0.15 [0.16]	-0.18*** [0.06]	0.25* [0.14]	-0.21* [0.11]	0.08 [0.11]	0.21* [0.12]
	Youth	-0.87*** [0.07]	-1.03*** [0.16]	-1.03*** [0.20]	0.04 [0.27]	0.35 [0.34]	-0.82*** [0.25]	0.01 [0.37]	0.04 [0.42]	-0.82*** [0.13]	0.06 [0.46]	-0.66** [0.33]	0.39 [0.52]	0.37 [0.36]

Statistical significance – (***) 1%, (**) 5%, (*) 10%

Table C.5: Estimates of the impact of migrants on the wages of the UK-born by model specification, position in the wage distribution and UK-born group

Native wage	Native group	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Impact of all migrants														
Mean	All	0.05 [0.04]	0.21** [0.10]	0.21*** [0.07]	-0.11 [0.10]	-0.79** [0.33]	0.09* [0.05]	-1.13** [0.55]	-0.11 [0.37]	0.09 [0.06]	-0.47 [0.46]	-0.20 [0.20]	0.11 [0.09]	-0.57* [0.31]
	High-skill	0.15*** [0.05]	0.10 [0.10]	0.10 [0.09]	0.22* [0.13]	0.51 [0.35]	0.17*** [0.06]	-0.33 [0.27]	0.22 [0.40]	0.17*** [0.06]	0.96* [0.58]	0.41 [0.26]	0.05 [0.10]	0.78** [0.33]
	Inter-skill	-0.06 [0.08]	0.07 [0.17]	0.07 [0.20]	-0.04 [0.22]	-0.52 [0.91]	0.03 [0.09]	-1.61 [1.24]	-0.04 [0.46]	0.03 [0.09]	-0.16 [0.95]	-0.07 [0.41]	0.16 [0.17]	-0.51 [0.95]
	Low skill	-0.62*** [0.15]	0.61* [0.32]	0.61*** [0.16]	-2.29*** [0.38]	-7.87*** [1.53]	-0.53*** [0.10]	-4.76*** [1.56]	-2.29 [2.02]	-0.53 [0.34]	-9.91*** [2.33]	-4.29*** [1.23]	0.33 [0.41]	-7.74*** [1.78]
	Older	0.00 [0.05]	0.17* [0.10]	0.17*** [0.05]	-0.19** [0.09]	-0.88*** [0.30]	0.04 [0.05]	-1.03* [0.62]	-0.19 [0.38]	0.04 [0.07]	-0.82** [0.38]	-0.35** [0.17]	0.08 [0.09]	-0.69** [0.29]
	Youth	-0.25*** [0.08]	0.18 [0.17]	0.18 [0.27]	-0.13 [0.29]	-1.10 [0.85]	-0.24* [0.12]	-2.85*** [0.71]	-0.13 [0.57]	-0.24*** [0.09]	-0.54 [1.25]	-0.23 [0.55]	0.11 [0.20]	-0.95 [0.86]
5th	All	-0.29*** [0.06]	-0.27** [0.13]	-0.27*** [0.07]	-0.20 [0.21]	0.06 [0.59]	-0.41*** [0.16]	0.18 [0.53]	-0.20 [0.39]	-0.41*** [0.07]	-0.87 [0.90]	-0.38 [0.41]	0.22 [0.18]	0.11 [0.61]
10th		-0.25*** [0.03]	-0.34*** [0.06]	-0.34*** [0.06]	-0.17 [0.11]	0.25 [0.27]	-0.32*** [0.08]	-0.19 [0.26]	-0.17 [0.23]	-0.32*** [0.04]	-0.74* [0.45]	-0.32 [0.21]	0.18** [0.08]	0.46* [0.25]
25th		-0.04 [0.03]	-0.06 [0.06]	-0.06 [0.06]	-0.05 [0.09]	0.01 [0.19]	-0.07 [0.06]	-0.69** [0.29]	-0.05 [0.23]	-0.07* [0.04]	-0.23 [0.38]	-0.10 [0.17]	0.08 [0.08]	0.12 [0.22]
50th		0.12*** [0.03]	0.15** [0.06]	0.15*** [0.05]	0.29* [0.15]	0.49 [0.35]	0.17*** [0.04]	-0.07 [0.19]	0.29 [0.22]	0.17*** [0.04]	1.24** [0.57]	0.54** [0.22]	0.09 [0.10]	0.58* [0.32]
75th		0.15*** [0.03]	0.19** [0.07]	0.19*** [0.05]	0.20*** [0.07]	0.23 [0.29]	0.21*** [0.05]	-0.44** [0.18]	0.20 [0.22]	0.21*** [0.04]	0.85** [0.34]	0.37*** [0.14]	-0.05 [0.12]	0.37 [0.27]
90th		0.27*** [0.04]	0.24*** [0.08]	0.24** [0.12]	0.16 [0.12]	0.07 [0.33]	0.33*** [0.06]	-0.16 [0.48]	0.16 [0.31]	0.33*** [0.05]	0.70 [0.59]	0.30 [0.26]	-0.04 [0.12]	0.38 [0.32]
95th		0.21*** [0.05]	0.06 [0.11]	0.06 [0.13]	0.20 [0.13]	0.43 [0.50]	0.31*** [0.09]	0.01 [0.57]	0.20 [0.39]	0.31*** [0.07]	0.87 [0.63]	0.38 [0.28]	0.04 [0.14]	0.84* [0.45]
Impact of EU migrants														
Mean	All	0.04 [0.12]	0.55** [0.24]	0.55*** [0.19]	-0.39 [0.32]	-0.90** [0.43]	0.24* [0.12]	-2.48*** [0.49]	-0.39 [0.59]	0.14 [0.16]	-0.66 [0.53]	-0.51 [0.40]	0.27 [0.62]	-0.59 [0.47]
	High-skill	0.28** [0.13]	0.23 [0.25]	0.23 [0.22]	0.18 [0.38]	0.19 [0.55]	0.45*** [0.17]	-2.03*** [0.59]	0.18 [0.61]	0.35** [0.16]	0.30 [0.65]	0.23 [0.51]	-0.45 [0.77]	0.65 [0.60]
	Inter-skill	-0.27 [0.22]	0.26 [0.44]	0.26 [0.61]	-1.45* [0.78]	-2.42 [1.48]	0.08 [0.25]	-3.38*** [1.13]	-1.45 [0.91]	-0.06 [0.26]	-2.44** [1.20]	-1.88** [0.80]	0.45 [1.11]	-2.72* [1.62]
	Low skill	-1.38*** [0.40]	1.60** [0.80]	1.60*** [0.46]	-0.86 [0.70]	-2.14* [1.12]	-1.43*** [0.29]	-2.39 [1.60]	-0.86 [4.38]	-1.31 [0.86]	-1.45 [1.13]	-1.12 [0.82]	3.53 [6.92]	-1.81* [0.98]
	Older	-0.09 [0.12]	0.43* [0.24]	0.43*** [0.14]	-0.42 [0.28]	-0.85** [0.38]	0.11 [0.14]	-2.22*** [0.58]	-0.42 [0.62]	0.01 [0.16]	-0.70 [0.45]	-0.54* [0.32]	0.40 [0.90]	-0.59 [0.41]
	Youth	-0.60*** [0.21]	0.57 [0.42]	0.57 [0.76]	0.11 [0.74]	-0.30 [0.98]	-0.65* [0.33]	-1.88* [1.14]	0.11 [0.97]	-0.52** [0.24]	0.18 [1.24]	0.14 [0.95]	0.23 [0.75]	0.00 [1.12]
5th	All	-0.78*** [0.16]	-0.73** [0.31]	-0.73*** [0.23]	0.13 [0.57]	0.60 [0.84]	-1.11*** [0.41]	0.80 [0.65]	0.13 [0.73]	-0.99*** [0.20]	0.21 [0.95]	0.16 [0.73]	0.28 [0.73]	0.80 [0.84]
10th		-0.73*** [0.08]	-0.88*** [0.16]	-0.88*** [0.16]	-0.63** [0.25]	-0.47 [0.35]	-0.87*** [0.22]	-0.48 [0.34]	-0.63 [0.42]	-0.85*** [0.11]	-1.06** [0.43]	-0.82** [0.39]	0.50 [1.04]	-0.16 [0.38]
25th		-0.23*** [0.08]	-0.23 [0.15]	-0.23 [0.15]	-0.50** [0.20]	-0.61** [0.26]	-0.18 [0.17]	-1.95*** [0.24]	-0.50 [0.52]	-0.26** [0.11]	-0.84*** [0.33]	-0.65*** [0.25]	0.17 [0.45]	-0.49 [0.31]
50th		0.24*** [0.08]	0.33** [0.16]	0.33*** [0.12]	0.18 [0.28]	0.08 [0.42]	0.47*** [0.09]	-1.57*** [0.37]	0.18 [0.38]	0.36*** [0.12]	0.30 [0.47]	0.23 [0.37]	-0.49 [1.10]	0.24 [0.48]
75th		0.32*** [0.09]	0.43** [0.18]	0.43*** [0.14]	0.51* [0.28]	0.57 [0.43]	0.56*** [0.12]	-1.74*** [0.24]	0.51 [0.36]	0.44*** [0.11]	0.86* [0.48]	0.67* [0.39]	-0.46 [0.88]	0.88** [0.44]
90th		0.66*** [0.10]	0.58*** [0.21]	0.58* [0.30]	0.15 [0.43]	-0.02 [0.54]	0.89*** [0.15]	-1.28** [0.60]	0.15 [0.44]	0.78*** [0.14]	0.26 [0.74]	0.20 [0.59]	-0.50 [0.80]	0.48 [0.68]
95th		0.47*** [0.14]	0.21 [0.27]	0.21 [0.29]	-0.72 [0.64]	-1.19 [0.91]	0.85*** [0.23]	-2.23** [1.09]	-0.72 [0.66]	0.71*** [0.21]	-1.21 [1.03]	-0.93 [0.71]	-0.27 [0.39]	-0.68 [1.00]

Statistical significance – (***) 1%, (**) 5%, (*) 10%

Earnings of the self-employed

C.26. In this annex we provide some more detail on the datasets used, and the analysis carried out, in Chapter 1 covering the self-employment earnings by nationality.

C.27. We would like to thank HMRC, particularly the Knowledge, Analysis & Intelligence Personal Taxes team, for their help and support with this project.

HMRC Self-Assessment data

C.28. The registered self-employed are required to submit a Self-Assessment (SA) tax return detailing a number of key financial details about their activities[13].

C.29. Those who have earned income through self-employment need to fill-in the SA103 pages of this tax return (or the online equivalent). Those with turnover above the VAT threshold must file the full return (SA103F[14]) and those below can fill in a short return (SA103S[15]). A SA103S/F return must be provided for each business an individual is declaring. In addition, if an individual also received income from employment they must detail this in the SA102 pages of the SA return. Alternatively, those self-employed with turnover below a certain threshold can be asked by HMRC to submit a short SA return (SA200[16]), which is a four-page simplified paper return. This cannot be filled out online.

C.30. Another way for an individual to be self-employed is by being a partner in a Business Partnership[17]. A Business Partnership is business that is owned and operated by several individuals who share in the business profits/loses. A tax return must be submitted for the Partnership as a whole but the individual partners must also submit individual tax returns (SA104F[18]/S[19]) detailing their share of the Partnerships profits/loses.

C.31. Finally there are a very small number of underwriting members of Lloyd's of London who we also include in our definition of self-employment and who are obliged to fill out and return SA103L forms detailing their Lloyd's activities.

C.32. Our final dataset includes earnings information from all of these sources (SA103F/S/L, SA200, SA104F/S). Individuals can submit multiple SA103 or SA104 forms (or online equivalents), but as we are interested in individuals rather than businesses we collapse this data over individual/tax-year combinations to produce a single observation for each individual who submitted a relevant SA return in a given tax-year.

---

[13] https://www.gov.uk/self-assessment-tax-returns/who-must-send-a-tax-return
[14] https://www.gov.uk/government/publications/self-assessment-self-employment-full-sa103f
[15] https://www.gov.uk/government/publications/self-assessment-self-employment-short-sa103s
[16] https://www.gov.uk/government/publications/self-assessment-short-tax-return-notes-sa210
[17] https://www.gov.uk/set-up-business-partnership
[18] https://www.gov.uk/government/publications/self-assessment-partnership-full-sa104f
[19] https://www.gov.uk/government/publications/self-assessment-partnership-short-sa104s

C.33. Other minor details about this dataset that are worth being mindful of are the way the region and industry variables are derived. The dataset we used only contained the latest and the previous postcodes associated with an individual and so there is the possibility that some observations, particularly earlier observations, are associated with a non-contemporaneous postcode. The industry variable is generated using from the self-reported "business description", relying on an Automated Classification Text Recognition tool to match descriptions to a 5 digit SIC 2007 code. Where an individual has multiple SA returns, the business description from the return reporting the highest earnings is used. This matching process is only available from 2001/02 onwards with the proportion of matches between descriptions and SIC codes starting at around 70 per cent, rising to around 98 per cent between 2007/08 and 2010/11 before falling to back to around 90-85 per cent for the rest of the period to 2016/17.

C.34. An important caveat which applies to the dataset we use is that it does not include those individuals who structure their self-employment through an incorporated company and pay themselves either as an employee or via dividends. This may be important as there is some evidence that the numbers of individuals doing so has increased over time. For example the OBR, as part of their November 2016 Economic and Fiscal Outlook report, published a chart[20] showing substantial growth in the number of companies with only one director between 2006/07 and 2013/14, and a fall in the numbers of companies with two or more directors. There is also evidence that the value of this activity is not trivial. In 2017, the ONS started using HMRC data to derive their estimates of the amount of dividend income households receive. This resulted in a significant increase in the amount of dividends it was estimated that households were receiving, from around £16 billion to £76 billion in 2015/16, which is at least in part associated with a better capturing of self-employment income[21]. There is little we can do about this omission other than bear it in mind when reviewing any outputs from this data.

C.35. Finally as our dataset relies on income declared for tax purposes there is always the issue of misreporting. HMRC estimates that the overall SA tax gap, the difference between tax due and tax paid, was 16.4% (£7.9 billion) in 2016/17. HMRC also noted that those with self-employment income were responsible for the majority of the SA tax gap (£5.1 billion)[22]. We do not attempt to adjust for misreporting.

---

[20] http://obr.uk/download/economic-and-fiscal-outlook-charts-and-tables-fiscal-november-2016/ - Chart C4.A

[21] http://obr.uk/download/forecast-evaluation-report-charts-tables-october-2017/ - Chart C2.A, also see para 1.12 from OBR Nov 2017 EFO (http://cdn.obr.uk/Nov2017EFOwebversion-2.pdf). The extent of the difference between the two estimates may also reflect differences in the extent of capturing dividend income brought forward into 2015/16 in advance of dividend taxation rate rises.

[22] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/715742/HMRC-measuring-tax-gaps-2018.pdf

Migrant Workers Scan (MWS)

C.36. A National Insurance number (NINo) is usually sent to children in the UK shortly before their 16th birthday, but adult overseas nationals resident in the UK[23] can also apply for a NINo. Information on those who do apply and are allocated a NINo is held on the Migrant Worker Scan (MWS)[24].

C.37. The MWS includes data on the nationality at point of registration of applicants, as well as their date of registration and their date of arrival in the UK. We use this dataset to identify migrants within the HMRC SA data by matching records using the NINo.

C.38. Not all self-employed immigrants, defined by nationality, in the UK will have applied for a NINo – at least not as an adult. For example, if Child Benefit is claimed on behalf of a foreign national resident child, that child will be automatically allocated a NINo at the age of 15 and so will not appear on the MWS. The MWS does not record whether an individual holds more than one nationality, nor in most cases when a change in nationality occurs after the allocation of the NINo. The MWS is also subject to accurate collection and input of data during the registration process, for example there are individuals in the MWS where the country of nationality is missing or is unspecified. This results in a group of observations from the SA data for which no particular nationality can be assigned, this group represents around 3.5% of all observations in 1996/7 before falling to around 0.5% by the end of the series in 2016/17. Furthermore, time series of NINo registrations produced using the MWS tend to focus on the period from 2002 onward due to concerns that the data prior may be incomplete. We use the full dataset to make sure we can match as many SA records as possible.

C.39. The EU13, NMS and Non-EEA nationality groups we use are the same as the country of birth groups outlined earlier in Annex Y. Ultimately nationality is not our preferred measure of whether an individual is an immigrant as it can change over time.

C.40. We categorise individuals present in the Self Assessment datasets but who are not present in the MWS as UK nationals, and then also include Irish nationals identified in the MWS to create our "native" group – which we simply refer to as UK nationals in the main text. As stated above, absence from the MWS does not automatically imply an individual is a UK national however it is likely that the majority are.

---

[23] Some cases of applications prior to arrival to the UK are allowed.

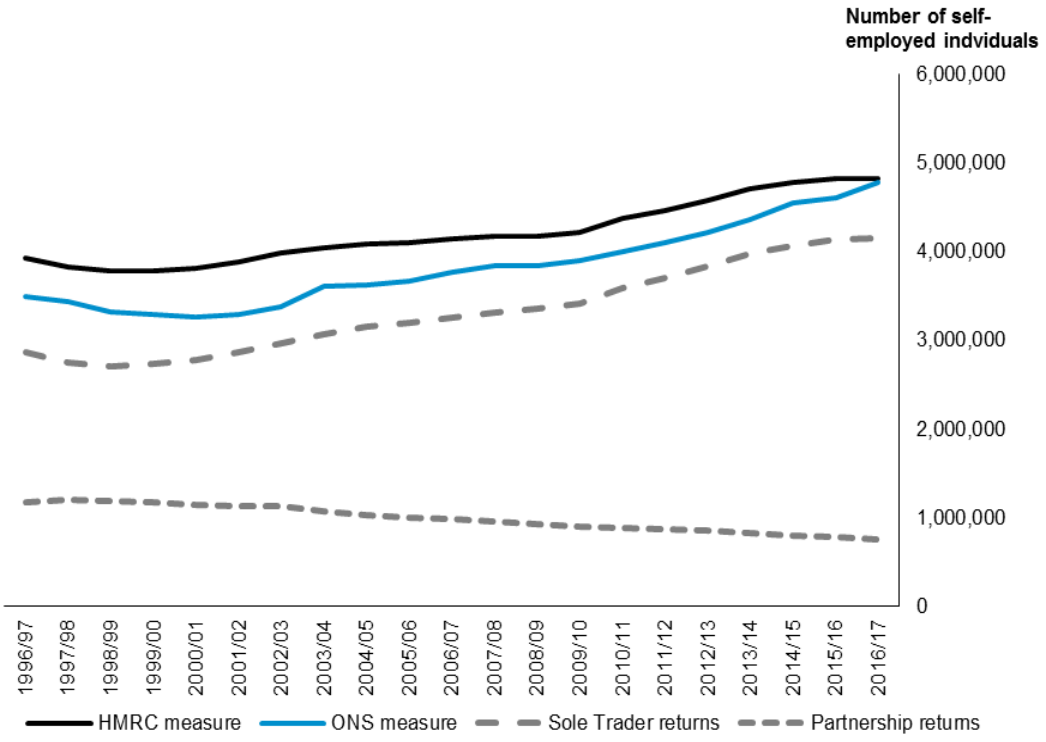
[24] See MWS user guide for a more in depth discussion of methodology and limitations - https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/615454/nino-allocations-background-information.pdf

Comparison with other data sources

C.41. Labour Force Survey already provides information on the numbers of types of immigrants who are self-employed. In the charts below we compare the total number of self-employed individuals by nationality from the latest published LFS figures[25] and from the dataset we have created using the SA and MWS data. We use a four quarter (Q2-Q1) average of the LFS to match the tax year nature of the SA data.

C.42. Figure C.6 shows the total numbers of self-employed individuals from both the LFS and SA data, as well as a breakdown of the SA data between number of individuals submitting sole-trader returns (SA103F/S and SA200) and those submitting Business Partnership returns (SA104F/S)[26]. The LFS measure shows fewer self-employed individuals than the SA data for most of the period covered, however by the end of the period the total numbers are very similar – at around 4.8 million individuals.

Figure C.6: Number of self-employed individuals – MAC analysis of HMRC vs ONS measures



Source: MAC analysis of HMRC data, ONS "Summary of Labour Market Statistics" 17th July 2018

[25] Table 3 of ONS publication "UK labour market; July 2018".

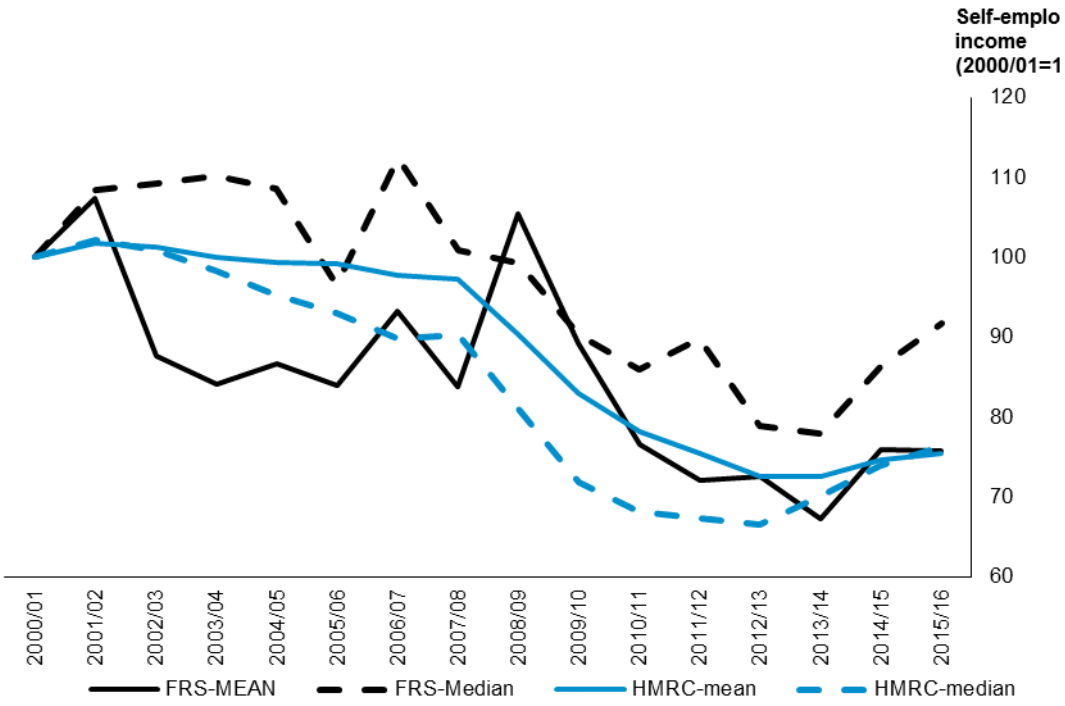
[26] There is a small overlap between these groups with 2.5% of observations over the full period recorded has having submitted both types of return.

C.43. While the Family Resources Survey (FRS) does provide some time series information on self-employment earnings the statistics presented in publications are given as average weekly earnings. It is not obvious that this can be directly compared to the total earnings for a given tax year. Instead, in Figure C.7, we compare the evolution of mean and median self-employment income (adjusted by CPI) from the FRS and HMRC data[27] indexed to 2000/01.

C.44. Comparisons between the two series are sensitive to the choice of base year and the FRS measure is more variable that our measure. However there is a broad agreement in the direction of self-employment income over time, which is reassuring.

C.45. Ultimately much more work could be undertaken to refine and validate this new dataset against existing ones, however there appears to be a broad agreement in high-level aggregate trends between datasets.

Figure C.7: Mean and median self-employment income – MAC analysis of HMRC vs Resolution Foundation analysis of Family Resources Survey



Source: MAC analysis of HMRC data, Resolution Foundation Earnings Outlook

C.46. The "adjusted" differentials between immigrant and "native" mean self-employment earnings was calculated using a linear regression model, estimated by Ordinary Least Squares. The variables of interest where three nationality dummies, one for each migrant group. Additional covariates introduced for the "adjusted" regression included sex, age, travel to work area,

[27] Sample is not restricted to 16-64 year olds as elsewhere in order to be more comparable with FRS series.

a variable indicating whether an individual also has employment income, a variable indicating whether an individual has declared positive employee costs, the year the individual first appears in the dataset and a variable measuring the number of years an individual appears in the dataset (across 1996/97 to 2016/17). All covariates were estimated as a series of binary (dummy) variables, as such the number of parameters are too numerous to report.

Regressions on zero hours contracts, union representation and recruitment

C.47. Analyses use the LFS and the APS and the migrant classification described in B.2. The sample year is 2017. Information on LFS and APS variables can be found in the LFS User Guide vol. 3, which is updated annually[28].

C.48. *Figure 1.11* – Regression analysis on the proportion of workers on zero hour contracts by migrant group using linear probability model. The dependent variable is flexw7 (whether respondent works zero hours contract), estimated using 2017 LFS. Adjusted estimates control for occupation, industry, region, age, gender, job tenure and education level.

C.49. *Figure 1.12* – Regression analysis on union representation by migrant group. The dependent variable is union (whether respondent is a member of a trade union or staff association), estimated using pooled years of the LFS from 2011-2017. Adjusted estimates control for occupation, industry, region, age, gender, job tenure and education level.

C.50. *Table 1.5* – The dependent variable (recruitment regression results) for each is a dummy variable for each different method of recruitment created from the howget variable. The controls in the regressions are occupation, industry, region, age, gender, job tenure and education level.

Regression on labour mobility

C.51. *Figure 1.13* – This uses the 2-quarter LFS with 5 rolling quarters pooled to increase sample size. We tracked respondents across quarters to find the proportion of people who moved industry. We used a linear probability model using the proportion of people who changed industry over the quarters. The controls used in this regression are also occupation, industry, region, age, gender, job tenure and education level.

[28] https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/methodologies/labourforcesurveyuserguidance

Annex D: Productivity, Innovation, Investment and Training Impacts

Key Figures

D.1. *Figure 2.1* – Gross Value Added (GVA) per hour of work, taken from ONS publication "Labour productivity, UK: January to March 2018".

D.2. *Figure 2.2* – Gross Domestic Product (GDP)per hour worked for G7 countries compared to the UK. Taken from latest ONS International Comparisons of Productivity publication.

D.3. *Figure 2.3* – Gross fixed capital formation as a percentage of gross domestic product, UK compared with Organisation for Economic Co-operation and Development nations, 1997 to 2017. Taken from ONS publication "An international comparison of gross fixed capital formation" published 2nd November 2017.

D.4. *Figure 2.4* – Proportion of working aged (16-64) individuals in work ("ILODEFR"=1) who report having received job related training or education in the last 3 months ("ED13WK"=1) by Country of Birth groups. Four quarter LFS averages used between 1997 and 2003 and the APS is used between 2004 and 2017.

D.5. *Figure 2.5* – Participants in Continuous Vocational Training courses as a per cent of persons employed in all enterprises. Taken from Continuous Vocational Training Survey 2015. Data download from Eurostat table "Participants in CVT courses by sex and size class - % of persons employed in all enterprises (trng_cvt_12s)".

D.6. *Figure 2.6* – UK share of global innovation outputs/inputs compared to UK share of global population. Selected estimates from Table 1.3 of "International Comparative Performance of the UK Research Base 2016"[29].

[29] https://www.elsevier.com/__data/assets/pdf_file/0018/507321/ELS-BEIS-Web.pdf

Annex E: Prices Impacts

Consumer Prices

E.1. There is a very limited literature on the impact of immigration on consumer prices. The best-known paper is probably Cortes (2008)[30] who investigates the effect of low-skilled immigration on US prices. She finds significant though modest impacts. Frattini (2014)[31] tried a similar approach for the UK and found little evidence of any impact of migration on prices.

E.2. A major difficulty in this area is to link elements of the CPI to some measure of migrant intensity in the production of those goods and services. Both Cortes and Frattini use industry and we also try this approach. But we also use occupation as this can also be tightly linked to some consumer prices in some cases. Both these approaches only naturally lend themselves to the analysis of the price of non-traded goods and services.

E.3. The consumer price data comes from the individual quote data that underlies the published CPI indices[32]. Each price quote has an identifier for the item e.g. "hiring a plumber for an hour" as well as some information about the origin of the price quote e.g. the type of store and the location. For traded goods one would not expect price to reflect local variation but for non-traded it will.

E.4. The price quote data also records the government office region (of which there are 12) so there is also regional variation.

E.5. We use the APS (after 2004) and the LFS (prior to 2004) to estimate the fraction of migrants by region by occupation by year. We then regress the mean log price by item by region by year on the migrant share for the linked occupation together with some fixed effects. In our baseline specification we include item, region and year fixed effects.

E.6. The basic specification is reported in column 1 of Table E.1 where the overall migrant share is the regressor. There is a significant negative effect of the

---

[30] Cortes, Patricia. (2008) "The Effect of Low-Skilled Immigration on U.S. Prices: Evidence from CPI Data." *Journal of Political Economy*, 116, 381–422.

[31] Tommaso Frattini "Impact of Migration on UK Consumer Prices", https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/328006/Impact_of_migration_on_UK_consumer_prices__2014.pdf

[32] Office for National Statistics. (2018). *Prices Survey Microdata, 1996-2018: Secure Access*. [data collection]. *8th Edition.* UK Data Service. SN: 7022, http://doi.org/10.5255/UKDA-SN-7022-8

migrant share on prices. One concern with this OLS estimate is that the migrant share is endogenous – it might be that migrants move to booming areas where there is upward pressure on prices. In this case the estimated effect will be too small. We use the standard shift-share instrument taking the number of migrants in each occupation-region cell in our baseline years of 1994-1996 and then applying aggregate growth in the number of migrants to work out a predicted migrant share. The first-stage is strong as shown in Table E.2. The results are shown in column 2 of Table E.1. As one might expect the estimated coefficient is more negative and is large in magnitude. Although it is significantly different from zero, the standard error is large. But the IV estimate also suggests that more migrants leads to lower prices for these services.

E.7. We next investigate the impact of different types of migrants, differentiating between EU15+, EEA migrants, NMS migrants (from the new member states of the EU) and non-EEA. In columns 3-8 of Table E.1 we investigate the migrant share of each group individually using both OLS and IV. The EU15 first stage is weak and there are no significant effects. For NMS migrants the first stage is weak and a very large though insignificant effect is found. For non-EEA migrants there seems a sizeable negative effect. Next, we include all 3 migrant groups – the first stages do not work well for all: again it seems there is a negative effect from non-EEA. Finally, we combine NMS and non-EEA migrants as a single group: the instrument is strong in this case and the IV estimates similar to that found when all migrant groups are combined (columns 1 and 2).

E.8. Table E.3 investigates the robustness of the results with Table E.4 showing the corresponding first stages for the IV estimates. The first column includes fixed effects for item*region a more demanding specification. Results are similar. Next we investigate whether there are differences by skill level, estimating models for professional and associate professionals, craft trades, personal services and manual trades. We find no effect for the prices of the higher-level services but larger effects for medium and low skill.

E.9. Our next set of estimates use industry rather than occupation as in Cortes (2008) and Frattini (2014). The methodology is similar to that followed in the occupational approach – identify items in the CPI that can be thought of as products of a single industry and use the migrant share at industry-region-year level as the variable of interest. The products matched to an industry are different from those matched to an occupation. For example, it seems natural to match “hiring a plumber for an hour” to the occupation of “plumber” but a stretch to match it to “construction” which is what one would have to do if one sought to match it to an industry. But there are also some products that are more naturally matched to an industry than a particular occupation. For example, many of the products matched in this part of the analysis are menu

items from restaurants – these are matched to the restaurant industry but use the inputs of chefs and wait-staff among others.

E.10. Table E.5 shows the results of specifications that are similar to those reported in Table E.1 though where products are matched to industry rather than occupation. The results are, however, similar: there is some evidence that non-EEA and NMS migrants reduce prices. Because so many products come from the same industry, one might want to cluster on industry rather than product – this has only small effects on the standard errors. The first stages remain strong even when region trends are introduced.

Table E.1: OLS and IV estimates of the Impact of Migrant Share on Consumer Prices: Dependent variable log prices												
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
All Migrant Share	-0.116	-0.567										
	[0.047]	[0.232]										
EU15 Migrant Share			-0.080	-3.035					-0.093	3.716		
			[0.051]	[7.473]					[0.050]	[7.925]		
NMS Migrant Share					0.012	-3.949			0.006	-2.893		
					[0.090]	[2.294]			[0.088]	[2.285]		
non-EEA Migrant Share							-0.160	-0.403	-0.162	-0.514		
							[0.057]	[0.206]	[0.057]	[0.396]		
non-EEA+NMS Migrant Share											-0.119	-0.590
											[0.053]	[0.234]
Observations	10,410	10,410	10,410	10,410	10,410	10,410	10,410	10,410	10,410	10,410	10,410	10,410
R^2	0.620		0.618		0.618		0.620		0.620		0.620	
Number of item_id	68	68	68	68	68	68	68	68	68	68	68	68

Table E.2: First stages for IV estimates of Table E.1								
VARIABLES	(1) All Migrant Share	(2) EU15 Migrant Share	(3) NMS Migrant Share	(4) non-EEA Migrant Share	(5) EU15 Migrant Share	(6) NMS Migrant Share	(7) non-EEA Migrant Share	(8) NMS+non-EEA Migrant Share
Predicted All Migrant Share	0.357							
	[0.065]							
Predicted EU15 Migrant Share		0.030			0.030	0.062	0.042	
		[0.045]			[0.044]	[0.064]	[0.068]	
Predicted NMS Migrant Share			0.059		0.008	0.060	0.195	
			[0.020]		[0.017]	[0.019]	[0.044]	
Predicted non-EEA Migrant Share				0.441	0.027	0.020	0.450	
				[0.092]	[0.027]	[0.024]	[0.092]	
Predicted NMS+non-EEA Migrant Share								0.366
								[0.063]
Observations	10,410	10,410	10,410	10,410	10,410	10,410	10,410	10,410
R^2	0.514	0.075	0.154	0.465	0.077	0.155	0.476	0.500
Number of item_id	68	68	68	68	68	68	68	68

Table E.3: Robustness Checks: Dependent variable log prices

VARIABLES	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV	(7) OLS	(8) IV	(9) OLS	(10) IV
			High-Skill		Medium-Skill (craft)		Medium-Skill (service)		Low-skill	
All Migrant Share	-0.023	-0.806	-0.067	-0.175	-0.134	-0.796	-0.138	-0.934	-0.202	-0.842
	[0.027]	[0.370]	[0.038]	[0.807]	[0.086]	[0.383]	[0.080]	[0.297]	[0.145]	[0.488]
Observations	10,410	10,410	1,876	1,876	3,315	3,315	3,339	3,339	1,880	1,880
R^2	0.761		0.713		0.565		0.732		0.566	
Number of id	815	815								
Number of item_id			9	9	27	27	21	21	11	11

Table E.4: First Stages for Table E.3

VARIABLES	(1) All Migrant Share	(2) All Migrant Share	(3) All Migrant Share	(4) All Migrant Share	(5) All Migrant Share
		High-Skill	Medium-Skill (craft)	Medium-Skill (service)	Low-skill
Predicted All Migrant Share	0.313	0.142	0.410	0.27	0.678
	[0.086]	[0.104]	[0.077]	[0.109]	[0.102]
Observations	10,410	1,876	3,315	3,339	1,880
R^2	0.108	0.269	0.589	0.696	0.641
Number of id	815				
Number of item_id		9	27	21	11

Table E.5: OLS and IV estimates of the Impact of Migrant Share on Consumer Prices: Industry Analysis. Dependent variable log prices

VARIABLES	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV	(7) OLS	(8) IV	(9) OLS	(10) IV	(11) OLS	(12) IV	(13) OLS	(14) IV
All Migrant Share	0.180	0.859												
	[0.036]	[0.207]												
EU15 Migrant Share			-0.197	-1.311					-0.208	-0.648				
			[0.071]	[0.844]					[0.072]	[5.113]				
NMS Migrant Share					-0.118	-13.750			-0.136	-24.876				
					[0.076]	[10.398]			[0.076]	[70.517]				
non-EEA Migrant Share							-0.183	-0.807	-0.187	1.617				
							[0.038]	[0.245]	[0.038]	[7.779]				
non-EEA+NMS Migrant Share											-0.175	-0.949	-0.179	-1.002
											[0.039]	[0.240]	[0.038]	[0.253]
Observations	13,043	13,043	13,043	13,043	13,043	13,043	13,043	13,043	13,043	13,043	13,043	13,043	13,043	13,043
R-squared	0.635		0.632		0.632		0.634		0.636		0.635		0.635	
Number of item_id	83	83	83	83	83	83	83	83	83	83	83	83	83	83

Table E.6: First-stages for IV estimates in Table E.5

VARIABLES	(1) All Migrant Share	(2) EU15 Migrant Share	(3) NMS Migrant Share	(4) non-EEA Migrant Share	(5) EU15 Migrant Share	(6) NMS Migrant Share	(7) non-EEA Migrant Share	(8) NMS+non-EEA Migrant Share	(9) NMS+non-EEA Migrant Share
Predicted All Migrant Share	0.334								
	[0.032]								
Predicted EU15 Migrant Share		0.227			0.202	0.013	0.135		
		[0.052]			[0.050]	[0.034]	[0.079]		
Predicted NMS Migrant Share			0.013		-0.008	0.023	0.176		
			[0.011]		[0.008]	[0.011]	[0.017]		
Predicted non-EEA Migrant Share				0.326	0.037	0.034	0.354		
				[0.045]	[0.009]	[0.010]	[0.046]		
Predicted NMS+non-EEA Migrant Share								0.308	0.305
								[0.033]	[0.031]
Observations	13,043	13,043	13,043	13,043	13,043	13,043	13,043	13,043	13,043
R-squared	0.607	0.193	0.176	0.534	0.199	0.179	0.543	0.567	0.575
Number of item_id	83	83	83	83	83	83	83	83	83

House Prices

E.11. This section contains an empirical investigation of the impact of migration on house prices. It seeks to replicate and extend the work of Sa[33] who used data for 170 English local authorities for the period 2003-2010. We extended this in the time dimension to the period 2001-2016, and in the cross-sectional dimension by using district and not county councils, and by adding Welsh councils (for a total of 346 local authorities). The sample size is increased from 1,190 to 5,536.

E.12. The basic equation estimated by Sa is of the form:

$$\Delta \ln\left(HP_{it}\right) = \beta_1 \frac{\Delta FB_{it}}{POP_{it-1}} + \beta_2 x_{it} + d_t + \theta_i + \varepsilon_{it}$$

E.13. A regression of the change in log house prices on the change in the foreign-born population as a fraction of the initial population, other regressors, year effects and local authority fixed effects. The latter effects imply systematic differences in house price growth across local authorities that would imply ever-widening price differentials across areas. Sa reports that results are similar when these fixed effects are removed though does not present those estimates.

E.14. Table E.5 reports some results for specifications similar to those reported by Sa.

E.15. Column (1) shows the estimates of the equation when fixed effects are omitted, column (2) when region fixed effects are included and column (3) when LA fixed effects are included. In all 3 specifications the impact of migration on house prices is small and insignificantly different from zero. But it may be the case that migration into an area is correlated with unobserved factors e.g. because they move to areas where housing is relatively cheap or to areas where economic opportunity is growing.

[33] http://ftp.iza.org/dp5893.pdf

Table E.7: OLS and IV estimates of impact of migration on house prices						
	(1)	(2)	(3)	(4)	(5)	(6)
Change in Migrant Share	0.011	-0.004	-0.009	0.923(***)	0.302	-0.371
	[0.015]	[0.016]	[0.016]	[0.150]	[0.263]	[0.387]
Constant	0.069(***)	0.064(***)	0.069(***)	0.062(***)	0.066(***)	0.072(***)
	[0.002]	[0.002]	[0.002]	[0.003]	[0.002]	[0.004]
Observations	5,536	5,536	5,536	5,536	5,536	5,536
R^2	0.791	0.794	0.796			
Fixed Effects	No	Region	LA	No	Region	LA
Method	OLS	OLS	OLS	IV	IV	IV

E.16. For these reasons, Sa instruments the change in the migrant share using predicted change based on the composition of the area in 2001 and subsequent national growth. Her IV estimates were large and negative leading to the conclusion that migration had reduced house prices. Our estimates are different – the estimate without any fixed effect implies an elasticity of house prices with respect to migration of about 1. The estimates with region and local authority fixed effects differ in sign but are both insignificantly different from zero. One problem here is that the instrument may be weak, a problem that is known to make IV estimates unreliable. First-stage results are reported in Table E.6.

Table E.8: First stage results: Dependent Variable: Change in Migrant Share			
VARIABLES	*(1)*	*(2)*	*(3)*
Migrant Share Instrument	0.788 (***)	0.264 (*)	0.455 (***)
	[0.158]	[0.142]	[0.171]
Constant	0.003	0.004 (**)	0.005 (**)
	[0.002]	[0.002]	[0.002]
Observations	5,536	5,536	5,536
R^2	0.013	0.018	0.006
Fixed Effects	None	Region	LA

Statistical significance – (***) 1%, (**) 5%, (*) 10%

E.17. Only without any fixed effects does the instrument pass the test for being strong enough to be reliable. The problem is that predicted change in migrant share is close to a trend in regions and local authorities.

E.18. One would expect that the impact of migration on prices depends on the elasticity of supply of housing with a migration leading to higher house prices in areas where house-building is more difficult. To investigate this we used a variable that has been used in other research[34] to measure the difficulty of building houses: the refusal rate on applications for major developments, averaged across years. This is included both as a separate regressor and interacted with the change in the migrant share. The OLS and IV results are reported in Table E.7 – we only report results without region or local authority fixed effects as the results are not reliable when these are omitted. These estimates are for England only as data for Welsh local authorities is not available.

Table E.9: OLS and IV estimates of impact of migration, interacted with refusal rate on house

	(1)	(2)
Change in Migrant Share	0.009	0.927 (***)
	[0.016]	[0.158]
Change in Migrant Share interacted with refusal rate	0.247 (*)	3.884 (**)
	[0.147]	[1.594]
Constant	0.071 (***)	0.070 (***)
	[0.002]	[0.004]
Observations	5,184	5,184
R^2	0.797	0.665
Estimation Method	OLS	IV

Statistical significance – (***) 1%, (**) 5%, (*) 10%

E.19. We also investigated whether there was a differential impact of EU and non-EU migration on house prices – we found no significant difference but the IV estimates were not reliable as the two instruments have insufficient independent variation.

[34]Paul Cheshire, Christian A.L. Hilber, Hans R.A. Koster, "Empty homes, longer commutes: The unintended consequences of more restrictive local planning", Journal of Public Economics,2018, https://www.sciencedirect.com/science/article/pii/S0047272717302086?via%3Dihub and Christian A. L. Hilber and Wouter Vermeulen "The Impact of Supply Constraints on House Prices in England", Economic Journal, 2016,https://onlinelibrary.wiley.com/doi/abs/10.1111/ecoj.12213

Annex F: Public Services Impacts

Health Impacts

F.1. *Figure 5.1* – Regression analysis using linear probability model estimated using 2017 APS. The dependent variable is qhealth1. Adjusted estimates control for age, sex, and employment status.

Panel Data Set Regressions

F.2. Using data from NHS England for GP satisfaction scores in Primary Care Trusts (PCTs) across England between 2009 and 2017[35] as well as data from the APS/LFS on migrant shares and age demographics in local areas we construct a panel data set. We aligned results from the GP Patient Survey with calendar years on the basis of when the fieldwork for the data was collected rather than the publication date. We use this to explore if there is an effect on GP satisfaction from increasing migration into a local PCT.

Table F.1: GP satisfaction in primary care trusts. OLS estimates. 2009-2017

Dependent variable: Very or fairly good satisfaction with GP services/care

Independent variables	(1) OLS	(2) Fixed effects on PCT, OLS
EU13+ share	0.34 (***)	0.29 (***)
NMS share	-0.42 (***)	-0.23 (***)
Non-EEA share	-0.20 (***)	-0.03
Observations	837	837
R^2	0.70	0.35

Statistical significance – (***) 1%, (**) 5%, (*) 10%

F.3. In Table F.1 we investigate the effect of migrant shares on GP satisfaction using an ordinary least squares regression with and without fixed effects for PCTs. We control for year fixed effects in both regressions, with each year proving to be significant. The fixed effects of PCTs controls for the within-PCT variation which may affect the outcome variable.

[35] https://www.gp-patient.co.uk/

Table F.2: GP satisfaction in primary care trusts. Instrumental variable regression estimates. 2009-2017		
Dependent variable: Very or fairly good satisfaction with GP services/care		
Independent variables	(1) IV regression, 2SLS	(2) IV regression, 2SLS, fixed effects on PCT
EU13+ share	0.46 (***)	1.15 (**)
NMS share	-0.36	-0.50
Non-EEA share	-0.24 (***)	-0.28
Observations	837	837
R^2	0.69	0.89

Statistical significance – (***) 1%, (**) 5%, (*) 10%

F.4. In Table F.2 we use an instrumental variable regression using a shift-share instrument. The migrant share instrument is constructed from 2001 census data, APS and NOMIS data. The IV regression is run with both fixed effects by PCT and non-fixed effects. We find our instrument is not weak once tested for in the first stages regression.

Education Impacts

F.5. *Table 5.1* – Using the household identifiers in the APS, it is possible to identify the relationship between respondents in a household. Children born in the UK living with parents/guardians in the household who are non-UK-born can be used as a proxy for second generation migrants. This can only be estimated for those still living with their parents/guardians. To do this analysis, we used the 2017 APS data. We used children aged 4-10 as a proxy for primary school children and children aged 11-16 for secondary school pupils. We grouped the CRYOX7 variable into the MAC migrant groups and used the household and person identifiers to create a matrix of relationships of respondents including their country of birth.

Impact of EAL pupils on the attainment of non-EAL pupils and school choice

F.6. Using public DfE data taken from gov.uk we constructed a basic panel dataset of KS2 and KS4 results by first language and applications and offers outcomes by Local Education Authority (LEA). We use this dataset to investigate the existence or otherwise of a correlation between the proportion of pupils with English as second language (EAL) in an LEA and the attainment of non-EAL pupils and also the proportion of non-EAL pupils getting their first choice of school. The data on pupil attainment covers the years 2006-2015, while the data on applications and offers covers the years 2008-2017. In total, we made use of 1,800 year/LEA combinations. We estimated this correlations in the presence of time and LEA fixed effects, no

additional covariates were included and no IV approach was utilised, as such no causal claim can be made. The results are summarised in Table F.3 below.

Table F.3: Relationship between proportion of EAL pupils and the attainment of non-EAL pupils and the proportion of applicants receiving their first choice of school.

Dependent variable:	**Attainment of non-EAL pupils**	**Proportion of applicants receiving first choice of school**
Independent variables	**(1) Fixed Effects**	**(1) Fixed Effects**
EASL Ratio	5.80(***)	0.92
Observations	1567	1430
R^2	0.92	0.15

Statistical significance – (***) 1%, (**) 5%, (*) 10%

Annex G: Communities Impacts

Estimating the Impacts of Migration on Crime

G.1. For data on population, we used data from 314 consistently identified local areas over the period 2002-2017. The data comprised of 301 individual local authorities and 13 Community Safety Partnership (CSP) areas which combine a number of local authorities into one area for police reporting purposes. Data on crime comes from recorded crime figures reported annually to the Home Office by each police force, population and migration data from APS.

G.2. The basic estimating equation is of the form:

$$\Delta\left(\frac{Crime}{Pop}\right)_{it} = \beta_1 \Delta\left(\frac{Migrants}{Pop}\right)_{it} + \beta_2 \Delta \ln(Pop)_{it} + \beta_3 \Delta X_{it} + T_t + \varepsilon_{it}$$

where *Crime* is the number of notified offences, *Pop* is the resident adult population, *Migrants* are the stock of EEA migrants, *X* denotes a set of local area level controls and *T* are a set of time dummies. Standard errors are clustered at the local authority level. Since we are estimating a first-difference model, this controls for area specific time trends. Results using an alternative fixed-effect specification are broadly similar. The model is as in Bell, Fasani and Machin (2013)[36].

Results

G.3. This section explains the results of the estimating equation above on the impacts of migration on property and violent crime.

G.4. We instrument the migrant share measure using the 1991 Census settlement patterns. Columns (1) and (4) of Table G.1 show that OLS estimates of equation (1) suggest no effect anywhere of EEA migrants on crime, either violent or property. Columns (2) and (5) present IV estimates. The F-Stat on the instrument is above the standard weak instrument level and the results now suggest a negative effect of EEA migrants on crime, though the effect is insignificant for property crime. Finally, in columns (3) and (6) we show IV estimates excluding the final two years of data. The results generate much smaller negative point estimates that are all insignificant. This highlights the somewhat fragile nature of IV estimates in this setting and suggests that the conservative reading of these results is that we can find no evidence of a substantial effect of EEA migration on crime and that at the margin the evidence points to a negative effect. It should be noted that no cut of the sample period generates significantly positive OLS or IV estimates. This conclusion is in line with BFM (2013).

[36] http://eprints.lse.ac.uk/59323/1/CEP_Bell_Fasani_Machin_Crime-and-immigration_2013.pdf

Table G.1: Panel Regressions for Crime (property and violent) and Migration						
	Violent (1)	Violent (2)	Violent (3)	Property (4)	Property (5)	Property (6)
$\Delta(\frac{\text{EU Migrant}}{\text{Pop}})$	0.001	-0.213	-0.021	-0.005	-0.256	-0.148
	(0.002)	(0.078)	(0.040)	(0.005)	(0.229)	(0.290)
OLS/IV	OLS	IV	IV	OLS	IV	IV
F-Stat, 1st Stage		17.1	15.3		17.1	15.3
Year Dummies	x	x	x	x	x	x
Sample Size	4393	4393	3768	4393	4393	3768
Sample Period	2003-2017	2003-2017	2003-2015	2003-2017	2003-2017	2003-2015

Notes: The dependent variable is $\Delta(Number\ of\ Crimes\ recorded/\ Population)$. All regressions include controls for ln(population), the benefit claimant rate and the share of young people (16-24) in the adult population. Regressions are run over the period 2003-2017 and are weighted by adult population. Standard errors in parentheses are clustered at the local authority level.

Results for Victimisation

G.5. This section reports on the data on victimisation, looking at the differences between UK born, EEA and non- EEA migrants.

G.6. Table G.2 shows the victimisation rates experienced by each group and was calculated using the public release versions of the Crime Survey for England and Wales, 2011/12 – 2016/17. At the descriptive level, migrants have a slightly higher probability of being a self-reported victim of crime over the prior 12 months, a slightly lower probability of being a victim of violent crime and a higher probability of being a victim of hate crime. The data allow us to identify UK, EEA and non-EEA migrants (defined as country of birth) but the public release version does not provide individual country of birth. We can use year of arrival in UK to identify post-03 NMS migrants.

Table G.2: Descriptive statistics of victimisation				
	UK and Ireland	EEA	Non-EEA	NMS
Crime Victim in last year (%)	16.9	17.5	17.6	16.8
Violent Crime Victim in last year (%)	2.4	2.1	2.0	2.0
Hate Crime Victim in last year (%)	0.6	1.0	1.3	1.1
Sample Size	193,957	8,052	18,365	4,577

Notes: Data are from pooled Crime Survey for England & Wales (2011/12 to 2016/17 waves). Figures are population-weighted.

G.7. Table G.3 then runs probits and reports marginal effects. Cols (1), (3) and (5) adjust the raw differences for time dummies and only hate crime remains as a strongly significant effect. When we control for a whole set of observables, (cols (2), (4) and (6)), migrants (both EEA and non-EEA) experience significantly lower victimization rates, except for hate crime where EEA migrants continue to experience higher rates. Note that in (6) we control for ethnicity which is why the non-EEA hate crime effect goes away.

Table G.3: Results of analysis on Victimisation- EEA and non-EEA						
	(1) Total Crime	(2) Total Crime	(3) Violent Crime	(4) Violent Crime	(5) Hate Crime	(6) Hate Crime
EEA	0.009*	-0.022***	-0.003	-0.013***	0.004***	0.002**
	(0.005)	(0.006)	(0.002)	(0.002)	(0.001)	(0.001)
Non-EEA	0.007**	-0.016***	-0.004***	-0.003	0.005***	0.001
	(0.003)	(0.005)	(0.001)	(0.002)	(0.001)	(0.001)
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes
Mean of Dependent Variable	0.170	0.170	0.023	0.023	0.006	0.006
Sample Size	220,374	188,631	220,374	188,631	220,374	188,631

Notes: Table reports estimated marginal effects from a probit where the dependent variable is a dummy equal to 1 if the individual reports having been a victim of crime in the last year and 0 otherwise. Personal controls include sex, age and age squared, ethnic group (5 categories), education (five categories), household income (seven categories), economic status (fifteen categories), housing tenure (three categories), health status (three categories) and government office region (10 categories). The sample covers ages 16 and over from the pooled Crime Surveys of England and Wales, 2011-2012 to 2016-2017. Regressions use individual sample weights. Robust standard errors in parentheses.
Statistical significance – (***) 1%, (**) 5%, (*) 10%

Data from the Police National Database (PNC)

G.8. The PNC database is a live operational system which is subject to continuous revision and updating: the data analysed represent annual snapshots of the information contained, which means they may subsequently have been revised. The figures only include the primary or principal offence (roughly speaking, the most serious) in each case. Where an offender receives convictions for multiple offences within the same court proceedings, only the principal offence is counted in our extract, which may mean that some less serious offence types are undercounted. This data relies on self-declared nationalities, which means that there is a possibility of some inaccuracies.

G.9. Table 6.1 of the main report uses data from the APS- 2012- 2016 inclusive to explain the share of population aged 16+ and share of population 16-29 male (who crime is most likely to be committed by). Data showing the average share of violent, property, robbery and drug offences was obtained from the PNC database from 2012-2016 all years inclusive.

Migration Advisory Committee report
https://www.gov.uk/government/organisations/migration-advisory-committee

ISBN: 978-1-78655-711-7