ESSAYS ON INTERNATIONAL TRADE AND LABOUR ECONOMICS

Anwar Seid Adem

Submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

in the Department of Economics

LANCASTER UNIVERSITY

MAY 2020

Declaration

I declare that this thesis has been composed solely by myself and that it has not been submitted, in whole or in part, in any previous application for a degree. Except where states otherwise by reference or acknowledgement, the work presented is entirely my own.

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisors, Maria Navarro Paniagua and Maurizio Zanardi, for their scholarly supervision, constructive criticism, and professional and technical guidance. Besides being a great privilege of working with them, their support of my research has continued throughout the years. Maria has offered patience and sincere advice that helped me navigate tricky situations, and stay focused and productive. Maurizio has provided me with opportunities and help at every turn. All of which have made me a better economist.

I also thank Jean-Francois Maystadt for his helpful comments and encouragements over the past years. I am also indebted to many other faculty at Lancaster University. I would like to especially thank Caren Wareing, Themis Pavlidis, Ian Walker, David Rietzke, Giorgio Motta, Marwan Izzeldin, and Pavel Chakraborty for their support at different stages of my study.

It is difficult to express my gratitude for the Economic and Social Research Council (ESRC). This study owes its realisation to the financial and institutional support provided by ESRC under NWDTC Standard Studentship and its host institutions.

I would have been much more at ease if only, words could explain my indebtedness for those who extended their hands whenever I am in need of one. Here, special thanks go to my colleagues and friends, Sakib Anwar, Habtamu Ali, Charlotte Edney and Konstantinos Protopappas for their friendship and unyielding support through this long ride.

Above all, my sincere thanks also go to my family for their priceless and unconditional moral support. The completion of this study would have not been a reality had it not been for their support.

ABSTRACT

ESSAYS ON INTERNATIONAL TRADE AND LABOUR ECONOMICS

Anwar Seid Adem

The importance of location in shaping the economic outcomes of different people is well documented. Thus, it lays in the intersection between different areas of research. More importantly, the empirical findings will have major policy implications, since the resilience or sensitivity of local labour markets to changes in economic fundamental is a function of their location and characteristics. And understanding this can be considered as the first step to remedy local economic problems.

In the first chapter, I investigate the causal effect of import shocks at local labour markets on the wage distribution using individual-level data from Great Britain in the period 1997-2010. In the analysis, I exploit regional variation in initial industrial structure and its concentration for identification, and apply a group IV quantile approach to estimate the effect of import shocks on workers at different points of the wage distribution. First, I find that the effect of an import shock generated by the increased imports from China is concentrated on the middle of the wage distribution. While the import shock negatively and significantly affects workers at the lower-middle range of the wage distribution, its effect on the very lower and upper part of the wage groups is positive but insignificant. Second, in trying to uncover the mechanism behind these results, I find that the labour adjustment process takes place through a reduction in the hourly wage rather than a decline in hours worked.

The second chapter aims at identifying the gains from imported inputs and foreign presence and to verify whether productivity gains from the two are either substitutes or complements. Understanding the relationship between these sources is crucial to evaluate the welfare implications of FDI promotion and trade liberalisation policies, which are particularly important for developing countries. To this end, I rely on a firm-level data-set

from Ethiopia for identification. After isolating the productivity gains from the numbers of inputs a firm chooses to import, I assess the role of FDI spillovers. I find evidence of positive gains from imported inputs for both domestic and foreign-owned firms with the magnitude being larger for the latter. I also find limited evidence on the substitutability or complementarity between the gains from imported inputs and FDI spillovers indicating that the productivity gains from the two sources are different in nature and do not interact.

The third chapter examines the effect of commuting on residential-mobility preference using data from the UK household longitudinal study. Together with preference to move, I also assess the impact of commuting on expectation to move. For identification strategy, I use a change in commuting time for those individuals who stay with the same employer and remain in the same place of residence. I find that commuting increases the individuals intent to relocate. The paper also finds commuting increases, besides preference to relocate, the expectation to move. The results contribute to the literature on the effect of commuting on residential choice which is crucial for labour market outcomes. Moreover, understanding the impact of commuting on individuals' preference to relocate have great policy implications, since the commuting and the corresponding decision surrounding it are considered the remedy to local economic problems.

Table of Contents

List of Tables

List of Figures

Dedication

This thesis is dedicated to my great grand mother, Zeyneba Yasin.

Chapter One

Distributional Effect of Import Shocks on the British Local Labour Markets

1.1 Introduction

In the past two decades, the labour market of most rich countries has been characterised by the decline in the share of manufacturing workers and a shrinking of middle-skill jobs (Autor, Dorn, and Hanson, 2016). In this regard, the disappearance of middle-skilled workers is particularly well documented. For instance, from 1995 to 2015, middle-skill jobs have shrunk from 49 to 40 per cent in the 23 OECD member states (OECD, 2017a). Here, the often mentioned causes are skilled-biased technological change, institutional setups and globalisation, that is, the rise in international trade and off-shoring (see Acemoglu and Autor, 2011; Goos, Manning, and Salomons, 2014). From a policy perspective, studies in this area not only help us answer questions related to a welfare effect of the above factors, that is, who gains and who loses, but also how to properly implement re-distributive policies.

This paper investigates the causal effect of import competition on workers at different parts of the wage distribution using worker-level data from Great Britain for the years 1997-2010. After accounting for changes in individual characteristics and return to those characteristics, I find that an increase in import exposure adversely affect those at the middle of the wage distribution. Thus, the paper contributes to both international trade and labour literature (Autor, Dorn, and Hanson, 2013; Chetverikov, Larsen, and Palmer, 2016; Felbermayr, Impullitti, and Prat, 2018).

The British labour market represents an interesting case to investigate the effect of trade shocks on wage inequality, as it has experienced both a rise in wage inequality and a substantial increase in import competition over the past three decades. On the one hand, unlike other European countries, the UK is a country with a high degree of inequalities among its regions: the inter-regional inequality of the UK is above the OECD average. In fact, it is the only European country with NUTS-2 regions[1] in all five quantiles of the EU GDP per-capita distribution (see Gibbons, Overman, and Pelkonen, 2010; McCann, 2016; Arellano and Bonhomme, 2017), with regional divergence being the phenomenon that began to accelerate in the 1990s (McCann, 2016). Figure 1.1a illustrates this by showing the trend in wage ratios, which have clearly increased over the past 30 years. On the other hand, and similar to other developed countries, UK's imports from China have grown rapidly following China's accession to the World Trade Organisation (WTO) in 2001, as shown in Figure 1.1b.

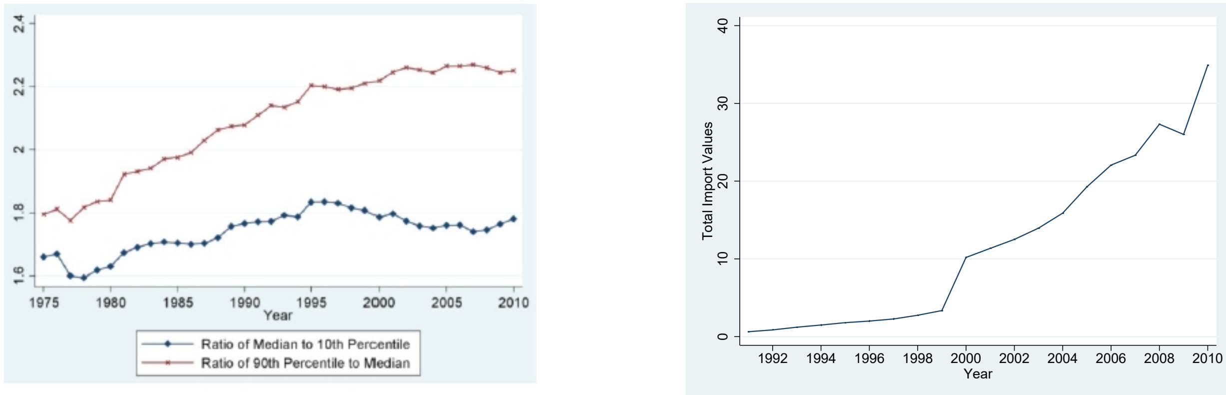
I use China's accession to the WTO as a natural quasi-experiment due to the rapid growth of the UK's imports from China during the period under consideration and the recent trend in the literature. In fact, it can be assumed to be unanticipated from the point of view of UK regions in general and UK firms within those regions in particular. Therefore, it is an exogenous shock for all regions regardless of their industrial composition. This identification strategy provides causal evidence on the role of trade shocks on local labour market outcomes. Moreover, thanks to the availability of micro datasets i.e., worker-level data, I am able to zoom into individual worker's outcomes in investigating the local or regional labour market effect of trade liberalisation (e.g., Kovak, 2013; Autor et al., 2013; Dix-Carneiro and Kovak, 2015).

Given that regions within a country differ in their industrial structure and concentration of activities, their exposure to trade shocks is also likely to differ. For instance, regions specialised in textiles would be affected more by increased import competition from low-wage countries than regions specialised in auto-mobile manufacturing. Thus, this regional variation in the degree of exposure to trade shocks is commonly used by studies to identify the effect of import surges and answer research questions related to the adjustment of local labour markets to trade shocks. This paper follows this literature and exploits a similar identification strategy to answer how import shocks affect the local labour dynamics.

From the analysis, I find that high import exposure contributes to the rise of wage polar-

[1]The UK is divided into 37 NUTS-2 regions for the period under consideration.

Figure 1.1 Trends in Total UK Import Value from China and Wage Inequality



(a) Trend in Wage Inequality (b) Trend in Import value

Notes: (a) Sources: LSE blogs using ASHE data (b) Author calculation using the WITS dataset from the World Bank. Values are in billions of 1987 pounds.

isation as it negatively affects workers at the lower-middle range of the wage distribution while leaving those at the very lower and upper end unaffected. The effect is particularly significant for those between the 30^{th} to 50^{th} quantiles of the wage distribution. In a further disaggregation of workers into subgroups, in the heterogeneity analysis, I find that the effect of import competition is mainly concentrated on manufacturing workers who are at the lower-middle part of the overall wage distribution. As for the mechanisms, the analysis demonstrates that most of the adjustment takes place through a change in the hourly wage rather than total hours worked.

This paper is related to the growing literature that examines the differential effects of trade shocks on local labour market outcomes in general and wage distribution in particular. I use the group instrumental variable (IV) quantile approach developed by Chetverikov et al. (2016) to identify the effect of import competition on the wages of workers at different quantile levels. By comparing these effects, I can verify the extent of the effect of trade liberalisation on wage inequality. Importantly, the approach allows me to account for the problem of endogeneity associated with the variable of interest. Here, similar to Autor et al. (2013), I use a Bartik type instrument to correct the endogeneity problem.

This paper contributes to the trade literature in different dimensions. First, it empirically extends the analysis of IV quantiles by accounting for individual characteristics such as age, gender and occupation. The findings demonstrate that the inclusion of these controls is of great importance, that is, the inclusion of these covariates increases the precision of the estimated coefficients at every quantile level. Second, in trying to undercover the underlying forces at work, it decomposes the weekly wage effect into its hourly wage and hours worked component. Given the richness of the data, it also investigates whether there are heterogenous effects across different sub-samples of workers. Third, it provides a comprehensive study of the effect of trade shock on the British local labour markets.

The remainder of the paper is organised as follows. Section 2 discusses related and recent developments in the literature. Section 3 describes the data sources and present descriptive statistics. Section 4 explains the methodological framework, and the main results are presented in section 5. Section 6 concludes.

1.2 Literature Review

The causal effect of globalisation on labour market outcomes has been an important research topic and has featured very highly in the policy debate. This is because while gains from increased imports are spread across the economy, losses are concentrated among the direct competitors, and hence the overall effect of import competition on employment and wages of the latter group is non-trivial (Greenland and Lopresti, 2016).

In the past, researchers have attributed the observed reduction in manufacturing workers and the rise in wage inequality in most developed countries to, *inter alia*, skill-biased technological progress, institutional setups, and trade. However, the emphasis on the former two channels overshadowed the attention given to the effect of trade shocks. Also, due to the limited evidence on local labour market effects of trade shocks, the causal link between the two has remained ambiguous for long (see Krugman, 2008; Bloom, Draca, and Van Reenen, 2016).

The previous literature in the area also concentrated on the effect of trade on economy-wide outcomes. However, the use of microeconomic data and different features of recent trade relations distinguish recent studies from their previous counterparts. Micro-data availability allows researchers to investigate the causal effect of trade shocks on local labour markets at more disaggregated levels. In this regard, one strand of the literature uses reduced-form analysis to study the impact of trade shocks on welfare, employment and wages. Similar to most previous studies, these studies find that trade-related demand shocks cause different effects on different sub-economies or groups. However, unlike previous studies whose analysis was limited to the effects of trade shocks on different owners of resources, that is, capital and labour, recent studies analyse trade shocks on different regions within a country (see Topalova, 2010; McCaig, 2011; Kovak, 2013), on different occupations (see Acemoglu and Autor, 2011; Ritter, 2014; Peri and Sparber, 2009), on different industries (see Revenga, 1992; Attanasio, Goldberg, and Pavcnik, 2004), on different occupations and industries (see Utar, 2016; Artuç and McLaren, 2015), and even among different age groups (see Artuç, 2012).

In her pioneering work, Topalova (2010) investigates the effects of variation in tariff reductions on poverty levels of districts in India. After constructing tariff reduction intensities for each district based on variation in sectoral composition, she finds that regions which are more exposed to trade, through higher tariff reduction, experience a slower decline in poverty and consumption. She also shows that the effect is severe for less mobile groups of workers such as those at the bottom of the income distribution. Likewise, Kovak (2013) exploits the regional variation in exposure to trade shocks to analyse local labour market effects in Brazil. His findings show the presence of negative, location-specific effects from trade shocks on wages and employment.

Similarly, Hakobyan and McLaren (2016) analyse the effect of tariff reduction in the US following the implementation of NAFTA on wages at the local labour market level between 1990 and 2000. Their reduced-form analysis finds a large negative effect of NAFTA on wages of unskilled workers in regions and industries that experienced a larger reduction of tariffs. Chiquiar (2008) also analyses the effect of NAFTA on Mexico's local labour market. He finds evidence of an increase in wage inequality and overall wage levels and a decline in the skill premium for highly exposed regions.

The second strand of literature uses structural models to investigate the dynamics of labour markets following trade shocks. Beyond offering insights into the causal link of exogenous and endogenous variables, these studies allow researchers to model the underlying mechanism through which causal relations operate. Therefore, they help us to answer a multitude of research questions (Reiss and Wolak, 2007). Unlike reduced-form studies, structural studies find mixed evidence on the effects of trade shocks on labour market outcomes, particularly on wage inequalities. On the one hand, there are studies that find a considerable impact of trade liberalisation on wage inequality. For example, after structurally estimating a heterogeneous trade model with an imperfect labour market of search and matching, Helpman, Itskhoki, Muendler, and Redding (2017) find a significant effect of Brazilian trade liberalisation on wage dispersion for the period spanning 1986-1998. Similarly, Dix-Carneiro and Kovak (2015) also use Brazilian data to show the presence of lags in adjustment following trade shocks and the cost of mobility. Egger, Egger, and Kreickemeier (2013) develop and estimate a structural model that combines a heterogeneous firm model and worker with fair-wage preferences, and find a non-negligible impact of openness on wage inequality using data from France and Balkan countries. On the other hand, Coşar, Guner, and Tybout (2016) find no evidence on the impact of trade liberalisation, *per se*, on wage inequality after analysing a Colombian trade and labour market reform. Felbermayr et al. (2018) find no evidence to conclude trade openness of Germany is the cause for the increase in wage inequality.

In addition to using different approaches and levels of disaggregation, studies also differ in their local labour market outcome of interest. Some studies analyse the effect on the employment level; others investigate the effect on wage and wage inequality; others still aim at analysing the effect of import competition on productivity, innovation and R&D related investments; and still, others are concerned with the welfare implications of trade shocks.

The recent increase in trade relations demands re-investigation since it might have different implications on the impact of trade shocks on labour market outcomes (Bloom et al., 2016). Particularly, the causal effects of trade shocks on the rise of wage inequality and the decline in manufacturing employment (see Helpman et al., 2017; Sampson, 2016; Felbermayr et al., 2018). In this regard, a growing literature uses China's accession to the WTO as an identification strategy.

A seminal paper by Autor et al. (2013) analyses the consequences of import competition from China on US commuting zones. They use the presence of variation in industrial structure and specialisation among these zones to exploit variations in trade exposure. The authors find that more exposed regions experienced lower wages, reduced employment prospects, and increased transfer payments from federal and state programs.

Another study for the US includes a work by Pierce and Schott (2016) which uses China's grant of Permanent Normal Trade Relation (PNTR)[2] status to exploit the impact of import competition on US employment and find similar results as Autor et al. (2013). However, recent studies find contrary results (e.g., Feenstra, Ma, and Xu, 2017; Wang, Wei, Yu, and Zhu, 2018).

Other recent studies which use a similar identification strategy include a study by Balsvik, Jensen, and Salvanes (2015) who highlight the negative effect of Chinese import competition on employment, particularly on those low-skilled ones using Norwegian data. Dauth, Findeisen, and Suedekum (2014); Dauth and Suedekum (2016) analyse the effect of the rise of imports from China and Eastern Europe on regional labour markets of Germany from 1988 to 2008 and 1990 to 2010. Mendez (2015) finds that highly exposed regions to import competition from China experience a larger reduction in the share of manufacturing employment and greater mobility of workers using Mexican data. This paper is also related to this strand of literature by examining the causal effects of trade shocks: trade liberalisation, expansion of exporters, and lower trade costs, on wage inequality.

Methodologically, a study by Han, Liu, and Zhang (2012), which analyses the causal effect of Chinese accession to the WTO on wage inequality of highly exposed and low exposed regions of urban China, is closely related to ours. By analysing the presence of significant changes at the 90^{th} and 10^{th} quantiles, they show how import competition exacerbates or cushions wage inequality. However, the present paper uses a recently developed measure of regional import exposure, and the analysis focuses on the British local labour market.

[2]PNTR is a legal status in the United States for free trade with a foreign nation. In the case of China, the principal impact of PNTR was to eliminate uncertainty from a potential increase in the US import tariffs due to politically contentious annual renewals associated with its temporary NTR status.

1.3 Data Description

To answer the main research question on the causal link between wage distribution and trade, a wealth of data are required which are discussed in this section (with further details relegated to the Appendix). To this end, I use five data sources: Annual Survey of Hours and Earnings (ASHE), Annual Business Inquiry (ABI), Business Register and Employment Survey (BRES), UN Comtrade, and OECD regional statistics and indicators. In short, the first data source provides the individual level variables, the next three allow me to construct the main variable of interest (i.e., region level exposure) while additional regional covariates are taken from the last one. A detailed description of the main variables of interest is available in Appendix A.

First, data on individual workers and their characteristics come from the ASHE dataset of the UK data service. The ASHE data is used by many researchers (e.g., Manning and Petrongolo, 2017; Elsby, Shin, and Solon, 2016). For details of this dataset see Pike (2011). Through this source, I have been given access to a 1% sample of employees from National-Insurance records for the years 1997 to 2010, which determines the analysis period. The sample is representative at regional-industry level.[3] The advantage of this dataset is its granularity at individual and regional level and its accuracy, given that the data is reported by employers to HM Revenue and Customs PAYE of employees. Most importantly, this dataset includes variables on both employee and employer characteristics. Variables referring to employee characteristics include wages, hours worked, age, gender, type of occupation in nine categories from managers to elementary occupations (that can be used as a proxy for education), manufacturing indicator (i.e. manufacturing and non-manufacturing) and full/part time status (ONS, 2017a).

Since the identification strategy involves changes in labour market outcomes before and after China's accession to the WTO in December 2001, the analysis exploits the changes between two periods: 1997 to 2002 and 2002 to 2010, with the changes weighted to represent decadal changes for ease of comparison.[4] Table 1.1 reports descriptive statistics

[3] In this paper, NUTS are used as regional classification. NUTS is a geocode standard representing the subdivisions of counties in European Union, and are often used for statistical purpose.

[4] Following Autor et al. (2013), I convert these changes into their decadal equivalence changes by multiplying changes by 10/5 and 10/8 for the years 2002 and 2010, respectively.

of the main outcome variables and individual covariates from the ASHE dataset for the three years under consideration, that is, 1997, 2002 and 2010.

Table 1.1 Descriptive Statistics

	Mean	Std. Dev.	Min	Max	Q1	Q2	Q3	Q4
Year 1997								
Real gross weekly earnings	379.29	322.05	0	>8,000	101.91	262.83	404.28	748.21
Real hourly earnings	10.17	8.68	0	>220	4.63	6.90	10.04	19.11
Average weekly paid hours worked	32.98	14.57	0	>125	17.62	35.60	39.07	39.62
Male	0.52	0.49	0	1	0.21	0.46	0.65	0.76
Full time	0.72	0.45	0	1	0.17	0.82	0.93	0.97
Manufacturing	0.21	0.40	0	1	0.10	0.22	0.27	0.24
Age	38.77	11.67	16	64	37.91	37.42	38.55	41.20
Observations	148,759				37,192	37,197	37,181	37,189
Year 2002								
Real average gross weekly earnings	449.59	424.62	0	>12,500	129.54	302.16	462.52	904.20
Real hourly earnings	12.68	11.34	0	>360	6.40	8.46	12.02	23.83
Average weekly paid hours worked	34.64	10.97	<1	>100	22.16	37.37	39.80	39.23
Male	0.51	0.49	0	1	0.23	0.47	0.63	0.73
Full time	0.76	0.43	0	1	0.23	0.86	0.95	0.98
Manufacturing	0.16	0.37	0	1	0.07	0.17	0.22	0.19
Age	39.55	11.80	16	64	38.04	38.72	39.67	41.76
Observations	151,472				37,870	37,868	37,867	37,867
Year 2010								
Real average gross weekly earnings	461.69	407.88	0	>9,150	132.48	305.20	473.95	935.16
Real hourly earnings	13.54	11.70	0	>380	7.51	8.93	12.66	25.07
Average weekly paid hours worked	33.13	11.74	0	>110	18.75	36.08	38.89	38.80
Male	0.50	0.50	0	1	0.26	0.44	0.57	0.67
Full time	0.71	0.45	0	1	0.13	0.82	0.93	0.96
Manufacturing	0.10	0.30	0	1	0.04	0.11	0.14	0.13
Age	40.05	12.21	16	64	37.42	39.10	40.66	43.02
Observations	165,544				41,386	41,386	41,386	41,386

Note: Real monetary units in 2010 Pound sterling; some of the minimum and maximum values are suppressed for disclosure avoidance; and a detailed explanation of the variables are provided in the appendix. Notice that the earnings are constructed weekly, that is, weekly earning is a multiple of weekly hours worked and hourly earnings. Occupation represents a categorical variable following the International Standard Classification of Occupations (ISCO-88) of the ILO, where, the lower value representing high skilled jobs.

The first four columns of Table 1.1 present statistics for the overall distribution while the last four columns show the average values of the same variables for four quartiles constructed based on real gross weekly earnings. A closer inspection of these last columns reveals two sources of variations with different implications. First, there is a difference in characteristics along the wage distribution in a given year, that is, in each year there are differences in individual characteristics across quartiles. For example, individuals in the higher quartile tend to have higher real hourly earnings, work more hours, are more likely to be male, and work in high skilled occupations, and almost always full time. Second, there is a change in characteristics across time periods for a given quartile. For instance, between 1997 and 2002, the share of females among the top quartile increases while that of manufacturing workers declines. These two facts clearly illustrate that individual characteristics vary across the wage distribution at a given time and also through time. This consideration is particularly important as it is important to account for both dimensions of the variations in the empirical analysis.

The employment figures at regional and industry level are used to construct the change in regional import per worker (I weight the change in import by the number of manufacturing workers). To this end, data from two sources, that is, ABI and BRES of official labour market statistics Nomis, are used (ONS, 2017b). These datasets are available at the 3-digit level of SIC 2003, which determines the levels of industrial disaggregation.

The fourth data source is import data from UN Comtrade. This database includes import data under different industry classifications. I use the 3-digit level of NACE Rev.1 because at this level it is identical to the SIC 2003 classification, which is used in ABI and BRES (WITS-UNSD, 2017).

Using these data, I calculate the regional change in import exposure to China for each of the 128 NUTS-3 regions of Britain for which there is also representative data from ASHE. Thus, these NUTS-3 regions are considered as the local labour market in this paper.[5]

Specifically, I use the region's share of employment in industry j and the change in imports per number of workers to calculate the change in import per worker of region r at time t (ΔIPW_{rt}). In other words, I sum changes in import values per regional employment across 93 industries and weight them by regional share of employment in each industry:

$$\Delta IPW_{rt} = \sum_j \frac{Emp_{rjt_0}}{Emp_{jt_0}} * \frac{\Delta IMP_{jt}}{Emp_{rt_0}}, \tag{1.1}$$

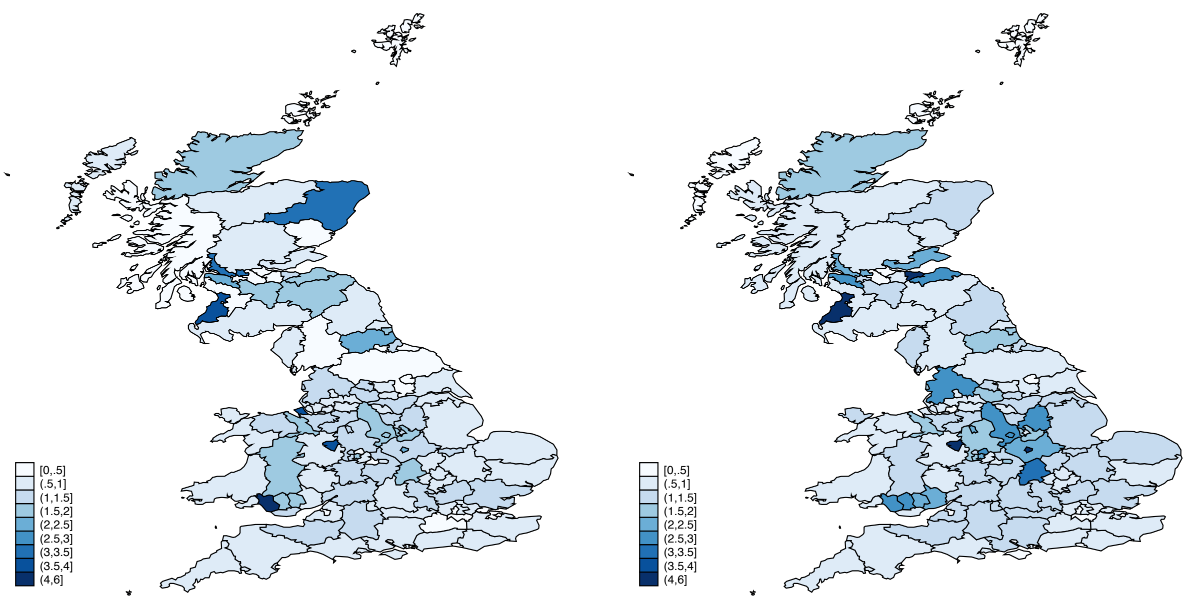
where Emp_{rjt_0} represents the start of period regional employment in industry j, Emp_{jt_0} stands for start of period total number of workers in industry j in Britain, Emp_{rt_0} is start

[5]Other studies use travel-to-work-areas (TTWAs) as local labour markets in the UK (e.g., Ottaviano, Peri, and Wright, 2018; Manning and Petrongolo, 2017). I do not consider travel to work since I do not have data on regional characteristics at that level of classification. NUTS-3 regions are slightly bigger (128 in our period of analysis versus 243 TTWAs) but not too big to be considered as an alternative. Considering the recent reduction in commuting costs and the increasing number of workers who commute to work, NUTS-3 can be considered a local labour markets.

of period number of employment in region r, and ΔIMP_{jt} stands for the change in value of industry $j's$ import by the UK from China (in £1000s).[6]

Figure 1.2 provides a geographical representation of ΔIPW_{rt} for 2002 and 2010. The figures clearly illustrate the extent of the substantial geographical variation in the two time periods. On average, the change in regional import per worker was £960 between 1997 and 2002, and it increased to £1,160 for the period between 2002 and 2010, which represents a 20.8% increase.[7]

Figure 1.2 Decadal Change in Regional Import per Worker



(a) 1997-2002 (b) 2002-2010

Notes: The figure shows a ten year equivalent change in import per worker for the period between 1997 and 2002 and 2002 and 2010.

Finally, data on other regional characteristics such as the proportion of female workers and the proportion of manufacturing workers is obtained from OECD regional statistics (OECD, 2017b).

[6] Although imports are at the UK level, this analysis is for Great Britain since manufacturing employment for Northern Ireland is absent from Nomis data.

[7] Table A.1 in the appendix presents the top five most exposed regions out of the most fifty highest regions in terms of their working-age population for the year 2002 and 2010. The table also presents the median, mean and standard deviation for the changes in import exposure.

1.4 Econometric Methodology

Since the objective of this study is to investigate the causal effect of trade shocks on the wage distribution, the analysis is based on a quantile regression approach. Quantile regression allows us to investigate the impact of trade shocks at different levels of the wage distribution. This is done by investigating the presence of significantly different effects at different parts of the wage distribution, which allows me to verify whether import competition exacerbates or cushions wage inequality.

Due to unobservable characteristics of the local labour market which are most likely correlated with the trade shocks, an ordinary least squares (OLS) estimator for the causal analysis between trade shocks and labour market outcomes would suffer from an endogeneity problem. To address this issue while conducting the analysis at the quantiles rather than at the average level, this paper exploits a group IV approach developed by Chetverikov et al. (2016). By focusing on the quantiles, this approach facilitates the identification of the causal effect of trade shocks along the wage distribution.

In the following, I start by providing an overview of the general econometric approach before engaging in the discussion of the empirical specification in more detail.

The general econometric model is given by:

$$q_{y_{ir}|v_{ir},x_r,\zeta_r}(\tau) = v_{ir}'\gamma(\tau) + x_r'\beta(\tau) + \zeta_r(\tau), \qquad \text{for all } \tau \in (0,1), \tag{1.2}$$

where $q(\tau)$ represents the τ^{th} conditional quantile; y_{ir} stands for the dependent variable (e.g. log weekly earnings) of an individual i in region r; v_{ir} represents individual-level covariates that affect the dependent variable; $\gamma(\tau)$ is the τ^{th} quantile coefficient estimates for individual covariates; x_r corresponds to regional level covariates; $\beta(\tau)$ is a coefficient for region level covariates; and $\zeta_r(\tau)$ represents region level unobservables.

Given the general model, the identification of the parameter of interest, $\widehat{\beta}(\tau)$, which is the region level treatment effect, takes two steps. The first step involves undertaking quantile regressions using the individual-level outcome as the dependent variable on individual characteristics for each region separately. This is given by:

$$y_{ir}(\tau) = v'_{ir}\alpha_r(\tau) + u_{ir}(\tau) \qquad \text{with} \qquad E[u_{ir}(\tau)|v_{ir}] = 0, \tag{1.3}$$

where $u_{ir}(\tau)$ is an individual-level iid error term and other variables are as defined above. The coefficient estimate of the quantile regression, $\hat{\alpha}_r(\tau)$, solves the following equation for each region r:

$$\hat{\alpha_r}(\tau) \equiv \underset{\alpha \in A}{\operatorname{argmin}} \frac{1}{n} \sum_{i=1}^{n} [\rho_\tau(y_{ir} - v'_{ir}\alpha)], \tag{1.4}$$

where $\rho_\tau(.)$ is known as the check function and can be rewritten as:

$$\rho_\tau(u_{\tau i}) = \begin{cases} \tau u_{\tau i}, & \text{if } u_{\tau i} \geq 0 \\ (\tau - 1)u_{\tau i} & \text{if } u_{\tau i} < 0. \end{cases} \tag{1.5}$$

The estimation of equation (1.3) using quantile regressions for each region provides me with a coefficient estimate for k^{th} individual level covariates, $\widehat{\alpha}_{r,k}(\tau)$, and the residual term.

The second step uses group level estimates from the first step as a dependent variable and regresses it on our variable of interest and other group level covariates, x_r, to recover estimates for the parameter of interest, $\beta(\tau)$. This step can employ either an OLS regression or an IV approach, which is the method I follow because of the variable of interest is endogenous.

This general econometric method is implemented in this paper by the following two steps. In the first step, I control for individual characteristics and estimate the changes between periods at different quantiles. In the second step, I use a two-stage least squares (2SLS) to identify the coefficient of interest. I now discuss these steps in more detail.

As noted when discussing the descriptive statistics in Table 1.1, there are two sources of variations. First, individual characteristics vary across quantiles in a given year, and it is important to account for this in order to compare changes between similar individuals. Second, there are also changes to the composition of workers' characteristics and their returns between time periods. Therefore, both of these problems need to be addressed in the first step. To control for individual characteristics, in correspondence with equation

(1.3), I estimate a Mincer-type wage equation for each region separately (i.e. 128 regions), which is given by:

$$\begin{aligned} lny_{irt} &= \alpha_1 + \alpha_2 age_{irt} + \alpha_3 age^2_{irt} + \alpha_3 male_{irt} + \alpha_4 occupation_{irt} + \alpha_5 fulltime_{irt} \\ &+ \alpha_6 manufacturing_{irt} + \varepsilon_{irt}, \text{ with } \mathbb{E}(\varepsilon_{irt}|v_{irt} = 0), \end{aligned} \tag{1.6}$$

where lny_{irt} indicates the dependent variable, that is, the log of weekly earnings of an individual i in region r at time t in the baseline case, and log of hourly earnings and log of hours worked in the later specifications to investigate the mechanisms behind the results. Individual level covariates include age, age squared, male dummy, which takes a value one if the individual is male, nine occupations, full time and manufacturing dummies and ε_{irt} represents standard regression residuals.

Although the residuals from the above regression isolate the effect of observed individual characteristics, they do not account for changes across time. To address this other concern simultaneously, I follow a decomposition method by Melly (2005) where changes between periods at different quantiles and across time can be attributed to changes in characteristics, coefficients and residuals. For instance, the decomposition between 1997 and 2002 is calculated as:

$$\begin{aligned} \hat{q}(\hat{\beta}^{02}, v^{02}) - \hat{q}(\hat{\beta}^{97}, v^{97}) = [\hat{q}(\hat{\beta}^{02}, v^{02}) - \hat{q}(\hat{\beta}^{m02,r97}, v^{02})] + [\hat{q}(\hat{\beta}^{m02,r97}, v^{02}) - \\ \hat{q}(\hat{\beta}^{97}, v^{02})] + [\hat{q}(\hat{\beta}^{97}, v^{02}) - \hat{q}(\hat{\beta}^{97}, v^{97})], \end{aligned} \tag{1.7}$$

where v represents the above mentioned individual covariates. The expression in the first square bracket of the right-hand side of (1.7) indicates the effect of changes in the residuals, thus it represents changes in residuals between the years, $\Delta\varepsilon_{rt}(\tau)$. The expressions in the latter two square brackets indicate changes in characteristics and changes to returns for them between the two periods. Therefore, our dependent variable for the second step is represented by the expression in the first bracket, that is, the difference in residuals for each region at every quantile level between 2002-1997 and 2010-2002.

In the second step, I employ a 2SLS estimation approach and regress the change in region specific residuals at given quantiles, $\Delta\varepsilon_{rt}(\tau)$, on the change in import per worker,

ΔIPW_{rt}, and other covariates. The empirical specification is given by:

$$\Delta\varepsilon_{rt}(\tau) = \Delta IPW_{rt}\beta(\tau) + x_r\gamma(\tau) + \delta_r(\tau) + \eta_t(\tau) + \zeta_r(\tau), \tag{1.8}$$

where $\Delta\varepsilon_r(\tau)$ indicates decadal equivalent changes in the τ^{th} quantile residual of region r; and x_r represents the beginning of period regional level covariates other than a measure of the change in import exposure. These covariates include the percentage of employment in manufacturing, the percentage of employment among women, and the percentage of employment in routine occupations.[8] The last three terms, that is, $\delta_r(\tau)$, $\eta_t(\tau)$ and $\zeta_r(\tau)$, respectively, indicate NUTS 1 region fixed effects, time dummy for the period 2002-2010 and the error term.

In the baseline empirical regressions, the dependent variable is weekly earnings. Unless specified, the main specifications are in changes; standard errors are clustered at the NUTS-2 regional level; regressions are weighted by their start of period population share, and all regressions include a constant term. Individuals are aggregated by region thus the number of observations in the regression tables reflect this aggregated figure. As the analysis is over two periods, the maximum sample size in the analysis is 256 (i.e. 128 regions observed over two periods). However, this final stage is reached by using around 150,000 observations for each change (i.e. over 1997-2002 and 2002-2010) over regions and quantiles.

Estimating equation (1.8) using OLS will lead to biased estimates. This is due to the possible correlation of unobserved demand shocks with both import demand and labour market outcomes. To address this endogeneity problem, Autor et al. (2013) introduce a Bartik type of instrument where they use change in imports from China in other developed countries as an instrument. They argue this external instrument is exogenous to labour market outcomes but is correlated with trade shocks to which the country of interest is exposed to.

Following Autor et al. (2013), I construct import exposure of seven developed coun-

[8] An index measuring routine task-intensity (RTI) of occupations for each region is calculated following Autor, Levy, and Murnane (2003) and Goos et al. (2014). That is, $RSH_{rt} = \sum_{k=1}^{k} E_{rkt} * 1[RTI_k > RTI^{1966}] \sum_{k=1}^{k} E_{rkt}^{-1}$, where E_{rkt} is region r employment in sector k at time t, and the indicator function identifies the set of occupations in the top third of employment weighted routine task intensity (RTI).

tries, namely, Australia, Canada, Japan, Korea, New Zealand, Singapore, and the United States of America.[9] Moreover, the regional share of manufacturing workers out of the UK's national employment is lagged by six years[10] to avoid reverse causality, that is, current wage and employment may respond to expected trade exposure. Formally, the instrument is given by:

$$\Delta IPW_{rt}^{OTH} = \sum_j \frac{Emp_{rjt_0-6}}{Emp_{jt_0-6}} * \frac{\Delta IMP_{jt}^{OTH}}{Emp_{rt_0-6}}, \tag{1.9}$$

where ΔIPW_{rt}^{OTH} is a change in import exposure of other developed countries, $\frac{Emp_{rjt_0-6}}{Emp_{ujt_0-6}}$ is the six years lagged share of industry j employment in region r out of the UK's national employment of industry j, ΔIMP_{jt}^{OTH} is a change in the import of industry j by other countries, and Emp_{rt_0-6} is six years lagged level of employment in region r.

1.5 Results

Before presenting the main regressions, it is instructive to consider an analysis at the average level. In fact, these results can be later compared with the findings of Autor et al. (2013). This exercise also allows me to test the strength of the instrument and show the underlying relationship between openness and the average wage.

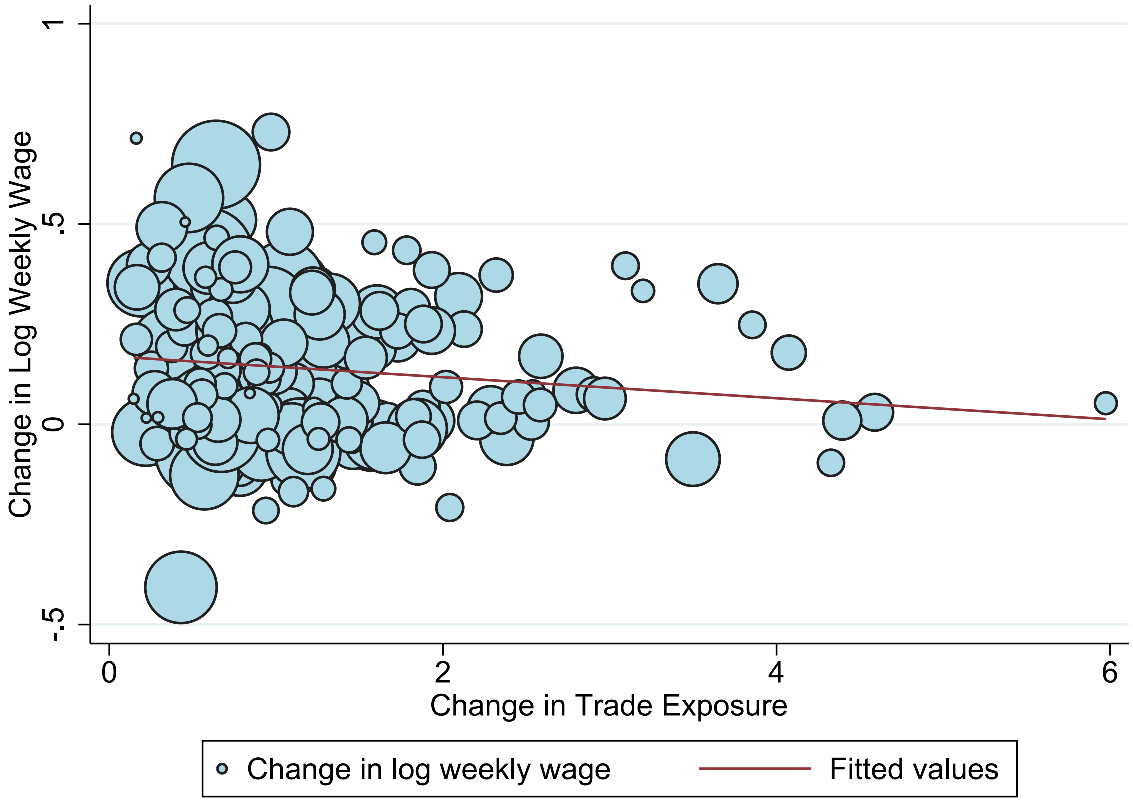
Beginning with a graphical illustration, Figure 1.3 provides a scatter plot with a fitted line for the change in decadal equivalent mean log weekly earnings and the change in regional imports per worker (in the pooled data). For this graph, regions are weighted by their start of decade population shares and the size of the bubble indicates their respective sizes. The slope of the fitted line is -0.039, indicating the inverse relationship between change in exposure and wage growth.

Moving to the econometric results of this first pass at the data, Table 1.2 provides the 2SLS estimation results of the change in mean weekly wage in a region on the change in import per worker. The four specifications differ in terms of included fixed effects and

[9] The results are robust to including Norway and Switzerland or removing the USA from the group of other developed countries. I do not include countries of the European Union to avoid correlation in demand and supply shocks with the UK labour market.

[10] Six years is the longest lag available in the data.

Figure 1.3 Change in Average Wage Response to Change in Import per Worker



Note: The size of the circle indicates the start of period region population share.

controls. Independently of the chosen specification, the signs and significance levels of the estimated coefficients of our key variables are unaffected (in both panels).

In particular, the lower panel shows the first stage results. As indicated by its statistical significance, our instrument (i.e. changes in import per worker of the other developed countries) is a highly significant determinant of changes in import exposure of the UK. The instrument explains a significant amount of variation in the endogenous variable as indicated by a relatively high partial R^2 and the F-tests are above critical values in all specifications, ensuring the absence of a weak instrumental variable problem.

The top panel of Table 1.2 presents the second stage results. Based on these estimates, we would conclude that there is no significant effect of the Chinese import shock on mean log wages (i.e. the estimated coefficient is negative but highly insignificant). This conclusion is not affected whether regional fixed effects are excluded (in column 1) or further

Table 1.2 First and Second Stage Results on Average Weekly Earnings

Second Stage	(1)	(2)	(3)	(4)
ΔIPW_{rt}	-1.549	-1.265	-1.714	-1.818
	(0.977)	(1.540)	(2.893)	(2.912)
Time dummy	-26.996***	-27.059***	-28.213***	-28.363***
	(2.666)	(2.743)	(3.958)	(3.944)
Lag female share			-66.255	-58.483
			(83.668)	(81.930)
Lag routine share			-25.364	-28.742
			(24.338)	(24.345)
Lag manuf share				110.820
				(114.305)
Region FE (NUTS1)	No	Yes	Yes	Yes
First Stage				
ΔIPW_{rt}^{OTH}	0.083***	0.075***	0.071***	0.071***
	(0.007)	(0.010)	(0.006)	(0.006)
Time dummy	-0.428***	-0.371***	-0.221***	-0.222***
	(0 .054)	(0.052)	(0.077)	(0.077)
Lag female share			-0.229	-0.141
			(1.029)	(1.074)
Lag routine share			2.308***	2.267***
			(0.711)	(0.717)
Lag manuf. share				1.258
				(2.857)
Region FE (NUTS1)	No	Yes	Yes	Yes
F-test	137.5	208.2	122.9	120.2
Partial R^2	0.588	0.507	0.438	0.436
Obs	256	256	256	256
R^2	0.536	0.546	0.552	0.553

Notes: Dependent variable is change in average import per worker for region r. Change in import per worker from other countries is used as instrument in the first stage. All regressions include constants and NUTS2 level clustered standard errors (in parenthesis).
* Denotes significance at the 10% level, ** Denotes significance at the 5% level, *** Denotes significance at the 1% level.

controls that affect regional labour market outcomes are included. In column 3, I add the beginning of period regional share of female workers and share of employment in routine jobs. Both variables seem to have no significant effect on average wages growth (but the share of routine jobs is significant in the first-stage regression). And finally, in column 4, I add lagged shares of manufacturing employment as an additional control. Similarly, lagged share of manufacturing wage has no significant effect on average wage change.

Although Autor et al. (2013) find a significant effect on mean weekly wage for the US, we do not find a significant effect for the UK. In fact, our results are more in line with studies by Balsvik et al. (2015) for the US and Edwards and Lawrence (2010) for Norway, who also fail to find a significant effect of import competition on average wages.

The scatter plot in Figure 1.3 shows us the general pattern, and the regression results at the average level help us to check for the validity of the instrument. However, it is important to go beyond these results, in exploiting the rich dataset available from ASHE and implement the necessary econometric tools to deal with the various issues mentioned earlier. This is the objective of the main analysis, which relies on quantile regressions, since the causal link between the wage growth and change in import per worker can be different at different parts of the wage distribution.

In the quantile estimation, as mentioned in the data descriptive section, two questions need to be addressed. First, there are observed individual characteristics that affect labour market outcomes of an individual. And second, there are changes in distribution and return to those characteristics between time periods. Below, I present the regression results after controlling for individual characteristics and accounting for changes in composition and returns to worker characteristics using Melly's method of decomposition.

Table 1.3 presents the baseline regression results of the change in import exposure on changes in weekly wages, which are graphically illustrated in panel a of Figure 1.4. Notice that for the whole of the quantile analysis I employ the same specification as in the last column of Table 1.2 (i.e. including 3 regional controls). There is evidence for a causal effect of import competition on polarisation where the middle income is affected the most. Specifically, Table 1.3 shows that the effect of the change in import exposure on the change in weekly wage is negative and significant for those individuals between the 35^{th} and 50^{th} percentile of the wage distribution. For instance, for those individuals at the 40^{th} percentile of the wage distribution, a £1000 increase in regional import per worker is estimated to reduce their weekly wage by 1.01 log points. The effects on the higher quantiles are not significantly different from zero. In comparison, previous studies which use worker level data find a higher effect at the lower part of the wage distribution. For example, using data from the US, Autor, Dorn, Hanson, and Song (2014) find a heterogeneous impact of import competition across workers with effects being concentrated among the low wage earners.

Table 1.3 Models After Controlling for Individual Characteristics

	Dependent Variable: Change in the log of weekly earning								
	(1) 10^{th}	(2) 20^{th}	(3) 30^{th}	(4) 40^{th}	(5) 50^{th}	(6) 60^{th}	(7) 70^{th}	(8) 80^{th}	(9) 90^{th}
ΔIPW_{rt}	0.396	-0.127	-0.799	-1.014**	-0.771*	-0.468	-0.519	-0.452	-0.081
	(0.952)	(0.623)	(0.552)	(0.430)	(0.411)	(0.361)	(0.343)	(0.348)	(0.528)
Lag manuf. share	-109.7**	16.24	-28.06	-33.93	-16.69	11.14	31.62*	67.14**	105.9***
	(53.347)	(43.837)	(34.095)	(29.318)	(25.868)	(18.348)	(17.390)	(26.080)	(40.520)
Lag female share	5.022	4.165	0.783	-5.911	-7.362	-2.193	-1.905	3.338	18.36
	(20.091)	(10.510)	(10.209)	(9.573)	(8.990)	(6.470)	(6.167)	(7.918)	(11.711)
Lag routine share	0.298	21.55***	29.55***	22.77***	14.77***	8.111*	3.470	-0.679	1.230
	(12.805)	(7.424)	(8.184)	(6.434)	(5.481)	(4.450)	(4.780)	(6.281)	(10.230)
Time dummy	-1.063	4.452***	4.430***	3.294***	2.143***	0.951**	-0.216	-1.008*	-1.432
	(1.108)	(0.795)	(0.726)	(0.544)	(0.511)	(0.429)	(0.379)	(0.544)	(0.981)
Region FE	yes	yes	yes	yes	yes	yes	yes	yes	yes
F-test	120.1	120.1	120.1	120.1	120.1	120.1	120.1	120.1	120.1
Partial R^2	0.436	0.436	0.436	0.436	0.436	0.436	0.436	0.436	0.436
R^2	0.0588	0.162	0.152	0.125	0.0823	0.0218	0.0241	0.0765	0.0734
Obs	255	255	255	255	255	255	255	255	255

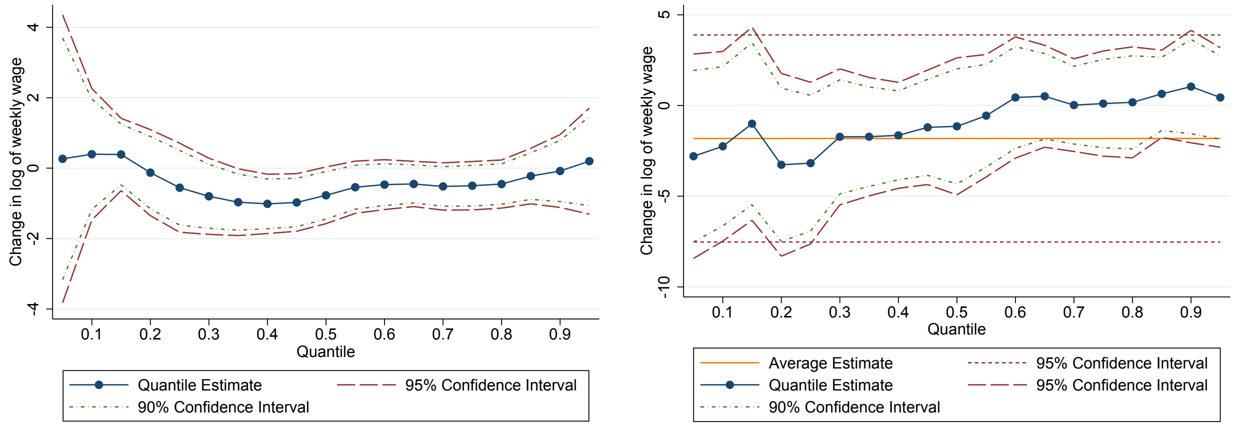
Notes: All regressions include constants and NUTS2 level clustered standard errors (in parenthesis). For all quantile regression, we control for start of period region characteristics such as the share of manufacturing employment, the share of female workers, the share of employment in routine works, region fixed effects and a time dummy for the period 2002-2010.
* Denotes significance at the 10% level, ** Denotes significance at the 5% level, *** Denotes significance at the 1% level.

The above decomposition makes this study distinct from the paper by Chetverikov et al. (2016). However, in order to see the effect of controlling for individual characteristics, Table 1.4 (and Panel b in Figure 1.4) reports the results when using the same group IV quantile methodology without controlling for individual characteristics. This specification closely follows the study by Chetverikov et al. (2016). While they find evidence for the causal effect of an increase in import competition on wage inequality for the US, the result for Britain is different in that there is no significance at any point of the distribution.

Furthermore, Table 1.4 shows that some coefficient estimates switch from negative to positive, as one moves from the lower to the higher quantiles of the wage distribution. Hence, the importance of using rich enough data sources, like ASHE, to control for individual characteristics.

In conclusion, summarising the results of Table 1.3 and Table 1.4, Figure 1.4 shows the plot of estimated coefficients of the change in regional import per worker at different quantiles (by 5 percentile increment) with respective 90% and 95% confidence intervals. At the same time, the straight line in Panel b shows the coefficient estimate at the average which corresponds to Autor et al.'s estimation result. The contribution of this paper is to point out that the average effects missed on important variation along the distribution and

Figure 1.4 Estimates at Different Quantiles with and without Controls



a. Quantile with individual controls   b. Quantile without controls & on average

Note: Estimation coefficients and confidence intervals for regression on quantile with and without controls and at average. The dependent variable is the change in the log of weekly wage and estimation is on all workers.

Table 1.4 Models Before Controlling for Individual Characteristics

	Dependent Variable: Change in log of weekly earning								
	(1) 10^{th}	(2) 20^{th}	(3) 30^{th}	(4) 40^{th}	(5) 50^{th}	(6) 60^{th}	(7) 70^{th}	(8) 80^{th}	(9) 90^{th}
ΔIPW$_{rt}$	-2.249	-3.269	-1.726	-1.644	-1.148	0.442	0.0205	0.174	1.041
	(2.666)	(2.570)	(1.912)	(1.490)	(1.926)	(1.703)	(1.303)	(1.560)	(1.581)
Lag manuf. share	222.0	207.2	123.7	121.1	120.2	136.2	144.8	123.3	210.0*
	(218.677)	(158.928)	(122.898)	(106.127)	(112.751)	(100.729)	(99.447)	(92.018)	(111.788)
Lag female share	36.48	7.300	11.89	12.90	50.03	44.44	10.84	-9.946	1.610
	(60.451)	(53.042)	(45.418)	(34.419)	(35.251)	(32.716)	(26.970)	(26.291)	(36.257)
Lag routine share	-103.5**	-60.73	-37.82	-4.920	2.840	-5.691	-9.599	-19.44	-18.13
	(45.669)	(40.286)	(32.665)	(26.377)	(26.126)	(20.088)	(17.155)	(14.294)	(13.095)
Time dummy	-43.17***	-33.25***	-24.78***	-21.50***	-20.83***	-20.82***	-21.35***	-22.47***	-24.49***
	(5.308)	(4.281)	(2.796)	(2.695)	(2.572)	(2.001)	(1.724)	(1.552)	(1.624)
Region FE	yes	yes	yes	yes	yes	yes	yes	yes	yes
F-test	125.0	125.0	125.0	125.0	125.0	125.0	125.0	125.0	125.0
Partial R^2	0.373	0.373	0.373	0.373	0.373	0.373	0.373	0.373	0.373
R^2	0.354	0.352	0.400	0.442	0.500	0.483	0.554	0.593	0.571
Obs	256	256	256	256	256	256	256	256	256

Notes: All regressions include constants and NUTS2 level clustered standard errors (in parenthesis). For all quantile regression, we control for start of period region characteristics such as the share of manufacturing employment, the share of female workers, the share of employment in routine works, region fixed effects and a time dummy for the period 2002-2010.
* Denotes significance at the 10% level, ** Denotes significance at the 5% level, *** Denotes significance at the 1% level.

that controlling for individual characteristics does substantially affect the results (at every quantile).

1.5.1 Transmission Mechanism

In order to understand the forces at work behind the result just established, I decompose the effects on weekly wage into hourly wage and total hours worked. This analysis helps us to identify the underlying mechanism of the wage effect by disentangling the wage effect into its price and quantity components. Again, the results presented in the following are obtained after controlling for individual characteristics and accounting for changes in composition and returns to characteristics.

Table 1.5 Models for Hourly Wage After Controlling for Individual Characteristics

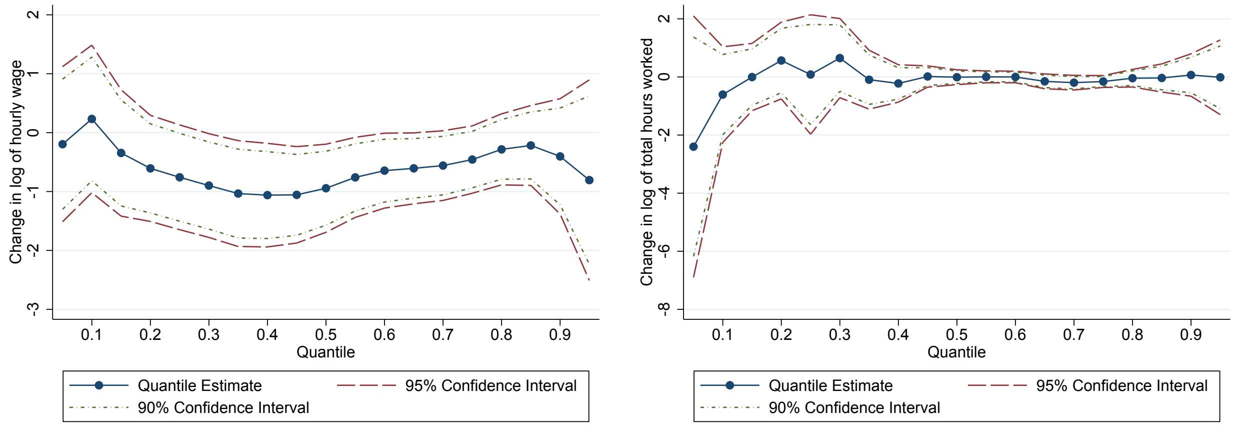
	Dependent Variable: Change in the log of hourly wage								
	(1) 10^{th}	(2) 20^{th}	(3) 30^{th}	(4) 40^{th}	(5) 50^{th}	(6) 60^{th}	(7) 70^{th}	(8) 80^{th}	(9) 90^{th}
ΔIPW_{rt}	0.233	-0.607	-0.899**	-1.061**	-0.944**	-0.645**	-0.560*	-0.284	-0.403
	(0.639)	(0.460)	(0.449)	(0.448)	(0.382)	(0.324)	(0.302)	(0.308)	(0.499)
Lag manuf. share	-53.99*	-16.80	3.403	2.519	0.821	0.898	20.48	50.83**	71.45*
	(30.767)	(18.954)	(16.731)	(15.273)	(12.971)	(14.789)	(20.029)	(24.666)	(39.588)
Lag female share	13.27	0.459	-6.327	-9.391	-9.279	-5.598	1.559	10.81*	17.24*
	(12.986)	(7.401)	(8.365)	(8.789)	(8.192)	(7.667)	(7.142)	(6.181)	(9.910)
Lag routine share	20.56**	17.14***	14.76***	13.31***	12.71***	10.84***	10.39**	10.87**	19.76*
	(8.080)	(5.589)	(4.932)	(4.381)	(4.173)	(4.039)	(4.624)	(5.484)	(10.382)
Time dummy	0.576	1.261***	1.501***	1.496***	1.323***	0.904**	0.384	0.0208	-0.691
	(0.653)	(0.339)	(0.447)	(0.521)	(0.508)	(0.437)	(0.393)	(0.432)	(0.747)
Region FE	yes	yes	yes	yes	yes	yes	yes	yes	yes
F-test	120.1	120.1	120.1	120.1	120.1	120.1	120.1	120.1	120.1
Partial R^2	0.436	0.436	0.436	0.436	0.436	0.436	0.436	0.436	0.436
R^2	0.121	0.134	0.101	0.0800	0.0683	0.0417	0.0452	0.0898	0.106
Obs	255	255	255	255	255	255	255	255	255

Notes: All regressions include constants and NUTS2 level clustered standard errors (in parenthesis). For all quantile regression, we control for start of period region characteristics such as the share of manufacturing employment, the share of female workers, the share of employment in routine works, region fixed effects and a time dummy for the period 2002-2010.
* Denotes significance at the 10% level, ** Denotes significance at the 5% level, *** Denotes significance at the 1% level.

Table 1.5 (and Panel a of Figure 1.5) presents the results for the change in the log of real average hourly earnings as a dependent variable. As the first row shows, the effects are negative and significant in the middle of the hourly wage distribution. Particularly, the table also shows that the effect of the change in import per worker is negative and significant for those from the 30^{th} to the 70^{th} quantile of the distribution. At the other quantiles of the hourly wage distribution, the effect is not significant. This implies import competition has more of polarising effect rather than increasing inequality per se.

Instead, Table 1.6 (and Panel b of Figure 1.5) presents the results when using the change in the log of total paid hours worked as a dependent variable. In this case, the

Figure 1.5 Estimates at Different Quantiles for Hourly Wage and Hours Worked



a. Quantile with individual controls   b. Quantile with individual controls

Note: Estimation coefficients and confidence interval from quantile regression with individual characteristics. The dependent variable is the change in the log of hourly wage and total hours worked. Estimation is on all workers.

Table 1.6 Models for Hours Worked After Controlling for Individual Characteristics

	Dependent Variable: Change in the log of total hours worked								
	(1) 10^{th}	(2) 20^{th}	(3) 30^{th}	(4) 40^{th}	(5) 50^{th}	(6) 60^{th}	(7) 70^{th}	(8) 80^{th}	(9) 90^{th}
ΔIPW_r	-0.608	0.565	0.649	-0.224	-0.009	-0.001	-0.197	-0.044	0.067
	(0.840)	(0.674)	(0.695)	(0.328)	(0.130)	(0.102)	(0.128)	(0.155)	(0.372)
Lag manuf. share	-103.9**	12.87	-30.89	-29.16**	-8.035	2.765	6.629	-13.05	-0.341
	(50.653)	(55.342)	(59.907)	(14.763)	(7.168)	(7.721)	(9.756)	(11.423)	(21.589)
Lag female share	-0.464	11.53	-12.54	-22.33***	-9.248***	-3.880*	-1.679	6.780*	6.957
	(15.773)	(16.766)	(17.032)	(5.700)	(3.058)	(2.323)	(2.347)	(4.041)	(7.565)
Lag routine share	-24.64	6.496	-5.546	5.230	1.731	2.466*	5.381***	-1.169	0.423
	(19.391)	(8.513)	(9.128)	(4.073)	(1.508)	(1.467)	(1.477)	(3.464)	(7.403)
Time dummy	-4.013**	0.612	-2.317***	-0.883**	-0.187	0.216	0.699***	1.354***	3.576***
	(1.872)	(0.888)	(0.730)	(0.449)	(0.178)	(0.138)	(0.193)	(0.329)	(0.650)
Region FE	yes	yes	yes	yes	yes	yes	yes	yes	yes
F-test	120.1	120.1	120.1	120.1	120.1	120.1	120.1	120.1	120.1
Partial R^2	0.436	0.436	0.436	0.436	0.436	0.436	0.436	0.436	0.436
R^2	0.116	0.0278	0.112	0.197	0.172	0.0802	0.130	0.282	0.351
Obs	255	255	255	255	255	255	255	255	255

Notes: All regressions include constants and NUTS2 level clustered standard errors (in parenthesis). For all quantile regression, we control for start of period region characteristics such as the share of manufacturing employment, the share of female workers, the share of employment in routine works, region fixed effects and a time dummy for the period 2002-2010.
* Denotes significance at the 10% level, ** Denotes significance at the 5% level, *** Denotes significance at the 1% level.

results are negative for those at the bottom, middle and upper-middle part of the distribution of hours worked but they are not statistically significant. Meanwhile, its effect on those who are at the top and lower middle quantiles is positive but still insignificant.

Generally, these last two tables (and their graphical representation in the two panels of Figure 1.5) indicate that the labour market adjustment to import shocks occurs through a reduction in hourly earnings rather than hours worked. In other words, the adjustment occurs on prices rather than quantity, which is an important observation to keep in mind when, for example, examining statistics on unemployment rates (which may hide part of the effects of a trade shock).

1.5.2 Heterogeneity and Robustness

Having established the main results of this paper, it is relevant to verify whether there is heterogeneity across different groups of workers. Thus, I now turn to conduct the analysis on relevant sub-samples to see if the effects of trade shocks are concentrated on a particular group of workers.

The various panels in Table 1.7 report results on different sub-samples of workers, namely: male, female, manufacturing, non-manufacturing, full time, and part time workers. I also show results excluding London from the analysis (becuase of its peculiarities) or using the year 2007 instead of 2010 as a reference year. The corresponding plots of the estimated coefficients are presented in the Appendix B.

As panels A.1 and A.2 of Table 1.7 show, there is no significant effect of the change in import per worker on the change in the wage of female and male workers throughout the wage distribution. This is possibly due to the fact that splitting the sample distorts the distribution, so that we are now comparing wages within gender. Although insignificant, the effect is negative for those at the 20^{th} percentile.

Panel A.3 and A.4 report the estimates for manufacturing and non-manufacturing workers. The results for manufacturing workers show a negative and significant effect for those from the 30^{th} to the 50^{th} quantile; whereas, for those above the 50^{th} percentile, the effect becomes insignificant although it remains negative. The result for non-manufacturing workers is not significantly different from zero throughout the wage distribution.

Furthermore, Panels A.5 and A.6 present the results for full time and part time workers. In this case, I find positive and significant effects (at the 10% level) for full time workers at the 20^{th} and 30^{th} quantiles, and positive but insignificant effects for the remaining percentile of the wage distribution. Meanwhile, the effect for part time workers is negative for those at the lower part of the wage distribution and positive for those above the 50^{th}

Table 1.7 Models After Controlling for Individual Characteristics on the Subgroups

	Dependent Variable: Change in the log of weekly wage								
	(1) 10^{th}	(2) 20^{th}	(3) 30^{th}	(4) 40^{th}	(5) 50^{th}	(6) 60^{th}	(7) 70^{th}	(8) 80^{th}	(9) 90^{th}
Panel A.1 Female Workers									
ΔIPW_{rt}	0.763	-0.056	0.773	0.746	0.680	0.521	0.495	0.555	1.023
	(1.022)	(0.547)	(0.612)	(0.559)	(0.426)	(0.347)	(0.359)	(0.502)	(0.908)
R^2	0.069	0.255	0.374	0.403	0.344	0.134	0.019	0.110	0.160
N	255	255	255	255	255	255	255	255	255
Panel A.2 Male Workers									
ΔIPW_{rt}	0.375	-0.478	0.206	0.326	0.445	0.431	0.415	0.332	0.616
	(1.104)	(0.587)	(0.571)	(0.477)	(0.405)	(0.344)	(0.315)	(0.324)	(0.463)
R^2	0.064	0.055	0.056	0.072	0.082	0.097	0.111	0.178	0.252
N	255	255	255	255	255	255	255	255	255
Panel A.3 Manufacturing Workers									
ΔIPW_{rt}	0.173	-0.749	-1.209**	-1.093**	-0.822*	-0.342	-0.194	-0.279	-0.121
	(0.935)	(0.582)	(0.570)	(0.468)	(0.425)	(0.370)	(0.361)	(0.413)	(0.525)
R^2	0.063	0.106	0.067	0.062	0.057	0.065	0.046	0.060	0.057
N	256	256	256	256	256	256	256	256	256
Panel A.4 Non-manufacturing Workers									
ΔIPW_{rt}	0.871	-0.370	-0.200	-0.114	0.025	0.234	0.130	0.031	0.456
	(0.839)	(0.598)	(0.557)	(0.464)	(0.390)	(0.337)	(0.337)	(0.322)	(0.392)
R^2	0.093	0.092	0.059	0.058	0.069	0.080	0.0605	0.087	0.163
N	256	256	256	256	256	256	256	256	256
Panel A.5 Full-time Workers									
ΔIPW_{rt}	1.946	1.225*	0.826*	0.678	0.584	0.713	0.680	0.488	0.080
	(1.485)	(0.668)	(0.481)	(0.517)	(0.595)	(0.648)	(0.692)	(0.666)	(0.646)
R^2	0.251	0.141	0.082	0.058	0.023	0.029	0.045	0.059	0.096
N	255	255	255	255	255	255	255	255	255
Panel A.6 Part-time Workers									
ΔIPW_{rt}	-0.060	-1.304*	-0.745	-0.401	-0.161	0.509	0.791	0.718	0.602
	(1.644)	(0.749)	(0.488)	(0.675)	(0.849)	(0.982)	(1.007)	(0.913)	(0.840)
R^2	0.076	0.086	0.163	0.223	0.203	0.175	0.142	0.130	0.110
N	255	255	255	255	255	255	255	255	255
Panel A.7 Excluding London									
ΔIPW_{rt}	0.562	-0.316	-0.850	-0.766*	-0.428	-0.203	-0.368	-0.558	-0.496
	(0.823)	(0.630)	(0.573)	(0.436)	(0.384)	(0.336)	(0.345)	(0.378)	(0.522)
R^2	0.063	0.165	0.159	0.119	0.077	0.031	0.029	0.0614	0.055
N	245	245	245	245	245	245	245	245	245
Panel A.8 With respect to 2007									
ΔIPW_{rt}	1.058	1.379	0.555	-0.270	-0.179	0.215	0.471	0.846	0.441
	(1.380)	(1.478)	(1.083)	(0.766)	(0.692)	(0.615)	(0.689)	(0.750)	(0.935)
R^2	0.075	0.161	0.136	0.091	0.056	0.032	0.033	0.058	0.065
N	254	254	254	254	254	254	254	254	254

Notes: All regressions include constants and NUTS2 level clustered standard errors (in parenthesis). For all quantile regression, we control for start of period region characteristics such as the share of manufacturing employment, the share of female workers, the share of employment in routine works, region fixed effects and a time dummy for the period 2002-2010.
* Denotes significance at the 10% level, ** Denotes significance at the 5% level, *** Denotes significance at the 1% level.

percentile. However, it is only significant at the 10% level for those at the 20^{th} quantile. In estimating these two regressions, I control for all the individual characteristics other than a full time indicator.

Panels A.7 and A.8 provide two robustness checks by excluding the four NUTS-3 regions of London and using the year 2007 as a reference, respectively. In both cases, I find a similar pattern (see Figure A.8b. and A.9b. in the appendix) as the main regression result where the negative effect is concentrated around the 40^{th} quantile, although the precision of the estimate declines and become insignificant for 2007.

It may be possible that migration from regions highly exposed to the trade shock to less exposed regions may provide another channel for the adjustment in the local labour markets. In order to verify if this could be the case, I regress a decadal change in import per worker of a NUTS-3 region on the region's change in working age population (results are available in Table A.3 of the Appendix).

Under the preferred model specification[11], the coefficient for the effect of the change in import per worker on the change in regional working age is insignificant. This is in line with previous empirical findings such as the one by Kovak (2013) who finds a sluggish labour market adjustment through workers mobility following labour market shocks. Thus, we can exclude intra-UK migration as a possible channel.

1.6 Conclusion

Unlike the well-documented aggregate effect of trade liberalisation, which finds that countries gain from trade liberalisation by specialising in areas of their comparative advantage, its distributional effect has been the focus of recent literature. Here, the main mechanism for the distributional pass-through is its heterogeneous effect on the labour market outcomes of different groups of workers. These groups can be classified by their skills, industries, gender, location, or even age. Therefore, the question of who gains and who loses from globalisation remains an empirical issue.

In this paper, I investigate the causal relationship between import shocks and labour market outcomes, with an special emphasis on the wage distribution. The descriptive results suggest that there are variations in changes in exposure to import competition among British local labour markets. And by combining these measures of variations with worker level data, I provide evidence on the different effect of the import shock caused by China's integration in the multilateral system on the outcomes of workers, depending on

[11]The preferred model includes region characteristics such as lagged share of manufacturing workers, lagged share of female workers, lagged share of worker in routine sector, region and year fixed effects.

their position on the wage distribution. Importantly, I show that it is important to control for individual characteristics and compositional changes across time.

From the analysis, I find differential effects depending on the position of workers in the wage distribution. An increase of £1000 import per worker reduces the wage of workers who are between the 30^{th} and the 50^{th} quantiles of the wage distribution, whereas its effect on those at very lower and upper quantiles are not significant. These findings suggest that import shocks can contribute to the rise in wage polarisation by negatively affecting the middle wage earners.

To disentangle the wage effect into its primary components, I consider the effect on the hourly wage and total hours worked. From this exercise, I can conclude that the effect of import exposure mainly manifests itself on the hourly wage of workers instead of their total hours worked. That is, an increase in regional import per worker causes a reduction in the hourly wages of those in the middle of the hourly wage distribution.

By splitting workers into various sub-groups, the analysis is enriched by showing relevant heterogeneity. In particular, I find that the effect of the change in import exposure is negative for manufacturing workers at lower parts of the wage distribution. The effect for female, male, full time and part time workers is different over the wage distribution, but it is not significantly different from zero.

To address the potential adjustment of local labour markets to import exposure by workers mobility, I show that changes in regional working age population are not explained by the trade shock, confirming that there is no evidence suggesting the adjustment of the labour market from high to low exposed regions through labour mobility. That is, change in import exposure does not significantly affect the change in the working age population in a region.

Results in this paper are important in order to understand how the trade shock is transmitted in the local labour markets. Since the adjustment takes place through prices (i.e. wages) instead of quantity (i.e. hours worked), it is important to consider both dimensions together to put in place policies that can cushion workers negatively affected by trade integration. And the heterogeneity displayed by various groups of workers further highlight the necessary level of detail for any policy intervention. This is all the more important at a time when globalisation is under threat.

Chapter Two

Imports, FDI Spillovers and Firm Performance

2.1 Introduction

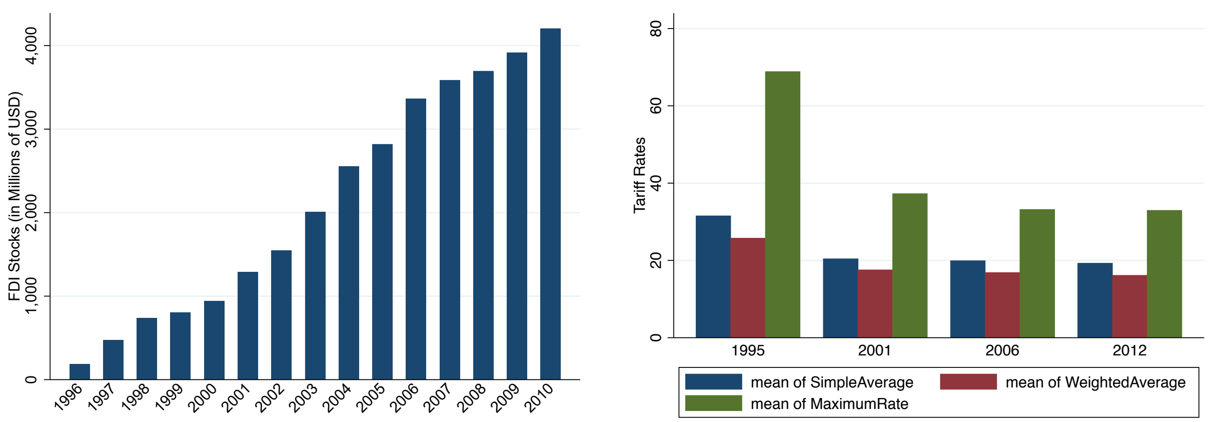
Technological adoption and productivity gains by a firm can come from various sources. And the question of how firms improve their productivity remains a topic of great interest among both policy-makers and researchers alike. The literature documents various sources of firm productivity improvement (Topalova and Khandelwal, 2011). In particular, productivity gains from imported intermediate goods and learning from other foreign firms, the so-called FDI-spillovers, have long occupied the central stage in the international trade literature. A number of studies find a positive effect of imported inputs on firm productivity (Amiti and Konings, 2007; Kasahara and Rodrigue, 2008; Zhang, 2017), whereas the effects of foreign direct investment (FDI) spillovers remain case-specific (Lu, Tao, and Zhu, 2017; Javorcik, 2004).

Although there is a rich literature that focuses on productivity gains from either imported inputs or from proximity to other more productive firms separately, the combined effect and the question of whether the gains from the two are complement or substitute remains unanswered. Answering this question directly shapes policies on whether to employ trade liberalisation policies, FDI promotion policies or a combination of both. Furthermore, it helps in answering related questions such as how much of the productivity improvement comes from spillovers through imitation of better management, how much from skill transfers through employees job switching as noted by Dunning (2015), and how much from the use of high-quality imported inputs. Most importantly, the literature on the underlining mechanism of productivity gain by a firm is recent and limited (Halpern, Koren, and Szeidl, 2015), particularly in a developing country context.

For the empirical analysis, I use firm-level panel data from Ethiopia over the 1996-2010 period. The data represent the population of Ethiopian manufacturing plants with at least

10 employees and use electric-powered tools in production. Ethiopia is a particularly interesting case for many reasons. First, like most developing countries, following policy advises from international organisations such as the World Bank (WB) and the International Monetary Fund (IMF), trade liberalisation and attracting FDI has been a top priority policy of the country for decades. The involvement of such organisations makes the policies exogenous shocks to the economy (Fiorini, Sanfilippo, and Sundaram, 2019). Second, besides being a developing country, Ethiopia went through rapid growth in the stock of FDI and a major trade liberalisation episode during the period under investigation. The following figures illustrate these facts. Figure 2.1a reports the FDI stock for the period from 1996 to 2010 and Figure 2.1b shows the trend in the tariff rate. FDI stock, for instance, increased from less than two hundred million USD to more than 4 billion USD between 1996 to 2010. Meanwhile, the average simple tariff declines from 31.5% to 19.3% during the same period.

Figure 2.1 Trends in Stock of FDI and Tariff Rates



(a) Trend in FDI Stock

(b) Trend in Tariff Rates

Notes: (a) Source: United Nations Conference on Trade and Development (b) WITS dataset of the World Bank.

This paper identifies the productivity gains from imported inputs and estimates FDI spillovers after accounting for gains from imports. To this end, first, I estimate total factor productivity (TPF) with and without controlling for imported input, and then estimate FDI spillovers. This, in turn, answers the main research question: to what extent imports of more inputs complement or substitute the effect of FDI spillovers on productivity.

From the analysis, I find evidence of productivity gains from imported inputs. Controlling for productivity gains from imported inputs is important in identifying the performance-enhancing effect of FDI. Specifically, I find a small difference in productivity spillovers before and after controlling for imported inputs indicating limited substitutability between the gains from imports and FDI spillovers. This implies that productivity gains from the two sources are different in nature. I also find a positive effect of imports on productivity for both domestic and foreign-owned firms. Meanwhile, I find a positive backward spillover and a negative horizontal and forward spillover.

This paper is related to several strands of literature. First, the recent international trade literature focuses on the gains from the import of intermediate inputs. These studies find evidence in favour of an import premium, that is, firms that import are more productive and perform better on arrays of firm performance measures (Bøler, Moxnes, and Ulltveit-Moe, 2015; Bas and Strauss-Kahn, 2014; Kasahara and Lapham, 2013). Specifically, studies show that importation of intermediate inputs promotes R&D investment, improves productivity, increases the volume and scope of exports, and affects technology choice (Smeets and Warzynski, 2013; Bøler et al., 2015; Bas and Berthou, 2017). The present paper is related to this line of research by showing the existence of firm-level productivity improvement from importing and identifying the gains from intensive margins.

Second, studies in the area of development economics emphasise the role of management practices and managerial human capital in improving the performance of manufacturing firms in developing countries (Harrison and Rodríguez-Clare, 2010; Bruhn, Karlan, and Schoar, 2010; Bloom, Eifert, Mahajan, McKenzie, and Roberts, 2013; Bloom et al., 2016). Domestic firms can learn from the organisational and managerial system of more efficient foreign firms. For instance, Arnold and Javorcik (2009), using Indonesian data from 1983 to 2001, attribute productivity improvements of acquired plants to the employment of organisational and managerial systems by foreign firms. Specifically, in a developing countries context, foreign-owned firms are better at using imported inputs, pay higher wages (Lipsey and Sjöholm, 2004), and are more efficient. Part of the foreign premium is attributed to their lower fixed cost of importing and thus imports of more products. This study contributes to this literature by focusing on a developing country for whom studies are limited.

Third, the paper contributes to the literature on FDI spillovers. Previous studies documented that firms are affected by the geographical presence of other firms (Gaubert, 2018; Greenaway and Kneller, 2008). Theoretically, the effect of the presence of foreign firms on the local economy can be either direct or indirect. The direct effect is towards the firm acquired by a foreign firm. Whereas, the indirect or spillover effect is towards domestic and other foreign firms which operate nearby. While the former is mostly positive, the latter can be either positive or negative (Girma, Gong, Görg, and Lancheros, 2015).

Depending on the position of firms in relation to a foreign firm, the indirect effect (spillovers) could be either horizontal or vertical. The latter further splits into backward and forward spillovers. Horizontal spillovers refer to a relationship between a foreign firm and firms within the same industry as a foreign firm. Here the mechanism of pass-through is mainly through demonstration and competition. However, these spillovers could be negative since these firms are direct competitors of the foreign firm. Backward spillovers occur when firms supply intermediary input to the foreign firm and through that process gain efficiency and production know-how. Meanwhile, forward spillovers occur when firms buy intermediary inputs from the foreign firm which increases their productivity (Javorcik, 2004). Thus far, the empirical literature on FDI spillovers does not reach a consensus regarding the overall effects of FDI. Some studies find a positive spillover (Haskel, Pereira, and Slaughter, 2007; Keller and Yeaple, 2009), others find mixed-effects (Javorcik, 2004; Lopez, 2008), and others find a negative spillover effect (Lu et al., 2017; Javorcik and Spatareanu, 2008; Fatima, 2016). Therefore, the extent and nature of FDI spillovers seem to remain case-specific.

The remaining part of the paper is organised as follows. In section 2, I describe the datasets I use for this study. Section 3 presents the econometric methodology and identification strategies. Section 4 discusses the empirical results, and the last section concludes.

2.2 Dataset and Descriptive Statistics

As the main data source, I use Ethiopian manufacturing data from 1996 to 2010. The data is collected by the Central Statistics Agency (CSA) on the universe of Ethiopian medium and large scale manufacturing firms[1] which hires more than 10 workers and uses electricity powered tools for production. The manufacturing firms in the dataset are classified into 4-digit ISIC Rev. 3 industry level.[2]

The data contain information to estimate firm productivity. Specifically, the data include firm characteristics such as sales, capital, investment, number of employees, material inputs, ownership status, trade status, number of imported inputs, year of establishment, and region of location. The final dataset constitute an unbalanced panel of around 1,500 firms or 15,958 observations over the period between 1996 and 2010.[3]

Table 2.1 Descriptive Statistics of Main Variables of Interest

	Obs	Mean	Std Dev	Min	Max
Total sales	15,612	1,310,428	5,241,826.89	69.38	1.30×10^{8}
Capital per worker	14,582	5,618.24	22,423.40	0	852,324.12
Output per worker	15,083	12,262.39	17,8629.52	0.05	16,374,833
Material per output	14,840	169.61	4,330.40	0	301,099.97
No. of permanent workers	15,503	91.03	260.20	1	7,909
Log of output	15,059	10.01	3.17	0	20.37
Log of labour	15,503	3.31	1.38	0	8.98
Log of capital	14,893	9.59	3.38	-7.32	19.81
Log of material	15,445	9.86	2.87	-2.52	19.09
Log of investment	15,508	-0.41	10.46	-23.08	17.89
Exporter dummy	15,898	0.04	0.21	0	1
Importer dummy	15,898	0.66	0.47	0	1
Import material share	15,817	0.34	0.39	0	1
No. imported inputs	15,826	2.85	2.88	0	12
Private dummy	15,737	0.88	0.33	0	1
Foreign dummy	15,898	0.04	0.20	0	1
Age	15,747	13.24	14.39	0	88

Note: Monetary units are in USD and 1996 is used as a base year.

[1] The data is at an establishment (plant) level, and I use a firm to represent these units of observations.

[2] Other works which use the same dataset include Abebe, McMillan, and Serafinelli (2018); Fiorini et al. (2019).

[3] Table B.1 in the appendix presents the mean values and number of observations for each year under consideration.

In the empirical analysis, I combine data from CSA with other datasets from a number of sources. Tariff data from WITS of the World Bank is used to calculate input and output tariffs (at 2-digit ISIC level). An Input-Output table is obtained from the Ethiopian Development Research Institute (EDRI) for the period 2005/06 (EDRI, IDS and IFPRI-Ethiopian Development Research Institute (EDRI) and Institute of Development Studies (IDS) and International Food Policy Research Institute (IFPRI), 2014). The table is used in the construction of FDI spillovers. Monetary values are deflated by firm level price indices[4]. Moreover, data on exchange rates come from IMF financial statistics to convert monetary units into US dollar.

Table 2.1 provides summary statistics of the main variables of interest for the sample period. On average, 4% of firms export whereas 66% of them import during the period considered. Moreover, imports accounts for 34% of the material input value with the average number of imported inputs being 2.85. From the perspective of ownership structure, 88% of firms are privately owned while 4% are foreign.

Descriptive statistics for the initial and final year in the sample provide evidence of the change in the structure of the Ethiopian economy during this time period. For instance, as Table 2.2 shows, the total number of firms increases from 617 to 1,958 which represents more than 200% increase in the number of establishments. Moreover, the performance of firms, as measured by output per worker increased by 91%. At the same time, the average sales and number of workers declined by 58% and 36% respectively. This can be explained by the relatively small size of new entrants.

A further descriptive analysis documents some empirical facts that distinguish importers and foreign-owned firms from their non-trading and domestic counterparts. Three stylized facts emerge from the summary statistics.

Stylized Fact 1: Importing and foreign-owned firms perform better

Table 2.3 reports several descriptive statistics. The first column presents descriptive statistics for non-importers, while the second column provides it for importers. The last two

[4] The computation of the firm-level deflator follows a study by Smeets and Warzynski (2013). In line with Fiorini et al. (2019), I make an adjustment to compensate for the missing product codes and repetitive product categories in the dataset. Appendix A1 provides more details on the construction of this price index.

Table 2.2 Evolution of Mean Values for Main Variables of Interest

	1996	2010	Change between 1996 & 2010 (%)
Total sales	1,393,620.84	1,448,954.48	4
Capital per worker	5,298.30	5,242.53	-1.1
Output per worker	8,657.87	35,006.25	304.3
Material per output	210.11	77.29	-63.2
No. of permanent workers	132.73	83.76	-36.9
Exporter dummy	0.04	0.04	0
Importer dummy	0.66	0.61	-7.6
Import material share	0.30	0.33	10
No. imported inputs	2.12	2.94	38.7
Foreign owned	0.04	0.05	25
Privately owned	0.74	0.94	27
No. establishments	617	1958	217.3

panels show descriptive statistics for domestic and foreign firms, respectively. As can be seen from Table 2.3, importing and foreign-owned firms perform better in terms of several performance measures. They hire more workers, have higher levels of sales, employ more capital per worker and workers are more productive as it is measured by output per worker.

Furthermore, in line with findings of the previous literature, importers and foreign-owned firms are on average larger (in terms of the number of workers and sales), more productive (output per worker), more capital intensive (capital per worker) and are more likely to be exporters.

Stylized Fact 2: Foreign-owned firms import more and are more likely to be importers
Table 2.4 shows the average values of an importer dummy variable and import shares of intermediate inputs used between domestic and foreign-owned firms through time. As the table shows, on average foreign-owned firms are more likely to be importers and use more imported products. For instance, in 2010, 57.8% and 78.3% of domestic and foreign firms import. Likewise, for the same period, 31.6% and 51.8% of material inputs of domestic and foreign firms are imported.

Table 2.3 Descriptive Statistics by Import and Ownership

	Non-importer	Importer	Mean Diff.
Total sales	607,841.85	1,662,734.70	-1,054,893***
Capital per worker	4,884.72	5,973.06	-1,088***
Output per worker	6,782.57	14,985.45	-8,203***
Material per output	44.05	234.73	-191**
No. permanent of workers	49.42	111.82	-62***
Exporter dummy	0.02	0.06	-0.031***
Importer dummy	0	1	-1
Import material share	0	0.50	-0.503***
No. of imported inputs	0	4.26	-4.263***
Private dummy	0.90	0.86	0.039***
Age	11.61	14.07	-2.456***
Observations	5,333	10,565	
	Domestic	Foreign	Mean Diff.
Total sales	1,248,395.65	2,805,388.97	-1,556,993***
Capital per worker	5,529.24	7740.04	-2,211**
Output per worker	12,213.27	13,456.55	-1,243
Material per output	123.03	1333.55	-1,210***
No. permanent of workers	89.95	115.30	-25**
Exporter dummy	0.04	0.10	-0.058***
Importer dummy	0.66	0.777	-0.117***
Import material share	0.33	0.52	-0.197***
No. of imported inputs	2.82	3.50	-0.679***
Private dummy	0.88	0.93	-0.057***
Age	13.01	18.54	-5.524***
Observations	15,239	659	

Note: Monetary units are in 2010 USD.

Stylized Fact 3: For a given size foreign-owned firms import more inputs

A simple ordinary least squares regression shows a positive correlation between the log of the number of imported inputs and a foreign dummy. Even after controlling for firm size, the correlation remains positive and highly significant. The estimated coefficients indicate a positive association between the two variables, which implies that foreign firms use more imported inputs. The literature also finds that foreign firms are efficient in using imported inputs (e.g. Arnold and Javorcik (2009)).

The main variable of interest is the number of varieties a firm chooses to import and its effect on productivity. Following the above stylized fact, in the analysis, I allow productivity gains to differ between domestic and foreign firms. However, due to the absence of

Table 2.4 Mean of Import and Import Share by Ownership and Year

Year	Import		Import Share	
	Domestic	Foreign	Domestic	Foreign
1996	0.657	0.696	0.297	0.403
1997	0.683	0.682	0.306	0.538
1998	0.734	0.684	0.326	0.433
1999	0.692	0.679	0.318	0.484
2000	0.726	0.786	0.339	0.521
2001	0.657	0.629	0.307	0.358
2002	0.648	0.791	0.327	0.509
2003	0.708	0.795	0.365	0.561
2004	0.707	0.792	0.347	0.493
2005	0.705	0.860	0.410	0.551
2006	0.684	0.745	0.361	0.509
2007	0.623	0.820	0.327	0.620
2008	0.635	0.800	0.325	0.562
2009	0.600	0.840	0.289	0.597
2010	0.578	0.783	0.316	0.518

information on the source country in the data, unlike the literature that commonly considers product-country pairs as a variety (Bas and Strauss-Kahn, 2014), I use the product as a variety. Similarly, the data does not distinguish between final and intermediary inputs, thus the analysis does not differentiate between the two.

2.3 Econometric Methodology

This section presents the baseline econometric specification and identification strategy. To this end, first, I discuss the estimation of productivity with and without accounting for the number of imported varieties. Next, I use the estimated productivity as a dependent variable to identify the effect of FDI spillovers on productivity before and after accounting for imported inputs.

In short, the analysis develops in two steps. First, I estimate productivity before and after isolating the effect of imports. Second, I analyse the impact of isolating productivity gains from imported inputs on FDI spillovers.

2.3.1 Productivity Estimation

The productivity level of each firm is estimated as a Solow residual from the production function. To correct for endogeneity, I followed the approach by Olley and Pakes (1996) in both cases. In order to account for imported inputs, I follow Halpern et al. (2015)'s approach. They suggest including the number of input varieties a firm chooses to import into the production function and controlling for the productivity effect of those inputs. This method accounts for the effect of each imported input on productivity and derives an estimate for imported inputs adjusted productivity measure, ω_{jst}.

Assuming a Cobb-Douglas production function, the production technology of a firm j at time t can be represented as:

$$Q_{jt} = \Omega_{jt} K_{jt}^{\alpha_k} L_{jt}^{\alpha_l} \prod_{i=1}^{N} X_{jit}^{\gamma_i}, \tag{2.1}$$

where Ω_{jt} stands for Hicks-neutral productivity term, K_{jt} represents capital, L_{jt} captures labour, X_{jit} stands for an intermediate input i and γ_i indicates the importance of the intermediate input for production. With X_{jitF} and X_{jitH} denoting imported and domestic inputs respectively, intermediate inputs enter the production function in a CES form,

$$X_{jit}^{\gamma_i} = [(B_{jit} X_{jitF})^{\frac{\theta-1}{\theta}} + (X_{jitH})^{\frac{\theta-1}{\theta}}]^{\frac{\theta}{\theta-1}}, \tag{2.2}$$

where B_{jit} and θ represent the input quality effect of the imported inputs relative to domestic inputs and the elasticity of substitution of domestic and imported inputs, respectively.

Halpern et al. (2015) show that by incorporating imported inputs, one can rewrite the production function, in natural logarithm, as[5]

$$q_{jt} = \alpha_0 + \alpha_l l_{jt} + \alpha_k k_{jt} + \gamma(m_{jt} - \rho) + \gamma a G(n_{jt}) + \omega_{jt} + \varepsilon_{jt}. \tag{2.3}$$

A simple rearrangement and substitution of δ for the product of γ and a gives us a value-

[5] Appendix A3 presents the mathematical derivation of equation (2.3).

added of equivalence of equation (2.3) as,

$$q_{jt} - \gamma(m_{jt} - \rho) = \alpha_0 + \alpha_l l_{jt} + \alpha_k k_{jt} + \delta G(n_{jt}) + \omega_{jt} + \varepsilon_{jt}, \tag{2.4}$$

where γ corresponds to the coefficient for material inputs (i.e., the total weight of all intermediate goods), a is per-product import gain, and $G(n_{jt})$ corresponds to the relative importance of imported inputs.

Following Gandhi, Navarro, and Rivers (2011), γ is calculated as the material share from total revenue. Having computed γ, the coefficient estimate, δ, can be used to calculate the per-product import gain, a. In the baseline specification, I assume this gain to be the same for all firms. In other model specifications, I let this gain vary between domestic and foreign firms.[6]

Given that the production function includes a productivity parameter which is unobservable by the researcher but observed by the firm, estimating equation (2.4) using OLS leads to biased estimates. To solve this problem of endogeneity that affects OLS estimation, researchers suggest a number of approaches. One of the most widely used approaches is by Olley and Pakes (1996) (OP henceforth) who suggest the use of an investment function, which embodies information on productivity, as a proxy. Levinsohn and Petrin (2003) (LP) suggest the use of material input instead of investment. To address the simultaneity bias in the labour coefficient, I adopt Ackerberg, Caves, and Frazer (2015)'s extension in the OP context and identify the labour coefficient in the second stage together with the capital coefficient.

The OP approach develops in two stages.[7] First, the OP model implies that investment is a strictly monotonic function of productivity and other state variables. From this, we can inverse the investment policy function and express productivity as a function of investment and other state variables, i.e., $\omega_{jt} = f_{jt}^{-1}(i_{jt}, l_{jt}, k_{jt})$. Thus, rewrite equation (2.4) as:

$$q_{jt} - \gamma(m_{jt} - \rho) = \alpha_0 + \alpha_l l_{jt} + \alpha_k k_{jt} + \delta G(n_{jt}) + f_{jt}^{-1}(i_{jt}, l_{jt}, k_{jt}) + \varepsilon_{jt}, \tag{2.5}$$

[6] Doing so involves assuming that the relative importance of each input, $G(n_{jt})$, takes different values and this, in turn, changes the value of δ and thus a.

[7] Studies also consider estimation of the survival decision as another stage to control for non-random exit of firms.

$$= \Phi(i_{jt}, k_{jt}, l_{jt}) + \delta G(n_{jt}) + \varepsilon_{jt}, \tag{2.6}$$

where $\Phi_{jt}(.)$ is parameterised as a third-order polynomial function of i_{jt}, l_{jt} and k_{jt}. The OLS regression on equation (2.6) provides the first stage of the OP estimation. This assumes a moment condition of $E[\varepsilon_{jt}|I_{jt}] = 0$, where I_{jt} is an information set that includes current and past productivity shocks. From this, I estimate the fitted value of $\Phi_{jt}(.)$, $\widehat{\Phi}_{jt}(.)$ and $\widehat{\delta}$ and express productivity as $\omega_{jt} = \widehat{\Phi}_{jt}(.) - \alpha_0 - \alpha_l l_{jt} - \alpha_k k_{jt}$.

For the second stage, the OP model assumes that the productivity shock, ω_{jt}, evolves according to a first order Markov process. This implies,

$$\omega_{jt} = E[\omega_{jt}|I_{jt-1}] + \xi_{jt} = E[\omega_{jt}|\omega_{jt-1}] + \xi_{jt} = h(\omega_{jt-1}) + \xi_{jt}, \tag{2.7}$$

where ξ_{jt} is an innovation term satisfying $E[\xi_{jt}|I_{jt-1}] = 0$.

After incorporating the above assumptions, a production function can be rewritten as

$$q_{jt} - \gamma(m_{jt} - \rho) = \alpha_0 + \alpha_l l_{jt} + \alpha_k k_{jt} + \delta G(n_{jt}) + h(\omega_{jt-1}) + \xi_{jt} + \varepsilon_{jt}, \tag{2.8}$$

$$q_{jt} - \gamma(m_{jt} - \rho) = \alpha_0 + \alpha_l l_{jt} + \alpha_k k_{jt} + \delta G(n_{jt}) + h(\widehat{\Phi}_{jt-1}(.) - \alpha_0 - \alpha_l l_{jt-1} - \alpha_k k_{jt-1}) + \xi_{jt} + \varepsilon_{jt}. \tag{2.9}$$

The second line comes from substituting ω_{jt-1} with its lagged equivalence, and the $h(.)$ function takes a simple linear functional form. The conditional moment condition required for the second stage is, $E[\xi_{jt} + \varepsilon_{jt}|I_{jt-1}] = 0$. By estimating equation (2.9) using GMM or non-linear least squares, I identify the unbiased coefficient estimates for inputs. Then, I use these coefficient estimates to construct productivity for each firm as $TFP_{jt} = Exp[q_{jt} - \widehat{\alpha}' x_{jt}]$, where $\widehat{\alpha}$ stands for a vector of estimated parameters and x_{jt} denotes inputs.

2.3.2 FDI Estimation Specification

Having measured productivity with and without controlling for gains from imported inputs, I proceed with estimating the effect of FDI spillovers. While the former follows Halpern et al. (2015) method as discussed above, the latter measure of productivity used the traditional OP approach. And by comparing the coefficient estimates from the two models, I can argue if controlling for imported inputs dampens or intensifies FDI spillovers.

Here, following the commonly used approach by Javorcik (2004), I define three types of FDI spillovers as follows:

$$Horizontal_{st} = \frac{\sum_{j\forall j\in s} ForeignShare_{jt} \times Q_{jt}}{\sum_{j\forall j\in s} Q_{jt}}, \tag{2.10}$$

$$Backward_{st} = \sum_{k if k\neq s} \alpha_{sk} Horizontal_{kt}, \tag{2.11}$$

$$Forward_{st} = \frac{\sum_{m if m\neq s} \sigma_{sm}\left[\sum_{j\forall j\in m} ForeignShare_{jt} \times (Q_{jt} - EX_{jt})\right]}{\left[\sum_{j\forall j\in m}(Q_{jt} - EX_{jt})\right]}, \tag{2.12}$$

where $ForeignShare_{jt}$ is the share of foreign ownership in firm j at year t, Q_{jt} stands for output, EX_{jt} represents exports, and α_{sk} and σ_{sm} correspond to proportions of output supplied by s to sector k and input purchased by sector s from sector m respectively.

To assess how controlling for imported inputs affects the narrative of FDI spillovers and argue whether the two are complementary or substitute with one another, I test the equality of coefficient estimates from the two regressions. In other words, I include a dummy variable to indicate which method is used to measure productivity and regress over the same vectors of spillovers as independent variables. The interaction of spillovers with the dummy represents the effect of accounting for imported inputs on productivity spillovers. The regression results have equivalent interpretation as using seemingly unrelated regression (SUR). In particular, the specification is

$$\begin{aligned}\omega_{jst} = \;& \eta_0 + \eta_1\text{HKS Dummy} + \eta_2 Backward_{st} + \eta_3 Backward_{st}\times\text{HKS Dummy} \\ & + \eta_4 Horizontal_{st} + \eta_5 Horizontal_{st}\times\text{HKS Dummy} + \eta_6 Forward_{st} \\ & + \eta_7 Forward_{st}\times\text{HKS Dummy} + X'_{jst}\lambda + \gamma_s + \gamma_r + \gamma_t + \xi_{jst}.\end{aligned} \tag{2.13}$$

On the right-hand side, I include horizontal, backward and forward FDI spillovers, an index for distinguishing the productivity measures, vectors of firm characteristics as denoted by X_{jst}, industry fixed effects, γ_s, region fixed effects, γ_r, time fixed effects, γ_t, and an iid error term, ξ_{jst}.

2.4 Results

In this section, I present the results from the main analysis. The next subsection presents the results from estimating the production function after controlling for imported inputs. Next, I present the results without accounting for imported inputs. After that, I demonstrate how FDI spillover differs across the two cases.

2.4.1 Baseline Results for Productivity

Studies in international trade literature finds that firms benefited from the technologies embodied in imported inputs (Bas and Strauss-Kahn, 2014). In Table 2.5, I present the regression results of the production function where I account for gains from imported inputs in estimating productivity. This model specification helps us to isolate the productivity gains from imports from overall productivity.

Table 2.5 Coefficient Estimates of Inputs with Domestic and Foreign

	(1) Baseline Estimate	(2) Different gains of import		(3) Including Export
		Domestic	Foreign	
Capital (α_k)	0.115***	0.115***		0.143***
	(0.0140)	(0.0136)		(0.017)
Labour (α_l)	0.211***	0.211***		0.217***
	(0.0350)	(0.0328)		(0.050)
Material (γ_m)	0.476***	0.476***		0.476***
	(0.031)	(0.031)		(0.030)
Export (E)				0.555**
				(0.223)
Per-product Import Gain (a)	0.228*	0.204	0.598	0.233*
	(0.126)	(0.126)	(0.500)	(0.132)
Optimal Import Share (S)	0.944***	0.909***	1.129***	0.944***
	(0.0396)	0.0539	0.0457	(0.040)
Import Efficiency (A)	1.250***	1.214***	1.820	1.256***
	(0.153)	(0.166)	(1.131)	(0.163)
Elasticity of Substitution (θ)	13.637	13.081		13.384
Obs	9,465	9,465		9,465

* Denotes significance at the 10% level, ** Denotes significance at the 5% level, *** Denotes significance at the 1% level.
Robust standard errors in parentheses

The first column of Table 2.5 shows the baseline regression result. As we can see, all the coefficient estimates have the expected sign and magnitude. The coefficient estimates of the inputs are interpreted as elasticities.

In the baseline specification, the coefficient estimate for per-product gain from imports is positive and significant. The per-product import gain, a, of 0.228 implies that the combined use of imported and domestic inputs is 25.6%, [exp(0.228)-1], more efficient for each dollar spent than using only domestic ones. The column also presents the estimates for the price-adjusted quality advantage of imported products relative to domestics ones, A. For the baseline specification, this estimate is 1.25 which implies foreign goods are about 25% better than their domestic counterparts for each dollar of expenditure.

In column two, I estimate the model by allowing foreign and domestic firms to have different gains from imported inputs. The magnitude of the per-product import gain is greater for foreign firms indicating that foreign firms gain more from imports compared to domestic firms. This finding is in line with previous studies which show the foreign firms are better in using imported inputs (Lipsey and Sjöholm, 2004). For comparison, Table 2.6 reports the coefficient estimates from OLS and OP regression results. Column 1 presents coefficient estimates of inputs from the OLS. In column 2 of Table 2.6, I report the estimate results where the production function is measured using the traditional OP method. In both cases, the results have expected signs and magnitudes and are highly significant. OLS coefficient estimates for labour seems slightly smaller in magnitude than that of OP. Mean-

Table 2.6 Models Estimation from OLS and OP Methods

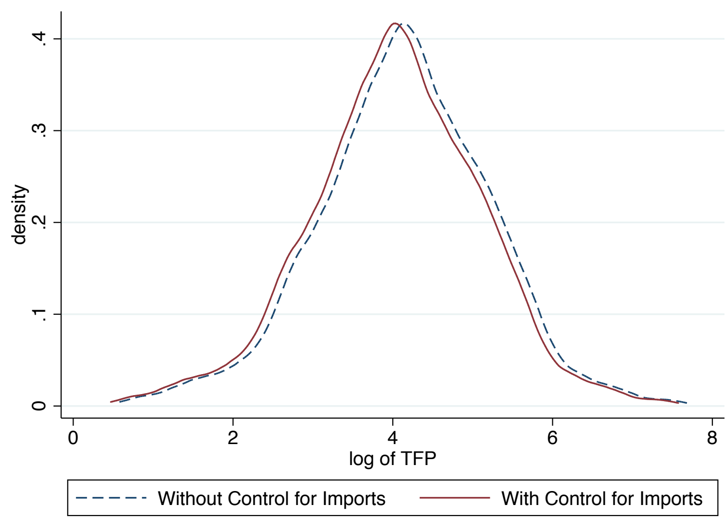
	All firms	
	(1)	(2)
	OLS	OP
Capital(α_k)	0.110***	0.116***
	(0.004)	(0.0138)
Labour (α_l)	0.243***	0.209***
	(0.008)	(0.0337)
Material (γ_m)	0.710***	0.473***
	(0.005)	(0.0304)
R^2	0.893	
Obs	14,506	9,465

* Denotes significance at the 10% level, ** Denotes significance at the 5% level, *** Denotes significance at the 1% level. Robust standard errors in parentheses. Year, region and industry fixed effects are included in the regression.

while, the OLS overestimates the coefficient for capital and material inputs.

After estimating input coefficients, we proceed with estimating productivity as a Solow residual with and without accounting for imported inputs. Figure 2.2 shows the distribution of productivity from the two estimation results, namely, with (i.e., column 1 of Table 2.5) and without (column 2 of Table 2.6) accounting for imported inputs. As the figure shows, accounting for imported inputs shifts the productivity distribution to the left. This is expected since a portion of the productivity gain comes from imported inputs. Furthermore, a t-test for the mean difference of the two distributions indicates there is a statistically significant difference between the two means.

Figure 2.2 TFP With and Without Accounting for Imports



Note: Estimated productivity before and after accounting for imported inputs.

2.4.2 Computing FDI Spillover

Having estimated productivity using the above two methods, namely, with and without accounting for imported inputs, I can proceed with estimating FDI spillovers. To this end,

Table 2.7 Models Estimation of Productivity on FDI Spillover from HKS and OP

	HKS						OP					
	Domestic firms			All firms			Domestic firms			All firms		
	Pooled	FE	Lag	Pooled	FE	Lag	Pooled	FE	Lag	Pooled	FE	Lag
Backward	0.197	0.304		0.151	0.224		0.195	0.299		0.150	0.221	
	(0.179)	(0.212)		(0.172)	(0.203)		(0.176)	(0.209)		(0.170)	(0.200)	
Forward	-1.782***	-1.331***		-1.823***	-1.390***		-1.754***	-1.310***		-1.794***	-1.368***	
	(0.481)	(0.514)		(0.463)	(0.491)		(0.474)	(0.507)		(0.456)	(0.484)	
Horizontal	0.0168*	0.00248		0.0160	-0.000284		0.0166*	0.00246		0.0158	-0.000261	
	(0.010)	(0.012)		(0.010)	(0.011)		(0.010)	(0.012)		(0.010)	(0.011)	
L.Backward			0.386			0.154			0.381			0.152
			(0.237)			(0.230)			(0.233)			(0.226)
L.Forward			1.245*			1.242*			1.229*			1.227*
			(0.656)			(0.650)			(0.646)			(0.641)
L.Horizontal			0.0495***			0.0514***			0.0487***			0.0505***
			(0.013)			(0.013)			(0.013)			(0.013)
ABC	7.335***	7.329***	7.308***	7.334***	7.329***	7.308***	7.438***	7.438***	7.416***	7.443***	7.438***	7.417***
	(0.002)	(0.004)	(0.003)	(0.002)	(0.004)	(0.003)	(0.004)	(0.004)	(0.003)	(0.002)	(0.004)	(0.003)
Foreign Share				0.000614	0.00153	0.00222				0.000602	0.00150	0.00218
				(0.002)	(0.003)	(0.002)				(0.002)	(0.003)	(0.002)
Firm FE		yes			yes			yes			yes	
Industry FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Region FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
R2	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
N	13256	13256	9058	13766	13766	9424	13256	13256	9058	13766	13766	9424

* Denotes significance at the 10% level, ** Denotes significance at the 5% level, *** Denotes significance at the 1% level.
ABC stands for absorptive capacity as measured by the distance between a firm's productivity from maximum. Robust standard errors in parentheses.

I used the estimated productivity as a dependent variable and estimate the model specifications presented in equation (2.13). Following Javorcik (2004), I estimate the contemporary and lagged horizontal and vertical spillovers as a proxy for foreign presence. The models in changes are after controlling for industry concentration, industry, region and year fixed effects.

Table 2.7 presents the coefficient estimates of backward, forward and horizontal spillovers. The first six columns show the results after accounting for gains from imports, whereas, the last six columns report results without. Furthermore, in each case, the first three columns are on domestic firms, the second three columns present regression results for all firms. First two columns show the result from pooled OLS and fixed effects regression, whereas, the regression result in every third column indicates the regression result where the dependent variable is lagged by a year.

From the estimated result, it seems domestic firms do not benefit from supplying to an industry where there is foreign presence as a proxy by backward spillover. Although the coefficient is positive, it is not statistically significant on both contemporary and lagged specifications. The sign of the result is in line with most findings in the literature (Javorcik and Spatareanu, 2008; Lu et al., 2017).

Lagged foreign presence at the same level of the industry appears to be beneficiary for all firms. Similarly, lagged results from foreign presence in the industry where firms buy their inputs seems to have a positive effect, while its contemporary effect on productivity appears to be not beneficiary. The former result is in line with most findings in the literature (Javorcik and Spatareanu, 2008; Lu et al., 2017; Bwalya, 2006), whereas the latter result is rather limited in the literature (Lopez, 2008).

For the purpose of comparison, in the remaining columns of Table 2.7, I present FDI spillover estimations where I do not account for productivity gains from imported inputs. This approach is similar to estimating productivity using the Olley and Pakes (1996) approach.

When we compare the productivity spillover results after and before accounting for imported inputs, it appears that there is a very small change in the magnitude of the coefficient estimates. The sign and significance of the results, however, remain intact.

This implies that accounting for productivity gains from imports do not seem to affect the spillover from foreign presence. This, in turn, indicates trade liberalisation and FDI promotion policies are unrelated.

In Table 2.8, I report the regression results on the productivity growth by estimating foreign presence in first, second and fourth differences. All specifications include industry, region, year fixed effects and control for a firm's absorptive capacity, measured as the distance between a firm's productivity and its frontier. To measure industry concentration, I include a Herfindahl index. Moreover, in specifications for all firms, we account for the firm's foreign shares.

In all model specification cases, the estimates indicate that horizontal spillovers have a negative and significant effect on productivity growth of firms. Meanwhile, in all specifications, backward spillovers seem to be positive but not statistically significant. Lastly, forward spillovers are negative and significant for the first and second differences but not significant for the fourth difference.

The above results hold for both domestic and all firms. Most importantly, the result seems to be the same when we measure productivity before and after accounting for imported inputs. This result reinforces the result we find from contemporary and lagged regression results.

Table 2.9 shows the result of FDI spillovers and the impact of controlling for imported inputs. This is done by including a dummy variable that separates the two estimation methods of productivity. The dummy variable index represents an indicator variable which takes a value of zero if productivity is measured using OP and one if Halpern et al. (2015) is used.

In Table 2.9, the coefficient estimate for the spillovers indicates how accounting for import affects productivity spillovers. Here, the main variables of interest are the interaction terms. In the case of backward and horizontal spillovers, the coefficient estimates for the interaction terms are negative. However, these estimates are not statistically different from zero implying estimating productivity after accounting for imported inputs do not change the above spillovers. Although, the interaction term for forward spillovers is positive, the estimate is still insignificant having a similar interpretation as backward and horizontal

Table 2.8 Models Estimation of FDI Spillover on Productivity Growth for HKS and OP

	HKS						OP					
	Domestic firms			All firms			Domestic firms			All firms		
	D1	D2	D4	D1	D2	D4	D1	D2	D4	D1	D2	D4
D.Backward	0.403			0.766			0.407			0.769		
	(4.058)			(3.884)			(4.057)			(3.883)		
D.Forward	-12.29*			-12.44*			-12.29*			-12.43*		
	(7.325)			(7.044)			(7.325)			(7.044)		
D.Horizontal	-0.724***			-0.800***			-0.724***			-0.800***		
	(0.230)			(0.224)			(0.230)			(0.224)		
D2.Backward		3.180			4.294			3.183			4.297	
		(4.257)			(4.058)			(4.257)			(4.057)	
D2.Forward		-16.87**			-15.77**			-16.87**			-15.77**	
		(7.896)			(7.707)			(7.896)			(7.707)	
D2.Horizontal		-0.814***			-0.868***			-0.814***			-0.868***	
		(0.236)			(0.229)			(0.236)			(0.229)	
D4.Backward			3.369			3.753			3.369			3.753
			(4.620)			(4.393)			(4.620)			(4.393)
D4.Forward			-3.404			-1.666			-3.404			-1.666
			(10.101)			(9.697)			(10.101)			(9.697)
D4.Horizontal			-0.893***			-0.898***			-0.893***			-0.898***
			(0.283)			(0.274)			(0.283)			(0.274)
ABC	3.081***	2.493***	3.152***	3.038***	2.424***	3.120***	3.128***	2.530***	3.199***	3.084***	2.460***	3.166***
	(0.089)	(0.143)	(0.522)	(0.086)	(0.139)	(0.503)	(0.090)	(0.145)	(0.530)	(0.087)	(0.141)	(0.511)
HHI Index	-0.791***	-0.402	-1.970	-0.789***	-0.329	-2.272	-0.791***	-0.402	-1.970	-0.789***	-0.329	-2.272
	(0.222)	(0.385)	(1.524)	(0.215)	(0.377)	(1.495)	(0.222)	(0.385)	(1.524)	(0.215)	(0.377)	(1.495)
Foreign Share				-0.00625	-0.0122	0.108				-0.00627	-0.0122	0.107
				(0.037)	(0.064)	(0.235)				(0.037)	(0.064)	(0.235)
Industry FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Region FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
R^2	0.277	0.0951	0.0249	0.275	0.0920	0.0235	0.277	0.0951	0.0249	0.275	0.0921	0.0235
N	8632	6315	3503	8989	6596	3691	8632	6315	3503	8989	6596	3691

* Denotes significance at the 10% level, ** Denotes significance at the 5% level, *** Denotes significance at the 1% level. Robust standard errors in parentheses.

spillovers.

Table 2.9 Models Estimation of Productivity on FDI Spillovers from OP and HKS

	Domestic		All Firms	
	Pooled	FE	Pooled	FE
HKS Dummy	0.00142**	0.00134***	0.00141**	0.00133***
	(0.001)	(0.000)	(0.001)	(0.000)
Backward	0.259*	0.339**	0.219	0.268*
	(0.143)	(0.148)	(0.138)	(0.142)
HKS Dummy×Backward	-0.0139	-0.0139	-0.0115	-0.0115
	(0.132)	(0.114)	(0.127)	(0.110)
Forward	-1.953***	-1.477***	-1.994***	-1.535***
	(0.465)	(0.429)	(0.447)	(0.412)
HKS Dummy×Forward	0.300	0.300	0.294	0.294
	(0.644)	(0.556)	(0.619)	(0.535)
Horizontal	0.0248***	0.00950	0.0240***	0.00675
	(0.009)	(0.009)	(0.009)	(0.009)
HKS Dummy×Horizontal	-0.0128	-0.0128	-0.0127	-0.0127
	(0.011)	(0.009)	(0.010)	(0.009)
Absorptive Capacity	7.443***	7.437***	7.443***	7.438***
	(0.002)	(0.002)	(0.002)	(0.002)
HHI Index	0.0153**	0.0113	0.0166***	0.0136*
	(0.006)	(0.007)	(0.006)	(0.007)
Foreign Share			0.000536	0.00145
			(0.001)	(0.002)
Industry FE	yes	yes	yes	yes
Region FE	yes	yes	yes	yes
Year FE	yes	yes	yes	yes
R^2	0.999	0.999	0.999	0.999
N	26512	26512	27532	27532

* Denotes significance at the 10% level, ** Denotes significance at the 5% level, *** Denotes significance at the 1% level.
Robust standard errors in parentheses

2.5 Conclusion

A firm's productivity gain can come from a multitude of sources. In this regard, the most commonly studied source is learning from other more productive and nearby firms or the so-called "FDI spillovers". Most recently, however, researchers started to investigating gains from imported inputs. Given the different nature and sources of these gains, it is indicative to ask whether the gains from the two are complement or substitute when they simultaneously happen. In doing so, the study analyses the effect of accounting for imported inputs on productivity spillovers. This understanding will have policy implications on trade liberalisation and FDI promotion policies.

For the empirical analysis this present study focuses on Ethiopia. Ethiopia is a good case since the country experiences both trade liberalisation and rapid increase in FDI inflow during the period considered. For instance, from the descriptive analysis we find that between the year 1996 and 2010, average tariff declines from 38% to 20% where as FDI stock increases from less that 1 million to over 4 billion USD. Moreover, during the same period the import material share increases by 10%. Furthermore, foreign-owned firms import more products on average and are more likely to be importers.

From the econometric analysis, I find that imports have a positive and significant impact on firm productivity. That is, firms tend to benefit from each imported input. Meanwhile, in line with previous studies the positive gains are bigger for foreign firms relative to their domestic counter parts.

Likewise, a separate analysis on the gain from FDI spillover suggests positive backward and horizontal spillover, although the former is insignificant. Forward spillover seems to be negative. This is true at levels and lags. Furthermore, it appears horizontal and forward spillovers have a negative and significant effect on productivity growth. The effect of backwards spillovers on growth seems positive. The spillover effects are higher for domestic firms.

Given a positive gain from imported inputs, accounting for imported inputs and see how that affects gains from FDI spillover is important for productivity analysis. Subsequently, I repeat the analysis after accounting for imported inputs and I find that accounting for imported inputs seems to have no effect on FDI spillovers. This implies, the productivity gains from the two sources are different in nature and are unrelated on either level or growth productivity. That is, they neither reinforce nor crowd-out each other.

Most developing countries open their economies partly due to policy advises from international organisations such as the World Bank and International Monetary Fund (IMF) on the importance of foreign direct investment (FDI). Given these policies happen simultaneously, understanding if they complement or crowd out one another is of great policy importance. The results from this study indicate that, in the case of Ethiopia, FDI promotion and trade liberalisation policies need to be considered separately.

Furthermore, the productivity-boosting effects of imported inputs are established suggesting a bigger gain from imports. However, the gains from FDI spillovers depend on the location of the industry of foreign firms vis-a-vis a domestic firm. Thus further analysis is required to identify those areas to benefit from.

Likewise, since the time of adopting a market economy, the Ethiopian government has been craving for attracting foreign investors. To do this the government increases infrastructure investment in areas where there are foreign investors. Revising investment policies in favour of investors (e.g. Investment policy has been modified more than 4 times in the last 20 years) Which leads to a substantial increase in FDI inflows over the past decade.

Chapter Three

Commuting and Residential Mobility: Evidence from the UK

3.1 Introduction

Different labour markets respond differently to changes in economic situations. One of the main determinants for how sensitive labour markets are to a given shock at the local level is their geographic location with respect to other labour markets (Monte, Redding, and Rossi-Hansberg, 2018). This interconnectedness and the subsequent difference in sensitivity are highly shaped by the relative easiness of commuting and mobility of workers between markets. Furthermore, workers preference to migrate and the role of commuting vis-a-vis individuals' location decisions help us understand the short and long term adjustment process of workers to local labour market shocks. This, in turn, has significant policy implications on the labour market, housing market, infrastructure and other public investments.

People choose where to live and where to work depending on different factors. Studies mention amenities, wages, and housing prices as the major determinants (Monte et al., 2018). Others emphasise the important role of transportation in general and commuting cost in particular since it serves as a bridge that links two locations. Besides the aforementioned factors, other variables that affect an individual's residence and workplace decisions include: individuals' preferences, family characteristics, professional characteristics and life-cycle stage (Haas and Osland, 2014).

Recently, researchers in economic geography, labour economics, and international trade have been interested in understanding the impacts of commuting on a multitude of outcomes. Although there are studies that analyse the effects of commuting on well-being and labour market outcomes of individuals (Roberts and Taylor, 2016; Jacob, Munford, Rice, and Roberts, 2019), there is limited evidence on the causal link between commuting and individuals' propensity to relocate. This paper questions whether commuting is causally linked to the preference for internal residential mobility. Specifically, it asks whether commuting induces people to move.

In the literature, there are few studies that analyse the relationship between commuting and residential decisions (Romani, Suriñach, and Artiís, 2003; So, Orazem, and Otto, 2001). The decision of households and individuals to either commute or move depends on the interaction between labour and housing markets. Many variables play a (positive or negative) role in shaping either of the decisions. Therefore, understanding the under-investigated role of commuting in a subsequent residential mobility remains an empirical issue.

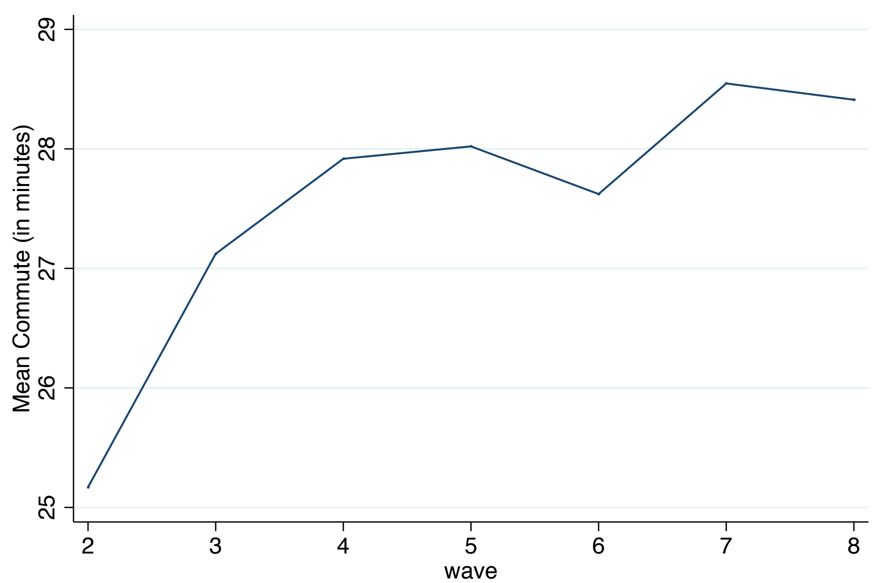
From an individual perspective, commuting induces residential mobility since it involves pecuniary cost (Schmidt, 2014). Besides pecuniary cost, commuting involves non-pecuniary cost. Ma and Ye (2019) shows that longer commuting lowers productivity. For instance, studies show commuting is associated with absenteeism (Goerke and Lorenz, 2017; Ma and Ye, 2019), reduced well-being and increased risk of ill-health (Künn-Nelen, 2016; Roberts, Hodgson, and Dolan, 2011). Moreover, commuting can induce mobility since it lowers the cost of moving by reducing job and housing search costs related with mobility (Haas and Osland, 2014).

Meanwhile, commuting reduces the need for residential mobility for different reasons. First, commuting increases the earnings of individuals by better matching the individual skill to a specific job via alleviating the tightness of the market (Haas and Osland, 2014). Second, commuting allows individuals to benefit from environmental amenities by independently choosing where to live and work (Kim, Horner, and Marans, 2005). Lastly, Brueckner and Šťastná (2019) argue that commuting does not increase the job-related gains (increase in wages in a new location), unlike those new immigrant, hence does prevent people from moving.

To answer whether commuting induce mobility, this research uses a 10-year long panel dataset from the UK Household Longitudinal Study (Understanding Society) (UKHLS, 2018). The UK is a good case for studying the effects of commuting since it is one of the countries which is characterised by a high commuting rate. For instance, full time workers in the UK has a commuting time of more than 40 minutes per day which is above the OECD average of 38 minutes (OECD, 2011). Moreover, except Northern Ireland, all regions in the UK have experienced an increase in commuting time over the last decade (Scott, John, and Alun, 2016). Figure 3.1 shows the time for a single journey between a person's home and their usual workplace. As the figure shows, there is an overall increasing trend in commuting time in the UK.[1]

From a theoretical perspective commuting time affects the individual's indirect utility function for a given residential location decision. In other words, change in commuting time affect residential location preference. However, other factors such as job characteristic including relative wage difference between locations and residential characteristics such as living cost and amenities also affect residential preference. Thus, by restricting the analysis on those individuals who do not change their job, stayed with the same employer and remain in the same place of residence across the waves, I isolate the effect of change in commuting time on preferences and intent to move (Gutiérrez-i Puigarnau and van

Figure 3.1 Commuting Time by Wave



Source: Author computation using UK Household Longitudinal Study data

[1]The waves are corresponding to the years from 2010 to 2018.

Ommeren, 2015; So et al., 2001). Specifically, the variation in our dependent variable is obtained from individuals answer on their willingness and expectation to move from their residence.

From the analysis, I find that commuting increases the likelihood of preferring to move. Moreover, the paper documents that commuting increases, besides preference to relocate, the expectation to move. Specifically, I find that the odds-ratio of preference to move and expectation to move increases as a result of a commuting shock. Therefore, commuting is considered by many as something that is undesirable and people prefer and intend to avoid it, if possible. Therefore, from a policy perspective, reducing commuting time is something to consider to increase people's welfare. In a further robustness check, I undertake an analysis by splitting the sample by gender and between those whose commuting time increases and decreases.

The remaining of the paper is organised as follows. The next section describes the data used in this paper. Section 3 presents the empirical set-up and methodology. Section 4 discusses the main results, subsequent discussions and robustness checks. Section 5 concludes.

3.2 Data Description

The study is based on data from Understanding Society, the UK Household Longitudinal Study (UKHLS, 2018). It is a representative sample of UK households at the national level and covers information similar to its predecessor, BHPS.[2] The empirical analysis of this study is based on data from wave two to wave eight. Since each wave represents approximately a year, the sample in the data corresponds to the period from 2010 to 2018. The data are rich in terms of variables and include information about individual characteristics, such as, individual and household demography, socioeconomic status, general health condition, employment status, earnings, commuting time, and residential information. The initial dataset consists of a total of 373,615 individual-wave observations. However, the

[2]The British Household Panel Survey started in 1991 by following the same representative samples of individuals. The survey interviews every adult household member for multi-purpose study. As part of wave 18, BHPS participants were asked if they would consider joining Understanding Society and 83.75% did so.

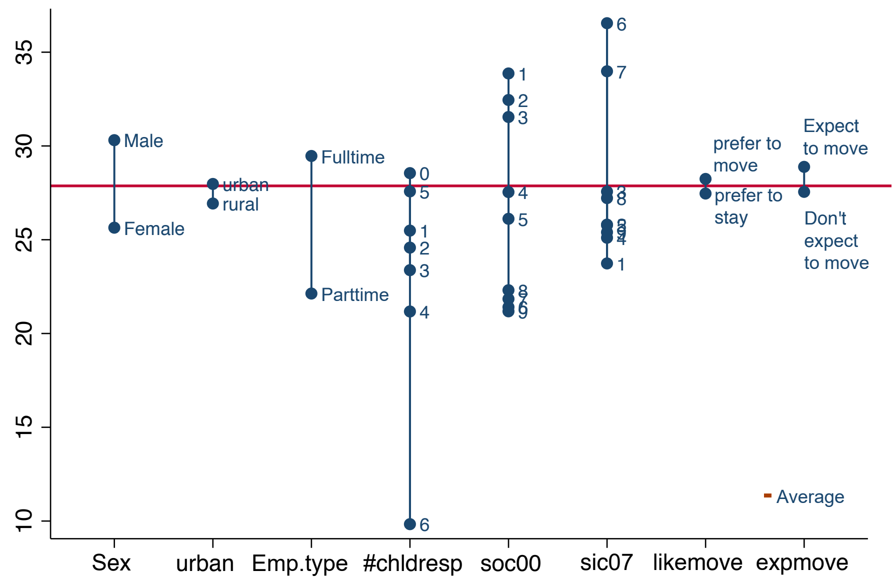
Table 3.1 Descriptive Statistics on the Overall and Estimated Sample

	All Sample				Estimated Sample			
	Mean	Std	Min	Max	Mean	Std	Min	Max
Age	47.283	18.588	16	104	42.576	12.083	16	87
Single	0.378	0.485	0	1	0.290	0.454	0	1
Female	0.541	0.498	0	1	0.546	0.498	0	1
Rural	0.235	0.424	0	1	0.223	0.416	0	1
Part-time	0.276	0.447	0	1	0.238	0.426	0	1
Commute	25.773	21.630	0	180	27.389	22.060	0	180
No. of children	0.307	0.760	0	10	0.358	0.755	0	6
Net income	1797.899	3404.253	0	26231	2007.680	3652.936	0	26231
Same Employer	0.922	0.269	0	1	1	0	1	1
Same residence	0.985	0.122	0	1	1	0	1	1
Obs	373,615				75,352			

Note: Income is in 2015 British pounds. Income is censored at the top and bottom 1%. Commuting is also censored at the top 1%.

final number of observations for the analysis is different because of the following procedures. First, following Jacob et al. (2019), I only keep workers who are observed at least for two consecutive waves in the dataset (355,487 obs). Second, I keep those who are employed or self-employed (174,456 obs), have not changed their place of residence (172,458 obs) and job in consecutive waves (110,375 obs). Third, I keep those individuals who experience a change in commuting time. Here, to allow for significant changes and reduce possible measurement error, I drop observations whose change in commuting time is less than five minutes. After the above adjustments, the analysis is based on a dataset consisting of a total of 75,352 observations. Since our main focus is on those individuals whose commuting time changes, it does not create an issue of representativeness. Table 3.1 shows the descriptive statistics for the initial and final sample. As shown in Table 3.1, individuals in the final sample, on average, tend to be younger, non-single, female, live in an urban area and work full time. They are also, on average, responsible for more children, earn higher net income and commute longer time to work. Using the final dataset, in Figure 3.2, I plot the average commuting time for different groups of individuals. From the Figure, on average, men, urban dwellers, full-time employees, individuals who look after more children and who work in skilled occupations tend to invest more time in commuting. In addition, those who prefer to move and expect to move tend

Figure 3.2 Mean Commute Time Among Groups



Source: Author computation using UKHLS

to commute longer time. Heterogeneity in these figures motivate the main analysis and robustness tests.

3.3 Methodology

The aim of the study is to identify the causal effect of a change in commuting on individuals preference and expectation for residential mobility. Therefore, the analysis proceeds in three parts. The first part analyses the causal link between commuting and preference for relocation. The study uses preference for relocation together with an expectation for relocation to investigate the impact of commuting on preference and expectation. To this end, I introduce a categorical variable that combine the two. This will also extend the analysis since the expectation to move indicates a stronger desire for relocation and hence closer to realised relocation. This part then further splits into two by first ignoring the ordinal nature of the categorical variable and later by accounting for it.

The first part of the analysis uses a discrete choice model. To be specific, I use a panel probit model. The dependent variable is a dummy which is obtained from people's answer to the question "Would you like to relocate?".

Here, the model can be derived from a latent variable model: $y^* = x\beta + \epsilon$, where y^* is unobserved and we only observe $y = 1[y^* > 0]$, that is, whether an individual would like to relocate or not. Thus, we can write the probability as:

$$Pr(y = 1|x) = Pr(y^* > 0|x) = Pr(\epsilon > -x\beta|x)$$

$$= 1 - G(-x\beta)$$

$$= G(x\beta) = \Phi(x\beta/\sigma)$$

The last line come from the fact that the probit model assumes that the error term, ϵ, follows a normal distribution and the symmetric nature of the normal distribution.[3]

The corresponding empirical specification for the latent variable is given as:

$$y^*_{it} = \beta_0 + \beta_1 logC_{it} + \mathbf{X}'_{it}\alpha + \delta_r + \gamma_t + \epsilon_{it},$$

where y_{it} takes a value of one if individual i at time t prefers to move and zero otherwise, $logC_{it}$ indicates the log of commuting time, X_{it} represents covariates that affect preferences to move including commuting, δ_r and γ_t indicates region and wave fixed effects respectively and ϵ_{it} is a normally distributed iid error term.

The second and third parts of the analysis use a multinomial and ordinal regressions respectively. In this case, the dependent variable is constructed from individuals' answer to the question "Would you like to relocate?" and "Would you expect to move?", which help us to classify individuals into four categories (groups). These groups are composed of those who would like to move and expect to move (1), those who would like to move but are not expecting to move (2), those who would like to stay but expect to move (3), and finally those who would like to stay and expecting to stay (4). Here, I rank the categories based on how close the individual is to actual relocation.

The multinomial regression assumes there is no natural order between the above categories. Following this assumption and taking the first category as a reference group, we

[3]For ordinal logit model the error term, ϵ, assumes to follow a logistic distribution.

can estimate the probability of choosing j can be given by:

$$Pr(y=j|x) = \frac{exp(x\beta_j)}{1+\sum_{h=2}^{J} exp(x\beta_j)} \ \ and$$

$$Pr(y=1|x) = \frac{1}{1+\sum_{h=2}^{J} exp(x\beta_j)}$$

From the above probabilities, we can easily derive the partial effect and the odds-ratio (McFadden, 1978).

The third and last part of the analysis uses ordinal regression model. Here, one might argue that there is order element to the above categories. Unlike the multinomial model which ignores the ordinal nature of the categories, ordinal models account for this.

Similar to probit, ordinal regression also assumes the observed variable, y can be considered as a latent function of another underlying continuous variable, y^*, that is not measured. However, unlike probit, the observed variables can take more than two values. The values for these variables depend on whether we have crossed a particular cut-off (threshold) of y^*. Therefore, the model involves grouping an underlying continuous variable y_i^* using cut-points in to j categories.

$$y_i = \begin{cases} 1, & \text{if } y_i^* \leq \kappa_1, \\ 2, & \text{if } \kappa_1 \leq y_i^* \leq \kappa_2, \\ 3, & \text{if } \kappa_2 \leq y_i^* \leq \kappa_3, \\ 4, & \text{if } y_i^* \geq \kappa_3. \end{cases} \tag{3.1}$$

Given the underlying or latent variable, $y_i^* = x_i^{'}\beta^* + \epsilon_i$, the probability of y taking a value of j or less is given by:

$$\gamma_{ij} = Pr\{y_i < j\} = Pr\{y_i^* < \kappa_j\} = Pr\{\epsilon_i < \kappa_j - x_i^{\prime}\beta^*\}$$

$$= \Phi(\kappa_j - x_i^{\prime}\beta^*)$$

The last line is derived from the assumption that the error term, ϵ_i, follows a normal distribution. The inverse of the last equation links probabilities to the real line.

The following empirical specification is used to estimate the latent variable,

$$y_{it}^{*} = \beta_1 logC_{it} + \mathbf{X}_{it}'\alpha + \delta_r + \gamma_t + \xi_{it},$$

where y_{it} stands for the values of the four categories as defined above, ξ_{it} denotes type I extreme-value iid error term and the remaining variables have the same definition as above.

3.4 Results

This section discusses the results from the econometric analysis. The section starts with discussing marginal effect results from the probit model where I use stated preference to move as a dependent variable. After that, I present the results from a multinomial logit. Here, I used a categorical variable, which indicates preference and expectation to move, as a dependent variable. Lastly, I present the results from ordinal logit and probit models after assigning order to the categories.

Table 3.2 shows the effect of commuting on the probability of preferring to move. The first three Columns report results from a linear probability model, whereas the last three Columns show results from the probit regression model. As we can see, in all specification cases, commuting has a positive and significant effect on the probability of preference for residential mobility. All specifications include household, region and wave fixed effects. Meanwhile, the columns differ in the set of controls and other fixed effects they include.

Column 1 and 4 report regression results where I only account for commuting time, household, region and wave fixed effects. Whereas, Column 2 and 5 show results after controlling for individual-level covariates. Under these specifications, commuting has a positive and significant effect on the likelihood of preference to move. Other covariates also have expected signs, which are in line with previous studies on residential mobility (Lee and Waddell, 2010). Specifically, commuting time, being single, working full time and being an urban dweller have a positive and significant effect on preference to move. The

Table 3.2 Linear Probability and Probit Models

	(1) LPM	(2) LPM	(3) LPM	(4) Probit	(5) Probit	(6) Probit
log(Commute)	0.00421***	0.00769***	0.00853***	0.0106**	0.0195***	0.0192***
	(0.002)	(0.002)	(0.002)	(0.00541)	(0.00601)	(0.00611)
Age		-0.000466	-0.000649		0.00302**	0.00267*
		(0.001)	(0.001)		(0.00147)	(0.00149)
Age^2		-0.0000321**	-0.0000307**		-7.35e-05***	-6.96e-05***
		(0.000)	(0.000)		(1.74e-05)	(1.76e-05)
Male		-0.0137***	-0.00903*		-0.00945*	-0.00714
		(0.004)	(0.005)		(0.00570)	(0.00620)
Single		0.0274***	0.0275***		0.0306***	0.0308***
		(0.005)	(0.005)		(0.00591)	(0.00596)
HRP		0.00946**	0.0107***		-0.00177	-0.000820
		(0.004)	(0.004)		(0.00535)	(0.00540)
Full time		0.0456***	0.0450***		0.0392***	0.0387***
		(0.006)	(0.006)		(0.00620)	(0.00630)
No. children		0.00869***	0.00820***		0.00636*	0.00637*
		(0.003)	(0.003)		(0.00379)	(0.00382)
Urban		0.0715***	0.0717***		0.0759***	0.0745***
		(0.005)	(0.005)		(0.00627)	(0.00635)
log(Net income)		-0.0105***	-0.00966***		-0.00928***	-0.00871***
		(0.002)	(0.003)		(0.00265)	(0.00280)
Education Level	no	yes	yes	no	yes	yes
Travel mode	no	yes	yes	no	yes	yes
House ownership	no	yes	yes	no	yes	yes
Occupation FE	no	no	yes	no	no	yes
Industry FE	no	no	yes	no	no	yes
Household FE	yes	yes	yes	yes	yes	yes
Region FE	yes	yes	yes	yes	yes	yes
Wave FE	yes	yes	yes	yes	yes	yes
Obs	67,796	57,660	56,697	67,796	57,660	56,703

Notes: Dependent variable is a dummy which takes a value of one if an individual prefer to move house. The coefficient estimates indicate marginal effects. HRP represents Household Reference Person. Robust standard errors in parentheses. * Denotes significance at the 10% level, ** Denotes significance at the 5% level, *** Denotes significance at the 1% level.

specifications also account for variables such as age, gender, income, number of dependent children under 16 responsible for, level of education, home-ownership status and mode of transportation. In Column 3 and 6, in addition to accounting for the covariates as Column 2 and 5, I also control for individual's sector and industry fixed effects at a 1-digit level. Despite accounting for these controls, the effect of commuting on preference to move remains positive and significant. Other covariates also have expected sign. Thus, Column 6 would be our preferred specification.

The average marginal effect results from the preferred specification, Column 6, indicates a ten-fold multiplicative increase in commuting time leads to a 0.02 increase in the likelihood of preference to move. This indicates, increase in commuting time might cause a dis-utility and thus increases individual's preference to relocate.

After considering the answers to preference to move together with expectation to move, I construct a categorical variable. This construction helps us to extend the analysis to multinomial models. Table 3.3 presents the results from a multinomial logistic regression. Here, individuals' answer "prefer to stay and expect to stay" is used as a base category. Therefore, the results from the table are interpreted with respect to or in comparison to this category and all results in the table report the odds-ratio.

Columns 1 to 3 report results where I control for the region and wave fixed effects. As we can see from these Columns commuting significantly increases the odd of preferring and expecting to move. Specifically, each additional unit of commuting multiplies the odd of preferring and expecting to move by 1.042 relative to preferring and expecting to stay. That is, commuting pushes people to preferring and expecting to move. Columns 4 to 6 report the result after controlling for other covariates such as age, age squared, gender, full-time indicator, urban, net pay and number of dependent children under 16 the individual is responsible for. The result from these Columns also indicates that commuting increases the odds of moving even among those who are not expecting to move. In Columns 6 to 9, in addition to the above covariates, I control for individual's occupation and industry at 1-digit level. These additional controls do not seem to change the result that commuting increases people preference and expectation to relocate. From the last column, a ten-fold increase in commuting time multiplies your odds of preferring and expecting to move by 1.057 relative to preferring and expecting to stay. These results are in line with previous literature which shows a positive association between commuting and migration in a different country and methodological setting (Brueckner and Šťastná, 2019).

Moving to the ordered regression, Table 3.4 presents the results for OLS, ordered probit,

Table 3.3 Multinomial Logit Model

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Prefer	To Stay	To move	To move	To Stay	To move	To move	To Stay	To move	To move
Expect	To Move	To Stay	To Move	To Move	To Stay	To Move	To Move	To Stay	To Move
log(Commute)	1.011	1.010	1.042***	1.045*	1.033***	1.067***	1.044*	1.038***	1.057***
	(0.022)	(0.008)	(0.012)	(0.026)	(0.009)	(0.014)	(0.027)	(0.009)	(0.014)
Age				0.898***	1.024***	0.964***	0.899***	1.024***	0.960***
				(0.015)	(0.007)	(0.0095)	(0.016)	(0.007)	(0.010)
Age^2				1.001***	1.000***	1.000	1.001***	1.000***	1.000
				(0.0002)	(7.40e-05)	(0.0001)	(0.0002)	(7.47e-05)	(0.0001)
Male				1.025	0.921***	0.989	1.096	0.948**	0.991
				(0.0653)	(0.0215)	(0.0331)	(0.0758)	(0.0243)	(0.0361)
Single				2.098***	1.067***	1.425***	2.099***	1.060**	1.435***
				(0.135)	(0.0264)	(0.0485)	(0.137)	(0.0265)	(0.0495)
HRP				1.103	1.061***	1.031	1.097	1.071***	1.026
				(0.0729)	(0.0236)	(0.0340)	(0.0737)	(0.0240)	(0.0343)
Fulltime				1.125	1.256***	1.271***	1.165*	1.272***	1.228***
				(0.101)	(0.0369)	(0.0592)	(0.108)	(0.0380)	(0.0584)
No. children				0.901*	1.076***	0.959*	0.892**	1.079***	0.957*
				(0.0481)	(0.0169)	(0.0240)	(0.0485)	(0.0171)	(0.0242)
Urban				1.032	1.496***	1.173***	1.035	1.477***	1.176***
				(0.0781)	(0.0398)	(0.0458)	(0.0800)	(0.0398)	(0.0466)
log(Net income)				1.102***	0.919***	1.024	1.081**	0.935***	0.999
				(0.0394)	(0.0117)	(0.0192)	(0.0421)	(0.0129)	(0.0204)
Education Level	no	no	no	no	no	no	yes	yes	yes
House ownership	no	no	no	no	no	no	yes	yes	yes
Travel mode	no	no	no	no	no	no	yes	yes	yes
Occupation FE	no	no	no	no	no	no	yes	yes	yes
Industry FE	no	no	no	no	no	no	yes	yes	yes
Region FE	yes	yes	yes	yes	yes	yes	yes	yes	yes
Wave FE	yes	yes	yes	yes	yes	yes	yes	yes	yes
Obs	67,735	67,735	67,735	57,598	57,598	57,598	56,625	56,625	56,625

Notes: Dependent variable is a categorical variable which takes a value of one to four if an individual prefer and expect to stay, prefer to stay but expect to move, prefer to move but expect to stay and prefer and expect to move house respectively. HRP represents Household Reference Person. Robust standard errors in parentheses. * Denotes significance at the 10% level, ** Denotes significance at the 5% level, *** Denotes significance at the 1% level.

ordered logit and corresponding odds-ratio for ordered logit. Here, an assumption of natural ordering about the above categories is made. The natural ordering is in terms of the closeness to realised mobility or intensity in the need to move. However, the interval between any two categories is not the same, which justifies the use of ordinal regression for the analysis. Here, it worth noting that the multinomial model ignores the ordinality.

In Column 1 of Table 3.4, I report the pooled OLS result for comparison. This regression considers the categories as continuous variables. Columns 2 and 3 show the results of ordered probit and logit regressions. Lastly, Column 4 shows the odds-ratio for the ordinal logit model.

In all specification cases, I find that commuting increases the likelihood of residential mobility (being in the higher category). The odds-ratio is interpreted as an increase in commuting increases the odds of being in a higher category, that is, in the prefer and expect to move category. Specifically, if an individual were to increase its commute by one point, its odd of being in a higher category would increase by 1.059, ceteris paribus. Moreover, the cut-off points represent the line where two consecutive categories separate from each other. Together with the predicted probability of an individual, they determine the category the individual likely to be in.

3.4.1 Robustness

This section discusses whether the results are robust to the different specification choices and across different subgroups. The analyse starts by investigating the effect of change in commuting time in different groups of individuals. This helps us not only to uncover the presence of heterogeneous effect of commuting but also serves as a robustness check for our results.

First, I analyse the data by splitting the sample by gender. It is well documented that there is a difference in commuting pattern and effect between the two genders (Roberts et al., 2011). This informs us of the sensitivity of the two genders for change in commute.

The result by gender, as shown in Table 3.5 and Table 3.6, indicates that that female tend to be sensitive to commuting. The result shows that the odd for preferring and

Table 3.4 Ordered Models

	(1) OLS	(2) Ordered Probit	(3) Ordered Logit	(4) Odd Ratio
log(Commute)	0.031***	0.034***	0.057***	1.059***
	(0.00595)	(0.00659)	(0.0112)	(0.0118)
Age	-0.015***	-0.017***	-0.022***	0.978***
	(0.00287)	(0.00322)	(0.00546)	(0.00534)
Age^2	5.43e-05	4.40e-05	2.87e-05	1.000
	(3.33e-05)	(3.78e-05)	(6.41e-05)	(6.41e-05)
Male	-0.020*	-0.020	-0.037*	0.964*
	(0.0117)	(0.0129)	(0.0218)	(0.0210)
Single	0.092***	0.113***	0.177***	1.194***
	(0.0115)	(0.0125)	(0.0211)	(0.0252)
HRP	0.023**	0.023**	0.042**	1.043**
	(0.0103)	(0.0115)	(0.0193)	(0.0202)
Full time	0.108***	0.120***	0.213***	1.238***
	(0.0137)	(0.0154)	(0.0260)	(0.0322)
No. children	0.009	0.004	0.015	1.015
	(0.00749)	(0.00826)	(0.0139)	(0.0141)
Urban	0.144***	0.160***	0.291***	1.338***
	(0.0118)	(0.0135)	(0.0230)	(0.0308)
Net income (in log)	-0.019***	-0.020***	-0.037***	0.963***
	(0.00611)	(0.00684)	(0.0116)	(0.0111)
Education Level	yes	yes	yes	yes
House ownership	yes	yes	yes	yes
Travel mode	yes	yes	yes	yes
Occupation FE	yes	yes	yes	yes
Industry FE	yes	yes	yes	yes
Region FE	yes	yes	yes	yes
Wave FE	yes	yes	yes	yes
Observations	56,756	56,756	56,756	56,756
Constant cut1		-0.019	0.119	1.126
Constant cut2		0.042	0.219	1.245
Constant cut3		0.998***	1.881***	6.560***

Notes: Dependent variable is a categorical variable which takes a value of one to four if an individual prefer and expect to stay, prefer to stay but expect to move, prefer to move but expect to stay and prefer and expect to move house respectively. HRP represents Household Reference Person. Robust standard errors in parentheses. * Denotes significance at the 10% level, ** Denotes significance at the 5% level, *** Denotes significance at the 1% level.

expecting to move significantly increases with commuting for females. Meanwhile, the result for male commuters is not significant indicating less sensitivity to commuting for them. These results are in line with previous studies which find a larger welfare effect of

Table 3.5 Ordered Models for Female

	(1) OLS	(2) Ordered-probit	(3) Ordered-Logit	(4) Odd-Ratio
log(Commute)	0.047***	0.053***	0.088***	1.091***
	(0.00821)	(0.00912)	(0.0154)	(0.0169)
Age	-0.0199***	-0.0227***	-0.0315***	0.969***
	(0.00405)	(0.00457)	(0.00770)	(0.00746)
Age^2	0.000104**	0.000110**	0.000131	1.000
	(4.78e-05)	(5.44e-05)	(9.18e-05)	(9.18e-05)
Single	0.0571***	0.0709***	0.115***	1.121***
	(0.0160)	(0.0174)	(0.0294)	(0.0329)
HRP	0.0550***	0.0606***	0.102***	1.108***
	(0.0147)	(0.0163)	(0.0275)	(0.0305)
Full time	0.102***	0.113***	0.200***	1.221***
	(0.0159)	(0.0179)	(0.0302)	(0.0369)
No. children	0.00862	0.00339	0.0147	1.015
	(0.00814)	(0.00900)	(0.0151)	(0.0153)
Urban	0.134***	0.146***	0.269***	1.308***
	(0.0158)	(0.0180)	(0.0307)	(0.0402)
Net income (in log)	-0.0136	-0.0113	-0.0248	0.975
	(0.00908)	(0.0102)	(0.0171)	(0.0167)
Education Level	yes	yes	yes	yes
House ownership	yes	yes	yes	yes
Travel mode	yes	yes	yes	yes
Occupation FE	yes	yes	yes	yes
Industry FE	yes	yes	yes	yes
Region FE	yes	yes	yes	yes
Wave FE	yes	yes	yes	yes
Observations	30,996	30,996	30,996	30,996

Notes: Dependent variable is a categorical variable which takes a value of one to four if an individual prefer and expect to stay, prefer to stay but expect to move, prefer to move but expect to stay and prefer and expect to move house respectively. HRP represents Household Reference Person. Robust standard errors in parentheses. * Denotes significance at the 10% level, ** Denotes significance at the 5% level, *** Denotes significance at the 1% level.

commuting on females but not males (Jacob et al., 2019).

Table 3.6 Ordered Models for Male

	(1) OLS	(2) Ordered-probit	(3) Ordered-Logit	(4) Odd-Ratio
log(Commute)	0.0117	0.0125	0.0218	1.022
	(0.00868)	(0.00960)	(0.0163)	(0.0166)
Age	-0.0106**	-0.0103**	-0.0125	0.988
	(0.00419)	(0.00469)	(0.00796)	(0.00786)
Age^2	3.82e-06	-2.41e-05	-7.60e-05	1.000
	(4.79e-05)	(5.41e-05)	(9.21e-05)	(9.21e-05)
Single	0.122***	0.152***	0.232***	1.261***
	(0.0176)	(0.0191)	(0.0323)	(0.0407)
HRP	0.00180	0.000849	0.00219	1.002
	(0.0153)	(0.0169)	(0.0286)	(0.0286)
Full time	0.108***	0.119***	0.218***	1.244***
	(0.0298)	(0.0338)	(0.0575)	(0.0715)
No. children	0.0439	0.0512	0.0866	1.090
	(0.0456)	(0.0494)	(0.0822)	(0.0897)
Urban	0.158***	0.179***	0.321***	1.378***
	(0.0178)	(0.0204)	(0.0349)	(0.0481)
Net income (in log)	-0.0213**	-0.0231**	-0.0420***	0.959***
	(0.00832)	(0.00930)	(0.0157)	(0.0151)
Education Level	yes	yes	yes	yes
House ownership	yes	yes	yes	yes
Travel mode	yes	yes	yes	yes
Occupation FE	yes	yes	yes	yes
Industry FE	yes	yes	yes	yes
Region FE	yes	yes	yes	yes
Wave FE	yes	yes	yes	yes
Observations	25,760	25,760	25,760	25,760

Notes: Dependent variable is a categorical variable which takes a value of one to four if an individual prefer and expect to stay, prefer to stay but expect to move, prefer to move but expect to stay and prefer and expect to move house respectively. HRP represents Household Reference Person. Robust standard errors in parentheses. * Denotes significance at the 10% level, ** Denotes significance at the 5% level, *** Denotes significance at the 1% level.

In a second robustness check, I split the sample between groups with an increase in commuting time and decrease commuting time and undertake the analysis for these groups

Table 3.7 Ordered Models Increase in Commute time

	(1) OLS	(2) Ordered-probit	(3) Ordered-Logit	(4) Odd-Ratio
log(Commute)	0.0299***	0.0318***	0.0524***	1.054***
	(0.00747)	(0.00767)	(0.0130)	(0.0137)
Age	-0.0122***	-0.0133***	-0.0167***	0.983***
	(0.00326)	(0.00362)	(0.00613)	(0.00603)
Age^2	1.56e-05	3.86e-06	-3.90e-05	1.000
	(3.80e-05)	(4.27e-05)	(7.24e-05)	(7.24e-05)
Male	-0.0201	-0.0215	-0.0366	0.964
	(0.0134)	(0.0145)	(0.0245)	(0.0237)
Single	0.0961***	0.116***	0.182***	1.200***
	(0.0131)	(0.0140)	(0.0237)	(0.0284)
HRP	0.0299**	0.0297**	0.0552**	1.057**
	(0.0118)	(0.0130)	(0.0219)	(0.0231)
Full time	0.102***	0.113***	0.200***	1.222***
	(0.0156)	(0.0173)	(0.0292)	(0.0357)
No. children	0.00593	0.00223	0.0118	1.012
	(0.00855)	(0.00932)	(0.0156)	(0.0158)
Urban	0.140***	0.149***	0.271***	1.311***
	(0.0136)	(0.0152)	(0.0259)	(0.0339)
Net income (in log)	-0.0166**	-0.0164**	-0.0297**	0.971**
	(0.00695)	(0.00766)	(0.0129)	(0.0125)
Education Level	yes	yes	yes	yes
House ownership	yes	yes	yes	yes
Travel mode	yes	yes	yes	yes
Occupation FE	yes	yes	yes	yes
Industry FE	yes	yes	yes	yes
Region FE	yes	yes	yes	yes
Wave FE	yes	yes	yes	yes
Observations	43,943	43,950	43,950	43,950

Notes: Dependent variable is a categorical variable which takes a value of one to four if an individual prefer and expect to stay, prefer to stay but expect to move, prefer to move but expect to stay and prefer and expect to move house respectively. HRP represents Household Reference Person. Robust standard errors in parentheses. * Denotes significance at the 10% level, ** Denotes significance at the 5% level, *** Denotes significance at the 1% level.

separately. This besides serving as a robustness check and informing us of the effect of the two commuting shocks, it also shades light on the consequences of different policies. An increase in commute can come from congestion and change in the mode of transport, a decrease in commute can be caused by an increase in infrastructure investment. Table 3.7 and Table 3.8 show that commuting has a positive effect for both groups regardless of the type of change in commuting time.

Table 3.8 Ordered Models Decrease in Commute time

	(1) OLS	(2) Ordered-probit	(3) Ordered-Logit	(4) Odd-Ratio
log(Commute)	0.029**	0.032**	0.056**	1.058**
	(0.0120)	(0.0138)	(0.0236)	(0.0250)
Age	-0.0273***	-0.0300***	-0.0439***	0.957***
	(0.00630)	(0.00728)	(0.0124)	(0.0118)
Age^2	0.000199***	0.000204**	0.000294**	1.000**
	(7.15e-05)	(8.35e-05)	(0.000142)	(0.000142)
Male	-0.0175	-0.0179	-0.0368	0.964
	(0.0243)	(0.0278)	(0.0475)	(0.0458)
Single	0.0813***	0.106***	0.162***	1.176***
	(0.0241)	(0.0271)	(0.0463)	(0.0544)
HRP	-0.000156	-0.000986	-0.00583	0.994
	(0.0213)	(0.0245)	(0.0417)	(0.0414)
Full time	0.129***	0.147***	0.265***	1.303***
	(0.0288)	(0.0336)	(0.0572)	(0.0745)
No. children	0.0122	0.00815	0.0243	1.025
	(0.0157)	(0.0179)	(0.0302)	(0.0310)
Urban	0.173***	0.200***	0.368***	1.445***
	(0.0246)	(0.0292)	(0.0505)	(0.0729)
Net income (in log)	-0.0309**	-0.0310**	-0.0646**	0.937**
	(0.0131)	(0.0153)	(0.0261)	(0.0244)
Education Level	yes	yes	yes	yes
House ownership	yes	yes	yes	yes
Travel mode	yes	yes	yes	yes
Occupation FE	yes	yes	yes	yes
Industry FE	yes	yes	yes	yes
Region FE	yes	yes	yes	yes
Wave FE	yes	yes	yes	yes
Observations	12,806	12,806	12,806	12,806

Notes: Dependent variable is a categorical variable which takes a value of one to four if an individual prefer and expect to stay, prefer to stay but expect to move, prefer to move but expect to stay and prefer and expect to move house respectively. HRP represents Household Reference Person. Robust standard errors in parentheses. * Denotes significance at the 10% level, ** Denotes significance at the 5% level, *** Denotes significance at the 1% level.

3.5 Conclusion

The decision of commuting or residential mobility lays in the crossroad between housing and labour markets. Although commuting opens the opportunity for individuals to live and work in different places, and thus increase their utility, it also imposes substantial pecuniary and non-pecuniary costs to individuals. These costs range from reducing subjective well-being to productivity. Similarly, from an economic perspective, while commuting improves the overall labour market by reducing the impact of shocks on local labour market, it involves an environmental cost and requires infrastructural investment.

In this paper, I analyse the causal effect of commuting on individuals and household residential mobility or intent to move. From the analysis, I find that an exogenous commuting shock increases the probability of preferring to relocate. The paper also documents that an increase in commute time, besides increasing preference to relocate, increases the expectation to move. Specifically, I find that the odds-ratio of preference to move and expectation to move increases following an increase in commuting time.

The research finds a ten-fold multiple increases in commuting time increase the odds of preference of moving by 0.02. A separate analysis of male and female commuters shows the effect on female commuters is bigger than their male counterparts for whom the effect is not statistically different from zero. This is in line with the literature that finds higher dis-utility of commuting for females (Jacob et al., 2019). The propensity of residential relocation is also affected by the presence of children, stage of the life cycle, and age among others. These results are in line with previous studies which use different methodology and are conducted in a different setting (Lee and Waddell, 2010).

The multinomial analysis, which categorise preference and expectation to move, gives us another dimension to look at the effect of commuting on residential mobility. Here, the results show that an increase in commuting time increases the odds of preferring and expecting to move over preferring and expecting to stay by about 5%.

From a policy perspective knowing which factors affect residential mobility decisions is crucial for infrastructure and residential investments. In addition, understanding, the role of commuting in shaping residential mobility help policymakers to better understand the impact of shocks at the local labour market. This understanding also informs decisions on transportation investment. For an individual perspective, it has implications in terms of residential investments and understanding how commuting plays a role in their decisions.

One possible area for future extension of the study is to incorporate local labour market characteristics of residential areas. This includes unemployment, crime rate and price of housing. Although region fixed effects capture some of the variations in this regard, accounting for lower-level region characteristics should also be considered. I am planning of extending this once I have granted access to the secure version of the data.

References

Abebe, G., McMillan, M. S., Serafinelli, M., 2018. Foreign Direct Investment and Knowledge Diffusion in Poor Locations: Evidence from Ethiopia. Tech. rep., National Bureau of Economic Research.

Acemoglu, D., Autor, D., 2011. Skills, Tasks and Technologies: Implications for Employment and Earnings. Handbook of Labor Economics 4, 1043–1171.

Ackerberg, D. A., Caves, K., Frazer, G., 2015. Identification Properties of Recent Production Function Estimators. Econometrica 83, 2411–2451.

Amiti, M., Konings, J., 2007. Trade Liberalization, Intermediate Inputs, and Productivity: Evidence from Indonesia. American Economic Review 97, 1611–1638.

Arellano, M., Bonhomme, S., 2017. Quantile Selection Models with an Application to Understanding Changes in Wage Inequality. Econometrica 85, 1–28.

Arnold, J. M., Javorcik, B. S., 2009. Gifted Kids or Pushy Parents? Foreign Direct Investment and Plant Productivity in Indonesia. Journal of International Economics 79, 42–53.

Artuç, E., 2012. Workers' Age and the Impact of Trade Shocks .

Artuç, E., McLaren, J., 2015. Trade Policy and Wage Inequality: A Structural Analysis with Occupational and Sectoral Mobility. Journal of International Economics 97, 278–294.

Attanasio, O., Goldberg, P. K., Pavcnik, N., 2004. Trade Reforms and Wage Inequality in Colombia. Journal of Development Economics 74, 331–366.

Autor, D. H., Dorn, D., Hanson, G. H., 2013. The China Syndrome: Local Labor Market Effects of Import Competition in the United States. The American Economic Review 103, 2121–2168.

Autor, D. H., Dorn, D., Hanson, G. H., 2016. The China Shock: Learning from Labor-Market Adjustment to Large Changes in Trade. Annual Review of Economics 8, 205–240.

Autor, D. H., Dorn, D., Hanson, G. H., Song, J., 2014. Trade Adjustment: Worker-level Evidence. The Quarterly Journal of Economics 129, 1799–1860.

Autor, D. H., Levy, F., Murnane, R. J., 2003. The Skill Content of Recent Technological Change: An Empirical Exploration. The Quarterly Journal of Economics 118, 1279–1333.

Balsvik, R., Jensen, S., Salvanes, K. G., 2015. Made in China, Sold in Norway: Local Labor Market Effects of an Import Shock. Journal of Public Economics 127, 137–144.

Bas, M., Berthou, A., 2017. Does Input-trade Liberalization affect Firms' Foreign Technology Choice? The World Bank Economic Review 31, 351–384.

Bas, M., Strauss-Kahn, V., 2014. Does Importing more Inputs Raise Exports? Firm-level Evidence from France. Review of World Economics 150, 241–275.

Bloom, N., Draca, M., Van Reenen, J., 2016. Trade Induced Technical Change? The Impact of Chinese Imports on Innovation, IT and Productivity. The Review of Economic Studies 83, 87–117.

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., Roberts, J., 2013. Does Management Matter? Evidence from India. The Quarterly Journal of Economics 128, 1–51.

Bøler, E. A., Moxnes, A., Ulltveit-Moe, K. H., 2015. R&D, International Sourcing, and the Joint Impact on Firm Performance. American Economic Review 105, 3704–39.

Brueckner, J. K., Št'astná, L., 2019. Commuting and Migration.

Bruhn, M., Karlan, D., Schoar, A., 2010. What Capital is Missing in Developing Countries? American Economic Review 100, 629–33.

Bwalya, S. M., 2006. Foreign Direct Investment and Technology Spillovers: Evidence from Panel Data Analysis of Manufacturing Firms in Zambia. Journal of development economics 81, 514–526.

Chetverikov, D., Larsen, B., Palmer, C., 2016. IV Quantile Regression for Group-Level Treatments, With an Application to the Distributional Effects of Trade. Econometrica 84, 809–833.

Chiquiar, D., 2008. Globalization, Regional Wage Differentials and the Stolper–Samuelson Theorem: Evidence from Mexico. Journal of International Economics 74, 70–93.

Coşar, A. K., Guner, N., Tybout, J., 2016. Firm dynamics, Job Turnover, and Wage Distributions in an Open Economy. The American Economic Review 106, 625–663.

Dauth, W., Findeisen, S., Suedekum, J., 2014. The Rise of the East and the Far East: German Labor Markets and Trade Integration. Journal of the European Economic Association 12, 1643–1675.

Dauth, W., Suedekum, J., 2016. Globalization and local profiles of economic growth and industrial change. Journal of Economic Geography 16, 1007–1034.

Dix-Carneiro, R., Kovak, B. K., 2015. Trade Reform and Regional Dynamics: Evidence from 25 years of Brazilian Matched Employer-employee Data. Tech. rep., National Bureau of Economic Research.

Dunning, J. H., 2015. The Eclectic Paradigm of International Production: a Restatement and Some Possible Extensions. In: *The Eclectic Paradigm*, Springer, pp. 50–84.

Ethiopian Development Research Institute (EDRI) and Institute of Development Studies (IDS) and International Food Policy Research Institute (IFPRI), 2014. Ethiopia Social Accounting Matrix (SAM): 2005-06.

Edwards, L., Lawrence, R. Z., 2010. US Trade and Wages: The Misleading Implications of Conventional Trade Theory .

Egger, H., Egger, P., Kreickemeier, U., 2013. Trade, Wages, and Profits. European Economic Review 64, 332–350.

Elsby, M. W., Shin, D., Solon, G., 2016. Wage Adjustment in the Great Recession and Other Downturns: Evidence from the United States and Great Britain. Journal of Labor Economics 34, S249–S291.

Eslava, M., Haltiwanger, J., Kugler, A., Kugler, M., 2004. The effects of structural reforms on productivity and profitability enhancing reallocation: evidence from colombia. Journal of development Economics 75, 333–371.

Fatima, S. T., 2016. Productivity Spillovers from Foreign Direct Investment: Evidence from Turkish Micro-level Data. The Journal of International Trade & Economic Development 25, 291–324.

Feenstra, R., Ma, H., Xu, Y., 2017. The China Syndrome: Local Labor Market Effects of Import Competition in the United States: Comment .

Felbermayr, G., Impullitti, G., Prat, J., 2018. Firm Dynamics and Residual Inequality in Open Economies. Journal of the European Economic Association 16, 1476–1539.

Fiorini, M., Sanfilippo, M., Sundaram, A., 2019. Roads: from Trade Liberalization to Firm Productivity. Robert Schuman Centre for Advanced Studies Research Paper No. RSCAS 32.

Gandhi, A., Navarro, S., Rivers, D., 2011. On the Identification of Production Functions: How Heterogeneous is Productivity? Tech. rep., CIBC Working Paper.

Gaubert, C., 2018. Firm Sorting and Agglomeration. American Economic Review 108, 3117–53.

Gibbons, S., Overman, H. G., Pelkonen, P., 2010. Wage Disparities in Britain: People or Place?

Girma, S., Gong, Y., Görg, H., Lancheros, S., 2015. Estimating Direct and Indirect effects of Foreign Direct Investment on Firm Productivity in the Presence of Interactions between Firms. Journal of International Economics 95, 157–169.

Goerke, L., Lorenz, O., 2017. Commuting and Sickness Absence.

Goos, M., Manning, A., Salomons, A., 2014. Explaining Job Polarization: Routine-biased Technological Change and Offshoring. American Economic Review 104, 2509–26.

Greenaway, D., Kneller, R., 2008. Exporting, Productivity and Agglomeration. European economic review 52, 919–939.

Greenland, A., Lopresti, J., 2016. Import exposure and human capital adjustment: Evidence from the US. Journal of International Economics 100, 50–60.

Gutiérrez-i Puigarnau, E., van Ommeren, J. N., 2015. Commuting and Labour Supply Revisited. Urban Studies 52, 2551–2563.

Haas, A., Osland, L., 2014. Commuting, Migration, Housing and Labour Markets: Complex Interactions.

Hakobyan, S., McLaren, J., 2016. Looking for Local Labor Market Effects of NAFTA. Review of Economics and Statistics 98, 728–741.

Halpern, L., Koren, M., Szeidl, A., 2015. Imported Inputs and Productivity. American Economic Review 105, 3660–3703.

Han, J., Liu, R., Zhang, J., 2012. Globalization and Wage Inequality: Evidence from Urban China. Journal of International Economics 87, 288–297.

Harrison, A., Rodríguez-Clare, A., 2010. Trade, Foreign Investment, and Industrial Policy for Developing Countries. In: *Handbook of development economics*, Elsevier, vol. 5, pp. 4039–4214.

Haskel, J. E., Pereira, S. C., Slaughter, M. J., 2007. Does Inward Foreign Direct Investment Boost the Productivity of Domestic Firms? The review of economics and statistics 89, 482–496.

Helpman, E., Itskhoki, O., Muendler, M.-A., Redding, S. J., 2017. Trade and Inequality: From Theory to Estimation. The Review of Economic Studies 84, 357–405.

Jacob, N., Munford, L., Rice, N., Roberts, J., 2019. The Disutility of Commuting? The Effect of Gender and Local Labor Markets. Regional Science and Urban Economics 77, 264–275.

Javorcik, B. S., 2004. Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers through Backward Linkages. American economic review 94, 605–627.

Javorcik, B. S., Spatareanu, M., 2008. To Share or Not to Share: Does Local Participation Matter for Spillovers from Foreign Direct Investment? Journal of development Economics 85, 194–217.

Kasahara, H., Lapham, B., 2013. Productivity and the Decision to Import and Export: Theory and Evidence. Journal of international Economics 89, 297–316.

Kasahara, H., Rodrigue, J., 2008. Does the Use of Imported Intermediates Increase Productivity? Plant-level Evidence. Journal of development economics 87, 106–118.

Keller, W., Yeaple, S. R., 2009. Multinational Enterprises, International Trade, and Productivity Growth: Firm-level Evidence from the United States. The Review of Economics and Statistics 91, 821–831.

Kim, T.-K., Horner, M. W., Marans, R. W., 2005. Life Cycle and Environmental Factors in Selecting Residential and Job Locations. Housing studies 20, 457–473.

Kovak, B. K., 2013. Regional Effects of Trade Reform: What is the Correct Measure of Liberalization? The American Economic Review 103, 1960–1976.

Krugman, P. R., 2008. Trade and Wages, Reconsidered. Brookings Papers on Economic Activity 2008, 103–154.

Künn-Nelen, A., 2016. Does Commuting Affect Health? Health economics 25, 984–1004.

Lee, B. H., Waddell, P., 2010. Residential Mobility and Location Choice: a Nested Logit Model with Sampling of Alternatives. Transportation 37, 587–601.

Levinsohn, J., Petrin, A., 2003. Estimating Production Functions using Inputs to Control for Unobservables. The review of economic studies 70, 317–341.

Lipsey, R. E., Sjöholm, F., 2004. FDI and Wage Spillovers in Indonesian Manufacturing. Review of World Economics 140, 321–332.

Lopez, R. A., 2008. Foreign Technology Licensing, Productivity, and Spillovers. World Development 36, 560–574.

Lu, Y., Tao, Z., Zhu, L., 2017. Identifying FDI Spillovers. Journal of International Economics 107, 75–90.

Ma, L., Ye, R., 2019. Does Daily Commuting Behavior Matter to Employee Productivity? Journal of Transport Geography 76, 130–141.

Manning, A., Petrongolo, B., 2017. How Local are Labor Markets? Evidence from a Spatial Job Search Model. American Economic Review 107, 2877–2907.

McCaig, B., 2011. Exporting out of Poverty: Provincial Poverty in Vietnam and US Market Access. Journal of International Economics 85, 102–113.

McCann, P., 2016. The UK Regional–National Economic Problem: Geography, Globalisation and governance. Routledge.

McFadden, D., 1978. Modeling the Choice of Residential Location. Transportation Research Record .

Melly, B., 2005. Decomposition of Differences in Distribution Using Quantile Regression. Labour Economics 12, 577–590.

Mendez, O., 2015. The Effect of Chinese Import Competition on Mexican Local Labor Markets. The North American Journal of Economics and Finance 34, 364–380.

Monte, F., Redding, S. J., Rossi-Hansberg, E., 2018. Commuting, Migration, and Local Employment Elasticities. American Economic Review 108, 3855–90.

OECD, 2011. Divided We Stand: Why Inequality Keeps Rising. Tech. rep., OECD Publishing, Paris.

OECD, 2017a. OECD Employment Outlook 2017.

OECD, 2017b. OECD Regional Statistics and Indicators. Organisation for Economic Cooperation and Development.

Olley, S., Pakes, A., 1996. The Dynamics of Productivity in the Telecommunications Equipment Industry. Econometrica 64.

ONS, 2017a. Annual survey of hours and earnings, 1997-2016: Secure access. UK Data Service, ed $=12^{th}$.

ONS, 2017b. Business Register and Employment Survey, 1990-2017: Secure Access. UK Data Service, ed $=6^{th}$.

Ottaviano, G. I., Peri, G., Wright, G. C., 2018. Immigration, Trade and Productivity in Services: Evidence from UK Firms. Journal of International Economics 112, 88–108.

Peri, G., Sparber, C., 2009. Task Specialization, Immigration, and Wages. American Economic Journal: Applied Economics 1, 135–169.

Pierce, J. R., Schott, P. K., 2016. The Surprisingly Swift Decline of US Manufacturing Employment. The American Economic Review 106, 1632–1662.

Pike, R., 2011. Patterns of Pay: Results of the Annual Survey of Hours and Earnings 1997 to 2010. Economic & Labour Market Review 5, 14–40.

Reiss, P. C., Wolak, F. A., 2007. Structural Econometric Modeling: Rationales and Examples from Industrial Organization. Handbook of Econometrics 6, 4277–4415.

Revenga, A. L., 1992. Exporting Jobs? The Impact of Import Competition on Employment and Wages in US Manufacturing. The Quarterly Journal of Economics 107, 255–284.

Ritter, M., 2014. Offshoring and Occupational Specificity of Human Capital. Review of Economic Dynamics 17, 780–798.

Roberts, J., Hodgson, R., Dolan, P., 2011. "It's Driving Her Mad": Gender Differences in the Effects of Commuting on Psychological Health. Journal of health economics 30, 1064–1076.

Roberts, J., Taylor, K., 2016. Intra-household Commuting Choices and Local Labour Markets. Oxford Economic Papers 69, 734–757.

Romani, J., Suriñach, J., Artiís, M., 2003. Are Commuting and Residential Mobility Decisions Simultaneous?: The case of Catalonia, Spain. Regional Studies 37, 813–826.

Sampson, T., 2016. Assignment Reversals: Trade, Skill Allocation and Wage Inequality. Journal of Economic Theory 163, 365–409.

Schmidt, C., 2014. Optimal Commuting and Migration Decisions under Commuting Cost Uncertainty. Urban Studies 51, 477–492.

Scott, L. V., John, P., Alun, H., 2016. Commuting Trends in England 1988 to 2015.

Smeets, V., Warzynski, F., 2013. Estimating Productivity with Multi-product Firms, Pricing Heterogeneity and the Role of International Trade. Journal of International Economics 90, 237–244.

So, K. S., Orazem, P. F., Otto, D. M., 2001. The Effects of Housing Prices, Wages, and Commuting Time on Joint Residential and Job Location Choices. American Journal of Agricultural Economics 83, 1036–1048.

Topalova, P., 2010. Factor Immobility and Regional Impacts of Trade Liberalization: Evidence on Poverty from India. American Economic Journal: Applied Economics 2, 1–41.

Topalova, P., Khandelwal, A., 2011. Trade Liberalization and Firm Productivity: The Case of India. Review of economics and statistics 93, 995–1009.

UKHLS, 2018. Understanding Society: Waves 1-8, 2009-2017 and Harmonised BHPS: Waves 1-18, 1991-2009: Special Licence Access.

Utar, H., 2016. Workers Beneath the Floodgates: The Impact of Low-Wage Import Competition and Workers' Adjustment.

Wang, Z., Wei, S.-J., Yu, X., Zhu, K., 2018. Re-examining the Effects of Trading with China on Local Labor Markets: A Supply Chain Perspective. Tech. rep., National Bureau of Economic Research.

WITS-UNSD, 2017. Comtrade: Import and Exports. Retrieved from World Integrated Trade Solution of the World Bank.

Zhang, H., 2017. Static and Dynamic Gains from Costly Importing of Intermediate Inputs: Evidence from Colombia. European Economic Review 91, 118–145.

Appendix

Appendix A

Chapter 1 Supplementary Figures & Tables

In the following section, I provide some details and descriptive statistics on the data used for the analysis.

Table A.1 Description of Variables

Variable	**Explanation**
Year	Year for which the inquiry was conducted. The inquiry reference date is always in April of the year.
Hourly wage (He)	Average hourly wage for the reference period (Gpay / Thrs)
Hours worked (Thrs)	Average total paid hours worked during the reference period (Bhr + Ovhrs)
Weekly earning (Gpay)	Average gross weekly earnings for the reference period 1997-2003 2004 strata1 definition (Bpay + Ipin + Ipop + Sppay + Ovpay) 2004 definition (Bpay + Ipin + Sppay+ Ovpay +imputed Othpay) current definition (Bpay + Ipayin + Sppay +Ovpay + Othpay)
Occ90	Occupation based on Standard Occupational Classification 1990. (1997 – 2001 only)
Occ	Occupation based on Standard Occupational Classification 2003. (2002 onwards)
Sex	Male =1 , female=0
Age	The age at the survey reference date. The dataset only contains people aged 16 and over at the survey reference date.
Full time (Ft)	Full time = 1, part time=0

Table A2 shows the calculated measures of decadal changes in import from China per worker among the top 50 most populous regions in terms of their working age population. For instance, for the ten year equivalence period until 2002, Durham CC is the highest change with a change of £2,094 per worker. Whereas for the period ending in 2010 Northamptonshire experienced the highest change with a change of £3,4499 per worker.

Table A2 also presents weighted median, mean and standard deviation of changes in import exposure for the 128 NUTS-3 regions of the UK. A ten year equivalence in 2002 (between 1997 and 2002), the mean region's growth of import from China is equivalent to £960 per worker. While, a ten year equivalence for the period between 2002 and 2010, the mean region's growth of Chinese import becomes £1,160 which is 20.8 percent increase from the previous decade.

Table A.2 Change in Exposure per Worker Among the Top 50 Most Populous Regions (in £1000)

Rank	Year	Regions	$\Delta Import/Worker$
1	2002	Durham CC	2.094
2	2002	Walsall and Wolverhampton	1.933
3	2002	South and West Derbyshire	1.729
4	2002	Northamptonshire	1.605
5	2002	Leicestershire CC and Rutland	1.449
		Median	0.800
		Mean	0.960
		Std. Dev.	0.550
1	2010	Northamptonshire	3.499
2	2010	South and West Derbyshire	2.799
3	2010	Leicestershire CC and Rutland	2.383
4	2010	Walsall and Wolverhampton	2.286
5	2010	Durham CC	1.930
		Median	0.950
		Mean	1.160
		Std. Dev.	0.770

In Figure 1 I showed that the overall value of import from China rises in the UK. And, the next natural question would be whether this is also the case for each industry. In Figure A.1, I illustrate the trends in the value of imports at the 2-digit NACE industry classification for the period 1990-2015. As one can see, there is a variation in the trend of import among industries. Here, it is worth noting the scales in the y-axis are different. The figure illustrates the presence of an upward trend for most industries after the turn of the millennium.[1] Specifically, the increase in import value is pronounced in industries like

[1]Note that China joined the World Trade Organisation at the end of 2001.

Manufacture of Machinery, not elsewhere specified (n.e.s.) (29); Manufacture of Office Machinery and Computers (30); Manufacture of Radio, Television and Communication Equipment and Apparatus (32); and Manufacture of Furniture, n.e.s. (36). The numbers representing NACE 2-digit code (see Table 2 of the appendix). Meanwhile, Manufacture of Tobacco Products (16) and Manufacture of Coke, Refined Petroleum Products and Nuclear Fuel (23) experience a modest increase in import between 2001 and 2008. After 2008, imports decline for most products. Particularly, for the Manufacture of Wearing Apparel; Dressing and Dyeing of Fur (18) import value decline from more than 8 billion in 2008 to 6 billion pounds in 2015.

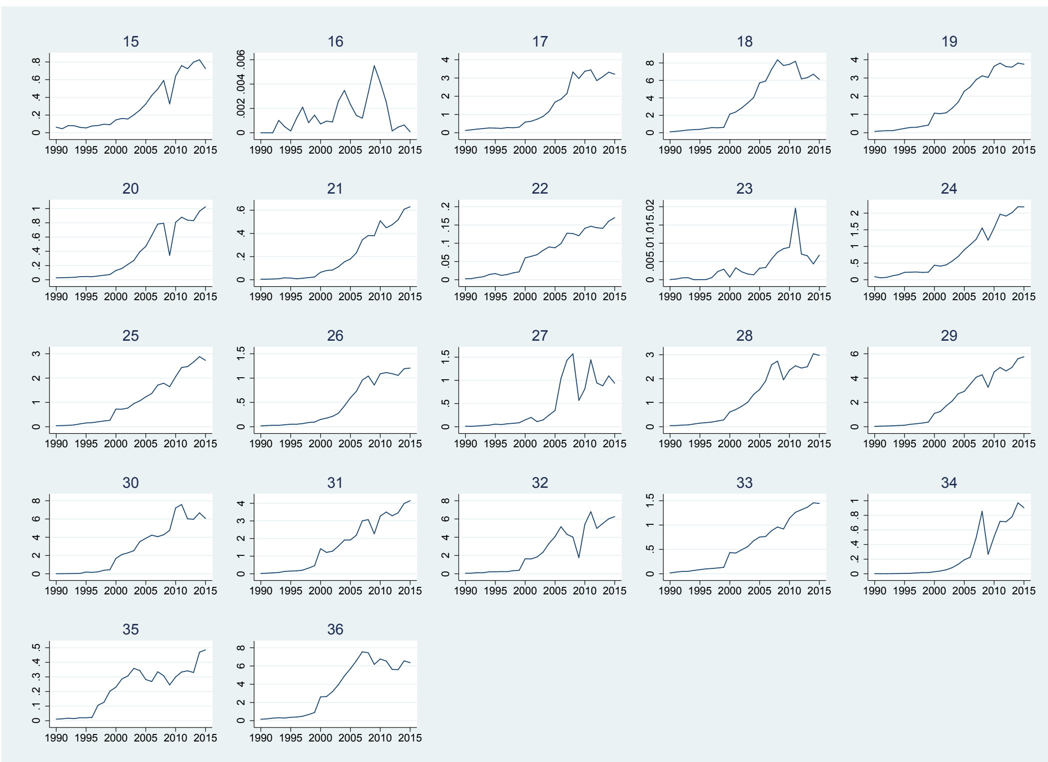


Figure A.1 NACE 2 Level Total Import Value in Billions of 2015 Pounds

As illustrated in Figure 1, the number of manufacturing workers in the UK declines over the past two decades. The aggregate figure, however, does not tell us if the decline is true for all regions. Therefore, in Figure A.2, I present the trend in the total number of manufacturing workers at NUTS-2 regional level for the period 1991 to 2015. Note that the vertical axis is scaled differently for each region. The figure presents three features.

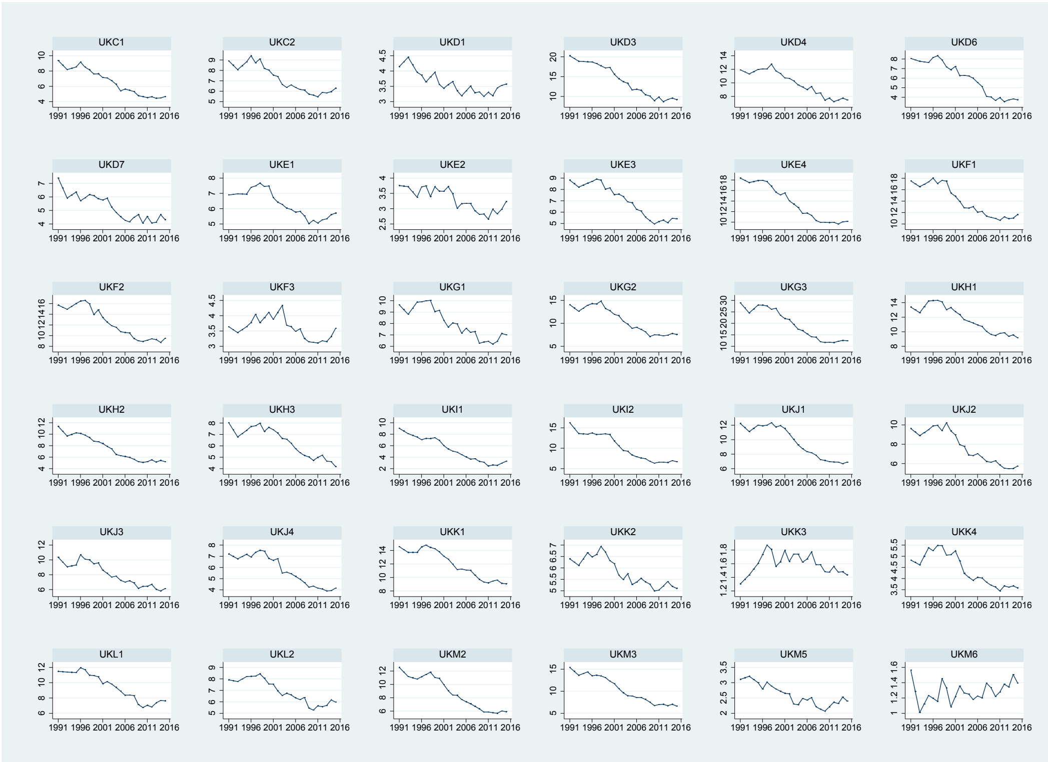


Figure A.2 Regional Trend in Number of Manufacturing Employment

First, the declining trend in the number of manufacturing workers is universal among regions. Second, for most regions the decline in manufacturing workers start to accelerate after the turn of the second millennium. Third,there is variation among regions regarding the extent of the decline, that is, for some regions the decline is steep whereas for others it is gradual. For example, in Tees Valley and Durham (UKC1) the number of manufacturing workers declines from around 90,000 in the early 1990s to a little more than 40,000 in 2015. The largest decline in number of workers during the period of investigation is for West Midlands (UKG3) where the number declines almost by 200,000 from 300,000 to 100,000 (i.e. a 67% decline). In this regard, the exceptions, by showing a more gradual decline, are Cumbria (UKD1), North Yorkshire (UKE2), Lincolnshire (UKF3), Cornwall and Isles of Scilly (UKK3), Highlands and Islands (UKM6) and North Eastern Scotland (UKM5).

A.1 Change in Working Age Population

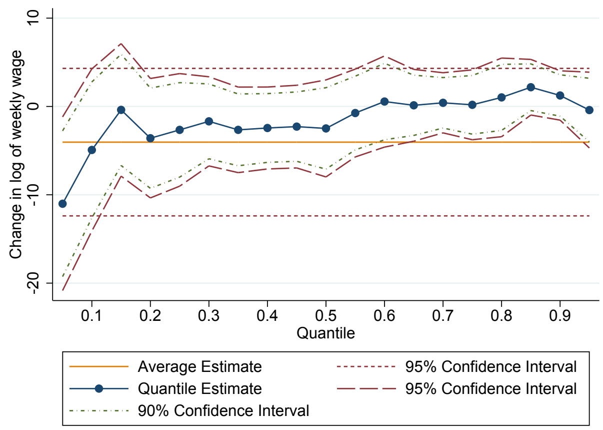
Table A.3 Second Stage Regression Results for the Change in Regional Working-age Population

Dependent Variable: decadal change in the log of working age (in log points)			
	(1)	(2)	(3)
ΔIPW$_r$	-2.693**	-0.933*	-0.986
	(1.284)	(0.550)	(0.854)
Time dummy	2.561***	2.192***	1.542
	(0.517)	(0.708)	(1.622)
Lag manuf. share			40.46
			(57.359)
Lag female share			-16.56
			(13.937)
Lag routine share			-13.38
			(18.533)
Region FE	No	Yes	Yes
F-test	135.9	207.4	121.5
Partial R^2	0.589	0.507	0.437
R^2	0.0818	0.438	0.448
Obs	251	251	251

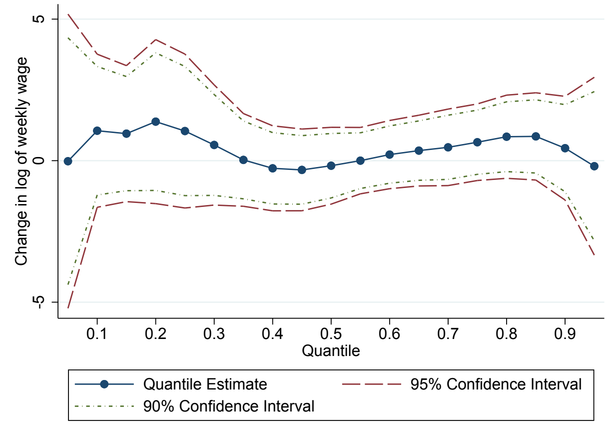
Notes: All regressions include constants and NUTS2 level clustered standard errors (in parenthesis). For all quantile regression, we control for start of period region characteristics such as the share of manufacturing employment, the share of female workers, the share of employment in routine works, region fixed effects and a time dummy for the period 2002-2010.
* Denotes significance at the 10% level, ** Denotes significance at the 5% level, *** Denotes significance at the 1% level.

A.2 With Year 2007

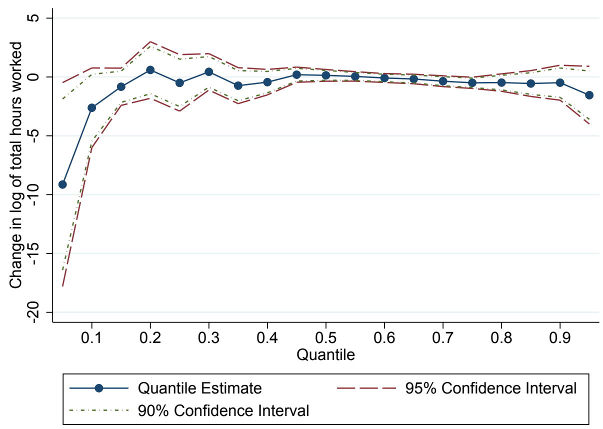
Figure A.3 Estimates With Year 2007



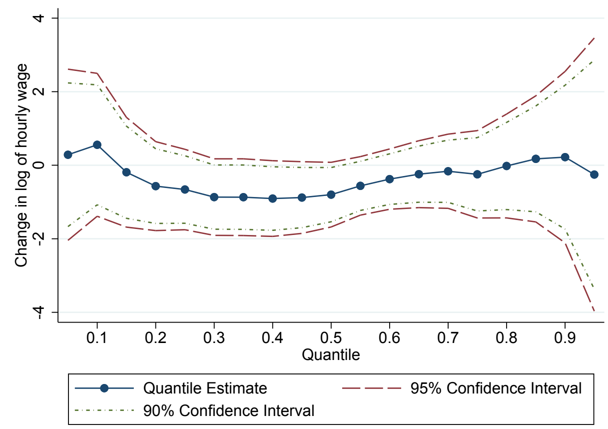
a. Change in log of weekly wage



b. Quantile with individual controls



c. Quantile with individual controls

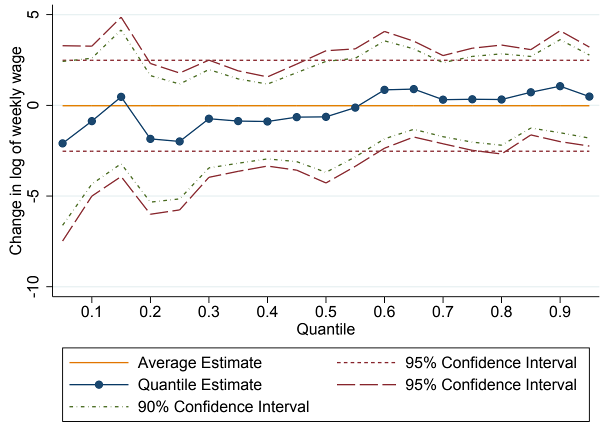


d. Quantile with individual controls

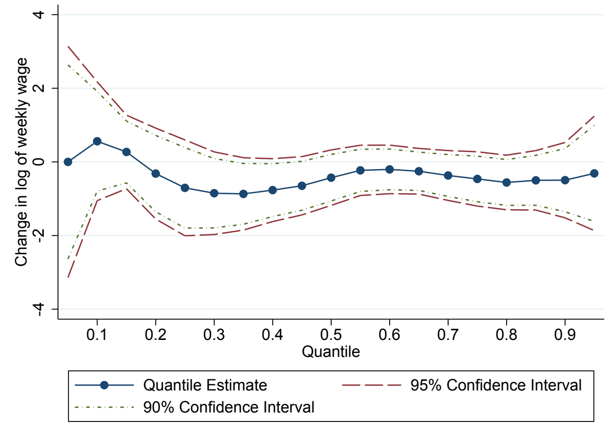
Note: Estimation coefficients and confidence interval for regression on average and quantile with and without controls for individual characteristics. The dependent variable is the change in the log of weekly wage for the upper panels and total hours work and hourly wage for the lower panel. Estimations are on all workers.

A.3 Excluding London

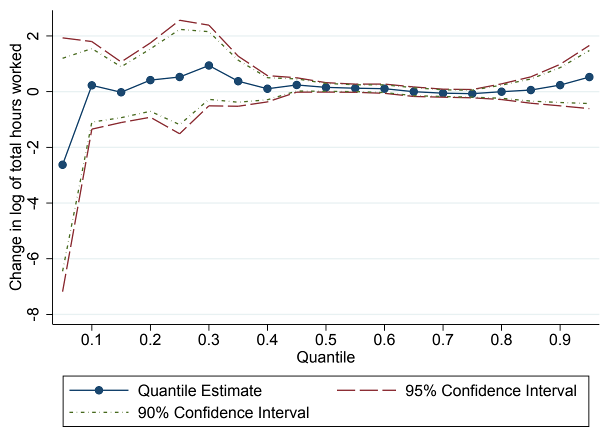
Figure A.4 Estimates Without London



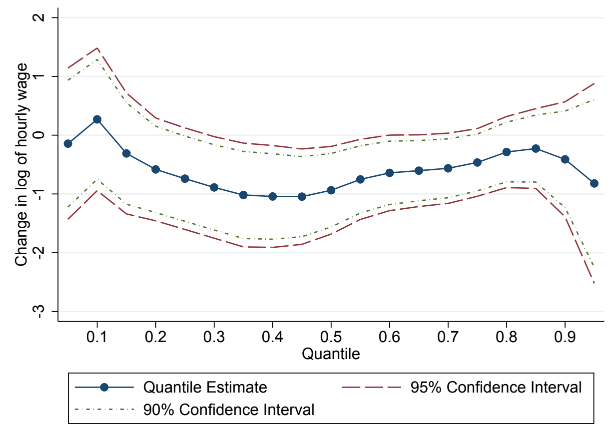
a. Average and quantile without individual controls



b. Weekly wages



c. Total hours worked

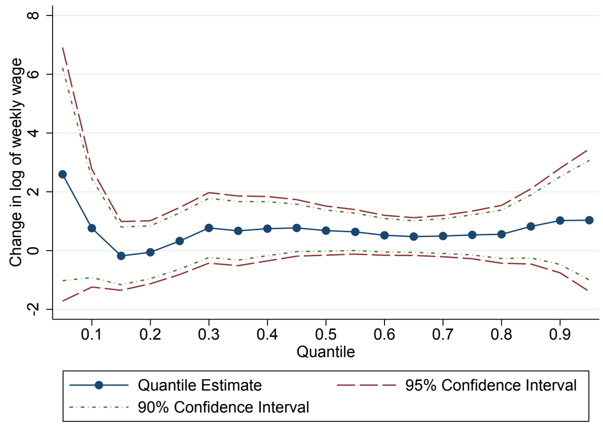


d. Hourly wage

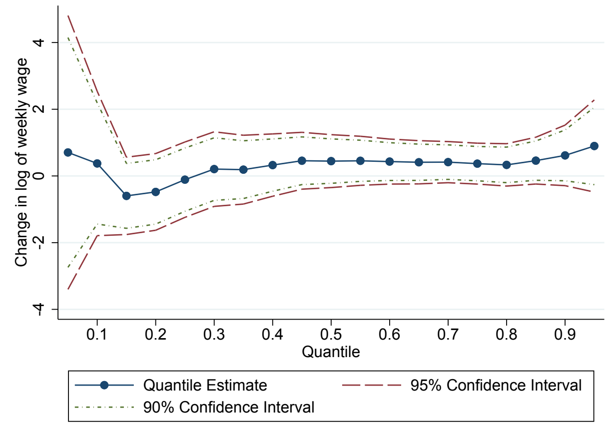
Note: Quantile IV regression on weekly wage, hourly wage and total hours worked after excluding London. The figure shows the effect of a 1000 pound per worker increase in imports from China on the conditional wage distribution.

A.4 On Sub-samples (Heterogeneous Analysis)

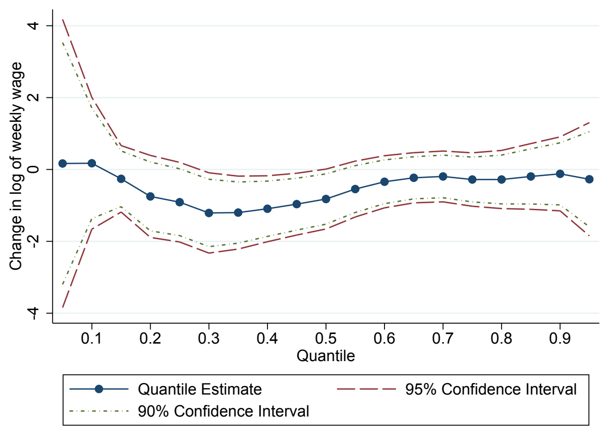
Figure A.5 Estimates on Sub-samples



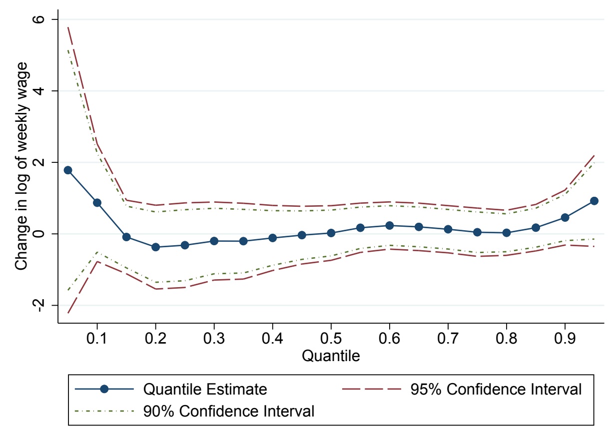
a. Female



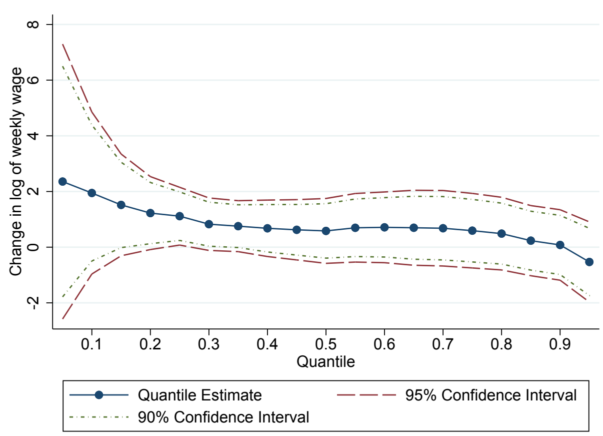
b. Male



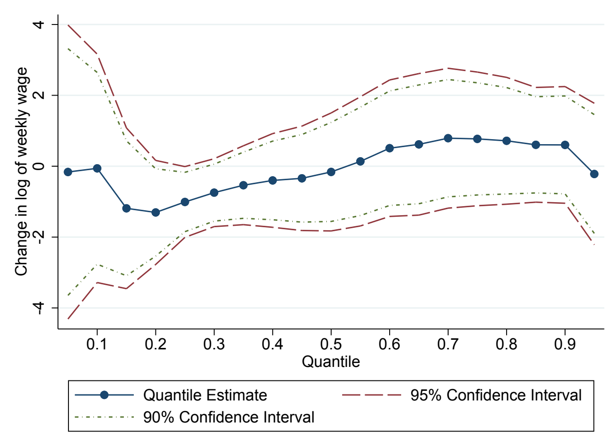
c. Manufacturing



d. Non-manufacturing



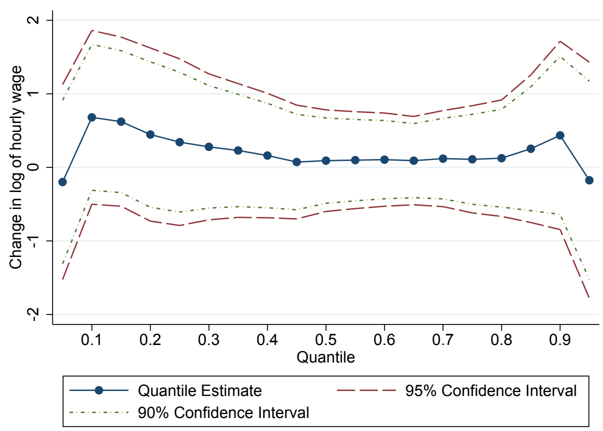
e. Full time



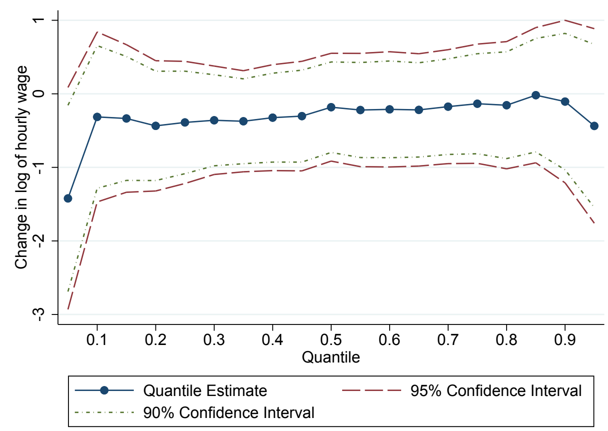
f. Part time

Note: IV Quantile regression on subsample of workers.

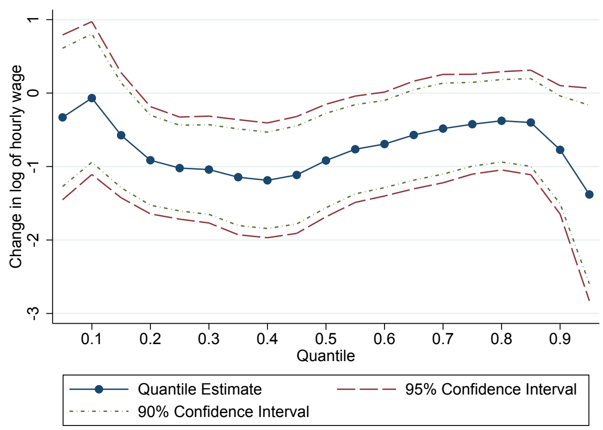
Figure A.6 Estimates on Hourly Wage and Subsamples



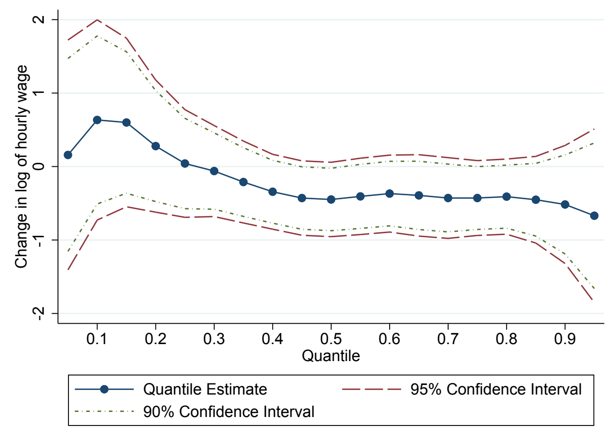
a. Female



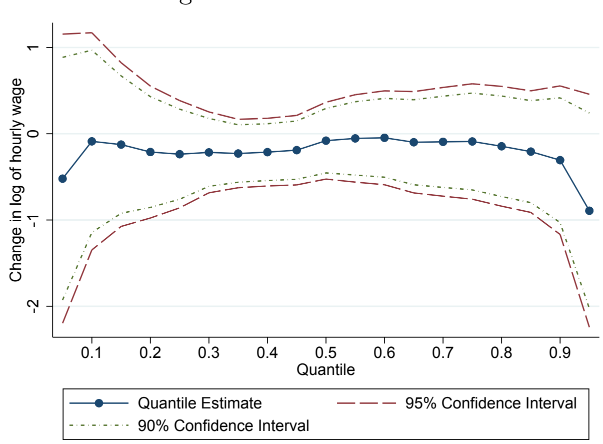
b. Male



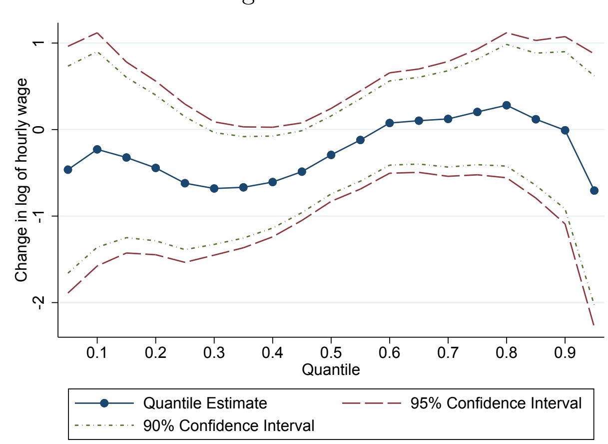
c. Manufacturing



d. Non-manufacturing



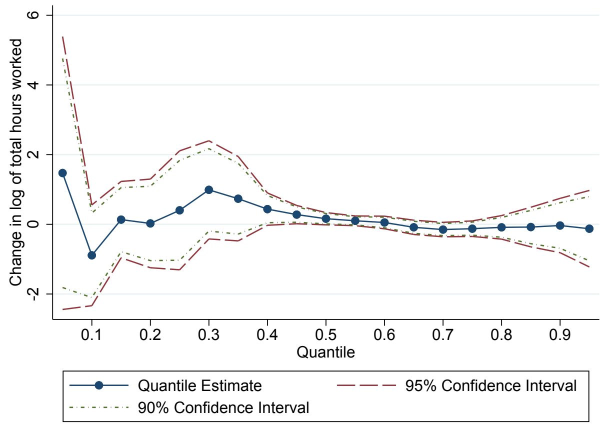
e. Full time



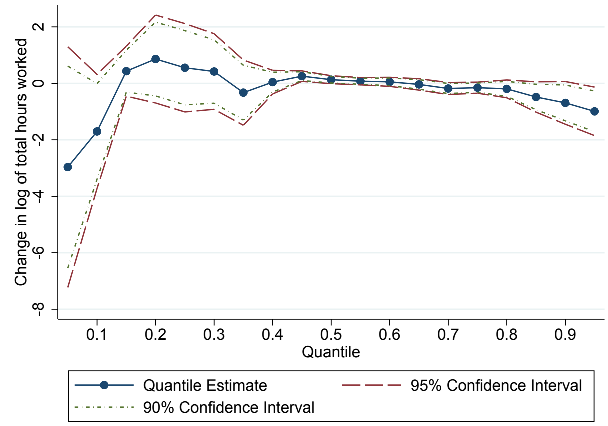
f. Part time

Note: IV Quantile regression on subsample of workers. The figure shows the effect of a 1000 pound per worker increase in imports from China on hourly wage distribution.

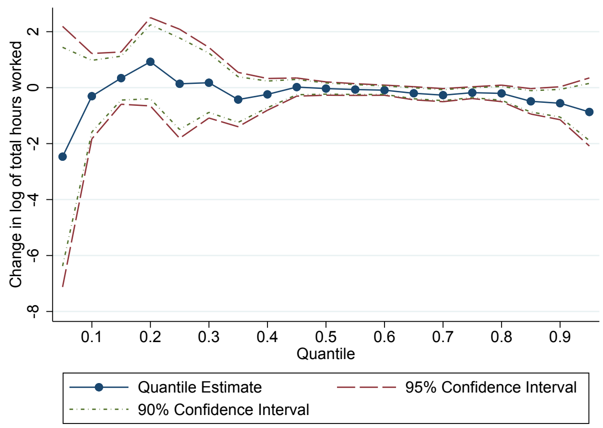
Figure A.7 Estimates on Hours Worked and Subsamples



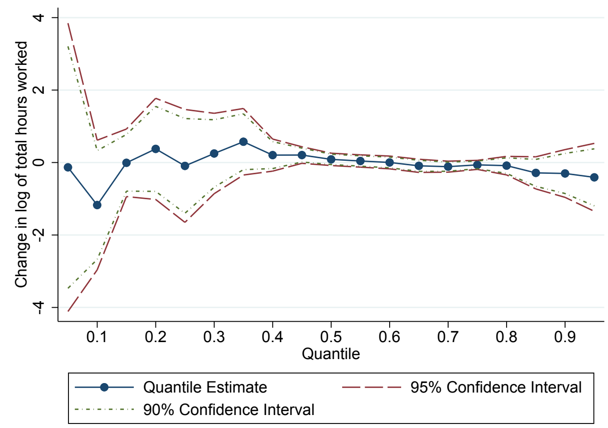
a. Female



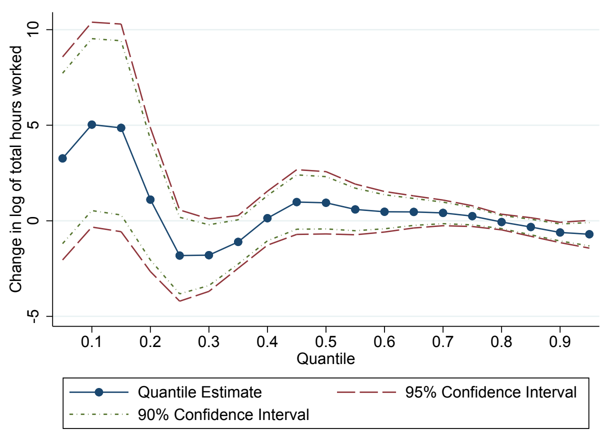
b. Male



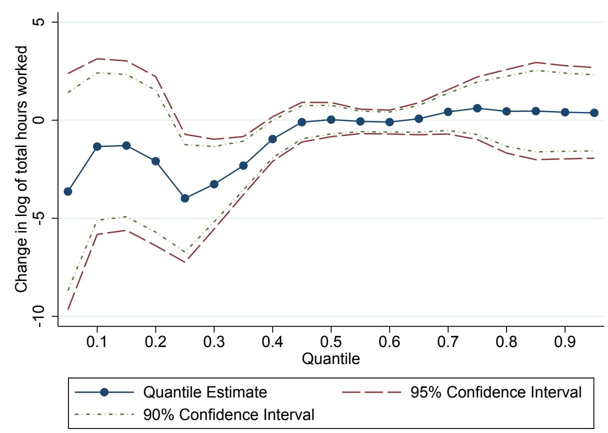
c. Manufacturing



d. Non-manufacturing



e. Full time



f. Part time

Note: IV Quantile regression on subsample of workers. The figure shows the effect of a 1000 pound per worker increase in imports from China on total hours worked distribution.

Appendix B

Chapter 2 Supplementary Figures & Tables

B.1 Price Index

To address measurement error that arises due to the use of aggregate deflator, I used plant level price index. Following Eslava, Haltiwanger, Kugler, and Kugler (2004) and Smeets and Warzynski (2013), I used Tornquist indices in the construction of price, that is, a weighted average of the growth in firm's prices. However, because the product code in the data changes across time for a given firm, I used data on the average price of the main products a firm produced.

Specifically, for a given average price charged by firm i at time t, P_{hit}, I calculate the weight as,

$$\Delta P_{it} = \sum_{h=1}^{H} [\frac{s_{hit} + s_{hit-1}}{2}].[ln(P_{hit}) - ln(P_{hit-1})],$$

where s_{hit} stands for the share of the average product from the total firm's sales.

After that, I used 1996 as the base year by setting $P_{i,1996} = 1$ and add the computed change to the price level. Thus the price indices are given by,

$$P_{it} = P_{it-1} + \Delta(P_{it})$$

In the case of missing values for prices, I used the industry average. I also used the industry average for the first year for firms that enter after 1996 and follow the above procedure for the remaining periods the firm observed.

Moreover, to address missing and repetitive recorders of products within a firm and year in the dataset, I followed Fiorini et al. (2019). Thus, I aggregate the missing values as a separate product category and I used a unit of measurement to correct the issue of repetitive categories.

B.2 Mean Values

Table B.1 Mean Values for Main Variables of Interest by Year

Year	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Labour	133	110	115	112	110	105	95	90	89	115	89	85	66	67	84
Capital per L	5298	7396	9444	9751	8104	9913	7130	5908	5557	6592	5064	3947	3211	1990	5243
Output per L	8658	9879	11297	10641	11290	20450	9577	8050	9890	11290	8235	8481	6191	5289	35006
Material per Q	210	158	146	189	187	186	489	66	235	367	223	137	116	76	77
Exporter dummy	0.04	0.04	0.04	0.05	0.05	0.05	0.04	0.04	0.05	0.07	0.05	0.04	0.04	0.04	0.04
Importer dummy	0.66	0.68	0.73	0.69	0.73	0.66	0.65	0.76	0.71	0.72	0.69	0.63	0.64	0.61	0.61
Import share	0.30	0.31	0.33	0.32	0.35	0.31	0.34	0.37	0.36	0.42	0.37	0.34	0.33	0.30	0.33
No. import	2.12	2.45	2.59	2.56	2.68	2.34	2.54	2.85	2.73	3.63	3.35	2.90	3.15	2.86	2.94
Foreign dummy	0.04	0.03	0.03	0.04	0.04	0.05	0.05	0.04	0.05	0.07	0.04	0.04	0.03	0.04	0.05
Private dummy	0.74	0.80	0.80	0.81	0.82	0.84	0.85	0.86	0.86	0.83	0.88	0.91	0.94	0.95	0.94
Observations	617	693	721	714	724	715	883	939	997	763	1153	1339	1734	1948	1958

B.3 Mathematical Derivation

Halpern et al. (2015)'s method of accounting for imported inputs can be estimated as follow. First, I assume a production function of labour, capital and intermediary inputs. The intermediary inputs can be further decomposed into imported and domestic inputs and enter into production function in CES form. An approach by Halpern et al. (2015) which accounts for the effect of imported inputs. To this end, I include the number of input varieties a firm imported into the equation below and control for the productivity effect of imported inputs. In the processes disentangling the effect of imported inputs on productivity from the crude measure of productivity. This renders an estimate for import adjusted productivity, ω_{ist}. Assuming Cobb-Douglas production function,

$$Q_j = \Omega_j K_j^{\alpha_k} L_j^{\alpha_l} \prod_{i=1}^{N} X_{ji}^{\gamma}, \tag{A1}$$

where $X_{ji}^{\gamma} = [(B_{ji}X_{jiF})^{\frac{\theta-1}{\theta}} + (X_{jiH})^{\frac{\theta-1}{\theta}}]^{\frac{\theta}{\theta-1}}$.

Quality-adjusted price gain of imported input is given by $A = B_{ji}\frac{P_{iH}}{P_{iF}}$. And minimizing input cost with respect to intermediate input use gives us the price of a composite intermediate input, $P_{ji} = P_{iH}[1 + A^{\theta-1}]^{\frac{1}{1-\theta}}$. From this equation, we can drive per-product import gain as log(percentage) reduction in the cost of the tradeable composite good i when imports are also used. Mathematically this is given as $a = lnP_iH - lnP_ij$ and it is equivalent to $a = \frac{log[1+A^{\theta-1}]}{\theta-1}$.

Halpern et al. (2015) assume the relative importance of imported, $G(n_j) = \frac{\sum_{i=1}^{n_j}\gamma_i}{\sum_{i=1}^{N}\gamma_i} = \frac{\sum_{i=1}^{n_j}\gamma_i}{\gamma}$, has a parametric function form of

$$G(n) = \begin{cases} \bar{G}(1 - [1 - (\frac{n}{\bar{n}})^{\lambda}]^{\frac{1}{\lambda}} & \text{if } n \leq \bar{n} \\ \bar{G} & \text{if } n \geq \bar{n}. \end{cases} \tag{A2}$$

where, $\bar{G}$ corresponds to the maximum share of tradeable from total inputs, n stands for the number of inputs and λ represents the curvature.

Here it is worth to note that in estimating different gains from imported inputs, I separately estimate the above function for domestic and foreign firms. This, in turn, allows the relative importance, $G(n)$ to take different values for domestic and foreign firms.

Once we drive the relative importance of each imported input from the above function, the import demand is given by the product of the relative importance of imported inputs and optimal import share, $\frac{M_j^F}{M_j} = S \times G(n_j)$. Then we use the information on import share to represent the left-hand side and use OLS regression to estimate optimal import share, S.

Given Cobb-Douglas production function, expenditure on intermediate inputs is given by:

$$M_j = \Gamma \prod_{i=1}^{N} P_{ji}^{\frac{\gamma_i}{\gamma}} \prod_{i=1}^{N} X_{ji}^{\frac{\gamma_i}{\gamma}},$$

$$= \Gamma \prod_{i=1}^{N} P_{iH}^{\frac{\gamma_i}{\gamma}} \prod_{i=1}^{n_j} exp(-a\frac{\gamma_i}{\gamma}) \prod_{i=1}^{N} X_{ji}^{\frac{\gamma_i}{\gamma}},$$

$$= \Gamma \prod_{i=1}^{N} P_{iH}^{\frac{\gamma_i}{\gamma}} \prod_{i=1}^{n_j} exp(-aG(n_j)) \prod_{i=1}^{N} X_{ji}^{\frac{\gamma_i}{\gamma}},$$

$$= \Gamma exp(-aG(n_j)) \prod_{i=1}^{N} P_{iH}^{\frac{\gamma_i}{\gamma}} \prod_{i=1}^{N} X_{ji}^{\frac{\gamma_i}{\gamma}},$$

$$M_j^{\gamma} = \Gamma^{\gamma} exp(-aG(n_j)\gamma) \prod_{i=1}^{N} P_{iH}^{\gamma_i} \prod_{i=1}^{N} X_{ji}^{\gamma_i},$$

$$\prod_{i=1}^{N} X_{ji}^{\gamma_i} = M_j^{\gamma} \Gamma^{-\gamma} exp(aG(n_j)\gamma) \prod_{i=1}^{N} P_{iH}^{-\gamma_i},$$

Substituting the last line into: $Q_j = \Omega_j K_j^{\alpha_k} L_j^{\alpha_l} \prod_{i=1}^{N} X_{ji}^{\gamma}$, and taking a natural logarithm give us:

$$q_j = \alpha_0 + \alpha_l l_j + \alpha_k k_j + \gamma(m_{jt} - \rho) + \gamma a G(n_j) + \omega_j + \varepsilon_j, \tag{A3}$$

As discussed in the main text, to solve this problem, Olley and Pakes (1996) (OP) suggest the use of investment as a proxy.

Table B.2 Models Estimation from OP and HKS Methods on Lead FDI

	OP		HKS	
	Domestic firms	All firms	Domestic firms	All firms
F.Backward	0.266	0.308*	0.270	0.313*
	(0.187)	(0.181)	(0.190)	(0.184)
F.Forward	-1.857***	-1.855***	-1.885***	-1.882***
	(0.467)	(0.457)	(0.474)	(0.464)
F.Horizontal	0.0413***	0.0385***	0.0419***	0.0391***
	(0.010)	(0.010)	(0.011)	(0.010)
Absorptive Capacity	7.356***	7.357***	7.247***	7.248***
	(0.003)	(0.003)	(0.003)	(0.003)
Foreign Share		0.00393**		0.00399**
		(0.002)		(0.002)
R^2	0.999	0.999	0.999	0.999
N	8924	9299	8924	9299

* Denotes significance at the 10% level, ** Denotes significance at the 5% level, *** Denotes significance at the 1% level. Robust standard errors in parentheses. Year, region and industry fixed effects are included in the regression.

Appendix C

Chapter 3 Supplementary Figures & Tables

C.1 Probit Regression

Table C.1 Probit Model

	Model 1	Model 2	Model 3
log(Commute)	0.019*	0.038***	0.038***
	(0.010)	(0.012)	(0.012)
Age		0.013**	0.012*
		(0.006)	(0.006)
Age^2		-0.001***	-0.001***
		(0.000)	(0.000)
Male		-0.039	-0.027
		(0.025)	(0.027)
Single		0.135***	0.135***
		(0.025)	(0.026)
HRP		-0.008	-0.003
		(0.023)	(0.023)
Full time		0.169***	0.167***
		(0.027)	(0.028)
No. children		0.025	0.026
		(0.016)	(0.017)
Urban		0.343***	0.338***
		(0.029)	(0.029)
log(Net income)		-0.042***	-0.039***
		(0.011)	(0.012)
Education Level	no	yes	yes
Travel mode	no	yes	yes
House ownership	no	yes	yes
Occupation FE	no	no	yes
Industry FE	no	no	yes
Household FE	yes	yes	yes
Region FE	yes	yes	yes
Wave FE	yes	yes	yes
Obs	67,881	57,726	56,749

Notes: Dependent variable is a dummy which takes a value of one if an individual prefer to move house. HRP represents Household Reference Person. Robust standard errors in parentheses. * Denotes significance at the 10% level, ** Denotes significance at the 5% level, *** Denotes significance at the 1% level.

C.2 Mobility by Commuting type (Multinominal)

Table C.2 Multinomial Logit Model: Decrease in Commuting

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Prefer	To Stay	To move	To move	To Stay	To move	To move	To Stay	To move	To move
Expect	To Move	To Stay	To Move	To Move	To Stay	To Move	To Move	To Stay	To Move
log(Commute)	1.020	1.019	1.011	1.027	1.044**	1.042	1.015	1.046**	1.035
	(0.054)	(0.017)	(0.025)	(0.061)	(0.0197)	(0.03)	(0.062)	(0.020)	(0.030)
Age				0.922*	1.001	0.927***	0.920*	1.003	0.927***
				(0.039)	(0.014)	(0.021)	(0.039)	(0.015)	(0.021)
Age^2				1.001	1.000	1.000	1.001	1.000	1.000
				(0.0005)	(0.0002)	(0.0003)	(0.0005)	(0.0002)	(0.0003)
Male				1.062	0.900**	1.003	1.085	0.929	1.037
				(0.165)	(0.045)	(0.076)	(0.184)	(0.051)	(0.085)
Single				2.562***	1.016	1.480***	2.467***	1.004	1.472***
				(0.391)	(0.0549)	(0.112)	(0.385)	(0.0548)	(0.113)
HRP				1.070	0.995	1.010	1.094	1.002	1.013
				(0.170)	(0.047)	(0.074)	(0.178)	(0.048)	(0.075)
Fulltime				0.993	1.387***	1.226*	1.103	1.405***	1.205*
				(0.212)	(0.089)	(0.130)	(0.246)	(0.092)	(0.130)
No Children				1.011	1.096***	0.962	0.981	1.104***	0.949
				(0.120)	(0.037)	(0.054)	(0.121)	(0.038)	(0.054)
Urban				1.007	1.599***	1.263***	1.010	1.594***	1.265***
				(0.181)	(0.093)	(0.111)	(0.188)	(0.094)	(0.113)
Log(Net income)				1.185*	0.867***	1.083*	1.168*	0.875***	1.026
				(0.103)	(0.025)	(0.047)	(0.109)	(0.027)	(0.048)
Education Level	no	no	no	no	no	no	yes	yes	yes
House ownership	no	no	no	no	no	no	yes	yes	yes
Travel mode	no	no	no	no	no	no	yes	yes	yes
Occupation FE	no	no	no	no	no	no	yes	yes	yes
Industry FE	no	no	no	no	no	no	yes	yes	yes
Region FE	yes	yes	yes	yes	yes	yes	yes	yes	yes
Wave FE	yes	yes	yes	yes	yes	yes	yes	yes	yes
Constant	0.0134***	0.380***	0.123***	0.0224***	0.355***	0.479	0.0167***	0.294***	0.725
	(0.00740)	(0.0476)	(0.0236)	(0.0254)	(0.130)	(0.260)	(0.0218)	(0.120)	(0.442)
Obs	14,833	14,833	14,833	12,913	12,913	12,913	12,747	12,747	12,747

Notes: Dependent variable is a categorical variable which takes a value of one to four if an individual prefer and expect to stay, prefer to stay but expect to move, prefer to move but expect to stay and prefer and expect to move house respectively. HRP represents Household Reference Person. Robust standard errors in parentheses. * Denotes significance at the 10% level, ** Denotes significance at the 5% level, *** Denotes significance at the 1% level.

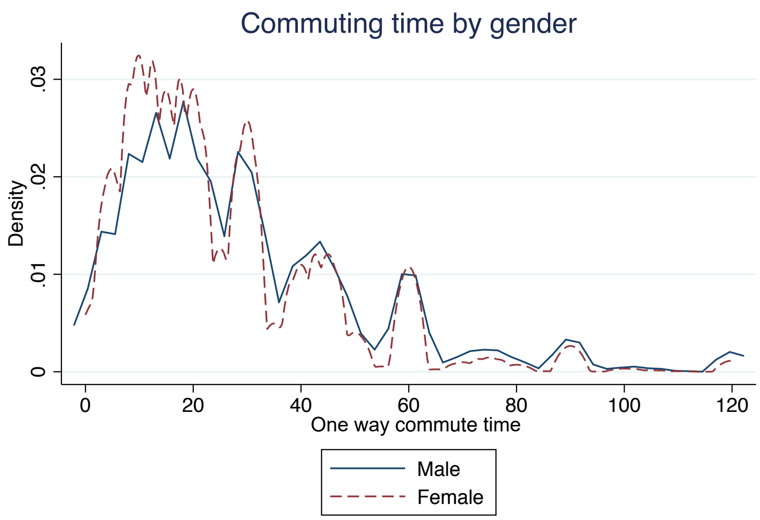
Table C.3 Multinomial Logit Model: Increase in Commuting

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Prefer	To Stay	To move	To move	To Stay	To move	To move	To Stay	To move	To move
Expect	To Move	To Stay	To Move	To Move	To Stay	To Move	To Move	To Stay	To Move
log(Commute)	0.989	1.004	1.037***	1.042	1.026**	1.071***	1.042	1.031***	1.060***
	(0.024)	(0.009)	(0.013)	(0.029)	(0.011)	(0.016)	(0.030)	(0.011)	(0.016)
Age				0.896***	1.031***	0.972**	0.899***	1.029***	0.968***
				(0.017)	(0.007)	(0.011)	(0.017)	(0.0075)	(0.011)
Age^2				1.001***	1.000***	1.000	1.001***	1.000***	1.000
				(0.0002)	(8.35e-05)	(0.0001)	(0.0002)	(8.44e-05)	(0.0001)
Male				1.015	0.927***	0.982	1.095	0.955	0.977
				(0.0710)	(0.0244)	(0.0368)	(0.0830)	(0.0276)	(0.0399)
Single				1.999***	1.082***	1.412***	2.017***	1.077***	1.426***
				(0.142)	(0.030)	(0.054)	(0.145)	(0.030)	(0.055)
HRP				1.113	1.081***	1.036	1.101	1.092***	1.028
				(0.0811)	(0.0272)	(0.0384)	(0.0814)	(0.0278)	(0.0386)
Fulltime				1.156	1.225***	1.281***	1.183*	1.240***	1.232***
				(0.115)	(0.041)	(0.066)	(0.120)	(0.042)	(0.065)
No Children				0.876**	1.070***	0.956	0.873**	1.072***	0.957
				(0.053)	(0.019)	(0.0268)	(0.053)	(0.019)	(0.027)
Urban				1.037	1.470***	1.150***	1.038	1.448***	1.153***
				(0.087)	(0.044)	(0.050)	(0.088)	(0.044)	(0.051)
Log(Net income)				1.089**	0.933***	1.011	1.068	0.951***	0.993
				(0.0428)	(0.013)	(0.021)	(0.046)	(0.015)	(0.023)
Education Level	no	no	no	no	no	no	yes	yes	yes
House ownership	no	no	no	no	no	no	yes	yes	yes
Travel mode	no	no	no	no	no	no	yes	yes	yes
Occupation FE	no	no	no	no	no	no	yes	yes	yes
Industry FE	no	no	no	no	no	no	yes	yes	yes
Region FE	yes	yes	yes	yes	yes	yes	yes	yes	yes
Wave FE	yes	yes	yes	yes	yes	yes	yes	yes	yes
Constant	0.0517***	0.458***	0.117***	0.271***	0.177***	0.341***	0.157***	0.133***	0.538**
	(0.00907)	(0.0302)	(0.0117)	(0.125)	(0.0312)	(0.0879)	(0.0866)	(0.0266)	(0.156)
Obs	52,902	52,902	52,902	44,685	44,685	44,685	43,878	43,878	43,878

Notes: Dependent variable is a categorical variable which takes a value of one to four if an individual prefer and expect to stay, prefer to stay but expect to move, prefer to move but expect to stay and prefer and expect to move house respectively. HRP represents Household Reference Person. Robust standard errors in parentheses. * Denotes significance at the 10% level, ** Denotes significance at the 5% level, *** Denotes significance at the 1% level.

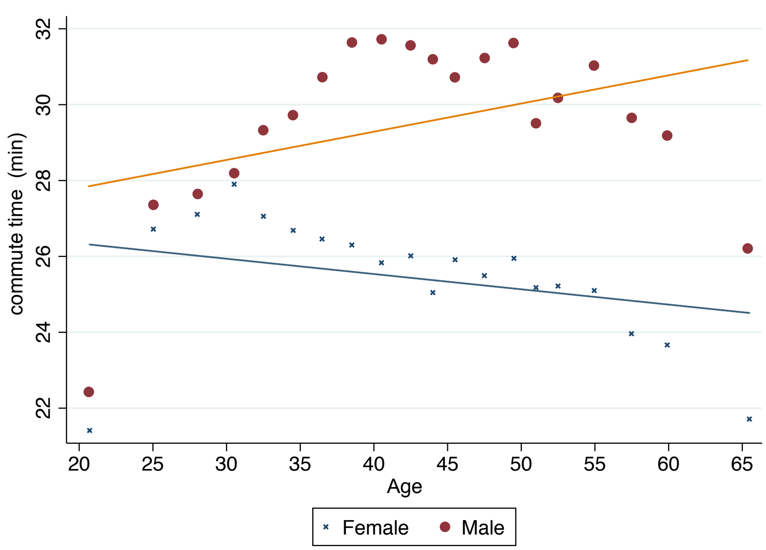
C.3 Commuting Time Data Patterns in the UK

Figure C.1 Commute time by gender



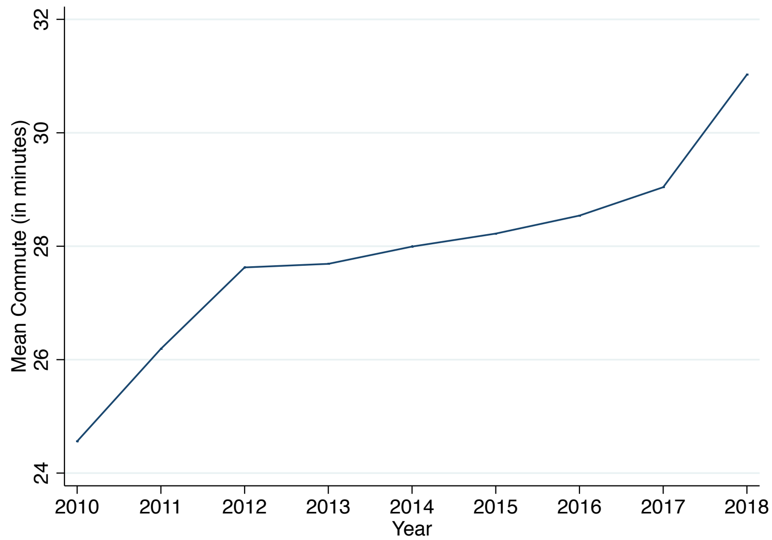
Source: Author computation using UKHLS

Figure C.2 Commute by Gender and Age



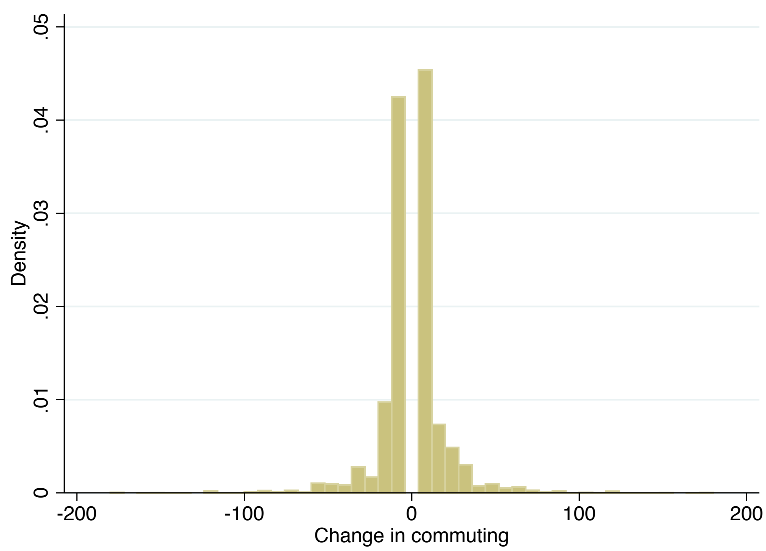
Source: Author computation using UKHLS

Figure C.3 Commute Time by Year



Source: Author computation using UKHLS

Figure C.4 Distribution of Change in Commuting Time



Source: Author computation using UKHLS