Spatial wage inequality in North America and Western Europe:
changes between and within local labour markets 1975-2019

Luis Bauluz
Paweł Bukowski
Mark Fransham
Annie Lee
Margarita López Forero
Filip Novokmet
Sébastien Breau
Neil Lee
Clément Malgouyres
Moritz Schularick
Gregory Verdugo

July 2023

WORLD INEQUALITY LAB

WID.WORLD
THE SOURCE FOR
GLOBAL INEQUALITY DATA

Spatial wage inequality in North America and Western Europe: changes between and within local labour markets 1975-2019

L. Bauluz[1], P. Bukowski[2,8], M. Fransham[3,8], A. Lee[4], M. López Forero[5], F. Novokmet[6],

S. Breau[4], N. Lee[8], C. Malgouyres[9], M. Schularick[10], G. Verdugo[5]

[1]CUNEF Universidad; World Inequality Lab

[2]University College London; Polish Academy of Sciences

[3]University of Oxford

[4]McGill University

[5]Université Paris-Saclay, Univ Evry

[6]University of Bonn; World Inequality Lab

[8]London School of Economics

[9]Paris School of Economics

[10]Kiel Institute; Sciences Po

Abstract

The rise of economic inequalities in advanced economies has been often linked with the growth of spatial inequalities within countries, yet there is limited comparative research that studies the relationship between national and subnational economic inequality. This paper presents the first systematic attempt to create internationally comparable evidence showing how different countries perform in terms of geographic wage inequalities. We create cross-country comparable measures of spatial wage disparities between and within similarly-defined local labour market areas (LLMAs) for Canada, France, (West) Germany, the UK and the US since the 1970s, and assess their contribution to national inequality. By the end of the 2010s, spatial inequalities in LLMA mean wages are similar in Canada, France, Germany and the UK; the US exhibits the highest degree of spatial inequality. Over the study period, spatial inequalities have nearly doubled in all countries, except for France where spatial inequalities have fallen back to 1970s levels. Due to a concomitant increase in within-place inequality, the contribution of places in explaining national wage inequality has remained fairly constant over the 40-year study period, except in the UK where we document a significant increase. Whilst common global social, economic and technological shocks are important drivers of spatial inequality, this variation in levels and trends of spatial inequality opens the way to comparative research exploring the role of national institutions in mediating how global shocks translate into economic disparities between places.

Acknowledgments

1. **Introduction**

While global income inequality has been falling over the last few decades, income inequality within most of the world's leading economies has been on the rise (Alvaredo et al. 2018). In parallel, a new, more unequal and polarized geography of prosperity and opportunity has become the defining feature of the first two decades of the 21st century, giving way to what has been dubbed the "great divergence" (Moretti 2012) or "great inversion" (Storper 2018). Social and economic crises including the 2007-2009 global financial crisis and the 2020-2022 coronavirus pandemic further compounded these spatial inequalities. Discontent with this uneven geography of opportunity is apparent in the rise of populist politics across Europe and the United States, challenging the stability of democratic societies (Rodríguez-Pose 2018). A fundamental question is thus to what extent the rise of national income inequalities is connected to the evolution of spatial inequalities within countries, especially the sharp divide between "superstar' global cities and "left-behind" post-industrial towns?

Comparative research on income inequalities across nations has been highly influential through projects such as the World Inequality Lab and the Distributional National Accounts approach (Alvaredo *et al.* 2016). However there is limited comparative data for spatial income inequalities that makes the link between national and subnational economic inequality and enables the comparison of the level of geographic inequality in different countries over time. Existing research into income inequality has tended to be focused either on national income inequalities or the spatial distribution of income, with little systematic research to link these two perspectives. One good example of this disconnect is the absence of a single chapter dealing with spatial or regional income inequality in the latest Handbook of Income Distribution (Atkinson and Bourguignon 2014), which otherwise covers the full range of topics related to income inequality. One important reason for this gap in the literature is the lack of consistent and comparable datasets on national inequality decomposed by sub-national regions.

This paper is a first systematic attempt to create internationally comparable evidence to show how different countries perform in terms of geographic income inequalities. We present cross-country and cross-time comparable measures of spatial wage disparities between and within local labour market areas (LLMAs) for Canada, France, (West) Germany, the United Kingdom and the United States since the 1970s, assessing their importance for national inequalities. To ensure comparability we use similarly defined LLMAs across countries, exploit high-quality administrative datasets where possible, and employ the same definition of income – weekly labour market earnings for adult full-time workers. The estimates presented in this paper are a first step in the construction of a global database of spatial income inequality that can be used to describe long-term trends over time, acting as a resource for researchers to study the drivers and consequences of geographic income inequality.

We show that today, the United States has the highest national wage inequality followed by Canada, Germany, the UK, and France. We reproduce a well-known fact (e.g., Piketty 2021, Guvenen et al. 2022) that national disparities grew between 1980 and the financial crisis of 2008 but have stagnated or even somewhat declined after. France experienced a relatively small increase in wage inequality, moving from being the most unequal of the three European countries in 1975 to being the most egalitarian today. Canada and the US have had strong and sustained rises in wage inequality throughout the period, whereas for the UK and Germany the substantial increases in inequality were limited to the periods 1980-1995 and 1995-2010 respectively.

Our novel set of results consider trends in spatial inequalities within countries. We find that by the end of the 2010s, spatial inequalities in LLMA mean wages are similar in Canada, France, Germany and the UK; the United States is by far the most unequal by this measure. Most countries experienced a near doubling of the variance of log mean area wages over the period. The exception is France, where spatial inequalities grew in the earlier part of the period but have since fallen back to 1970s levels. In all countries except France there is a strong trend of increased dispersion in wages paid at the top of the distribution between LLMAs, but for most countries some convergence in the lowest wages paid across areas.

How important are spatial inequalities for national inequality trends? We show that the overall importance of place in the total variance of wages is small - in the UK, the country with the biggest role of place, LLMAs explain today around 7% of the total variation in wages; in Canada, the country with the lowest importance, less than 3%. Although the US is the most spatially unequal country in our study, its higher degree of national wage inequality means that the relative contribution of spatial inequality to total national wage inequality is very similar to France. In most countries there has been little change in the contribution of place to national wage inequality; the UK is the only country which has experienced a substantial increase in the importance of place since 1975.

Figure 1 provides a visual representation of the scope of this paper, showing inequality measures for LLMAs in our five study countries in the 1980s and the 2010s. Here we show the share of total LLMA wages that were paid to the top 20% of earners in that LLMA - a measure of within-place inequality. It shows that North America has higher national wage inequality than Europe, and that in most countries wage inequality has increased over time. It is striking to observe a rise in wage inequality across much of Germany, while in contrast, many areas in France have experienced falls in the top 20% share over the period. In North America, we can see a concentration of the most unequal areas along the West and East coasts.
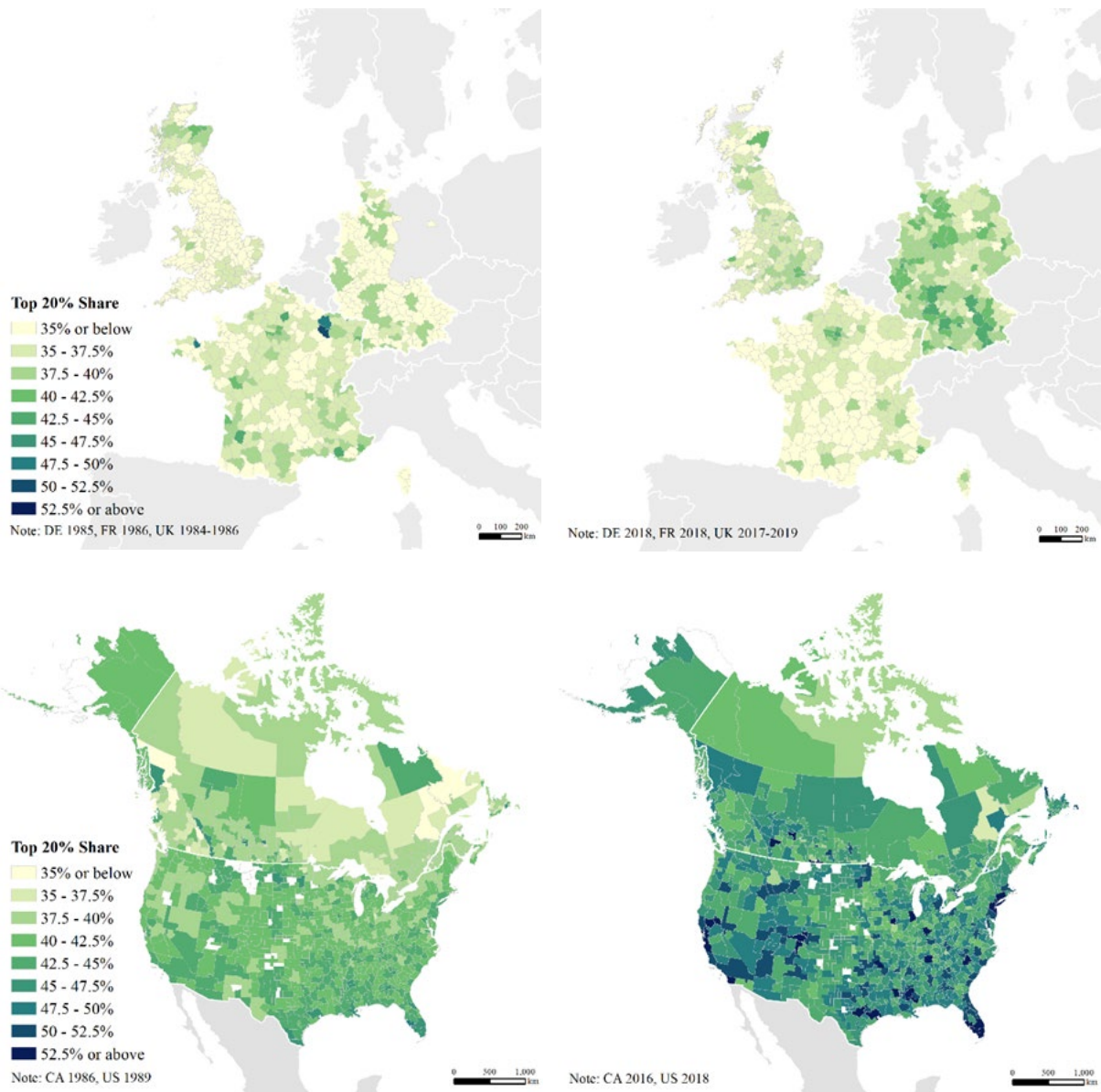
Figure 1: Share of wages of top quintile by local labour market, 1980s v 2010s

Note: Full-time workers, 20 +, weekly/daily earnings. Source: CA: CCP; DE: SIAB; FR: DADS; UK: NES/ASHE

Spatial income inequalities in all our study countries are higher today than they were four decades ago. However, differences in average wages between areas contribute relatively little to national wage inequality. We illustrate this through a counterfactual exercise which recalculates national wage inequality trends after equalising average wages across LLMAs; these counterfactual series of p90/p10 wage inequality are very close to the observed series.

To summarise by country, the US stands out as the most spatially unequal; Canada as having US levels of national wage inequality but European levels of spatial wage inequality; France as a country with very low spatial inequality in low and high wages across LLMAs; the UK in the unusually large

contribution of spatial wage inequality to national wage inequality; and Germany in the very low and stable contribution of spatial inequality to (the relatively high) national wage inequality. The results have important policy implications as they help governments find ways to spread prosperity more evenly across both their population and their territories.

The paper is organised as follows. Section 2 discusses the literature on spatial and national income inequalities. Section 3 presents our data and methodology. Next, Section 4 shows the trends in national wage inequality, and Section 5 shows our cross-country comparable measures of spatial inequalities. In Section 6, we combine these two sets of estimates and assess how important are spatial wage inequalities for the national ones. Finally, in Section 7 we conclude.

2. National, between-area and within-area inequality: the existing literature

This paper links two distinct bodies of literature: studies of national income inequality, including research on the polarization of labour markets and the evolution of top incomes; and the analysis of the spatial distribution of income, for instance, explorations of the divides between prosperous and declining sub-national regions.

The large literature on national income inequality has partially explained the observed increase in inequality since the 1970s by changes in the distribution of labour income. A popular account has used the 'demand and supply of skills' framework to explain changes in wage inequality, stressing the role of skill-biased technological change, globalization, or skill-complementary capital accumulation, among others (e.g., Autor et al. 2003, 2008; Krussel et al. 2003; Goldin and Katz, 2009; Autor et al. 2013; Dauth et al. 2014; Autor 2019; Acemoglu and Restrepo 2022; Machin 2011; Verdugo 2014; Guillot et al. 2020). On the other hand, the research advocating the institutional mechanism for wage inequality development has emphasized the role of non-market factors, such as the minimum wage regulation, unionization, liberalization, or market concentration in product and labour markets (e.g., Lee 1999; Card et al. 2004; Fortin et al. 2012; Egger et al. 2019; Furceri and Loungani 2018; Deb et al. 2022; Aznar et al., 2017, Autor et al. 2019).

However, explanations based only on the developments in wage inequality, such as that of biased technological change, are inadequate to account for the recorded secular trends in *total* income inequality as well as for divergent trajectories across countries.[1] The top incomes literature (Piketty, 2001, 2014; Piketty and Saez, 2003; Atkinson and Piketty, 2007, 2010; Atkinson et al. 2011) has used

---

[1] It has been pointed out that countries at the similar level of technological progress have experienced notably divergent increase in inequality (e.g. continental European countries and Japan vs. Anglo-Saxon countries; Atkinson et al, 2011).

administrative fiscal data to chart the long-run patterns of total income inequality and pointed to the critical role of capital income (itself reflecting changing patterns in the concentration of wealth) in shaping secular inequality trends. More generally, this research has emphasized the importance of broad coverage of all income sources for the analysis of the income distribution,[2] which would allow an appraisal of different mechanisms as determinants of inequality. This has correspondingly led to the development of the Distributional National Accounts (DINA) methodology, covering the entire income distribution and all types of income (Alvaredo et al. 2018; Piketty et al. 2018; Garbinti et al. 2018).

Nevertheless, national income inequality studies provide limited insight into whether growing inequality is an economy-wide phenomenon, equally affecting all regions within a country. This is important, as, for instance, the recent wave of anti-system voting seen in many countries around the world and its strong territorial grounds have drawn greater public attention to geographic disparities within countries. The Economist, for example, stated in 2017 that "regional inequality is proving too politically dangerous to ignore." Back in the mid-2000s, Kanbur and Venables (2005) had already pointed out a strong link between the spatial dimension of inequality and political tensions. In line with this, the seminal paper by Rodríguez-Pose (2018) argues that regions in economic decline with persistent poverty and lack of opportunities, the so-called "places that don't matter" turn to the anti-system vote as an act of revenge. Dijkstra, Poelman and Rodríguez-Pose (2019), who study the "Geography of EU Discontent" by mapping the anti-EU vote in electoral outcomes in the EU-28, support the lagging behind territories' "revenge" hypothesis. They show that industrial decline, coupled with unemployment and a less educated population, are major factors driving the anti-EU sentiment.

Research on regional inequality has gained traction and has demonstrated the rising disparity between sub-national regions.[3] Three main strands of the literature are worth noting here. The first research field analyses convergence and divergence patterns in incomes per capita across sub-national regions and over long periods. The bulk of these studies on geographic inequality is based on GDP per capita, a measure that is generally available across countries and at large sub-national levels and over long periods (NUTS1/TL1 or NUTS2/TL2 level). These studies provide support for a regional Kuznets curve (Kuznets, 1955 and Williamson, 1965) in which spatial inequality as a function of economic development - as countries move from an agricultural to an industrial economy - follows an inverted U shape (Barrios and Strobl 2009). Furthermore, they find that spatial inequality rises again when the level of economic development is high. Thus, these findings suggest an N-shaped curve for spatial inequality that resembles trends in national income distribution: divergence occurring at the first stages

---

[2] For example, Bukowski and Novokmet (2021) point to the importance of social transfers for divergent post-communist inequality trajectories between Poland and Russia.

[3] In the same vein, some sectors disproportionately contribute to national income inequality. Growing wages in the financial sector increasingly play a role in income inequality in the United States, the United Kingdom, and in France (Philippon and Reshef 2012, Bell and Van Reenen 2010, Godechot 2012). Besides, concentration of financial jobs in major financial centers and its interaction with globalization and technological change reinforces the spatial disparities favoured by the latter phenomena (Godechot 2013).

of development, convergence during periods of industrialization, diverging again in the latest phases of development (Lessman 2014; Lessman and Seidel 2017). While this literature has the advantage of mapping mean regional incomes to the national level, it faces two obvious limits: first, only average incomes are analysed, and not inequalities within areas; second, regions are generally defined according to political-administrative boundaries and do not reflect comparable and economically meaningful areas neither within nor across countries (e.g., Ganong and Shoag 2017; Rosés and Wolf 2018; Roses and Wolf 2021).

The second strand of this literature is represented in a large body of works from urban economics and economic geography, utilising administrative micro-level data to analyse the trends, and their sources, in productivity and wages dispersion across areas (e.g., Chetty et al. 2014; Stansbury et al. 2023; Sommeiller and Price 2015; Gaubert et al. 2021; Bonnet, d'Albis and Sotura 2021; Kemeny and Storper 2020; Manduca 2019; Breau and Saillant 2016). Some key stylised facts have been established, including that declining manufacturing production in rich countries has led to rising concentration of national output into fewer, successful local labour markets (Bauluz 2018). One implication of this trend is a decline in urban wage premium for low- and mid-skilled workers (Autor 2019). Consistent with this, the literature has also found large evidence of an assortative matching pattern, namely that high-skilled workers also tend to concentrate in large, highly productive and high-wage locations (see, for example, Card et al. 2021 or Moretti 2012 in the case of the US; de la Roca and Puga 2017 in the case of Spain; Dauth et al. 2018 in the case of Germany).

The third strand focuses on inequalities within local labour markets and cities (e.g., Moretti 2013; Diamond 2016; Faggio et al. 2017; Duranton and Puga 2005; Desmet and Rossi-Hansberg 2009; Venables 2018; Boeri et al. 2019). The evidence from the US shows that inequality is higher in larger cities, which also attract the most talented (Glaeser et al. 2009; Florida and Mellanders 2016), with similar results found for Canada (Bolton and Breau 2012) and the UK (Lee et al 2016; McCann, 2019). In addition, there are a few studies tackling changes in the local income distribution across US and Canadian regions (e.g., Bolton and Breau 2012; Baum-Snow and Pavan 2013; Albouy and Zabek 2016), finding that inequalities grew significantly more in the largest cities, potentially due to agglomeration economies becoming more biased towards the most skilled workers (Baum-Snow et al. 2018).

Yet, comparative cross-country analyses are still lacking, as the main shortcoming of administrative data is that for confidentiality reasons its access is usually restricted and that concepts are not harmonized across countries. As a result, the existing literature is unable to answer two fundamental questions that are the focus of this paper: (i) which countries have higher spatial inequalities and how do trends compare over time?[4] (ii) how has spatial inequality contributed to the rise of national

---

[4] Several studies compared regional inequality across countries using household surveys (e.g., Ezcurra et al. 2007; and Jesuit, et al. 2003; analysed incomes at the NUTS1 level and Ayala et al. 2020 at the TL2 level). The obvious limit of surveys is that by construction, the possibility to produce smaller geographic breakdowns is usually constrained due to small sample sizes –

inequality in the past decades? To the best of our knowledge, the only attempts to carry out a cross-country comparison of spatial inequality using administrative data have been made by the OECD: Boulant et al. (2016) and a forthcoming study by Königs et al. (2023). The former uses micro-aggregated administrative data to compare income inequality within metropolitan regions across 11 countries. Nevertheless, this study faces important limitations that we begin to overcome in this paper: they do not analyse medium and small-size cities and rural areas; they cover short periods; and they do not systematically relate metropolitan to national inequalities. Finally, Königs et al. (2023) provide income inequality measures both across and within relatively small geographic regions (TL3) in 19 OECD countries, using administrative microdata over the early and mid-2000s until 2020. Regarding the contribution of spatial inequality to national inequality, a few country-specific studies have looked at it based on varying definitions of regions, income concepts, time periods and methods. While lacking a unified approach to the question, these studies tend to find a relatively modest role of spatial inequality in overall national-level inequality (see Disslbacher and Mosser 2022 for the US; Gibbons et al., 2014 for the UK, Combes et al. 2008 for France and Briskar et al. 2022 for Italy).

The existing literature on both national and regional inequalities has been unable to provide a holistic view of income disparities and consistently tackle the fundamental questions outlined in the introduction. A cross-country comparative analysis of identically defined local labour markets and sub-national income distributions offers the opportunity to disentangle these difficult interactions between national and local inequalities. In this paper we will proceed by establishing comparable measures of labour income inequality using consistent definitions of income and geographic units for Canada, France, Germany, the United Kingdom and the United States.

3. **Estimating comparable measures of spatial inequality: data, income definition and geographic units**

There are two main methodological challenges when estimating cross-country and cross-time comparable measures of spatial wage inequality. First, the definitions of labour income and of target population differ substantially across data sources, even within the same country. For instance, while part-time workers are part of the sample in the UK for the entire considered periods, in Germany only a subset of them is captured before 1999. On the other hand, Germany has data on annual earnings for

---

and hence they are not representative at granular geographic scales. In general, cross-country data on household incomes and inequality are available for large geographic aggregations and are collected by Eurostat and by the OECD at the NUTS2 and TL2levels.

the entire period of analysis, while in the UK such data is available only since 1999 (before only weekly and hourly wages).

Second, the international comparison must be based on a common definition of geographic areas. One of the major methodological contributions of our paper is to analyse spatial inequality using geographic units that are theoretically and empirically coherent across country contexts. In contrast, comparative analyses of spatial inequality in other studies are frequently conducted using the aggregate data made available by national statistical institutes, typically provided for an unsatisfactory mix of political-administrative areas rather than a theoretically coherent concept.

In the following section we first discuss our main data sources. Next, we explain our baseline definition of income, which is consistent across countries and time. Finally, we provide details on the construction of consistent geographic units, and the way we have conceptualised and measured spatial inequality.

3.1. Data

Given that the growth in national income inequality has been driven in large part by growing earnings inequality (Atkinson et al. 2011), in this paper we focus on spatial inequalities in labour income. We draw the data from the existing from matched employee-employer registries (Germany, France and UK) and census data (Canada and the United States), which provide large representative samples of workers with a good coverage of incomes starting from the 1970s. The data provide a rich set of worker characteristics, including basic demographics, occupations, employment types and industries. Importantly, all the considered data sources provide detailed information on the place of work and/or residence. Appendix 1 provides a more detailed description of the data sources.

The data on German (before 1990 only West German) labour income come from the Sample of Integrated Labour Market Biographies (SIAB), which is a 2% panel of the universe of active population, covering the period 1975-2014. For the entire panel, workplace location is available at the *District* level (401 of them). The French equivalent is *Déclaration Annuelle de Données Sociales* (DADS Panel), which is a 4% sample of the universe of active population between 1976 and 2001, and an 8% sample since 2002. DADS includes detailed information on workplace location at the level of 36,000 *Communes*. The UK survey of employees is the New Earnings Survey/Annual Survey of Hours and Earnings (NES/ASHE), which is a 1% panel of the universe of workers available since 1975. It provides workplace location defined for around 100 Work Areas, and since 1997, for detailed Local Authorities,

Parliamentary Constituencies or Travel to Work Areas. The data for Canada come from the micro-files of the Census of Population, which provide a 20% sample of the population over the period 1986 to 2016 (with the Census collected every 5 years). The data includes place-of-residence data down to the Census Subdivision level. For the United States, we use the decennial Census of Population (CP) and its continuation American Community Survey (ACS) sourced the from Integrated Public Use Microdata Series samples (IPUMS).

The availability of data for Germany means that it is only possible to construct a consistent time series for the area covered by pre-1990 West Germany. Consequently we restrict our analysis to this area, and in the remainder of the paper we use 'Germany' to mean 'West Germany'.

3.2. Income definition

Our baseline definition of income is pre-tax labour income, which refers to the sum of earnings flows going to labour, but before other taxes and transfers and excluding bonuses and non-wage compensations. The unit of observation is the full-time employee aged 20 or above. We use daily (France, Germany) or weekly (Canada, UK, US) earnings as only this temporal aggregation is available for all time periods and countries. As we restrict our analysis to full-time employees our results are insensitive to whether daily or weekly wages are used in the analysis. Full details of the income data for each country is provided in the appendix.

3.3. Geographical units

The existing literature on spatial inequalities uses a mixture of administrative geographies, economic geographies and metropolitan geographies to measure inequality between and within places. A common international geographical system used in comparative work is the OECD Territorial Levels (TL), which has prima facie appeal as a hierarchy that provides a common framework across countries at varying spatial scales from TL1 (the largest geographical units) to TL3 (the smallest). However, the TL system is constructed from existing administrative geographies, the differing size of which makes comparative analysis very challenging. For example, despite their comparable population sizes, Germany is divided into 401 TL3 units compared to 52 TL3 units in Spain, determined by their administrative subdivisions and not their economic geography.

From a theoretical perspective, we would wish to analyse spatial inequality in wages using a geographic unit that approximates the spatial extent of the labour market within which someone could seek work without migrating, namely some form of labour market area. There is a large literature in regional science on the conceptualization of labour market areas, methods for delineation and criteria for evaluating them (Casado-Diaz and Coombes 2011; Fowler and Jensen 2020; Goodman 1970). The two principal approaches to constructing labour market areas differ in whether they start with an urban core and identify suburban areas with a strong connection to this core; or whether they are defined to minimise the commuting flows across area boundaries, that is, they approximate self-contained labour markets. A good example of the first type are the OECD/EU 'Functional Urban Areas', which are "identified as densely populated local units (urban centres) and surrounding local units connected to the urban centres by high travel-to-work flows" (Dijkstra et al. 2019). Examples of the second type include Commuting Zones in the US or Travel-To-Work-Areas in the UK.

For our purposes an advantage of the second type of area is that they cover the whole spatial extent of a country, whereas the first type covers only cities and their commuting hinterlands. As we are interested in decomposing national inequality to subnational inequality it is important for us to have full coverage of the territory, and therefore, we use extant commuting-zone-like areas, which we call local labour market areas or LLMAs. These existing areas - 266 self-contained labour market areas in Canada, 306 zones d'emploi (employment zones) in France, 223 Arbeitsmarktregionen (labour market regions) in Germany, 228 travel-to-work areas in the UK and 741 commuting zones in the US – aim at a similar self-contained labour market area concept, but are constructed using different geographic building blocks and slightly different methods. Nonetheless, we can see from Figure 2 below that our LLMAs have a broadly similar distribution of demand self-containment and population size. A small number of areas in France and Germany have self-containment rates less than 50%, the majority of which are very small LLMAs (<50,000 population). Canada has a relatively large number of small LLMAs in the remote northern regions of the country (as does the US, to a lesser extent). To reduce the effect of these small LLMAs, we weight our measures of geographic inequality between areas by population size.
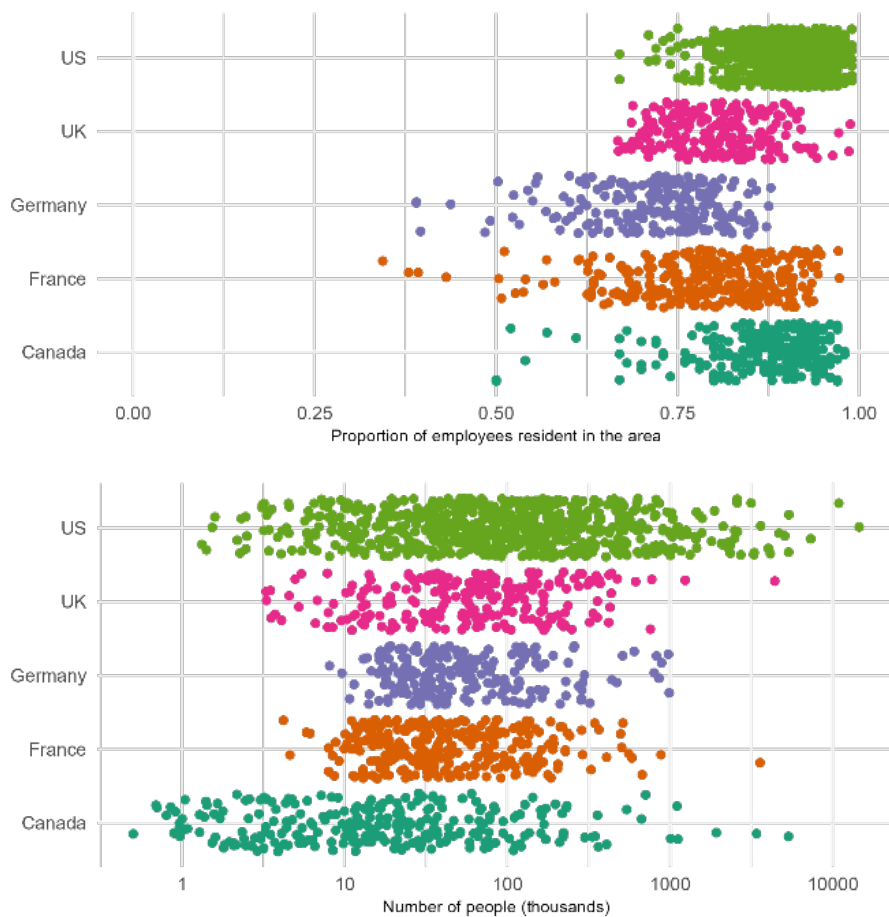
Figure 2: Demand self-containment and population size for local labour market areas

For some countries we have LLMAs that were defined at different points in time (for example 1981, 1991, 2001 and 2011 travel to work areas in the UK), which would in principle allow us to change the geographic areas used over time. To avoid problems of interpretation that would arise from using different geographic areas at different time points, we use a fixed set of LLMAs, the most recent that are available. [5] Depending on the geographical identifiers reported in our datasets, the LLMA in which a particular individual lives or works cannot always be uniquely identified; in these cases, observations are apportioned fractionally to different LLMAs in which they could appear in a manner similar to that used in Autor et al. (2013). Further details for each country are provided in the appendix.

---

[5] The exception is the US, where we follow the work of David Dorn harmonizing the definition of Commuting Zones since the 1970s, and which uses 1990 as its base year.

3.4. Conceptualising and measuring spatial inequality

Having defined a common measure of labour income and a consistent geographic area, there are choices to be made about how to conceptualise and analyse geographic inequalities in labour income. A first choice is whether the focus should be on inequality between areas or inequality within areas. Both have potentially important bearings on social welfare. Inequality in wages between areas means that residents of low-wage areas are potentially systematically disadvantaged in the labour market unless they are willing to bear the social and economic costs of migrating to a different area. These disparities and perceived spatial injustices may have effects on cohesion between different parts of a country – witnessed in regional political cleavages in Spain, the United Kingdom and Italy, for example. Inequality in wages within areas may mean that residents in high-inequality areas have wide disparities in their disposable income, potentially presenting cost barriers for lower earners in accessing housing and other amenities. In brief, a comprehensive view of spatial inequalities demands an analysis of between area *and* within area inequalities.

An analysis of wage inequality between areas amounts to an analysis of the differences in the wage distributions of different labour market areas, and those differences could be characterised by analysing different aspects of the wage distributions. Commonly it is the centre of the distribution that is compared across areas – the mean or the median wage – which provides a measure of the wage prospects of the average worker in an area. We also have to decide whether our unit of analysis should be areas or individual people living and working in those areas. These offer different perspectives. An analysis of inequality trends in the central wage of each area provides a view of whether areas are becoming more or less dispersed in terms of their wages. An analysis of inequality trends in the central local area wage experienced by individual people provides a summary of the between area inequalities experienced by workers. From a local economic development point of view, we might be more interested in the former; from a social welfare point of view, we might be more interested in the latter. In this paper our main concern is with individual social welfare and so we use individual workers as our unit of analysis. An ancillary benefit is that this allows us to account for the varying population sizes of local labour markets (see figure 2), such that large urban labour markets will be weighted more heavily than small rural areas[6].

---

[6] Remote LLMAs with a small population might have a very particular composition of workers, for instance, high-wage oil-workers in Alberta, Canada. These locations represent a tiny fraction of the overall population, but, if unweighted, could seriously distort the levels and changes in the aggregate measures of spatial inequality. Second, estimates of means and percentiles for LLMAs with small population can be noisy due to small sample sizes. We thus put less weight on them in constructing the aggregate measures.

Whilst an analysis of disparities in the centre of the wage distribution experienced by workers is important, it tells only a partial story. The mean wage does not provide information about the experience of lower wage or higher wage workers and whether they are relatively disadvantaged by the labour market in which they live and work. In addition, changes in mean wage inequality could be driven by changes in low wage inequality and/or high wage inequality.

Consequently, to represent the important – and potentially differing - inequalities between areas in the bottom, middle and top of the local distributions, we present measures that summarise the inequalities in the 10th percentile, mean and 90th percentile wages that are paid in each area. Finally we must choose a statistical measure to summarise these inequalities. For our purposes – given the very different levels of wages between different countries and over time within countries – it is important that our measure should be mean-independent. There are a number of established candidates for this including the variance or standard deviation of logs, the mean logarithmic deviation and the coefficient of variation. As working with log wages is standard in the labour economics literature we work with the variance of log wages; we could just as easily present the standard deviation of logs but this monotonic transformation would not affect the substantive interpretation of our findings. As our unit of analysis is the individual worker, in the calculation of the variance of log area wages we weight each area by the number of workers.

Measuring within area inequalities is important to understand the differences between areas in the local wage inequality experienced by workers. Once areas have been defined, measuring inequality within areas presents the same theoretical and empirical challenges as measuring national inequality. One approach is to focus on the width of the distribution; alternatively, as inequality is typically driven by the top of the distribution we could focus on inequality between the top and the rest. We adopt complementary measures that address both of these, namely the p90/p10 ratio and the share of wages earned by the top 20% of the distribution.

Finally, we take a holistic view by comparing the scale of between and within area inequalities through two different strategies. First, we decompose the variance of the national log wage distribution into the variance between areas and the variance within areas i.e. an analysis of variance approach. Second, we imagine a counterfactual scenario in which there was no between area inequality but within area inequality is preserved at its observed level; we recalculate national wage inequality under this counterfactual to observe how much inequality changes if between area inequality is removed. The methods for achieving this are explained below. Both approaches are aimed at understanding the relative contribution of between area and within area inequality to national wage inequality.

15

4. Setting the scene: national wage inequality

Before we present our measures of spatial wage inequality, we set the scene by describing how national wage inequalities have evolved over our study period. We focus on the percentile ratios of the pre-tax weekly earnings for adult full-time workers.
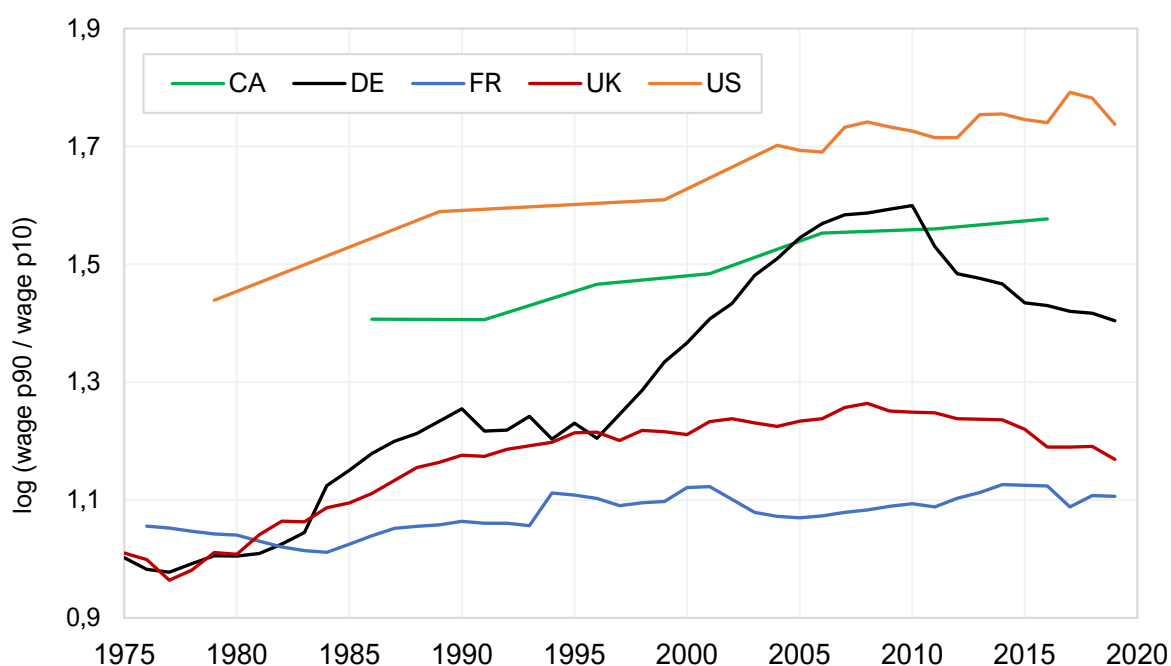

Figure 3: National wage inequality, log(p90/p10) 1975-2019

Note: Full-time workers, 20 +, weekly/daily earnings. Source: CA: CCP; DE: SIAB; FR:DADS;UK: NES/ASHE

The five countries have had different levels of and trends in wage inequality over the last four decades. Figure 3 shows the log(p90/p10) ratio for Canada, Germany, France, the UK and the US over the period 1975 to 2019. We use the additively decomposable log(p90/p10) so that we can examine in Figure 4 below whether these changes are driven by the top or the bottom of the distribution. The log(p90/p10) in our chart ranges from around 1 (p90/p10 ≈ 2.7) through 1.4 (p90/p10 ≈ 4) to around 1.8 (p90/p10 ≈ 6). The United States has had the highest national wage inequality throughout, and the level of inequality rose substantially during the study period. Canada is observed for a shorter window but follows a very similar trend to the US at a lower level of inequality, albeit higher than all the European countries for most of the period. The three European countries had very similar levels of national wage inequality to each other in the 1970s and all experienced an increase in inequality during the 1980s and early 1990s, although the increase in France was lower than the other two countries. Up to around 1997, the UK and Germany had similar levels of wage inequality, but during the 2000s the German p90/p10

ratio rose much higher, whereas it was relatively stagnant in the UK. Since the early 2010s, wage inequality in Germany has fallen but it remains higher than France or the UK.

In summary we reproduce a well-established fact (e.g., Piketty 2021; Guvenen et al. 2022) that national disparities grew most strongly between 1980 and the financial crisis of 2008, with evidence of stagnation or decline thereafter. In UK there is a steady increase in the p90/p10 ratio from the start of the period until the mid-1990s, after which there is a small decline after the Great Recession of 2008-09. France had a similar but much smaller rise in inequality up to the mid-1990s with little change since then. In Germany there is also a rise in national inequality that is sharpest between 1997 and 2010 and declines thereafter. National wage inequality rose in Canada and the United States throughout the period but the increases during the 2010s were relatively modest.
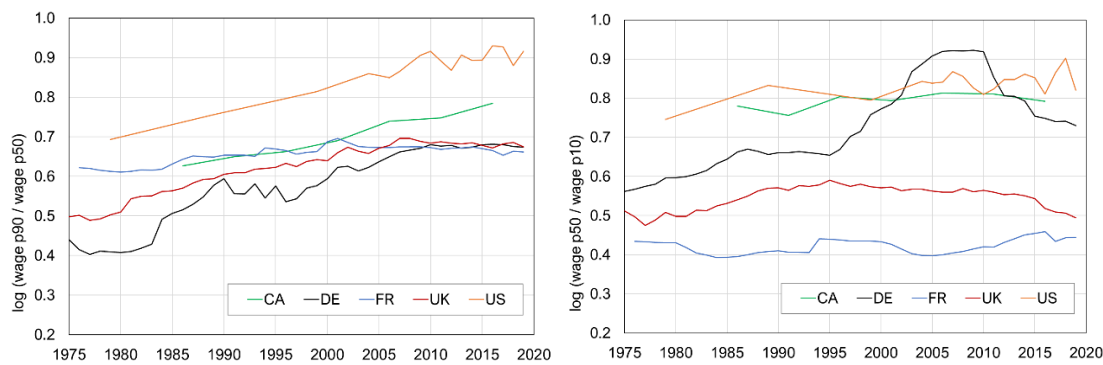

Figure 4: National wage inequality, log(p90/p50) (LHS) and log(p50/p10) (RHS) 1975-2019
Note: Full-time workers, 20 +, weekly/daily earnings. Source: CA: CCP; DE: SIAB; FR: DADS;UK: NES/ASHE

The countries differ with respect to whether the documented rise of inequality was driven by the increased dispersion of wages at the top or bottom half of the distributions. Figure 4 shows the decomposition of the log(p90/p10) ratio for each country into the log(p90/p50) and the log(p50/p10). In our figure, the log(p90/50) and the log(p50/p10) have a similar range, from around 0.4 (px/py ≈ 1.5) through 0.7 (px/py ≈ 2.0) to a maximum of around 0.9 (px/py ≈ 2.5). In the US and Canada, we observe secular rises in the p90/p50 ratio that are sustained over the bulk of the study period. At the start of the period, inequality in the top of the distribution was quite different in France, Germany and the UK but the level of inequality converged over the study period and has been very similar and stable since 2009. The p90/p50 has been stable in France since the mid-1990s after a short rise in the 1970s and 1980s.

Germany is a clear outlier in the inequality trends in the bottom half of the distribution. We observe two ten-year periods of increase in inequality in the mid-1970s to mid-1980s and again in the mid-1990s to the mid-2000s (for an analysis of these trends, see, for instance, Dustmann et al., 2009 or Card et al.,

2013). After the Great Recession, we observe a sharp decline towards the 1990s levels. By contrast, there is relatively little change in the p50/p10 ratio in our other countries. Canada and the US have very similar and stable levels of p50/p10 ≈ 2.2 throughout the period. The p50/p10 in France has fluctuated around a stable level over the period. In the UK, there is a clear inverse U shape, with inequality in the bottom half of the distribution rising up to the late 1990s, thereafter falling back to levels last seen in the 1970s.

From this decomposition analysis, we find that the differences in p90/p10 wage inequality between the three European countries are caused by differences in inequality in the bottom half of the distribution. By contrast, in North America, the US has a more unequal wage distribution than Canada due to the wider range of wages at the top of the distribution. Having set out some of the trends in national wage inequality over the period, we can investigate how trends in spatial wage inequality have evolved before drawing on measures that explicitly link the evolution of national and spatial wage inequality together.

5. Spatial wage inequalities

5.1. Between-area wage inequality

Now we present measures that investigate how wage inequality between local labour market areas has evolved over the study periods in our five chosen countries. We focus on two sets of measures, which capture the extent to which local labour market areas are similar in terms of wages and how these areas converged or diverged through time. All statistics shown in this section are weighted by the LLMA population, such that we place lower weight on places with smaller populations.

In most countries trends in spatial differences in the average wage across places mimic the national trends in the p90/p10 ratio. Figure 5 presents the variance of the log mean local labour market (weighted by population) as a measure of dispersion of the mean wages by area. Just as in the national inequality trends, we observe secular rises in wage inequality in Canada and the US throughout the period. The two separate decade-long rises in inequality we observe in Germany's p90/p10 ratio are also evident in concurrent rises in inequality between the mean wages of LLMAs. In the UK, we observe the same strong rise in inequality in the 1970s, 1980s and 1990s, with some convergence in mean wages since then. In France the trend is different from the national inequality trend. There was a modest rise in the variance of log mean wages during the 1970s and 1980s, peaking in the mid-1990s - mirroring the

national inequality trend – but there has since been a fall in spatial mean wage inequality back to 1970s levels.

In terms of levels, the main surprise is that Canadian spatial inequalities are relatively low throughout the period despite the relatively high levels of national inequality. This suggests that more of the wage inequality is within LLMAs than between them in Canada, compared to our other countries - which we will explore further below. By the end of the period, the European countries and Canada have very similar levels of spatial inequality in mean wages by this measure. The US, by contrast, has experienced the largest increase in spatial inequality in mean wages over the period and, by this measure, is by far the most unequal of the five countries by the end of the 2010s.
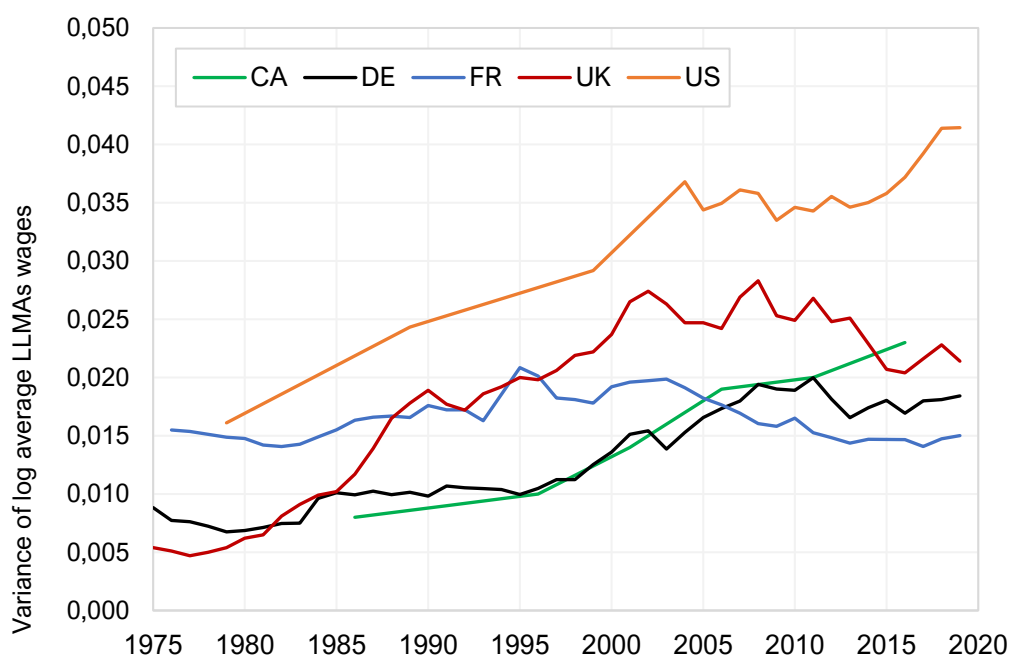

Figure 5: LLMA wage inequality, variance of log mean wage 1975-2019

Note: Full-time workers, 20 +, weekly/daily earnings. Source: CA: CCP; DE: SIAB; FR: DADS; UK: NES/ASHE

The trends in inequality between areas in the average wages mask interesting heterogeneity depending on which parts of the within-area wage distribution we consider. We illustrate this by looking at the dispersion across LLMAs of different percentiles of the log wages, similar to Gaubert *et al.* (2021).[7] More specifically, in Figure 6, we report the variance of the log of the 90th percentile and the log of the 10th percentile from each local distribution of wages. These series can be interpreted as measures of how similar high wages (p90) and low wages (p10) across labour markets are. The left-hand panel focuses on the variance of the log(p90). The first message arising from Figure 6 is an increase in the

---

[7] Gaubert et al. 2021 present the standard deviation of log income percentiles across states in the US as a measure of dispersion of percentiles across sub-regions, whereas we report the variance.

dispersion of top incomes (p90) across labour markets, with the only exception of France. At the same time, the dispersion of the bottom incomes has either decreased or remained relatively stable. These results are in line with Gaubert et al. (2021), who identify a similar spatial pattern across states in the US, which they describe as a concentration of affluence and democratization of poverty.

Although this series is noisier than the series based on the mean, we can see that in Canada, Germany, UK and US there is a rise in inequality between areas in the level of higher wages that are paid. In Canada and US, this trend continues through most of the study period, whereas in Germany and the UK, the between-area p90 inequality peaked in the early 2000s (other than noisy fluctuations) and has been stable since. In other words, in these countries, the highest wages paid across labour market areas have become increasingly dissimilar over the period; yet in Canada, although a striking rise is observed, the level of inequality between areas is relatively low. France once again shows an entirely different pattern – any trend is very unclear, with perhaps even a slight decline. The decoupling between the three European countries is very striking. In 1980 the three countries had very similar levels of dispersion in p90 wages across their local labour market areas; by 2015 they were markedly different, with Germany and the United Kingdom having a much higher degree of spatial inequality in high wages than France.
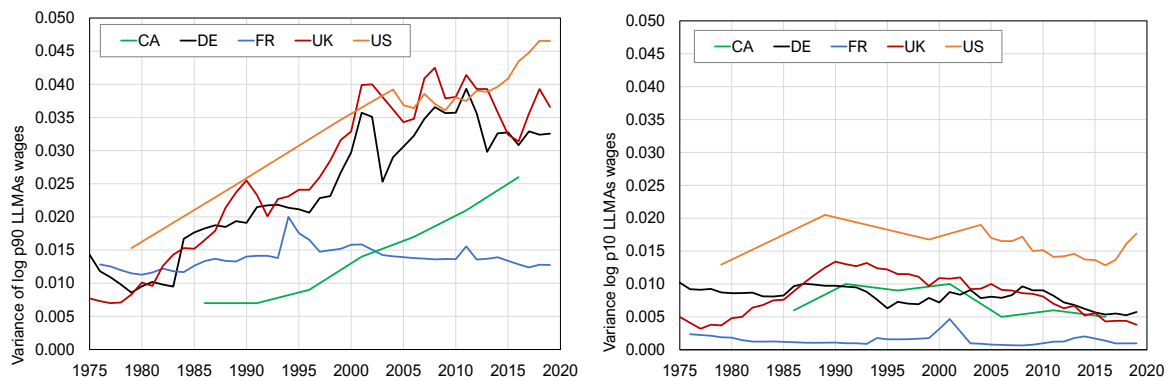
Figure 6: LLMA wage inequality, variance of log p90 (LHS) and log p10 wage (RHS) 1975-2019

Note: Full-time workers, 20 +, weekly/daily earnings. Source: CA: CCP; DE: SIAB; FR: DADS; UK: NES/ASHE

The right-hand panel shows the variance of the log(p10) wage i.e., the dispersion between LLMAs in the wages paid at the bottom of the distribution. The first observation is that there is much less contemporary dispersion in low wages than in high wages, as might be expected. However, the dispersion in p10 wages was of a similar magnitude to the dispersion in p90 wages in all countries at the start of the period; it is the growth in the dispersion of top wages between areas since the 1970s that

has resulted in very different levels of dispersion at the top and bottom. The United States has the highest dispersion between areas in the lowest wages paid throughout the period and France has the lowest. The dispersion in low wages in France is remarkably close to zero, showing that it has markedly little variation in p10 wages across its local labour market areas. The United Kingdom shows an interesting inverse U-shaped pattern, with divergence in p10 wages across LLMAs in the 1970s and 1980s, followed by convergence back to 1970s levels by the end of the 2010s. We also observe a convergence in p10 LLMA wages in Germany since the late 2010s.

4.2 Within-area inequality

We turn now to an analysis of how wage inequality within LLMAs has changed over time. Figure 7 shows, for each LLMA in each country at two time points: the 10th percentile wage, the mean wage and the 90th percentile wage. These charts build upon a similar visualisation from Overman and Xu (2022) for the UK, overlaying these distributions in the 1980s (in purple) with the distributions in the 2010s (in green). The exact time points used in each country vary according to data availability; 10 years of data had to be used in the UK to satisfy statistical disclosure rules.[8] The LLMAs are ranked according to mean wage at that time point (so the same LLMA may appear at different points on the chart for different dates). Wages are measured in the local currency unit (LCU, e.g. GBP for UK) deflated to 2015 values.

The charts show that, for all our countries, there is greater dispersion in wages within LLMAs in the 2010s compared to the 1980s, something which will be a significant contributor to the overall growth in national wage inequality. These presages work we do in section 6 below, in which we partition the growth in national wage inequality into between-area and within-area inequality. Some other features serve to re-illustrate patterns we have noted earlier, though we should note that the charts in Figure 7 do not take into account the differing population sizes of these LLMAs, unlike the charts in section 4.1 which are population-weighted. In most places we can see that the variation in mean wages - illustrated on the chart as the slope of the mean wage line - has increased between the 1980s and 2010s, with the notable exception of France. Likewise, we can observe the increased dispersion of areas in terms of the p90 wages over this period, again with France showing limited change. On the other hand, the dispersion

---

[8] Because of the disclosure rules we are also not able to report the percentiles for 31 (35) TTWAs in the period 2010-2019 (1980-1989).

of p10 wages across areas is much lower than the dispersion of p90 wages, and the changes we can observe visually are much more limited.
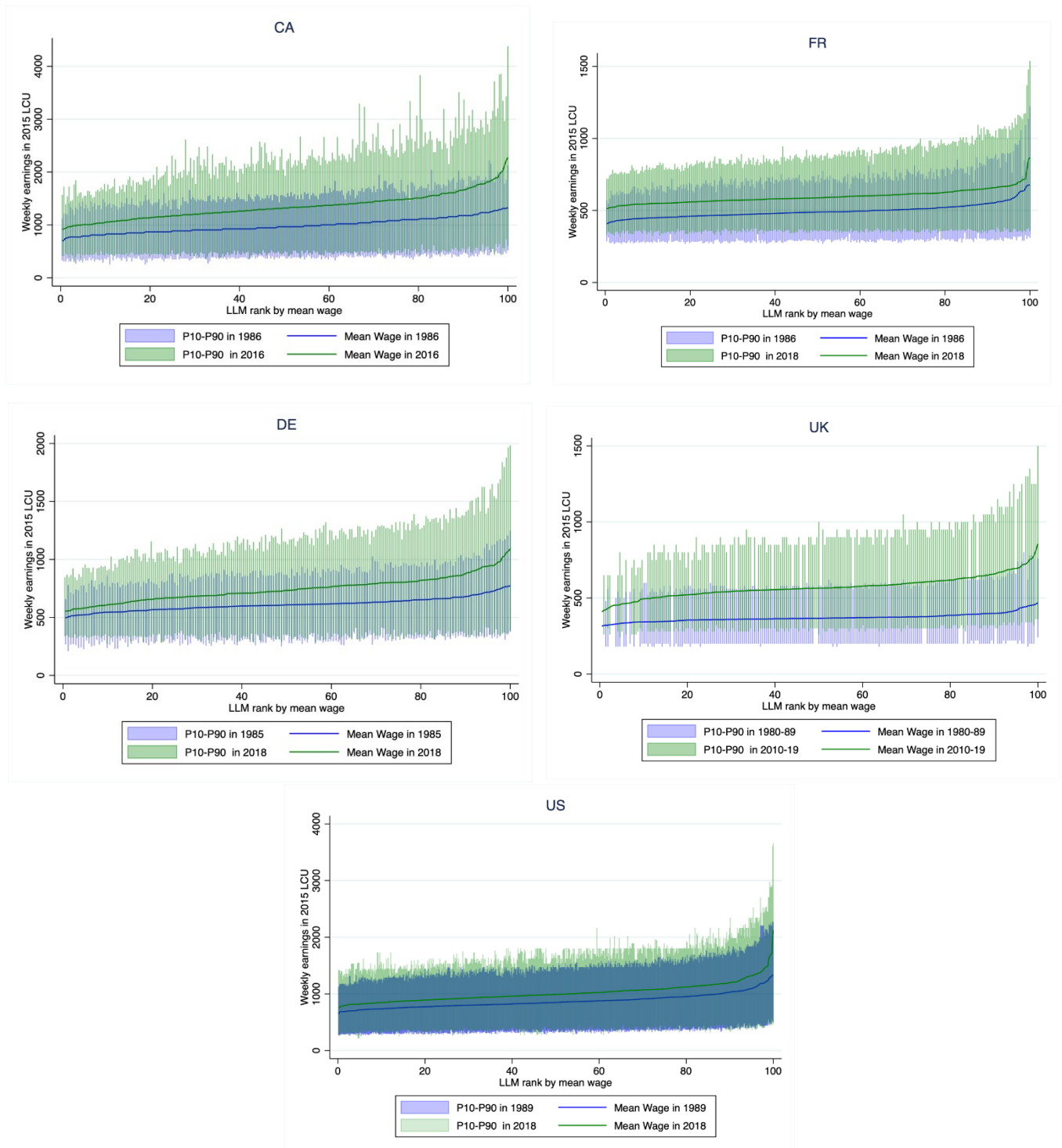

Figure 7: within-LLMA wage inequality in 1980s and 2010s: p10, mean and p90 wages by LLMAs ranked by mean wage

Note: Full-time workers, 20 +, weekly/daily earnings. Source: CA: CCP; DE: SIAB; FR: DADS; UK: NES/ASHE

Our final visualisation of the variations in between-area inequality uses the share of total LLMA wages that are paid to the top 20% of earners in that LLMA. In Figure 1, we visualise this on maps showing Europe and North America separately, noting that different scales are used (Europe would be relatively much smaller than North America if equal scales were used), and also that large areas on these maps (for example, Northern Canada) have very small populations. These maps show two things that we already know: that North America has higher national wage inequality than Europe, and that in most countries wage inequality has increased over time. We also see the striking rise in wage inequality across much of Germany, while in contrast, many areas in France have experienced falls in the top 20% share over the period. In North America, we can see a concentration of the most unequal areas along the West and East coasts.

6. Linking national and regional inequality

How important are spatial inequalities for national trends? Our final set of inequality measures aims at linking measures of national and spatial wage inequality, and in different ways, seeks to answer the question of how much national wage inequality is due to geographic wage inequality. Our first measure is the raw variance share (RVS; see Gibbons et al. 2014), which is the R-squared value from the following regression:

$$ln\ wage_{ir} = \alpha + \mu_r + \varepsilon_{ir}$$

where i denotes individual and r the local labour market area. μ_r denote LLMA fixed effects. We estimate this model each year and country separately to obtain time- and country- variant R-squared, which represents the percentage of variance in log wages that is accounted for by the variance of mean log wages between areas. It is a simple decomposition of the variance into the variance of log wages within areas (the residual variance from the regression) and the variance of log wages between areas.

We show that the overall contribution of mean LLMA wages to the total variance of wages is relatively small, explaining 2-8% of national wage variance over the period. Figure 8 below shows the trends in the raw variance share in our four countries over the study period. We observe some distinctive patterns that seem initially counterintuitive when compared to the measure of spatial inequality in mean wages presented in Figure 4. Most striking, whereas the US is the most spatially unequal country in our study measured by the variance of log mean LLMA wages, the raw variance share looks similar to most other countries in Figure 8. Conversely, the UK has similar levels of spatial inequality to the other three non-US countries but the 2010s RVS is the highest at around 7% compared to 3-5% for the other four

countries. To interpret these figures, it is important to remember that the RVS is a measure of inequality in mean wages between local labour market areas, *relative to total national wage inequality*. Although the United Kingdom had a lower degree of spatial inequality in mean wages compared to US in the 2010s (Figure 4), it also had a lower degree of national wage inequality as measured by the log(P90/P10) ratio (Figure 2). The US has a high degree of spatial inequality in mean wages but also a higher degree of national wage inequality. With the higher degree of national wage inequality in the denominator, the importance of dispersion in mean wages in the US is lower than that in the UK. In any case, these numbers are small compared to the total wage inequality that is 'explained' by within-area inequality; in the UK in the 2010s, around 7% of national wage inequality is accounted for by inequalities between areas in mean wages, leaving 93% to be accounted for by within-area inequality.
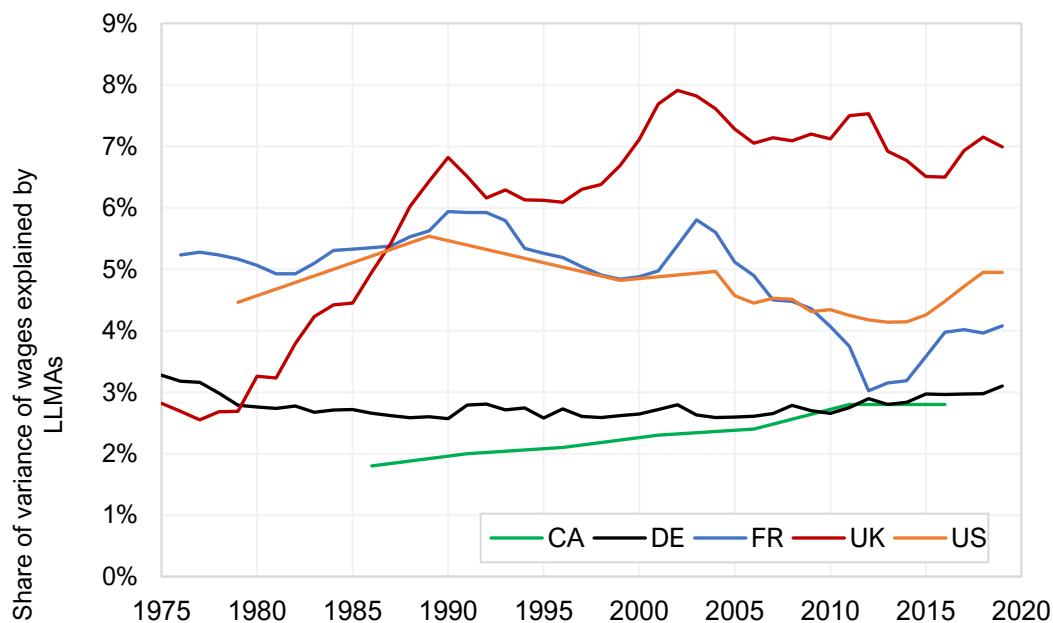
Figure 8: Share of variance of wages explained by LLMAs 1975-2019
Note: Full-time workers, 20 +, weekly/daily earnings. Source: CA: CCP; DE: SIAB; FR: DADS; UK: NES/ASHE

The trends over this period in the raw variance share are also of interest. In Germany, the share of log wage variance between local labour market areas has remained relatively constant and comparatively low over the period at around 3%, despite the large rise in spatial inequality described in Figure 4. Whilst there has been an increase in spatial wage inequality in Germany this has been matched by an increase in wage inequality within areas, such that the contribution of place to national wage inequality has been unchanged. A similar pattern is observed in Canada, where the importance of place has increased, but by a small magnitude. Throughout the period, France has a higher RVS than Germany, as it has a slightly higher degree of spatial inequality but a lower degree of national wage inequality. The series is rather volatile but there is evidence of a decline in the France raw variance share. Similarly, in the US, there is evidence of a decline in the raw variance share over the period, suggesting that

within-area inequality became slightly more important than between-area inequality over the period, though clearly, there have been increases in both. Overall, the relative importance of the between-area inequality in mean wages in explaining national wage inequality is very similar in France and the US, despite the very different levels of inequality they exhibit.

The UK is the only country that demonstrates a significant change in the importance of mean local labour market wages in explaining total national wage inequality over the period. In the 1970s, the raw variance share was very similar to Germany, but this rose sharply from 1980 to a peak in the early 2000s, since when it has been above all the other countries. Looking back to our charts of national wage inequality (Figure 2) and variance in log mean wages across local labour market areas (Figure 4) and comparing the UK to Germany, we can observe the root of this pattern. In the mid-1970s, Germany and UK had almost identical levels of national wage inequality and spatial wage inequality (as measured by the variance of LLMA log mean wages), and hence the importance of spatial mean wage dispersion was very similar. By the end of the 2010s, Germany had higher national wage inequality than the UK; whereas the UK had higher spatial wage inequality than Germany. Consequently, spatial wage inequality has become much more important relative to national wage inequality in the UK compared to Germany.
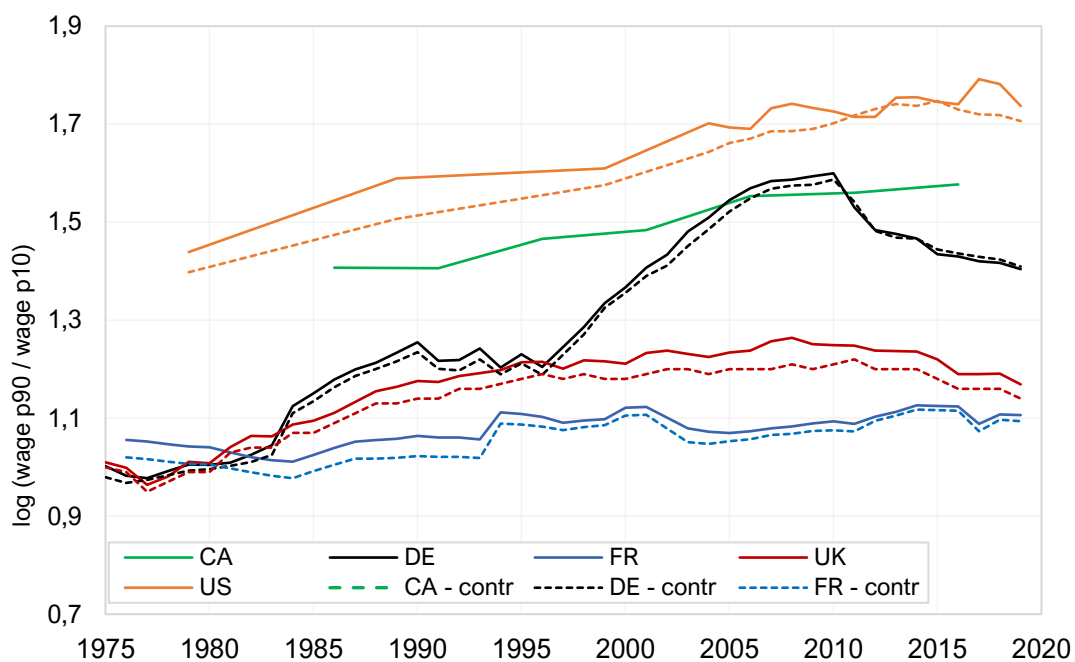

Figure 9: Counterfactual national log(P90/p10) ratio, 1975-2019
Note: Full-time workers, 20 +, weekly/daily earnings. Source: CA: CCP; DE: SIAB; FR: DADS; UK: NES/ASHE

The implication of the low estimates of the raw variance share is that spatial disparities in mean wages do not contribute a significant amount to national inequalities. We illustrate this through a comparison of the national log(p90/p10) wage inequality series (presented in Figure 2) to a counterfactual series in which there is no difference in average wages across LLMAs. We equalise the average wage across places by multiplying each individual living in area r, by the following factor:

$$p_r = \overline{wage} \Big/ \overline{wage_r}$$

where $\overline{wage}$ denotes the average wage, and $\overline{wage_r}$ the average wage in area r. We do that separately for each country and year and estimate measures of national inequality. This way, we preserve the within - LLMA inequality, but remove differences in mean income across the LLMAs. A policy equivalent of this experiment would be to increase the average income tax in high-wage places and decrease it in low-wage places, such that there is no spatial difference in average post-tax wages.

The observed and counterfactual series for the national log(p90/p10) ratio are shown in Figure 9, and they provide a similar conclusion to the analysis of raw variance share. We already know from the RVS measures that in Canada and Germany spatial inequality in mean area wages makes a small contribution to explaining overall national wage inequality. We can see this in Figure 9, where the counterfactual series for Canada is identical to the observed series, and in Germany, there is only a very small difference between the observed and counterfactual series. We can see that France has a large difference between the counterfactual and observed series at the start of the period, mirroring its large raw variance share in this period, and that this difference becomes smaller in the 2010s. The United Kingdom displays the opposite pattern, with a small difference between the series observed at the start of the period that grows larger than France towards the end. In the US, the difference between the counterfactual and observed series declines somewhat over the period.

There are clear differences in the extent to which differences between the mean wages of local labour markets explain national wage inequality, both between countries and within countries over time. However, this counterfactual series – and the analysis of raw variance share – emphasizes that wage inequalities within areas are much more important than wage inequalities between areas in driving overall national wage inequality. If the counterfactual series made a significant difference to national wage inequality, then we might expect the counterfactual 'effect size' to be comparable to the observed difference between countries; for example, making the counterfactual UK series closer to the observed

France series than the observed UK series. In Figure 7 above, we can see that all of the counterfactual series are relatively close to the observed series for the respective country.

7. Discussion and Conclusions

In conclusion we can state some stylized facts about the comparative levels and trends in spatial wage inequality in Canada, France, (West) Germany, the United Kingdom and the United States from the mid-1970s to 2019. In summary, we find that by the end of the 2010s, spatial inequalities in LLMA mean wages are very similar in Canada, France, Germany and the UK; the United States is by far the most unequal by this measure. All countries experienced an increase in spatial inequality over the period, though the degree of increase varies considerably from a limited amount in France to a rapid increase in the US. In all countries except France there is a strong trend of increased dispersion in wages paid at the top of the distribution between LLMAs, but for most countries some convergence in the lowest wages paid across areas. By country, the US stands out as the most spatially unequal; Canada as having US levels of national wage inequality but European levels of spatial wage inequality; France as a country with very low spatial inequality in low and high wages across LLMAs; the UK in the unusually large contribution of spatial wage inequality to higher national wage inequality; and Germany in the very low and stable contribution that between-area inequality in mean wages makes to national wage inequality.

We find that the United States is the most unequal country throughout the study period in terms of the dispersion of means wage across local labour market areas (LLMAs). In common with most other countries, it has experienced a strong rise in this geographic wage inequality and, measured by the variance of the log mean wage, it was by far the most unequal country by the end of the period. This parallels its rise in national wage inequality as measured by the p90/p10 ratio, which has had a secular rise over the period and is also much higher than all other countries. Despite this, the contribution of between-area inequality in wages has remained reasonably constant over time at approximately 5%, similar for most of the period to France (which has much lower spatial and national wage inequality). This is because both between area and within area wage inequality have risen to similar degrees so that their relative contributions to total inequality have remained relatively unchanged. The US also has the largest dispersion in top wages at the 90th percentile, but it is rather less remarkable in this respect, with inequality between areas in top wages being similar in level and trend to Germany and the UK. It is again quite remarkable in terms of the dispersion between areas in the lowest wages, which are much lower than the dispersion in high wages but much higher than the other four comparator countries.

Canada bears some similarities to its North American neighbour. It has levels and trends of national wage inequality that are similar to, though lower than, those in the United States. As in the US, the rise in national wage inequality has been driven by a rise in inequality at the top (p90/p50) of the distribution. However the level of inequality in mean wages between LLMAs is relatively low, although it has risen quite strongly over the period to be very similar to the spatial inequality observed in all three European countries by the 2010s. As, put crudely, Canada is a place with US levels of national wage inequality but European levels of spatial wage inequality, it stands out as a place where geographic wage inequality makes a particularly small contribution to national wage inequality, though it has increased by a small magnitude over the period. The inequality between LLMAs in top (p90) wages has risen sharply over the period but is still less unequal by this measure than all other countries except France. In contrast, there is evidence of a decline in the dispersion of p10 LLMA wages over the period, and Canada is very similar to the European countries on this measure. We find that, today, Canada is the country with the largest spatial wage inequalities across our four study countries, and there has been an increase in the dispersion of mean wages across local labour market areas over the period. Despite the relatively large between area inequalities we observe in Canada, they are relatively small as a proportion of total national wage inequality. Inequality between areas at the top end of the wage distribution – the between area dispersion of p90 wages - has grown very significantly over the study period.

Moving to Europe, since the mid-1980s France has had the lowest p90/p10 wage inequality, characterised by relatively low inequality in the bottom half of the wage distribution, and has experienced a much smaller increase in inequality by this measure. In the 1970s France had similar levels of dispersion in mean LLMA wages to the United States, but it has experienced only a small increase since then, such that it was much more equal than the US in the 2010s and at a similar level to Canada, Germany and the UK. It has lower levels of both national wage inequality and spatial wage inequality than the US but consequently spatial wage inequality accounts for a similar proportion of national wage inequality (~5%). France is quite remarkable when it comes to inequality in the bottom and top wages paid across LLMAs. In all other countries the dispersion of p90 wages across LLMAs has risen very strongly over much of the period, but in France it has actually fallen since the mid-1990s and by the 2010s was by far the lowest. This implies that LLMA wage distributions are becoming more similar to each other in the top wages paid. In general the dispersion of top wages is much larger than the dispersion of bottom wages, but the dispersion of p90 LLMA wages in France is actually lower in 2019 than the dispersion of p10 LLMA wages in the US. The dispersion of low (p10) wages across French LLMAs has been very low throughout much of the period, tending towards zero during the latter part of the 2010s.

The UK had a distinctive increase in the dispersion of mean LLMA wages between the late 1970s and the early 2000s, mirroring the increase in national wage inequality during this period. The level of dispersion in mean wages at the end of the study period was very similar to Canada, France and Germany. What makes the UK stand out is the sharp rise in the contribution of mean LLMA wage dispersion to total national wage inequality during the 1980s and 1990s. In the 1970s between-LLMA wage inequality accounted for around 3% of national wage inequality, similar to Germany but below France and the United States. In the 2000s and 2010s this stabilised at around 7%, far above the four other countries. The UK is unusual in the degree to which spatial wage inequality has contributed to higher national wage inequality. Alongside this, the dispersion in top LLMA wages has risen very strongly, to a similar degree observed in the US, while there has been a notable compression in p10 wages across LLMAs beginning in the 1990s and accelerating since.

Finally, Germany also experienced an increase in spatial disparities in mean LLMA wages over the period, with the strongest increase between 1995 and 2008, since when it has levelled off. This period of increase in geographic inequality occurred at the same time as the rapid increase in national wage inequality that was driven by increased inequality in the bottom of the wage distribution. There has been limited change in the dispersion of LLMA p10 wages however, with evidence of a compression of low wages across LLMAs in the 2010s. By contrast, and like most other countries, there was a large increase in p90 LLMA wage dispersion throughout much of the period up to the Great Recession. The contribution of between-area inequality in mean wages to national wage inequality has remained low and stable with a raw variance share of around 3% throughout the study period.

This paper considerably extends the previous scope of comparative spatial wage inequality studies. Our analysis demonstrates the importance of looking beyond differences in mean wages when analysing spatial inequality, and that there can be quite different things going on at different points in the distribution. In the United Kingdom, for example, previous literature has concluded that there was a considerable growth in spatial inequality in local mean wages during the 1980s and 1990s, with little change thereafter – mirroring the story that the real 'action' on income inequality in the UK took place under the Thatcher government of the 1980s. Using a comparative framework, our analysis confirms this conclusion but also shows that there are interesting patterns occurring at other points in the distribution. The compression of the local 10th percentile wages across local labour markets suggests that low wage workers are paid increasingly similar wages wherever they work in the UK. Consequently there may be limited incentives for people working in low wage sectors to seek work in other areas; if the lowest wages vary little, there is limited incentive to move to a high cost city to seek work. In France there has been consistently low levels of 10th percentile wage dispersion, and in Germany there is

evidence of a recent decline. The US has much higher levels of 10th percentile wage dispersion than other countries, though there is evidence of some decline in more recent periods.

At the other end of the wage distribution, we observe in all countries but France a strong trend throughout the period of increased dispersion of local 90th percentile wages. In the UK and Germany this trend is stronger and more sustained than the increased dispersion of mean wages. This means that areas are becoming increasingly dissimilar at the top end of the wage distribution; and that people earning the highest wages are becoming increasingly concentrated in certain local labour market areas. Whilst part of the story behind the increasing salience of geographic inequality (in the UK, for example) may be the persistence of high levels of mean wage inequality, it could also be explained by a sustained divergence at the top end of the distribution.

There are many theorised and empirically proven determinants of spatial inequalities within nations including first and second nature geography, technological change, agglomeration economies, globalisation and demographics. Many of these economic fundamentals have affected our study countries at similar times, yet our analysis shows that they have quite different patterns of spatial wage inequality. The degree of increase in spatial inequality varies considerably across nations and has occurred at different times in different countries; and similar trends in wage inequality at the local mean can be contrasted with quite different patterns in wage inequality at the bottom and the top of local wage distributions. Whilst the economic and geographic fundamentals really matter, there is clearly a large role for national institutional responses in shaping how economic trends drive spatial inequality. We might expect, for example, that national minimum wage policies have a role to play in the convergence of 10th percentile wages across local labour market areas, something we will investigate in later stages of this project.

Another clear conclusion from this paper is that differences in mean wages between local labour market areas are a relatively small contributor to total national wage inequality in all five countries. None of our counterfactual simulations where local mean wages are equalised make a significant difference to national wage inequality as measured by the ratio of 90th percentile to 10th percentile wages. If reducing national wage inequality is an important policy goal, focusing on reducing between-area wage inequality across local labour market areas will not make a significant difference. Within-area inequality is a much larger contributor to national wage inequality, and in countries where we observe a convergence in local 10th percentile wages but divergence in local 90th percentile wages, within-area inequality is itself becoming more dispersed, with some local labour markets becoming more unequal than others. These raise key economic and political questions about what goals it is desirable to pursue,

and what relationship different types of spatial inequality have with political outcomes, internal migration, social cohesion and wellbeing.

Although our analysis provides a rich first analysis of spatial wage inequalities in these five countries, it does only provide a partial picture of spatial economic inequalities. First, our analysis of inequality has largely rested upon the 90th percentile and the 10th percentile wages, and the ratio between them. We know from other work that income inequality is strongly driven by inequality in the top 1% and above, so our future analyses should investigate this as far as possible. Second, in this paper we have relied upon an analysis of daily or weekly earnings for full-time workers, so we are really analysing differences in the price of labour. This is an interesting approach from a labour economics point of view, and is a valuable contribution. For other audiences and purposes we are also interested in moving towards an analysis of welfare insofar as it can be measured by the labour market income of individuals. To analyse geographic inequalities in economic welfare from this perspective requires us to take into account the earnings of part time workers and periods of short and long term unemployment. Future work will therefore focus on an analysis of annual wages and unemployment of all workers. Third, and has been alluded to in our discussion of the convergence of 10th percentile wages, taking full account of the welfare that derives from labour market income across space requires an analysis of the differential cost of living, particularly arising from the cost of housing. Finally a full analysis of spatial economic inequalities demand analysis of total income from which households can draw, especially given the importance of business income and self-employment income in driving trends in inequality. Is to these tasks that we will now turn.

References

Acemoglu, D., & Restrepo, P. (2022). Tasks, automation, and the rise in US wage inequality. *Econometrica*, 90(5), 1973-2016.

Albouy, D., & Zabek, M. (2016). Housing Inequality. NBER Working Paper No 21916.

Alvaredo, F., Atkinson, A., Chancel, L., Piketty, T., Saez, E., & Zucman, G. (2016). Distributional National Accounts (DINA) guidelines: Concepts and methods used in WID.world. *WID.world WP, 2*.

Alvaredo, F., Chancel, L., Piketty, T., Saez, E., & Zucman, G. (Eds.). (2018). *World inequality report 2018*. Belknap Press.

Atkinson, A. B. (2005). Top Incomes in the UK over the 20th Century. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 168 (2), 325–343.

Atkinson, A. B., & Bourguignon, F. (Eds.). (2014). *Handbook of income distribution* (Vol. 2). Elsevier.

Atkinson, A. B. and Piketty, T., (2007). Top Incomes Over the Twentieth Century: A Contrast Between Continental European and English-Speaking Countries. Oxford: Oxford University Press.

Atkinson, A.B., Piketty, T., (2010) Top incomes: a global perspective. Oxford University Press, Oxford

Atkinson, A. B., Piketty, T., & Saez, E. (2011). Top incomes in the long run of history. *Journal of Economic Literature*, *49*(1), 3-71.

Autor, D. H. (2015). Why are there still so many jobs? The history and future of workplace automation. J*ournal of Economic Perspectives*, 29(3), 3-30

Autor, D. H. (2019). Work of the Past, Work of the Future. *AEA Papers and Proceedings* (Vol. 109, pp. 1-32). American Economic Association.

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(4), 1279-1333.

Autor, D. H., & Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the US labor market. *American Economic Review*, 103(5), 1553-1597

Autor, D., Dorn, D., & Hanson, G. H. (2013). The China syndrome: Local labor market effects of import competition in the United States. *American Economic Review*, 103(6), 2121-68.

Ayala, L., Martín-Román, J., & Vicente, J. (2020). The contribution of the spatial dimension to inequality: A counterfactual analysis for OECD countries. *Papers in Regional Science*, 99(3), 447-477.

Barrios, S., & Strobl, E. (2009). The dynamics of regional inequalities. *Regional Science and Urban Economics*, 39(5), 575-591.

Bauluz, L. E. (2018). Three Essays on the Long-term Accumulation of Wealth (Doctoral dissertation, Université Paris sciences et lettres).

Baum-Snow, N., Freedman, M., & Pavan, R. (2018). Why has urban inequality increased?. *American Economic Journal: Applied Economics*, 10(4), 1-42.

Baum-Snow, N., & Pavan, R. (2013). Inequality and city size. *Review of Economics and Statistics*, 95(5), 1535-1548.

Bell, B., & Van Reenen, J. (2010). Bankers' pay and extreme wage inequality in the UK.

Boeri, T., Ichino, A., Moretti, E., and Posch, J. (2019). Wage equalization and regional misallocation: evidence from Italian and German provinces, IZA Discussion Papers, No. 12279

Bolton, K., & Breau, S. (2012). Growing unequal? Changes in the distribution of earnings across Canadian cities. *Urban Studies*, 49(6), 1377-1396.

Bonnet, F., H. d'Albis and A. Sotura (2021), Income Inequality across French Departments over the Last 100 Years, *Economie et Statistique / Economics and Statistics* 526-527, pp. 49-69,

Boulant, J., Brezzi, M., & Veneri, P. (2016). Income Levels and Inequality in Metropolitan Areas. OECD Regional Development Working Papers 2016/06.

Breau, S. and R. Saillant (2016), Regional income disparities in Canada: exploring the geographical dimensions of an old debate, *Regional Studies, Regional Science*, Vol. 3/1, pp. 463-481

Briskar, J., Di Porto, E., Rodriguez Mora, J.V. & Tealdi, C. (2022) The Geographical Decomposition of Italian Inequality, *WorkINPS Papers*

Bukowski, P., & Novokmet, F. (2021). Between Communism and Capitalism: Long-Term Inequality in Poland, 1892-2015. *Journal of Economic Growth*, 26, 187–239.

Card, D., Heining, J., & Kline, P. (2013). Workplace heterogeneity and the rise of West German wage inequality. *The Quarterly journal of economics*, 128(3), 967-1015.

Card, D., Lemieux, T., & Riddell, W. C. (2004). Unions and wage inequality. *Journal of Labor Research*, 25(4), 519-559.

Card, D., Rothstein, J., & Yi, M., 2021. "Location, Location, Location," Working Papers 21-32, Center for Economic Studies, U.S. Census Bureau.

Casado-Díaz, J. M., and M. Coombes. (2011). "The Delineation of 21st Century Local Labour Market Areas: A Critical Review and a Research Agenda." *Boletín de la Asociación de Geógrafos Españoles* 57, 7– 32 & 393–396.

Chetty, R. et al. (2014), Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States, T*he Quarterly Journal of Economics*, Vol. 129/4, pp. 1553-1623,

Combes, P. P., Duranton, G., & Gobillon, L. (2008). Spatial wage disparities: Sorting matters!. *Journal of Urban Economics*, 63(2), 723-742.

Dauth, W., Findeisen, S., and 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(6), 1643-1675.

Dauth, W., Fuchs, M., & Otto, A. (2018). Long-run processes of geographical concentration and dispersion: Evidence from Germany. *Papers in Regional Science*, 97(3), 569-593.

de la Roca, J., & Puga, D. (2017). Learning by working in big cities. *The Review of Economic Studies*, 84(1), 106-142.

Deb, S., Eeckhout, J., Patel, A., & Warren, L. (2022). Market Power and Wage Inequality. US Census Bureau, Center for Economic Studies.

Desmet, K., & Rossi-Hansberg, E. (2009). Spatial growth and industry age. *Journal of Economic Theory*, 144(6), 2477-2502.

Diamond, R. (2016). The determinants and welfare implications of US workers' diverging location choices by skill: 1980-2000. *American Economic Review*, 106(3), 479-524.

Dijkstra, L., Poelman, H. & Veneri, P. (2019). The EU-OECD definition of functional urban area. *European Commmission, Department for Regional and Urban Policy*.

Dissblacher, F. and Moser, M. (2022). Regional, Spatial and Nationwide Income Inequality in the United States: Evidence from Novel Neighbourhood Inequality Data. *mimeo*

Duranton, G., & Puga, D. (2005). From sectoral to functional urban specialisation. *Journal of Urban Economics*, 57(2), 343-370.

Dustmann, C., Ludsteck, J., & Schönberg, U. (2009). Revisiting the German wage structure. *The Quarterly journal of economics*, *124*(2), 843-881.

Egger, P. H., Nigai, S., & Strecker, N. M. (2019). The taxing deed of globalization. *American Economic Review*, 109(2), 353-90.

Ezcurra, R., P. Pascual and M. Rapún (2007), The Spatial Distribution of Income Inequality in the European Union, *Environment and Planning A: Economy and Space*, Vol. 39/4, pp. 869-890,

Faggio, G., Silva, O., & Strange, W. C. (2017). Heterogeneous agglomeration. *Review of Economics and Statistics*, 99(1), 80-94.

Florida, R., & Mellander, C. (2016). The geography of inequality: Difference and determinants of wage and income inequality across US metros. *Regional Studies*, 50(1), 79-92.

Fortin, N., Green, D. A., Lemieux, T., Milligan, K., & Riddell, W. C. (2012). Canadian inequality: Recent developments and policy options. *Canadian Public Policy*, 38(2), 121-145.

Fowler, C. S., & Jensen, L. (2020). Bridging the gap between geographic concept and the data we have: The case of labor markets in the USA. *Environment and Planning A: Economy and Space*, 52(7), 1395–1414.

Furceri, D., & Loungani, P. (2018). The distributional effects of capital account liberalization. *Journal of Development Economics*, 130, 127-144.

Gaubert, C., Kline, P., Vergara, D., & Yagan, D. (2021). Trends in US spatial inequality: Concentrating affluence and a democratization of poverty. *AEA Papers and proceedings* (Vol. 111, pp. 520-525). American Economic Association.

Ganong, P., & Shoag, D. (2017). Why has regional income convergence in the US declined*? Journal of Urban Economics*, 102, 76-90.

Garbinti, B., Goupille-Lebret, J., & Piketty, T. (2018). Income inequality in France, 1900–2014: evidence from distributional national accounts (DINA). *Journal of Public Economics*, 162, 63-77.

Gibbons, S., Overman, H. G., & Pelkonen, P. (2014). Area Disparities in Britain: Understanding the Contribution of People vs. Place Through Variance Decompositions. *Oxford Bulletin of Economics and Statistics*, 76(5), 745–763. https://doi.org/10.1111/OBES.12043

Glaeser, E. L., Resseger, M., & Tobio, K. (2009). Inequality in cities. *Journal of Regional Science*, 49(4), 617-646.

Godechot, O. (2012). Is finance responsible for the rise in wage inequality in France?. *Socio-Economic Review*, 10(3), 447-470.

Godechot, O. (2013). Financiarisation et fractures socio-spatiales. *L'Année sociologique*, 63, 17-50.

Goldin, C. D., & Katz, L. F. (2009). *The race between education and technology*. Harvard University Press.

Goodman, J. (1970). "The Definition and Analysis of Local Labour Markets: Some Empirical Problems." *British Journal of Industrial Relations* 8, 179– 96.

Bozio, A., Breda, T., & Guillot, M. (2020). The Contribution of Payroll Taxation to Wage Inequality in France. halshs-01294599v2.

Guvenen, F., Pistaferri, L. and Violante, G.L. (2022), Global trends in income inequality and income dynamics: New insights from GRID. *Quantitative Economics*, 13: 1321-1360.

Jesuit, D., Rainwater, L., & Smeeding, T. (2003). 15. Regional poverty within the rich countries. *In Inequality, welfare and poverty: theory and measurement* (Vol. 9, pp. 345-377). Emerald Group Publishing Limited.

Kanbur, R., & Venables, A. J. (Eds.). (2005). Spatial inequality and development. OUP Oxford.

Kemeny, T., & Storper, M. (2020). Superstar cities and left-behind places: disruptive innovation, labor demand, and interregional inequality.

Kuznets, S., (1955). Economic growth and income inequality. *American Economic Review*, 45 (1), 1–28.

Kuznets, S., & Jenks, E. (1953). Shares of upper income groups in income and savings. *National Bureau of Economic Research*.

Königs, S, Vindics, A., Diaz Ramirez, M. & Veneri, P. (forthcoming). The geography of income inequalities in OECD countries - Evidence from national register data. OECD Social Employment and Migration Working Papers. OECD Publishing

Lee, D. S. (1999). Wage inequality in the United States during the 1980s: Rising dispersion or falling minimum wage? *The Quarterly Journal of Economics*, 114(3), 977-1023.

Lee, N., Sissons, P., & Jones, K. (2016). The geography of wage inequality in British cities. *Regional Studies*, 50(10), 1714-1727.

Lessmann, C. (2014). Spatial inequality and development—Is there an inverted-U relationship?. *Journal of Development Economics*, 106, 35-51.

Lessmann, C., & Seidel, A. (2017). Regional inequality, convergence, and its determinants-A view from outer space. *European Economic Review*, 92 , 110-132.

McCann, P. (2019). Perceptions of regional inequality and the geography of discontent: insights from the UK. *Regional Studies*, 1-12.

Machin, S. (2011). Changes in UK wage inequality over the last forty years. *The labour market in winter: The state of working Britain*, 155-169.

Manduca, R. (2019). The Contribution of National Income Inequality to Regional Economic Divergence. *Social Forces* 93, 622–648.

Moretti E (2012). The new geography of jobs. *Houghton Mifflin Harcourt*, Boston.

Moretti, E. (2013). Real wage inequality. *American Economic Journal: Applied Economics*, *5*(1), 65-103.

Munro, A., Alasia, A., Bollman, R. D., & Statistics Canada. (2011). Self-Contained Labour Areas: A Proposed Delineation and Classification by Degree of Rurality. Data and Definitions. *Statistics Canada*, 8, 2008008.

Novokmet, F., Piketty, T., & Zucman, G. (2018). From Soviets to oligarchs: inequality and property in Russia 1905-2016. *The Journal of Economic Inequality*, 16(2), 189-223.

Philippon, T., & Reshef, A. (2012). Wages and human capital in the US finance industry: 1909–2006. *The Quarterly Journal of Economics*, 127(4), 1551-1609.

Piketty, T. (2001). Les hauts revenus en France au XXème siècle. Grasset.

Piketty, T. (2014). Capital in the 21st Century. Harvard: Harvard University Press.

Piketty, T., & Saez, E. (2003). Income inequality in the United States, 1913–1998. *The Quarterly Journal of Economics*, 118(1), 1-41.

Rodríguez-Pose, A. (2018). The revenge of the places that don't matter (and what to do about it). *Cambridge Journal of Regions, Economy and Society*, 11(1), 189-209.

Rosés, J., & Wolf, N. (2018). Regional economic development in Europe, 1900-2010: a description of the patterns. LSE Economic History Working Papers No: 278/2018.

Rosés, J. R., & Wolf, N. (2021). Regional growth and inequality in the long-run: Europe, 1900–2015. *Oxford Review of Economic Policy*, 37(1), 17-48.

Sommeiller, E., & Price, M. (2015). The increasingly unequal states of America: Income inequality by state, 1917 to 2012.

Stansbury, A., Turner, D., & Balls, E. (2023). Tackling the UK's regional economic inequality: Binding constraints and avenues for policy intervention. *M-RCBG Associate Working Paper Series.*

Storper M (2018) Separate Worlds? Explaining the current wave of regional economic polarization*. Journal of Economic Geography* 18: 247-270.

Venables, A. J. (2018). Globalisation and urban polarization. *CEPR Discussion Paper Series* DP12877.

Verdugo, G. (2014). The great compression of the French wage structure, 1969–2008. *Labour Economics*, 28, 131-144.

Williamson, J., 1965. Regional inequality and the process of national development. *Economic Development and Cultural Change* 14, 3–45.

Appendix: data details

General information: The best data in Canada for the geographical analysis of weekly wages and incomes are the micro-files of the Census of Population, which provide a 20% sample of the population. Census data are provided every five years, specifically in years ending with '1' or '6.' The most recent dataset available is for the year 2016.

Population: Census data are collected in the spring (during the month of May) every five years and the target population consists of all Canadian residents (i.e., Canadian citizens, permanent residents, residents living in Indian reserves, non-permanent residents with work/student permits, and those with refugee status). Within the Census, information is gathered on individuals that are employed (both for full-time and part-time workers), unemployed as well as not in the labour force. As such, it does not restrict individuals by their age. Not included are Canadian citizens living in other countries or individuals living in collective housing such as hospitals, prisons or detention centres.

Collection: Starting in 1971, the Census adopted self-enumeration methods, where individuals and households complete the surveys by self-reporting their socio-demographic and economic information. The Census program includes mandatory short-form and long-form questionnaires. The entire population is required to complete the short-form questionnaire, whereas random samples of residents are selected to answer the long-form questionnaire. As of 2006, the long-form questionnaire gave respondents the option of allowing Statistics Canada to obtain income-based information directly from individuals' tax files. This new method of collection, whereby tax files are used for income data instead of self-reported incomes, was fully implemented in the 2016 Census. It is important to note that these wage and income figures are neither top-coded nor bottom-coded, and are reported based on the year prior to each Census (for instance, in the case of the 2016 Census, the income refers to that reported in 2015). Information for weekly wages and incomes are imputed using the *annual wage (or income)* and the *number of weeks worked during the reference year* information. Note also that in 2011, the Census long-form questionnaire was exceptionally replaced with the National Household Survey (NHS), a voluntary and self-reported survey

Geographic units: The regional unit of analysis in Canada is the self-contained labour market area (SLA). SLAs are territorial units that are delineated based on commuting flows. As such, they present an alternative to more traditional definitions of regions based on administrative boundaries (such as Census Divisions). While SLAs are still relatively new in the Canadian context, having been introduced just over 10 years ago (see Munro et al., 2011), they are adopted for the purposes of this study with 266 SLAs defined consistently across the country over the 1986 to 2016 period. To standardize the

boundaries of SLAs over time, we use the place-of-residence information reported at the Census subdivision level. Recall that a Census subdivision is the general term used by Statistics Canada for municipalities across provinces and territories. The Census also includes place-of-work information for individuals aged 15 or above, which provides details on an individual's usual workplace (down to the census subdivision level) and whether they work at home or in other countries.

Variables: Finally, in addition to detailed place-of-residence and place-of-work information, the Census also provides a rich collection of information on socio-demographic characteristics (e.g., age, gender, education), housing and economic circumstances such as labour status, occupation, industry, and employment types (e.g. self-employment, full-time/part-time). The data also collects labour market status information, such as whether individuals participate in labour activities, are self-employed in incorporated/unincorporated firms, are not in the labour force, or are unemployed. We use this information to identify individuals working more than 30 hours or more during the reference week (the first week of May) as full-time workers, whereas individuals working less than 30 hours during the reference week are classified as part-time workers.

France

General information: The best source providing a very precise image of wages and allowing to assess a detailed geographical wage distribution in France is *Déclaration Annuelle de Données Sociales* (DADS), which are different administrative files containing social contributions or payroll tax data in France. Alongside annual net and gross wages (which are not top coded), daily wages and (after 1993) hourly wages, a rich set of worker characteristics is available including detailed information on workplace and residential location at the level of 36,000 *Communes*. The sample structure (panel or cross-section) and population coverage (universe or sample and including or excluding civil servants) depends on the year of the DADS database or "file". In this paper, we rely on a representative sample, the "Panel DADS", which provides the longest series.

Population: First, the long series come from the "Panel DADS" and unlike the exhaustive database (DADS postes), this dataset allows following individuals over time and across jobs. This dataset is available from 1976 to 2019 and is a 4% sample (or 1/24th) of the universe of the private sector wage earners between 1976 and 2001 and an 8% sample (or 1/12th) since 2002, where individuals born in October every two years before 2002 are included and on a yearly basis thereafter. The unit of observation is the triplet individual-firm-year which are identified by individual and firm unique identifiers (nninouv and siren, respectively and where the former is anonymized while the latter is public). For the sake of comparability across time, civil servants are excluded from the analysis, as local civil servants, mail and hospital workers gradually entered the panel in the 1980's and then national

civil servants in 2008. We only keep individuals that are employees of public enterprises, who are in the panel since 1976. Managers are not included in the files when they are not employees as well.

Unemployed people only started to be included in the data in 2008. Additionally, even after 2008, unemployment in the DADS only accounts for those individuals who receive an unemployment benefit, which means that only unemployed individuals having had a "working period" – giving them the right to receive the unemployment insurance, appear in the files. Hence, it does not account for those who no longer receive it but are still unemployed, such as the long-term unemployed.

Depending on the focus of the analysis, the database can be aggregated at the level of the establishment, the level of the firm and the level of the individual. The file contains each individual's main job during the year, called the individual's "principal job", as well as all its other "annex jobs". Thus, the unit of observation is the job, a dyad linking an individual and a firm identifier. Hence, if a person has more than one job over the year, the person will appear more than once a year with all the different contractual relations with different firms.

Collection: The collection and statistical use is made by the national institute of statistics, the Institut national de la statistique et des études économiques (INSEE). The administrative declaration procedure of social data is annual and compulsory for all businesses that employ staff, which serves both fiscal and social administrative purposes.

Geographic units: In France, information is available at the most detailed geographic level, the *commune* or municipality. There are 65,200 municipalities and their boundaries have been quasi-fixed over time as their frontier have not changed since the French revolution. This allows aggregating geographic units which are comparable over time. We focus on Employment zones (Zones d'emplois), which are based on commuting patterns. They decompose France in 306 zones. An employment zone is a geographic zone within which most of the working population reside and work, and in which establishments find the bulk of the workforce filling the jobs offered. The 2020 definition has been harmonized with Eurostat procedures.

Variables: In the "Panel DADS", the data provides individual's and job's information such as gender, year and place of birth, whether the job pertains to the public or private sector, the occupation or socio-professional category (both 2-digit CS and 4-digit PCS-ESE classifications), permanent or temporary contract, full or part-time job status, the job start date, the number of days the individual worked in the firm during the year, total number of hours worked (since 1993) and its geographical (commune) place of residency and working place, whether the job is a non-annex job or not, gross and net wages, and in-kind benefits. Net wages are net of payroll taxes and gross of income tax. While the gross wage is the

net wage plus all employee payroll taxes. The dataset also provides information about the employing firm (and establishment within the firm) such as its unique identifier, the number employees, as well as the complete structure of employment from both firm and establishment, and naturally, the number of establishments within the firm. The wage measure that we use in this paper is the annual gross definition aggregated per individual across all jobs (annex and principal), which we divide by the number of days worked during the year in order to get a daily wage.

Known issues: Four important changes in the data production took place in 1993, 2002, 2009 and 2016; this may have caused some breaks in the series. Additionally, the following years also suffered from treatment issues - coming from different sources depending on the year: 1994, 2003; 2004 and 2005. Finally, due to a congestion in the statistical office, the DADS were untreated – and therefore there is no data, in the following years; 1981, 1983 and 1990. Likewise, data from the public sector –for those already integrated in the panel- is absent in 1979, 1981 and 1987.

Germany

General information: The data source used for regional wage analysis in Germany is the *Sample of Integrated Labour Market Biographies* (SIAB). This is a 2% random sample drawn from all individuals that have been registered in the German social insurance system at least once between 1975 and 2019. Due to the historical division of the country, SIAB only covers West Germany (including West Berlin) up until 1992. The administrative nature of the data entails highly accurate information on wages. Employers are required to report the annual earnings of their employees to accurately calculate social security contributions (and face strict penalties if they fail to comply). The downside is that wages in the dataset are top-coded due to existing assessment ceilings for all types of social security contributions. SIAB provides daily wages of working spells by dividing the reported annual earnings by the number of calendar days spent in employment.[9] The reported daily wage refers to the employee's gross wage, i.e., before any social security contributions or taxes paid by employees.

Population: SIAB is a 2% random sample of the *Integrated Employment Biographies* (IEB) representative both at the national and regional levels. IEB covers the population of full- and part-time workers subject to social security for the whole period from 1975 to 2019. However, marginal part-time workers whose earnings and hours worked are below certain thresholds are not (generally) subject to social security and therefore not covered before 1999. Thereafter, marginal part-time workers are fully

---

[9] An employment spell always runs from the start of the job until the end of the job or until the end of a year, depending on what happens first. This means that multiple employments within a year imply several employment separate spells in that year (while the start date in subsequent years is always January 1st). In contrast, unemployment spells can span multiple years from start to end. Further splitting of spells is done by IAB to facilitate dealing with parallel spells (e.g., multiple jobs at the same time).

covered as a result of a regulatory change. IEB covers unemployed individuals to a certain degree: recipients of benefits under *Social Code Book III* (SGB III) which concerns time-limited unemployment insurance for eligible individuals are covered for the entire period. In contrast, recipients of benefits under *Social Code Book II* (SGB II) which present basic security benefits for jobseekers are only covered from 2005 onwards.

Collection: IEB (and SIAB) are collected by the Institute for Employment Research (IAB). They create a data set of spell data that is accurate and exact to the day by merging various sources. For all employed individuals the data comes from the *Employee History* (BeH) whereas the data for unemployed individuals comes from the *Benefit Recipient History* (LeH) and the *Unemployment Benefit II Recipient History* (LHG). The BeH is a history of the notification procedure where all employers are required to submit notifications about their employees at least once a year to social security agencies. LeH and LHG are the histories of benefit recipients from the Federal Employment Agency (BA) and individual municipalities. SIAB provides are 2 percent random sample of the population where each recorded unemployment spell of the selected individuals is an observation.

Geographic units: SIAB contains time-consistent information on the workplace location on the level of districts (*kreisfreie Städe und Landkreise*). As of 2019, there are 401 districts in Germany that can be mapped to 223 Local Labor Markets (LLMs) following the definition of the *Bundesinstitut für Bau-, Stadt- und Raumforschung* (BBSR). For unemployment spells, the living location is available at the same level since 1999.

Variables: SIAB provides information on basic demographics (e.g., age, gender), education, detailed earnings (e.g., daily wages, benefits), employment type, occupation, and industry.

Known issues: The data is provided at the employment spell level, which requires aggregation of spells to obtain one observation per individual per year. While this is not an issue for the aggregation of wages, the location has to be selected when the workplace (or living) location changes during a year. We impute the top-coded wages by following the two-step procedure proposed by Dauth and Eppelsheimer (2020). In the first step, observations are clustered by year, education, and location. For each cluster, wages are predicted with a Tobit regression controlling for individual characteristics. In the second step, the regressions are repeated while including "leave-one-out" means calculated at the worker (per year) and the plant level as control variables. Coverage of marginal part-time and unemployment spells is incomplete and needs to be imputed with external sources (e.g., Microcensus). A change in the employment notification procedure in 1984 led to a significant increase in the share of wages above the upper contribution limit compared to the preceding period. Previously, one-time payments were not part of the reported annual earnings subject to social security contributions. The location of unemployment spells before 1999 needs to be imputed. A sensible assumption is to use the location of the last previous employment spell.

United Kingdom

General information: The UK survey of employees is the New Earnings Survey/Annual Survey of Hours and Earnings (NES/ASHE), which is a 1% panel of the universe of workers and available since 1975.

Collection: The NES/ASHE data are collected annually using an employer questionnaire and not taken directly from the administrative sources. The sample is 1% of the universe of workers subject to the Pay As You Earn (PAYE) system for collecting income tax. This means that employees working a small number of hours in one job will not be covered. For example in 2023 workers earning less than £123 per week are not covered by PAYE, equivalent to 12 hours at the national minimum wage. The sampling of workers is based on the last two digits of their National Insurance Number (i.e., all workers with the same last two digits are included). Since the digits are the same every year and it is possible to link individuals across the waves, NES/ASHE is effectively a panel.

Population: The target population are working individuals older than 16 years old and working in Great Britain (England, Scotland and Wales), and who are registered for the PAYE system. In practice, the target population of NES/ASHE covers well all workers whose annual labour contract exceeded time-variants thresholds,[10] but undercovers those with income below that threshold. In addition, employees of companies with a very low turnover are also not included, but they mostly overlap with those with contracted income below the threshold. As detailed below, the data coverage after 1996 is corrected using supplementary samples and weighting. The issue, however, remains for the first two decades of the data. Besides the under-coverage of low earners, a lower response rate for workers from high-wage occupations might also lead to a certain under-coverage of the top earners. Again, after 1996 weights can be used to account for this in the estimates.

Geographic units: We conduct our spatial analysis for the 2011 Census version of Travel to Work Areas (TTWA), which are approximations to self-contained labour markets. For ASHE (1997 onwards) we observe the postcode at the place of work so we derive the 2011 TTWA from the ONS Postcode Directory. For NES (prior to 1997) we only observe the 'Area' of place of work, administrative areas which do not directly map to 2011 Travel to Work Areas. Consequently we use the 1998 ASHE data – in which Areas and TTWAs are both observed – to calculate the probability that someone working in Area X was working in TTWA 1 or 2 (by calculating the proportion of observations in Area X that fall in TTWA 1 and 2). We then use this probability to weight the NES observations according to the probability that they worked in each TTWA, using the observed information on which Area they worked

---

[10] Our understanding is that at the time of sampling, a person must have a job contract with annual income exceeding the threshold to be registered for PAYE and constitute the target population. People who lost their jobs after the time of sampling, might be part of the sample and have the actual annual income below the threshold.

in. A similar method has been used in constructing estimates for commuting zones in the United States (Autor and Dorn 2013).

Variables: The data include information on wages and paid hours of work, allowing construction of various temporal definitions of labour income: hourly wages, weekly and annual earnings (available only since 1996). This is complemented with information about a type of job contract and a number of jobs. Pension arrangements are included as well. In terms of personal characteristics, basic variables, such as age, gender, occupation, education, are included for the entire time-window (1975-2020). For the same period basic firm-level variables, including industry, size, and revenue, can be obtained by linking workers with the BRS firm registries. Since 2003, it is, however, also possible to match the data with two rich firm-level datasets (ARD, ABS) providing detailed information on firm productivity, profits, and structure.

Known issues: A significant change in the methodology took place in 2004, when the old survey NES was replaced by the new ASHE. The ASHE methodology was applied retrospectively back to 1997. Although the sampling and methodology are the same across these two surveys, there are four important differences from the perspective of estimating spatial income inequalities: 1) Due to the differences in the employee questionnaire, the NES data has more variation than the ASHE. It is unclear, however, how this affects measures of variance; 2) As signalled above, supplementary samples, weighting[11] and imputation[12] were introduced in ASHE, which improves the coverage of the data. The impact of these is to widen certain measures of gender wage gaps, and disproportionally increases wages in London. We might therefore expect to observe a structural break in our measures of between and within inequalities around 1997; 3) In NES, workers who changed job between the sample selection and the survey collection are dropped from the sample, in ASHE these people are followed.; 4) NES has less precise and ad-hoc geographic areas, ASHE contains the NES definition and several others including Local Authorities, Parliamentary Constituencies, Travel to Work Areas and postcode.

United States of America

General information: The US data available since the 1970s are the decennial Census of Population (CP) and its continuation American Community Survey (ACS) sourced from the Integrated Public Use Microdata Series samples (IPUMS).

---

[11] The weighting takes into consideration that ASHE excludes people who are not registered for PAYE and corrects for lower response rate, which is more problematic among high earners. The latter usually implies that the weighted average wages should increase, compared to the unweighted estimates. This might have some spatial implication, as people in London and South East have lower response rate on average.

[12] The ASHE methodology introduces donor imputation of missing items. An important implication is that a significant number of high-wage worker, who were classified as part-time in NES are re-classified to full-time in ASHE. Consequently, if we focus on full-time workers only, we might observe an increase in the right rale of the distribution after 1996.

Population: We define full-time workers by following Baum-Snow & Pavan (2013), who rely on three criteria: i) weeks worked, ii) usual hours per week, and iii) wages in relation to the federal minimum wage. Specifically, a full-time worker is defined as an individual that worked at least 40 weeks with at least 35 usual hours per week, and who earns at least 50 percent of the federal minimum wage.[13] For 1970, usual hours per week are not available and we have only relied on the weeks worked and the wage earned.

Geographic units: The geographical units of analysis are commuting zones (CZs), that correspond to local labor markets in the United States. The limitation of used data sources (CPs and ACSs) is that the place of residence is defined only at Public Use Microdata Areas (PUMA), which has no one-to-one mapping with Commuting Zones. For 1990, 2000 and 2005-2021, the smallest identifiable geographic unit is the PUMA, containing at least 100,000 persons. For 1970 and 1980, the most basic geographic variable is the county group, which are a combination of counties or portions of counties that total 100,000 population (in 1980) and 250,000 population (in 1970). We match individual observations with Commuting Zones following the probabilistic mapping made available from David Dorn's website.[14]

Variables: Census of Population (CP) and American Community Survey (ACS) provide detailed information on income at the household and individual level (distinguishing different income sources, including wages and salaries), together with rich information on socio-demographic characteristics (e.g., age, gender, education), housing and economic circumstances such as labour status, occupation, industry, and employment types (e.g. self-employment, full-time/part-time).

Known issues: Depending on the year, wages are top-coded at different thresholds. For 1970 and 1980 all observations above the top-coded threshold are assigned that threshold. From 1990 onwards, top-coded wages are assigned the median value of all observations above the threshold in their respective state (thus retaining some information about the distribution at the top). Up until 2000 the top-coded thresholds of the decennial data are constant nation-wide. For the annual data starting in 2005, the top codes correspond to the 99.5th percentiles of each state. We adjust top-coded wages following Autor et al. (2008) and multiply all censored wage observations by a factor of 1.5.

---

[13] The federal minimum wage is defined on an hourly basis, so we have arrived at annual values by multiply it by 40 (minimum weeks to be a full-time worker) and 35 (minimum hours per week to be a full-time worker).

[14] See https://www.ddorn.net/data.htm. This approach is frequently used, e.g. Autor et al. (2013), Autor and Dorn (2013) and Autor (2015) or Albouy and Zabek (2016).