Import Competition, Industrial Decline, and Distributional Dynamics

A Reduced Form Analysis of the French Case (1990-2020)

A thesis presented for the degree of MSc in economics, by Andrea Fournel

Director: Pr. Thomas Piketty
Referee: Pr. Lucas Chancel

Ecole normale supérieure
Ecole d'économie de Paris – *Paris School of Economics*
2022

Es sey die Bestimmung des Staats, jedem erst das Seinige zu geben, ihn in sein Eigenthum erst einzusetzen, und sodann erst, ihn dabei zu schützen. [Fichte 1799, p. 399]

I want to thank Lucas Chancel for having accepted to referee this master thesis, Daniel Cohen for having provided some of the data on which it is based, but also Eric Monnet, Camille Hémet, Clément Malgouyres, Laurent Gobillon and Paul-Armand Veillon for having reviewed part of the findings. I take the occasion to express my gratitude towards Camille Hémet, Daniel Cohen and Marc Gurgand for having greeted me at the economics department of the ENS. It extends to every person in PSE or at the ENS who, one way or another, provided help and support for the completion of this thesis. There's not enough room there to fully express what this thesis owns to Thomas Piketty, who had the original idea, and closely followed the steps of its progress with his genuine benevolence. Findings and opinions expressed there are mine, with no connection with the institutions to which I am affiliated or for which I have worked. Replication files are available on demand.

Abstract

We want to take on the much debated question of the distributional effects of trade, with a special emphasis on the pre and post-redistribution income response. We implement a shift-share setting adapted from the most recent debates [Goldsmith-Pinkham, Sorkin, and Swift 2020; Adão, Kolesár, and Morales 2019; Borusyak, Hull, and Jaravel 2021] spawned by the seminal approach of [Autor, Dorn, and Hanson 2013], crossing French Census and employer-employee matched data at the level of the commuting zone (*ZE*), over 1990-2018.

Focusing first **(section 1)** on the import channel, with two major trade partners (China and Germany), using as measure of exposure the sectoral rise in imports normalized by the total local workforce and sectoral shares, we estimate that a marginal rise in Chinese import exposure is associated with a -4.02 pp decadal decline in total manufacturing employment within the ZE (compared to -1.56 for German imports). Overall, we are able to explain respectively one third and one fourth of the industrial decline of the last three decades through import competition from these two partners. Estimates are robust to the choice of alternative instruments and to regional and sectoral clustering, and exclusive to imports of final goods. The elasticity of nonmanufacturing employment to the shock is estimated at 1.37, a figure slightly inferior to the local multipliers of [Moretti 2010]. The induced decadal rise of the share of unemployed people within the adult population of the zone of $+0.12$.

Focusing now on Chinese imports, we fail to identify signs of a longer term recovery in the employment, unemployment and fiscal income response to the shock **(section 2)**, a finding at variance with the predictions of classical local demand shocks models [Blanchard and Katz 1992]. In a *between*-regions approach, the shock indeed cripples the fiscal income of more exposed zone (with a decadal marginal impact of -2.16pp), but leaves the income hierarchy of regions almost unaltered. In order to explore the *within*-region dimension, we implement a group-level IV quantile strategy in the spirit of [Chetverikov, Larsen, and Palmer 2016] **(section 3)**. We identify significant negative responses of wages (decadal marginal impact of -6.27pp) driven primarily by lower-paid manufacturing jobs. We investigate the distributional dimension of the employment response along the wage and skill distribution, substantiating the idea that the import shock has fostered job polarization [Mion and Zhu 2013; Malgouyres 2017a] and innovation decline among less productive firms [Aghion, Bergeaud, et al. 2021]. As a result, if the impact on disposable income is neutral (with a marginal response of -1.41pp), the fiscal income response is clearly regressive; precisely, neutral down to the third decile of the income distribution where it abruptly drops, the marginal impact for the first decile being about 5.5pp below the general response. The expected consequence is a rise in the shares of local incomes provided by redistribution ($+0.51$pp).

We hypothesize that this sharp response of transfers is fated to alter the social perception of redistribution and of the groups benefiting from it. We find that in more exposed zones, the residential strategies of native families become more averse to the presence of foreign-born populations in their neighbourhood **(section 4)**. Consistently, the import shock significantly bolsters the local vote shares of all conservative parties, whatever their agenda on trade and international integration **(section 5)**, with non-trivial marginal impacts (ranging from $+0.79$ to $+1.63$ for the last four second rounds of presidential elections).

Keywords: import competition, trade shocks, manufacturing decline, distributional impact of trade

JEL Classification: D72, E24, F14, F16, F66, L60, O47, R12, R23

Contents

1 Import competition and industrial decline – A shift-share strategy

1.1 Theoretical introduction

When the employment effects of growing trade competition from emerging economies first materialised, in the early 1990s, reactions from the academic and nonacademic literature alike remained widely optimistic.

At that time, the main theoretical framework to assess the impact of trade on inequalities remained the Heckser-Ohlin-Samuelson model, from which two classical predictions about the labour impact for Northern economies were derived: 1. Trade openness was doomed to bias Northern exports in favour of capital-intensive products and to increase the share of capital in the remuneration of factors (this is the well known Samuelson-Stolper condition[1]); 2. Openness would simultaneously lead to a redistribution of income in favour of high-skilled workers. In the classical account of [Samuelson 2004], this transfer results from a classical HOS factor-remuneration mechanism, coupled with inefficient redistribution at the national level. Contemporary literature, suffused in the spirit of the Krugman-Melitz models, ventured a more micro version of this argument: the main feature of models with Chamberlin-differentiation is that trade openness will force less productive firms to exit, [Melitz 2003] arguing that in late 20th c. U.S. context, these exiting firms were more intensive in low-skill labour [Burstein and Vogel 2017]. It has also been noted that managers have more comprehensive skills, and can more easily navigate between sectors, while blue-collar skills are more industry-specific, re-training, to them, being longer and more costly [G. M. Grossman, Helpman, and Kircher 2017].

However, econometricians of the 1990s found it extremely difficult to substantiate these predictions on available trade and labour data of the time. The dreary consequences of trade openness on inequality foretold by the HOS model were at most "elusive" [Krugman, Cooper, and Srinivasan 1995].

One explanation of these equivocal results involves data restrictions. In the early 1990s, import exposure remained relatively limited. It is mainly after the turn of the millennium that China rose to the status of leading manufacturing exporter of the World (in terms of share of international exports of goods, China grew by 1.7 points over 1990-2000, versus 6.6 points over 2000-2010; it surpassed the U.S. in 2006). Arguably, in the early 1990s, it should have been challenging to extract from labour data a massive negative effect of trade competition on unemployment and wage inequality using data from the 1980s. Available methodologies were mostly prospective, based on the factor content of trade and on estimates of the elasticity of substitution between skilled and unskilled labour [Borjas, Friedman, and Katz 1997; Rodrik 1997] or on the relative evolution of product prices [Slaughter 1998], and their conclusions were but mixed: in the great paper by Borjas et alii, trade competition from the South explained at most 6% of the wage premium obtained by college-educated workers in the U.S. over 1980-1995. Even in the maximal strategy of [Acemoglu 2003], who accounted for the indirect impact of trade on skill-biased technological change, import exposure explained 23% at most of these wage premia. This was one of Paul Krugman's leading line of argument at that time: the purported massive negative distributional impact predicted by HOS-type models and the Stolper-Samuelson theorem were nowhere to be found in actual income data:

> *When a country with a highly skilled labor force increases its trade with a country in which skill is at a greater premium, it can expect a decline in the real wages of its own unskilled workers. [...] All the evidence suggests, however, that this effect will be extremely small. While economic theory suggests that trade between the U.S. and Mexico should involve an exchange of skill-intensive for labor-intensive products, such a bias in trade against low-wage U.S. workers is surprisingly elusive in the actual trade data. Most notably, the widely cited study of NAFTA by Gary Hufbauer and Jeffrey Schott[2] finds that U.S. industries that compete with imports from Mexico pay almost exactly the same average wage as industries that export to Mexico. [...] This lack of evidence that trade really does worsen American income distribution is not unique to the Mexican case. Two economists who expected to find a significant effect of trade on wages have concluded that virtually none of the growth in wage inequality in the United States since 1979 is due to international factors[3]. A survey by Lawrence Katz reaches the same conclusion[4]. As a matter of theory, then, we must concede that NAFTA should be expected to hurt low-skill U.S. workers. In practice however, there is no evidence supporting this belief. [...] NAFTA will neither create nor destroy jobs, but it will make the existing North American labor force more productive. No serious study has failed to find that NAFTA will produce a small net gain for the United States.* [Krugman 1993]

As a consequence, at the end of the 1990s, it was widely assumed that trade had but a limited impact on inequality *within* one region or one country [Autor and Katz 1999]. The new generation of macro trade models indeed predicted a decline in low-skill employment in Northern economies, but it was assumed that these negative effects were largely offset by the positive macro gains in productivity, and that minor adjustments of the monetary and fiscal policy would be enough to tackle them. The consensus on that point was widespread [Wood 2018]; Postkenesians themselves were distancing from the usual left-wing critics of trade derived from structuralist approaches of prior decades in the spirit of [Kaldor 1940].

As emphasised by [Dorn and Levell 2021], the persistence of this consensus cannot be imputed exclusively to data availability issues, or to the fact that this literature was relying on old trade figures. Late replications of the factor content or price strategies [Bivens 2007; Edwards and R. Z. Lawrence 2010], even though they relied on data of the early 2000s, still failed to explain more than 10% of the widening in wage premia. The core issue was the specification used, most importantly the outcome variable of interest: the HOS predictions directed research towards an exclusive focus on wage premia for higher-skilled worker, obfuscating the wider context.

[1] The original wording of which implies very restrictive hypotheses, including perfect inner competition, constant returns to scale, and same number factors and of products.

[2] [Hufbauer et al. 1993]

[3] [R. Lawrence and Slaughter 1993]

[4] [Katz 1992]

In the early 2010s, the combination of a changing political atmosphere, of the rise of novel insights into inequality dynamics [Piketty 2013], and of a new focus in the academic literature on empirical identification [Angrist and Pischke 2010] prompted a new vein of research attempting to gauge the impact of trade openness on inequalities. Across the diverging methodologies of this literature, one might isolate at least three common points:

- *A shift towards more comprehensive definitions of trade exposure* – The factor content or price approaches, widely criticised [Burstein and Vogel 2017], have slowly been replaced by strategies which exploit the variability in exposure to trade across regions of a country, or sectors of a labour market. [Margalit 2011], who focused on trade-driven industrial layoffs in the U.S., or [Topalova 2010], who studied the 1991 Indian liberalisation, are well known precursors within the empirical literature, but concerns for sectoral and regional variability has also become an important feature of trade models [Caliendo, Dvorkin, and Parro 2015; Galle, Rodríguez-Clare, and Yi 2022];

- *The rise of quasi-experimental frameworks* – The credibility revolution in econometrics induced many scholars to turn away from all-encapsulating trade models to focus on specific trade policy events. Well studied episodes include trade liberalisations in emerging economies like Brazil [Kovak 2013], the fall of the USSR, the ensuing abnormally high trade surpluses of Russia and their inequality background [Novokmet, Piketty, and Zucman 2018], paralleled by a rise of competition from Eastern countries in Europe [Dauth, Findeisen, and Suedekum 2021]; more recently, the end of the Multi-fiber Agreement [Lopez-Acevedo and Robertson 2012], unexpected currency events like the appreciation of the Swiss Franc following the end of the Euro peg in 2015 [Kaufmann and Renkin 2018] or the fall of the British pound the day of the Brexit referendum [Costa, Dhingra, and Machin 2019], the protectionist policy of the Trump administration [P. D. Fajgelbaum et al. 2019], and even retrospective papers on key events of trade history, for instance to gauge the effect of the NAFTA [Hakobyan and McLaren 2016], of 19th c. colonial exclusive zones [Cogneau, Dupraz, and Mesplé-Somps 2018; Alvaredo, Cogneau, and Piketty 2021], or even of the *Zollverein* [Wolf 2009];

- *Looking beyond aggregate effects* – Through its careful exploitation of sectoral and regional variability, that literature was able to prove that, even if at the aggregate level the impact of trade on inequality was relatively moderate, the negative effects of liberalisation had been concentrated on some very specific manufacturing sectors and on a little number of left-behind regions [Helpman et al. 2016], and that these margins of trade liberalisation policies have played a critical role in the evolution of the social and political atmosphere of the last decade [Rodrik 2021].

Among this new trend of the literature, shift-share settings exploiting the differential sectoral structure of each region of a country – the most well known being [Autor, Dorn, and Hanson 2013] – encapsulated all of those three features: defining exposure at the regional level provided wide-ranging empirical opportunities, allowing to test the impact of the import channel on a large array of outcomes, well beyond the wage premia dimension, while Bartik instrumentation offered a straightforward solution to endogeneity issues.

1.2 Empirical framework

Measure of import exposure

In this type of shift-share setting, two types of data are combined: 1. Trade data, classically from the U.N. *Comtrade* base, which provides a detailed breakdown of exports and imports between countries by subtypes of exchanged products; 2. Regional data about the sectoral employment structure of each commuting zone. From this, it is straightforward to build an index of exposure to import competition for each region. Denoting zones with an i and industrial sectors with a j, that main index is the change in import exposure per worker of a zone over a period $t, t+1$, denoted $\Delta IPW_{it,t+1}$. In the setting of [Autor, Dorn, and Hanson 2013], the exposure index is computed for each sector, taking the imports from the trade partner over the time period, ΔM, times the share of region i in the total national workforce of sector j ($L_{ijt}/\sum_i L_{ijt}$) at time t (the beginning of the time period of interest), the grand total being normalised by the total number of workers in region i. Summing across sectors yields the individual loss per worker caused by imports over the period:

$$\Delta IPW_{it,t+1} = \sum_j \frac{L_{ijt}}{\sum_i L_{ijt}} \frac{\Delta M_{jt}}{L_{it}} \tag{1}$$

Limits and counfounding factors

A specification using ΔIPW as the main explanatory variable, and some measure of industrial decline as the dependent, implies several limitations: 1. Such a reduced-form framework focuses exclusively on one channel connecting a trade shock to employment dynamics, i.e., on the impact of competition from foreign products within the home market, ignoring home exports to the foreign partner, and competition from the foreign partner on other foreign markets that the home firms serve; 2. That specification relies on the strong assumption that there is zero mobility of labour between regions within the home country; i.e., that marginal effects that will be estimated will encapsulate employment decline, not outmigration, Autor, Dorn and Hanson arguing that even in U.S. context, it is a relatively realistic assumption, especially when it comes to blue-collar workers, which are known to be less mobile [Wozniak 2010]; 3. The reduced-form approach assumes a unique direct effect of exposure on job losses. However, as we'll see, in our setting, there is evidence of important retroactive impacts through wage decline (see our section 3.2.1.), and of spatial spillovers between sub-regions (see our annex H).

More important still are endogeneity issues. Since the explanatory variable essentially captures the local industrial structure, we rely on the assumption that this structure has been predetermined long before trade began with the foreign partner, and that the employment decline isolated is caused by exposure only, and does not encapsulate some region-specific or sector-specific employment shock.

The usual way out is a shift-share instrument strategy, which consists in rebuilding the main explanatory using trade data for a pool of advanced economies similar to the home country, under the assumption that there is no correlation between countries in import demand shocks.

Instrumentation strategy

In the Autor-Dorn-Hanson framework, these Bartik instruments are built like the ΔIPW index, with two slight differences: 1. Import figures ΔM are replaced with $\Delta \overline{M}$, the imports from the partner to a control group of advanced economies ; 2. The start-of-the-period labour force L is taken with a lag of one period (here, a decade). The authors use this lag to counter the simultaneity bias. Each instrument writes:

$$\Delta \overline{IPW}_{it,t+1} = \sum_j \frac{L_{ijt-1}}{\sum_i L_{ijt-1}} \frac{\Delta \overline{M}_{jt}}{L_{it-1}} \tag{2}$$

There is little questioning about the relevance of these instruments: in [Autor, Dorn, and Hanson 2013] and in this thesis alike (see our tables 1 and 3), the first stage yields very satisfying results. Using an alternative control group (see table 6) does not change the general picture.

On the contrary, the proper proof of the orthogonality of shift-share instruments is a contested issue, with a growing body of formal literature, some items of which are directly focused on the framework we'll be using, most importantly [Goldsmith-Pinkham, Sorkin, and Swift 2020], [Adão, Kolesár, and Morales 2019] and more recently [Borusyak, Hull, and Jaravel 2021]. Like most of the existing literature, we'll be clustering our standard errors at a meta-regional level, here, at the level of the INSEE's superzones or ZEAT (*Zones d'études et d'aménagement du territoire*); this should discard part of the risks at identifying a merely regional shock[5]. As to the risk involving sectoral shocks, we'll conduct at the end of our section 1 several robustness checks in the spirit of [Goldsmith-Pinkham, Sorkin, and Swift 2020] and [Adão, Kolesár, and Morales 2019].

Main specification

The main 2SLS specification is a very straightforward framework, in which the evolution of local manufacturing employment over the decade (ΔL) is regressed on the main exposure index, a time dummy for each decade and a vector of controls:

$$\Delta L_{it,t+1} = \beta_1 \Delta IPW_{it,t+1} + X'_{it}\beta_2 + \gamma_t + u_{it} \tag{3}$$

$\Delta IPW_{it,t+1}$ being instrumented by $\Delta \overline{IPW}_{it,t+1}$ as described above. It is not a first-difference model: the multiple decades are simply stacked with a specific time dummy.

Data

Applying this setting to French data is not straightforward, especially when it comes to combining Comtrade data with the sectorial classifications, all the more since we rely, for employment variables, not on social security declarations (DADS) but on the INSEE's Census. The main comparative features are the following:

- The period of estimation extends to three decades: 1990-1999, 1999-2008 and 2008-2018. Most of the U.S. literature uses as unit of interest the commuting zone (CZ), which has on average 400k inhabitants; in our setting, the main unit is the INSEE's *Zone d'emploi* (ZE, 2010 definition), with an average of 200k inhabitants, and sometimes the *département* (average of 600k inhabitants);

- Our instrument $\Delta \overline{IPW}_{it,t+1}$ is built on two non-EU economies (Japan and Switzerland) and two EU countries (Germany and Spain). We'll see in our robustness checks that using an alternative instrument based on non-EU countries only does not change the main results. Descriptively, over 1999-2008, exports from China to France have expanded a bit slower than exports to our instrumental group; once quantities are re-scaled to France's population, we get a +36.1 billion USD in imports to France versus +47.1 for the instrumental group. French imports are relatively inferior to the control group's trend for steelworks (+2.2 vs. +3.9) and computers (+11.1 vs +15.6), slightly superior for textiles (+8.4 vs +7.9).

- Employment figures are drawn from the *Recensement* of the INSEE. Prior to 2008, the Census relied on a classification of sectors known as NES-AE, which was specific to France, but closely connected to the United Nations' ISIC-rev.3 categorization. The 1999 Census provides, for each *commune*, the employment structure classified within successively 4, 36, 60 and 114 subsectors. The 1975, 1982 and 1990 issues provide a prior version of the 36 breakdown ; the 2006 & 2007 issues provide the 36 breakdown in its definitive version. The 2008 issue introduced the INSEE's new system, the NAF-rev.2, parallel to the ISIC-rev.4 and NACE classifications; from 2008 onward, the Census provides a breakdown in 38 activities, 16 of which belong to the industrial sector. We apply conversions between the rev.3 and 4 of the ISIC according to the concordance tables of the OECD [OECD 2017]. As to trade data, they are drawn from the UN Comtrade database; these are imports in value (in 2022 USD), classified using the 6-digit HS system. We rely then on concordance tables between HS and ISIC classifications provided by the WITS software of the World Bank.

Summary statistics and overview

The most interesting aspect of the whole picture is how quickly the structure of imports from China is shifting from light industry products to more technology-intensive goods: over 1999-2008, imports of computers and electrical devices grew 1.3 times faster than textile imports in value; over 2008-2018, it was 2.9 times faster, the shift being even more pronounced in the control group. The main drivers are, on the one hand, a sharp decline of textile imports from China after the Great Recession (partly because suppliers have offshored production units to more competitive countries), and massive hikes in imports of electronic devices. The average ΔIPW over the whole

[5] Clustering at an inferior level, like the pre-2015 *régions*, does not change the general results.

period 1990-2018 is slightly superior to 3,000 dollars per worker. In the U.S. framework, the median ΔIPW was 1.25 thousand USD per head over 1990-2000 and 2.9 over 2000-2007.

Overall, the most protected ZEs are mostly touristic resorts of the West and South (+146 in Corte in 1999-2008), regions with genuine product differentiation like wine valleys (+755 in Jonzac over 1999-2008), and the West, thanks to its specific mix of primary sector + STEM (566 over the whole period for Carhaix). On the other hand, the general portrait of the exposed regions tends to evolve over time. Consistent with our picture of Chinese exports swiftly transitioning to heavy industries, over 2008-2018, we see regions with a highly technology-intensive industrial base enter the top rankings of the most exposed zone (for instance, +1616 in Grenoble, +1283 in Saclay). Conversely, because of the decline of textile imports, many textile regions tend to transition from the top to the bottom of the rankings between the two periods. The rankings are but barely altered when we build our lagged instrument; for the first lagged decade (1982-1990), one of the main differences is the very high exposure of regions in which extractive industries had an historical role (especially among Northern or Northeastern mining or steelworks bastions).

1.3 Early results

1.3.1 A consistent negative impact of import exposure on manufacturing employment

From now on, and for much of this thesis, the trade partner of interest will be France's main developing partner, China. Some results will be reported for Germany (see table 6), for Turkey and Poland; but as we'll see, beyond the industrial dimension, the massive economic and social impact of the China shock that a shift-share framework can identify are extremely difficult to replicate for alternative partners.

The decade-per-decade estimation of model (3) for China is displayed in table 1. Unsurprisingly, the bulk of the depressing effect is concentrated in the second decade, the one characterised by the sharpest rise in imports from China.

Table 1: Exposure to import competition from China and change in manufacturing employment at the ZE level (I)

	Dep. : Decadal change in total manufacturing employ. (in pp)					
Type	OLS			2SLS		
	1990-1999	1999-2008	2008-2018	1990-1999	1999-2008	2008-2018
	(1)	(2)	(3)	(4)	(5)	(6)
Rise in imports from China per worker over the decade (in 2022 kUSD)	-2.99	-5.62^{***}	-2.79	-3.36^{**}	-6.28^{***}	-3.27
	(1.85)	(1.37)	(2.05)	(1.7)	(1.34)	(2.59)
R^2	0.09	0.22	0.02	0.09	0.21	0.02
F-stat	31.1^{***}	81.1^{***}	3.6^{*}	15.8^{***}	79.3^{***}	3.5^{*}
First stage: Instrumenting by the rise in imports to a group of control countries				0.41^{***}	0.93^{***}	0.58^{***}
				(0.06)	(0.09)	(0.05)
R^2				0.41	0.79	0.71
F-stat				203^{***}	1117^{***}	738^{***}
Obs.	304	304	304	304	304	304

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: The unit of observation is the ZE (*Zone d'emploi*, definition of 2010). The dependent variable is the change (in pp) of total manufacturing employment within the ZE. The main explanatory variable is the index ΔIPW, described herein above, which provides an estimation of the mean rise in Chinese imports per worker within each ZE. The instrument is the same ΔIPW, in which French trade data has been replaced by a control group of four countries (Japan, Germany, Spain, Switzerland) and all labour force variables are taken with a decadal lag. No other control is applied. Observations are weighted by the start-of-the-decade total Census population of the ZE. Standard errors are clustered at the level of the INSEE superzones.

1.3.2 Controlling for alternative theories of industrial decline (skill-biased technological change, offshoring, automation)

Aside of the usual regional controls about gender, ethnicity, education, insecure jobs[6] and the share of manufacturing within the total workforce, we'll be using throughout this thesis a set of three supplementary control variables meant to capture the great alternative explanations of industrial job decline ventured in the literature of the last three decades:

- *Automation and the rise of robots* – We replicate the methodology of one of the most classical articles on the issue of automation : [Acemoglu and Restrepo 2020]. The intuition is to use data about the stock of newly-bought robots within each sector in each country, from which an index of average penetration of robots in sector j is derived:

$$APR_{j,t_0,t_1} = \frac{R_{j,t_1} - Rj,t_0}{L_{j,1990}} - g_{j,t_0,t_1} \frac{Rj,t_0}{L_{j,1990}} \quad (4)$$

R being the stock of robots in industry j at time t, L total employment, and g the growth rate of the sector

[6] Namely, the share of internships, State-sponsored jobs, short-term *CDD* contracts, and contingent work within total employment, as reported in the INSEE's Census.

over the time interval. The average APR of a commuting zone i is then computed as:

$$\text{Exposure to automation}_{i,t_0,t_1} = \sum_j l_{j,1990} \times APR_{j,t_0,t_1} \quad (5)$$

Where l is the share of each sector in the total workforce[7].

- *Skill-biased technological change* – If, as we'll see, the seminal intuition of [Acemoglu 2003] about trade acting as a buffer of technological-change driven job polarisation has been substantiated many times since, even very recently on French employer-employee matched datasets [**harrigantoubal2020polaris**], the major earliest theorists of skill-biased technological change and polarisation tended to play down the role of trade competition, imputing much of the decline of industrial employment to the *routinisation* hypothesis, with the idea that, because of accelerated progresses in STEM sectors, jobs markets of the North had produced extra demand for either creative and innovative jobs on the one hand, or for social interaction jobs (*care* workers, delivery jobs) on the other, hence the decline of routine jobs like the intermediate office tasks or mechanised activities of the industry. One way to test that hypothesis is to include an index for routine jobs, as defined by [Autor, Levy, and Murnane 2003] and [Autor, Katz, and Kearney 2006][8];
- *Offshoring* – Some jobs are lost because of direct trade competition from foreign firms, but some other jobs are lost through offshoring from home firms. The most natural empirical strategy to implement this distinction, that we'll use at some points in this thesis, is to isolate within trade data the intermediate products, as opposed to final, consumption goods. A more straightforward approach relies on a specifically-built *offshorability* index of local jobs, taken from [Firpo, Fortin, and Lemieux 2011]. Those jobs are not easily offshorable that require face-to-face contact and direct on-site monitoring[9]. There's almost no correlation between the skill and routine dimension of a job and its offshorability: a highly qualified activity like coding softwares ranks among the most offshorable ones (because it does not require face-to-face contact or geographic proximity); conversely, nurses and first-aid workers are among the most protected;

Table 2: Raw correlation, at the employment zone (ZE) level, between key indexes (weighted by 1990 total pop.)

Exposure to imports from China	*1990-1999*	1.00							
à la [Autor, Dorn, and Hanson 2013]	*1999-2008*	0.44	1.00						
Share of routine jobs	*1990*	−0.03	0.01	1.00					
à la [Autor, Katz, and Kearney 2006]	*1999*	−0.03	0.01	0.86	1.00				
Offshorability of manuf. jobs	*1990*	0.16	0.25	−0.04	−0.04	1.00			
à la [Firpo, Fortin, and Lemieux 2011]	*1999*	0.23	0.33	−0.06	−0.07	0.06	1.00		
Expansion of automation	*1990-1999*	−0.15	−0.07	0.48	0.39	−0.09	−0.03	1.00	
à la [Acemoglu and Restrepo 2020]	*1999-2008*	−0.11	−0.05	0.57	0.49	−0.17	−0.05	0.51	1.00

These three indexes are relatively uncorrelated between one another. Over U.S. and French data alike, automation is slightly negatively connected with trade exposure and offshorability (compare table A3 in [Acemoglu and Restrepo 2020] and our table 2); textiles, typically, is a sector which has experienced high trade competition but limited automation. Conversely, at the ZE level, we find a positive correlation between routine jobs and automation.

In simple descriptive statistics, the variables which are most correlated with the decline of industrial employment at the ZE level evolve over time. Over 1990-1999, we find a slight negative correlation with the share of routine jobs and the penetration of robots; over 1999-2008 on the contrary, we find a strong correlation with offshorability and exposure to China trade. It is relatively consistent with our general picture of China imports being more and more skill & technology intensive as we move to more recent data[10].

[7]The original dataset used by [Acemoglu and Restrepo 2020] for the computation of APRs, the annual report of the International Federation of Robotics, is not available to the public; we are therefore forced to use the data provided by D. Acemoglu on his personal website; he computes sector-by-sector APRs over a set of five European countries used as a control group for the U.S., namely Denmark, Finland, France, Italy and Sweden. The only option we are left with is to use these APRs and to map them to the sectoral structure of manufacturing employment in each ZE in the INSEE's Census. Note that the original data from the IFR is not detailed at all, giving stocks of robots for 13 industrial sectors only.

[8]Autor and his coauthors have recourse to the US Department of Labor's *Dictionary of Occupational Titles*, defining routine jobs as those where repetitive tasks (cognitive or manual) are highly frequent. It is then possible to build a share of routine jobs within each industrial sector. Here, we rely on the averages by great industrial sector found in the replication provided on Daron Acemoglu's personal web-page for his article [Acemoglu and Restrepo 2020], mapping them to the industrial structure of each ZE in the INSEE's Census.

[9]We construct this index using the methodology described by [Autor and Dorn 2013]; they had recourse to the widely used O*Net base of the U.S. Dep. of Labor to build a standardised *offshorability* index of sectors. An activity is said to be offshorable when : 1. It does not require face-to-face contact; 2. It does not require on-site work. Following [Firpo, Fortin, and Lemieux 2011], they define face-to-face activities as the simple average of O*NET variables face-to-face discussions, establishing and maintaining interpersonal relationships, assisting and caring for others, performing for or working directly with the public, and coaching and developing others. On-site activities is similarly defined as the simple average of the variables inspecting equipment, structures, or material, handling and moving objects, operating vehicles, mechanised devices, or equipment, and the mean of repairing and maintaining mechanical equipment and repairing and maintaining electronic equipment. The next step consists in mapping each occupation of the SOC classification to the equivalent categories of the Census to get an index of the offshorability of local manufacturing employment for each zone. In our setting, we use concordance tables of the ILO to shift from the SOC to the international ISCO classification of the ILO. We then have recourse to the *Enquête emploi* (aggregated 1990-2002 issue) which provides, for a very large sample of 2.5 million workers, the ISCO code of their occupation, their industrial sector and their SES (the INSEE's PCS). Each worker is ascribed the offshorability index of its ISCO occupation; weighted averages are then computed along the INSEE's sectoral classifications and the INSEE's PCS scale; mapping to the Census values is then straightforward.

[10]We must heed to the fact that these differential correlations are also dependent on : 1. The quality of available data; 2. The choices made in the construction of each index. 1. Overall, our indexes of offshorability and of trade exposure are much more precise than those derived from D. Acemoglu's data. The original data of the IFR provides but a very coarse decompozition, with 13 subsectors,

If however we stack periods to ensure comparability with the existing literature, we find that exposure to China trade remains the key explanatory variable:

Table 3: Exposure to imports from China and change in manufacturing employment at the ZE level (II)

	Dep. : Decadal change in total manufacturing employ. (in pp)					
	1990-2018					
	(1)	(2)	(3)	(4)	(5)	(6)
Rise in imports from China per worker	-6.27^{***}	-6.47^{***}	-4.22^{***}	-4.35^{***}	-3.38^{*}	-4.02^{***}
over the decade (in 2022 kUSD)	(1.32)	(1.21)	(1.48)	(1.16)	(1.96)	(1.33)
Extra controls:						
Share of employ. in manufacturing		0.04	−0.04	−0.23	−0.02	−0.15
		(0.14)	(0.08)	(0.18)	(0.16)	(0.22)
Share of women in lab. force				-0.91^{**}		-0.77^{*}
				(0.38)		(0.35)
Share of foreign-born in pop.				-1.05^{***}		-1.09^{***}
				(0.25)		(0.24)
Share of higher educ. in pop.				0.27		0.31
				(0.21)		(0.19)
Share of insecure jobs				-0.57^{**}		-0.66^{**}
				(0.21)		(0.23)
Share of routine jobs					−0.11	-1.09^{*}
					(0.75)	(0.59)
Offshorability of manuf. jobs					-1.99^{*}	-1.36^{**}
					(1.13)	(0.51)
Penetration of robots					−0.58	−0.38
					(1.2)	(1.36)
Regional dummies			X	X	X	X
R^2	0.28	0.28	0.49	0.57	0.51	0.58
F-stat	115^{***}	86.6^{***}	32.1^{***}	37.9^{***}	30.1^{***}	35.7^{***}
First stage: Instrumenting by the rise	0.83^{***}	0.81^{***}	0.8^{***}	0.81^{***}	0.76^{***}	0.77^{***}
in imports to a group of control countries	(0.07)	(0.07)	(0.08)	(0.09)	(0.07)	(0.07)
R^2	0.89	0.9	0.89	0.9	0.91	0.91
F-stat	2565^{***}	1949^{***}	302^{***}	263^{***}	299^{***}	266^{***}
Obs.	912	912	912	912	912	912

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: The unit of observation is the ZE (*Zone d'emploi*, definition of 2010). The dependent variable is the change (in pp) of total manufacturing employment within the ZE. The main explanatory variable is the index ΔIPW, described herein above, which provides an estimation of the mean rise in Chinese imports (in value) per worker within each ZE. The instrument is the same ΔIPW, in which French trade data has been replaced by a control group of four countries (Japan, Germany, Spain, Switzerland) and all labour force variables are taken with a decadal lag. We stack two decades (1990-1990 and 1999-2008) and include a time dummy for the second decade. Observations are weighted by the start-of-the-decade total population of the ZE. Regional dummies denote the pre-2015 reform French *régions*. Standard errors are clustered at the level of the INSEE superzones.

In our preferred specification, i.e. column 6 of table 3, a $1000 rise in import exposure per worker within the ZE is associated within a decline of total manufacturing employment by 4.02 percentage points[11]. This is almost similar to the marginal impacts reported in table 5 of [Autor, Dorn, and Hanson 2013]. Compared to existing French estimates, it is but very slightly inferior to the effects reported in the regional strategy of [Malgouyres 2017a]. As far as comparison is possible, industry-based strategies find impacts which are half this magnitude [Aghion, Bergeaud, et al. 2021] while firm-level estimates are of the order of ten times lower [Biscourp and Kramarz 2007; Aghion, Bergeaud, et al. 2021] and in some specifications nonsignificant [**harrigantoubal2020polaris**]. Equivalent European regional strategies provide estimates slightly superior to ours [De Lyon and Pessoa 2021; Dauth, Findeisen, and Suedekum 2021]. Note that among our special control variables, the only one on which we find a consistent negative impact is the offshorability index, a results consistent with European estimates of the impact of offshoring on manufacturing employment [Biscourp and Kramarz 2007; Mion and Zhu 2013; Hummels et al. 2014] and with our intermediate goods estimates of section 1.3.5.

1.3.3 A third of manufacturing job losses could be explained by trade competition

To provide an intuition of these results, table 4 compares them to the Autor-Dorn-Hanson framework of 2013 (which we will abbreviate ADH from now on). Exposure of French industries came later, but was slightly more

compared with the 53 and 648 subsectors used to build the offshorability index over 1990-1999 and 1999-2008 respectively; 2. We rely on indexes which are widely used in the literature, but some of them might be criticised; the offshorability index of [Firpo, Fortin, and Lemieux 2011] for instance is highly efficient in isolating jobs which are indeed protected against offshoring (nurses, doctors), but it yields some surprising results when it comes to endangered jobs. Because of the prevalence of on-site work, many industrial jobs get an artificially low level of offshorability, while many STEM jobs get an artificially high level. This might explain the low-correlation with the decline of industrial employment over 1990-1999 and the higher one over 1999-2008.

[11] If we focus on the first two decades (see for instance table), or on the second only (see for instance figure 9), we find marginal impacts which are systematically superior.

pronounced. As a result, the predicted industrial decline to Chinese competition over the two decades is very similar. An important difference lies in the fact that, though the Hausman tests still lead us to use a 2SLS specification in all contexts, our 2SLS and OLS estimates are not as divergent as the ones of ADH; in the variance breakdown exercise (detailed in the corresponding annex) meant to isolate the supply-driven from the demand-driven dimension of the China shock, we ascribe a very tiny 7% share of the shock to demand-driven factors, meaning that our final estimate of the share of manufacturing employment destroyed by import competition from China is barely altered (from 34 to 32%) and overall superior to the U.S. figures.

Table 4: Comparing results of table 3 with those of [Autor, Dorn, and Hanson 2013]

	US Data	French Data
Main coefficient (marginal impact of exposure)	-4.23^{***}	-4.02^{***}
(tab. 5-col. 1 in origin. art.; here tab. 3-col. 6)	(1.05)	(1.33)
Av. rise in exposure to China trade per worker		
First decade	+$1,140	+$281
Second decade	+$1,839	+$2,029
Third decade		+$711
Total	+$2,979	+$3,021
Implied growth of manuf. empl. (1990-2008)	−12.6	
Actual growth of manuf. empl. (1990-2008)	−25.61	
Implied growth of manuf. empl. (1990-2018)		−12.44
Actual growth of manuf. empl. (1990-2018)		−36.33
Percentage explained (raw)	49%	34%
Percentage explained (supply-driven shock only)	24%	32%

1.3.4 Labour force impact: Lingering unemployment, multiplicative effects in nonmanufacturing

The depiction of the population impact of this industrial decline in [Autor, Dorn, and Hanson 2013], and in companion papers (most notably [Autor, Dorn, and Hanson 2016] and [Autor, Dorn, Hanson, and Song 2014]) hinges round the argument that the rise of employment in the service sector is not sufficient to offset the decline of local manufacturing jobs, that outmigrations are far too sluggish compared to the usual predictions of matching models, and that, as a result, the bulk of the adjustment is borne by unemployment and early retirement. Here, we'll be using the *Mobilité* database of the INSEE to check for migration patterns (i.e., we focus on persons having moved between 1990 and 1999, and 2001 and 2006 resp.).

Table 5: Exposure to imports from China and evolution of occupations within each ZE

Dep. : Decadal change in population log counts or shares of total adult population (1990-2008)								
	Population evolution			Adult population breakdown				
	Total change (1)	*Natural increase* (2)	*Migration increase* (3)	Working (manuf.) (4)	Working (tert.) (5)	Unempl. (6)	Retired (7)	Other inactivity (8)
Panel A. Change in log counts								
Rise in imports from China per worker:								
No controls:	2.19^{***}	−0.01	2.19^{***}					
	(0.56)	(0.09)	(0.55)					
Full vector of controls:	3.04^{***}	0.31	2.73^{***}					
	(0.62)	(0.22)	(0.71)					
Panel B. Change in shares of adult pop.								
Rise in imports from China per worker:								
No controls:				-0.14^{***}	-0.24^{**}	0.12^{*}	0.14	0.22
				(0.03)	(0.12)	(0.07)	(0.19)	(0.22)
Full vector of controls:				-0.22^{***}	-0.14^{**}	0.13^{**}	−0.02	0.38
				(0.05)	(0.07)	(0.05)	(0.17)	(0.25)
Obs.	608	608	608	608	608	608	608	608

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: The unit of observation is the ZE (*Zone d'emploi*, definition of 2010). We estimate model 3 with and without the full vector of controls mentioned in table 3, using as dependent variable a set of changes in population, expressed either in log counts, or in shares in adult (15 y.o. or more) population; data for these dependent variables are drawn from the *Mobilités* datasets of the INSEE. The main explanatory variable is the index ΔIPW, described herein above, which provides an estimation of the mean rise in Chinese imports (in value) per worker within each ZE. The instrument is the same ΔIPW, in which French trade data has been replaced by a control group of four countries (Japan, Germany, Spain, Switzerland) and all labour force variables are taken with a decadal lag. We stack two decades (1990-1990 and 1999-2008) and include a time dummy for the second decade. Observations are weighted by the start-of-the-decade total population of the ZE. Standard errors are clustered at the level of the INSEE superzones.

The main differences with the U.S. context might be summarised as such:

- *More exposed regions experience in-migrations* – It is a common feature of ADH and of its European equivalents, notably [De Lyon and Pessoa 2021], that an import exposure shock bolsters sectoral, but not regional mobility, i.e. workers adjusts to the shock by moving more frequently from one firm to another, but they rarely leave their home region. Here, we find a significant positive impact on the local population of exposed ZEs, even once controlled for a rich set of variables, and a positive impact which is driven primarily by in-migrations. Consistent with this figure, we also find that a marginal shock is associated with -0.22pp decline in the share of vacant accommodation ($t = 1.68$). This result is less surprising if we heed to the fact that more exposed regions are not traditional heavy-industry bastions: as we'll see in our next section, when the China shock materialised in the mid-1990s, the old manufacturing industries predominant in North-Eastern ZEs had been declining for decades; most manufacturing jobs had still been destroyed; in the *Nord* region, even for the first decade 1990-1999, Lille and Dunkerque are the only ZEs which are above the national average of the $\Delta IPW_{1990,1999}$ index; former steelworks or mining centres (Lens, Hénin, Bruay) are all far below. At the nation's level, more exposed ZEs are far from being the deprived and declining onesIn U.S. econometrical calibrations like [Caliendo, Dvorkin, and Parro 2015], California and New York are among the U.S. States with the highest $\Delta IPW.$. All in all, these marginal effects on in-migrations might encapsulate two types of reaction: 1. As we move on in time, Chinese exports become more technology-intensive, and ZEs like Saclay or Grenoble, known for their high shares of STEM jobs, enter the top tear of the exposure rankings. These are regions which, even if they are hurt at the margin, will benefit from trade exposure in many other ways, making them attractive to young workers (the average person which migrates to a ZE with a $\Delta IPW_{1999,2008}$ above national average is slightly younger and more educated than the national population); 2. Another type of mechanism, involving less attractive ZEs, is the default migration described by [Taffin and Debrand 2005; Vignal 2005; Davezies 2012], i.e., faced with an employment shock in their own district, low-income families, highly constrained on the housing market, are unable to move to the fastest-growing areas, and opt for a neighbouring region where the impact of local employment shocks has been less pronounced; descriptively, people who move to the 152 most exposed zones over 1990-2008 are more likely to come from the same region (63%) that migrants moving to the 152 least exposed ZEs (59%);

- *Persistent unemployment* – Contrary to ADH, we find little sign of increased transition to inactivity. The impact on local unemployment, on the contrary, is in the spirit of the American figures. The just-identified sector-by-sector instruments of [Goldsmith-Pinkham, Sorkin, and Swift 2020] allow for a sectoral analysis of this result. As mentioned above, the bulk of the employment dynamics of the shock is driven by a triad of sectors: microelectronics-textiles-steelworks; yet the highest marginal impacts on the job stock and on unemployment figures are found for all extractive industries and for plastics;

- *A sharp negative multiplicative impact on other sectors* – But one of the most interesting aspects is the multiplicative impact connecting the decline of industrial employment to job dynamics in the non-manufacturing, non-exposed sectors. Such spillovers are difficulty identified in [Autor, Dorn, and Hanson 2013], and [Autor, Dorn, and Hanson 2021] still fail to find any negative reaction of non-manufacturing employment to an import exposure shock. In European context, on the contrary, such multiplicative effects are found almost everywhere [Dauth, Findeisen, and Suedekum 2021; Citino and Linarello 2021], France included [Malgouyres 2017a]. In these estimates, or in the ones over DADS data in fig. 19, the implied local multipliers (the elasticity of local non-manufacturing employment to manufacturing employment shocks) are in the spirit of [Moretti 2010] (for the 2008-2018 DADS strategy for instance, it is 1.37 compared to Moretti's original U.S. figure of 1.57) and very similar to other European estimates. When we try to decompose this multiplicative effect, it seems like extractive industries in the earlier decades, and electronics later on, are the main drivers, light industries like textiles having a secondary role. The precise mechanism behind that multiplicative effect is however difficult to gauge: if we restrict ourselves to trade figures about final-consumption goods, the coefficient on manufacturing employment is multiplied by more than two (in tab. 38, col. 4) while the non-manufacturing coefficient becomes nonsignificant. There is a suspicion that the retroactive effect on service jobs is in fact capturing a foreign-for-domestic low-skill labour substitution mechanism through offshoring in the spirit of [Melitz 2003]. Our next section will attempt to determinate the specific nature of these multiplicative employment responses.

1.4 Robustness checks

Here's a brief overview of the robustness checks provided in the corresponding annex:

- *1.1. Testing for the relevance of shift-share instruments – Computation of the Rotemberg weights of [Goldsmith-Pinkham, Sorkin, and Swift 2020]* – Goldsmith-Pinkham and alii's initial argument about the assessment of the exogeneity assumption for Bartik instruments being reducible to a proof of the exogeneity of the initial shares (in this setting, the initial shares of each industry within each subzone) is applied to the Autor-Dorn-Hanson framework with the finding that STEM and technology-intensive-sectors account for an abnormally high part of the identifying variation, with the risk that the initial shares for these specific industries be correlated with important drivers of recent employment dynamics (typically here, with the over-representation of higher educated workers within each zone). We use the replication codes provided by [Goldsmith-Pinkham, Sorkin, and Swift 2020] to reconstruct the equivalent of their table 4 panel D. Their pivotal finding concerning the Autor-Dorn-Hanson framework was that, among the very large number of instruments (J industries times T years), 1% of them accounted for 49.5% of the absolute sum of the Rotemberg weights (in their setting, the final main coefficient is decomposed into a weighted sum of several just-identified instrumental variable estimators, one for each industry, allowing to gauge which industry contributes the more to the final aggregate marginal impact). In our replication, as far as comparison is possible, it is 11% of the instruments which account for 53% of the Rotemberg weights' sum. More important still, Goldsmith-Pinkham and alii's censured [Autor, Dorn, and Hanson 2013] for what they consider is an interpretative gap: Autor-Dorn-Hanson put the emphasis on light industries, among which are indeed found, in sheer descriptive statistics, sectors which have experienced the highest hikes in exposure to Chinese import competition (rubber, apparel, footwear...) and for which the exclusion restriction seems more than reasonable; yet once the breakdown is applied and the Rotemberg weights computed, these light industries account for a very low share of the final negative marginal impact, while the bulk of the effect is provided by technology-intensive sectors for which the exogeneity assumption is much more questionable. Once again, as far as comparison is possible, we find less

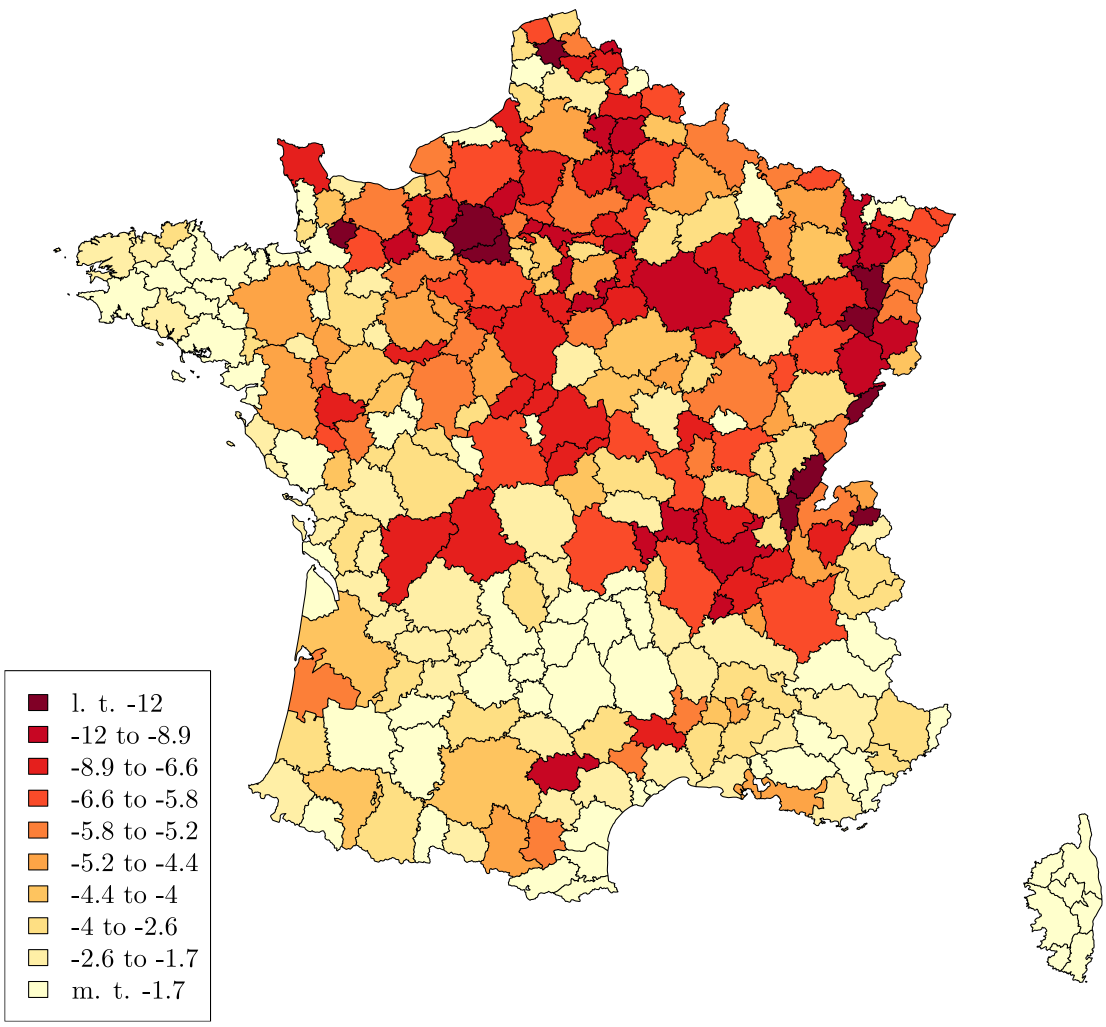


Figure 1: Industrial decline over 1990-2018 − Evolution of total manufacturing employment within each ZE (INSEE, *Zone d'emploi 2010*) as a ratio of the working age population, in percentage points [INSEE, *Recensement*]

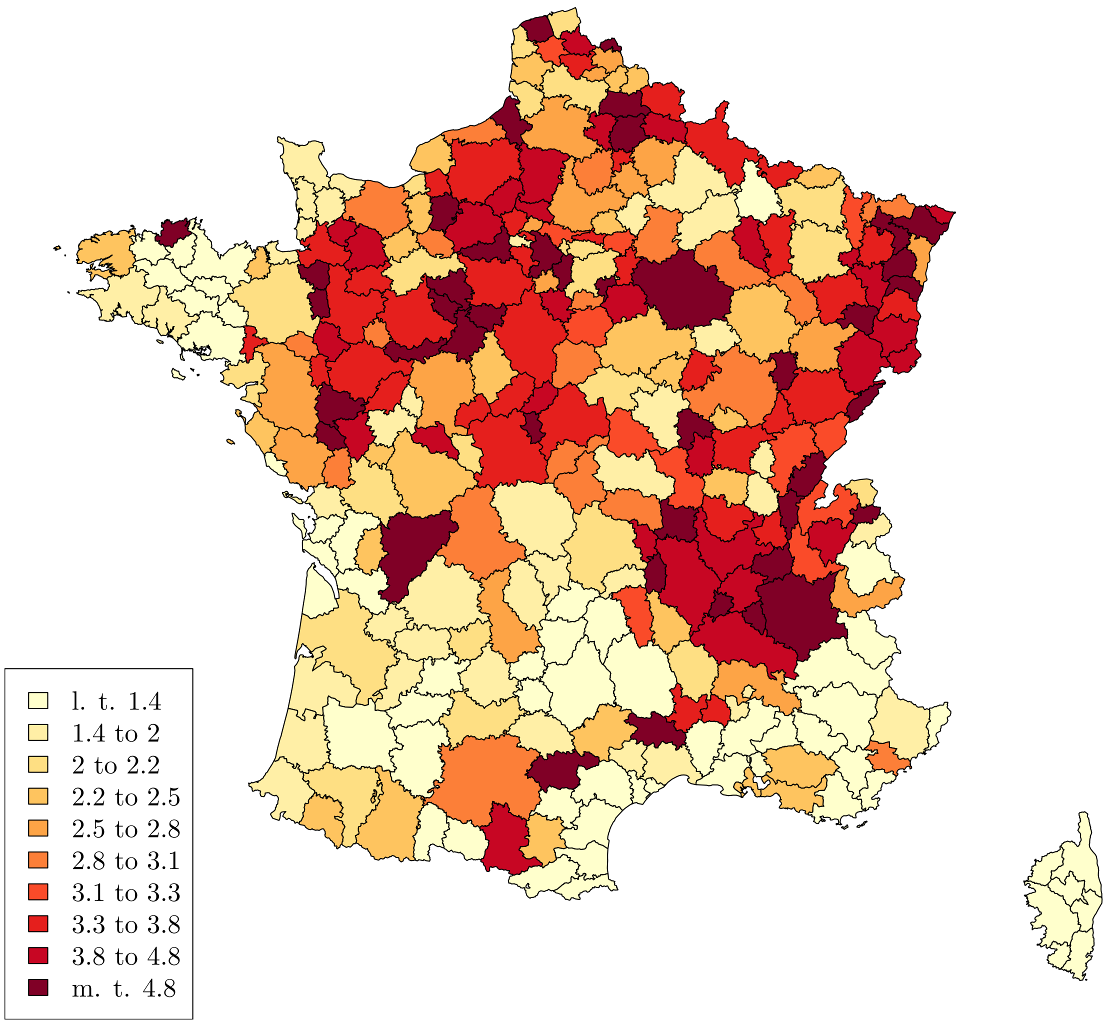
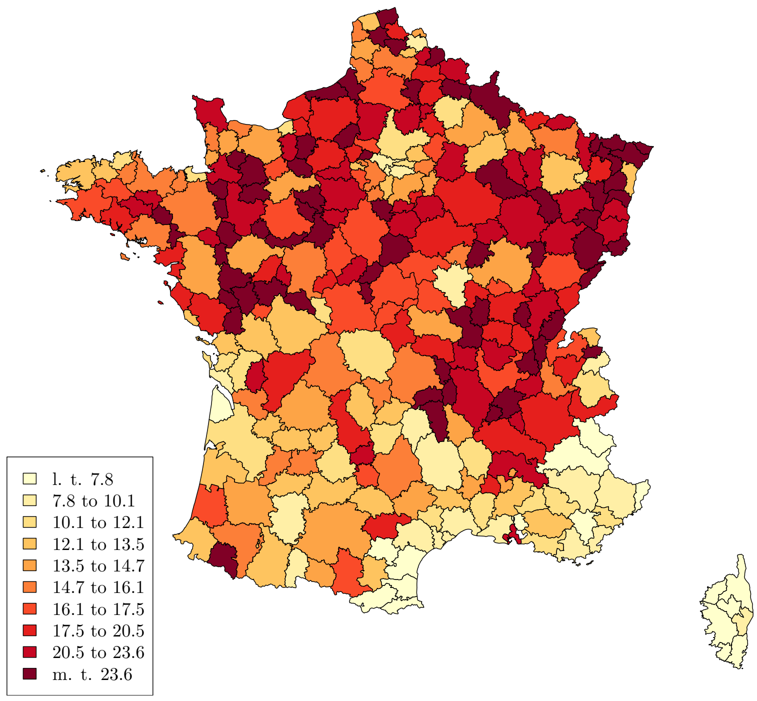


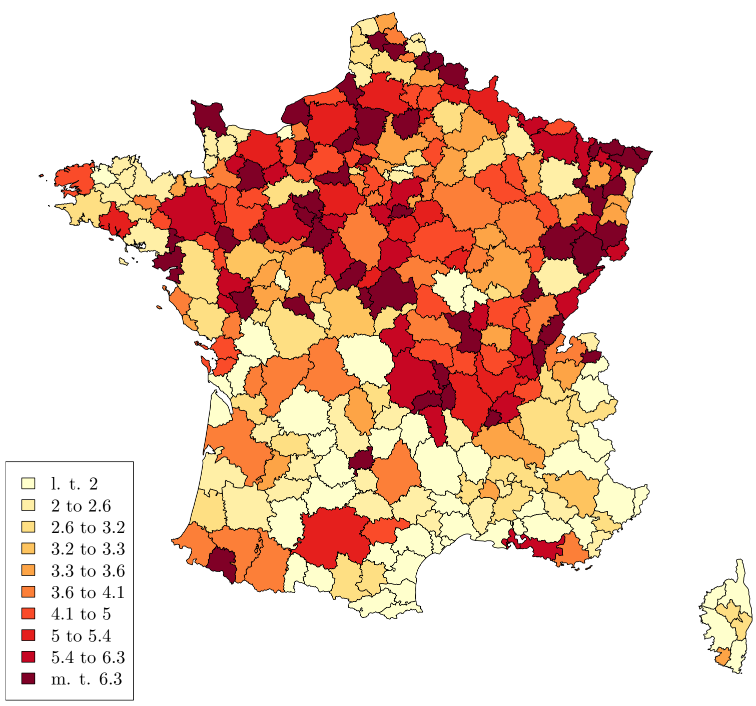
Figure 2: Main explanatory var.: Chinese imports competition exposure (1990-2018) − Av. loss of manuf. output per worker through Chinese imports (in th. of 2022 U.S. dollars) [UN, *Comtrade*; INSEE, *Recensement*]

Figure 3: Some major control variables

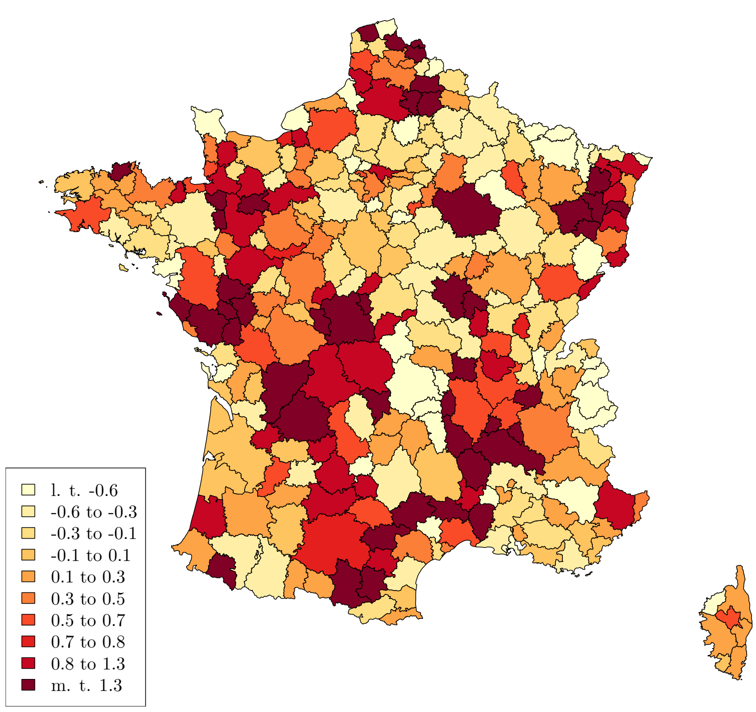
(a) Share of manufacturing in total employment in 1990 [INSEE, *Rec.*]



(b) Rise in the number of manufacturing robots per worker (1990-2018) [Acemoglu and Restrepo 2020; INSEE, *Rec.*]



(c) Offshorability index in 1990 [Autor and Dorn 2013; INSEE, *Rec.*; INSEE, *E.E.*]



extreme results in our replication: the top 3 instruments in terms of Rotemberg weight are in order (all three for the decade 1999-2008), "computers, micro-electronics and optics", "apparel, leather and footwear" and "steelworks (machinery and equipment excluded)"; in our setting, it seems there is more congruence between the descriptive statistics of the China shock and the estimated weights.;

- *1.2. Testing for the relevance of shift-share instruments – Alternative sectoral-regional clustering* – [Adão, Kolesár, and Morales 2019], constructing randomised sectoral placebo shocks and then applying a shift-share instrumentation strategy exactly similar to the Autor-Dorn-Hanson framework, find that the null hypothesis of zero employment effect of that random shock is frequently rejected. They venture two alternative methods to construct shift-share-designs dedicated standard errors. We applied these two methods, using the INSEE's superzone as the meta-regional level, and the ISIC rev. 4 two digits subdivisions as the meta-sectoral level; if we focus on the first coefficient in column (6) of tab. 3, the raw standard error without clustering is 0.72; the S.E. we use throughout this thesis, clustered at the superzone level, is 1.33; S.E. using their AKM and $AKM0$ methods are respectively 1.204 and 1.494, with corresponding p-values all below 0.05;

- *2. From 2SLS to OLS* – In [Autor, Dorn, and Hanson 2013], the authors find a strong difference between the OLS and 2SLS estimates (their main coefficient of interest, equivalent our first line in model (6) of table 3, is reduced from −0.596 to −0.17 when switching back to an OLS specification), indicating that there was a sizeable problem of endogeneity because of local demand shocks. In our setting, we rarely find such large gaps, even if we always reject the null in the Hausman test, indicating that 2SLS estimation is still to be preferred;

- *3.1. Placebo tests for reverse causality on ancient data* – The intuition is to regress *past* changes in manufacturing employment (1975-1990) on *future* trade exposure indexes (from 1990 to 2018); no correlation of any kind is observed;

- *3.2. Placebo tests for reverse causality on contemporary data* – The same exercise is applied on recent data with the first (1990-1999) and the second decade (1999-2008), finding no correlation;

- *4.1. Alternative sources* – The matched employer-employee DADS datasets provided by the INSEE are localised (in decreasing order of size) at the level of the *région*, of the *département*, and of the *Zone d'emploi - ZE*; yet the definition of these geographical unit has changed over time (regions in 2015 and ZEs twice, in 2010 and 2020). Using DADS data therefore implies, either a restriction to a shorter time period in order to remain at the ZE level (this is the choice of [Malgouyres 2017a]) or on the contrary the choice of a larger geographical unit of interest like the *département* in order to maintain a longer time period (this is our choice in section 3.2.1.). Yet throughout most of the results of this thesis, we'll be using Census instead of DADS data; issues of the *Recensement* provide since 1962 a sectoral decompozition at the *commune*-city level, from which it is possible to aggregate to any geographical unit of interest; besides, there is no more sectoral granularity over the DADS than over the Census data we access to;

- *4.2. Alternative explanatory and dependent variables* – [Autor, Dorn, and Hanson 2013] discuss the relevance of an alternative explanatory, i.e., the very same ΔIPW index, but normalised, not by the whole employment of the zone, but by manufacturing employment only, or to tell it plainly, the exposure, not by worker, but by industrial worker. As to alternative dependent variables, we tested the evolution of manufacturing employment divided by the total working-age population (the dependent of [Autor, Dorn, and Hanson 2013]) or by the total labour force, keeping the evolution of total industrial employment in pp or in log points to ensure comparability with existing European estimates. In our annex, we report a wide range of results using an alternative specification with the alternative ΔIPW described above (and the corresponding instrument) and, as dependent, the evolution of manufacturing employment divided by the total labour force;

- *4.3. Alternative instrument* – In tab. 6, we have recourse to an alternative instrument with exclusively non-UE countries (Japan, Australia, Canada & New Zealand). The estimation is unaltered;

- *4.4. Alternative trade partners* – Autor, Dorn & Hanson find similar negative impacts of trade exposure for other emerging economies which are key trade partners of the U.S., particularly Mexico. We therefore replicate our whole setting with the main European developing trade partner of France – Poland – and the main non-European developing partner behind China – Turkey. In both cases, we find some significant effects, but which are not robust across specifications. Since this might be due to the fact that these countries are minor partners in terms of sheer trade value compared to China, we replicate our setting with France's key trade partner: Germany (using our extra-EU instrument). Over 1990-2008, we find evidence of a significant negative marginal impact of exposure to imports from Germany, approximately half the size of our estimates for China. In decade-per-decade breakdown, we manage to explain 23% of the industrial decline over 1990-2008 through German competition on imports (see table 6). This computation is however somewhat artificial, since competition from European partners comes mainly from competition on export market shares.

- *5. Variance breakdown exercise* – Having computed OLS and 2SLS coefficients, we can implement a variance breakdown exercise to discriminate the share of variance in industrial decline explained by exogenous changes in trade exposure (the Chinese supply shock) from endogenous changes (the home demand channel); we find that the latter one accounts for 93% of the total for the 1990-2008 period;

- *6. Intermediate versus final goods* – Over our UN-Comtrade data, we have recourse to the BEC classification to distinguish imports of intermediate and of final consumption goods; the former one only can be considered as exemplifying true import competition, the latter one being a consequence of offshoring. When we replicate specification (6) of table 3 with final goods only, we get $\hat{\beta}_1 = -6.37$ ($t = 3.43$); on intermediate goods, we get a nonsignificant positive coefficient;

- *7. Spatial spillovers* – As opposed to [Autor, Dorn, and Hanson 2021] and concurrently with [Adão, Arkolakis, and Esposito 2019], we find in the spatial econometrics exercise of our last annex (see our tab. 65), some evidence of strong spillover effects across neighbouring ZEs when it comes to the employment effect of the China shock. [Dorn and Levell 2021] impute these reactions to a local suppliers' channel, i.e., the industrial decline in one region is doomed to hurt the suppliers of nearby zones.

Table 6: Robustness checks − Alternative instrument, alternative trade partner

Dep. : Decadal change in total manuf. employ. over 1990-2008 (in pp)						
	China				Germany	
	(1)	(2)	(3)	(4)	(5)	(6)
Rise in import exposure from specified exporter						
Panel A. OLS estimates						
β_1	-6.03^{***}	-3.78^{***}	-6.03^{***}	-3.78^{***}	-1.45^{**}	-1.31^{**}
SE	(1.39)	(0.89)	(1.39)	(0.89)	(0.62)	(0.51)
R^2	0.28	0.61	0.28	0.61	0.21	0.43
F-stat	116^{***}	27.5^{***}	116^{***}	27.5^{***}	75.9^{***}	45.2^{***}
Panel B. 2SLS estimates						
β_1	-6.83^{***}	-4.18^{***}	-6.81^{***}	-4.61^{***}	-1.57^{**}	-1.56^{***}
SE	(1.41)	(1.21)	(1.61)	(0.72)	(0.77)	(0.49)
R^2	0.27	0.61	0.27	0.604	0.19	0.61
F-stat	114.6^{***}	27.3^{***}	113.2^{***}	27.5^{***}	74.7^{***}	29.5^{***}
First-stage:						
Original instrument	0.96^{***}	0.88^{***}				
	(0.05)	(0.09)				
Extra-EU instrument			0.72^{***}	0.75^{***}	2.61^{***}	2.28^{***}
			(0.03)	(0.04)	(0.09)	(0.18)
R^2	0.91	0.92	0.92	0.92	0.86	0.91
F-stat	5801^{***}	210^{***}	5326^{***}	211^{***}	3686^{***}	191^{***}
Controls		X		X		X
Obs.	608	608	608	608	608	608
Av. national $\Delta IPW_{1999,2008}$	2.289	2.289	2.289	2.289	3.608	3.608
Explained share of manuf. decline		0.36		0.39		0.23

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: The unit of observation is the ZE (*Zone d'emploi*, definition of 2010). The dependent variable is the change (in pp) of total manufacturing employment within the ZE. The main explanatory variable is the index ΔIPW, described herein above, which provides an estimation of the mean rise in Chinese imports (in value) per worker (in 2022 kUSD) within each ZE. The original instrument is the same ΔIPW, in which French trade data has been replaced by a control group of four countries (Japan, Germany, Spain, Switzerland) and all labour force variables are taken with a decadal lag; the extra-EU instrument is similarly built with a control group made out of Japan, New Zealand, Australia & Canada. Observations are weighted by the start-of-the-decade total population of the ZE. Standard errors are clustered at the level of the INSEE superzones. The shares explained are computed using the variance breakdown exercise reported in the corresponding annex.

2 The longer term impact – A simple New Keynesian model

2.1 Formal setting

Theoretical debates about the convergence paradigm at the turn of the millennium

If the academic and nonacademic literature of the early 1990s was suffused into a general atmosphere of optimism on the impacts of trade over inequality *within* a country a region, it was equally optimistic as to the *between*-region impact of trade shocks; the new classical theme of *convergence* pervaded the research of that time, best exemplified by the famous set of articles of R. Barro and X. Sala-i-Martin about income convergence across U.S. states [Barro and Sala-i-Martin 1990; Barro and Sala-i-Martin 1992] with the general wisdom that poorer states were growing faster.

There were arguably concerns about the possibility of some zones declining so sharply that they would permanently remain left-behind and depressed. The combination of offshoring strategies of international firms on the one hand, and of increased trade competition on the other, was anticipated to put extra pressure on declining industrial bastions still badly hurt by the crises of the last two decades; in highly centralised European economies, such regional shocks were doomed to be converted into sizeable local unemployment hikes. This is a common fear in the official reports of that time:

> *In the age of globalisation, international firms are urged to swiftly adapt their production networks in order to achieve an optimal use of the economies of scale of each territory. In a sense, they are to monitor a sort of international competition between regions [....]. Noncompetitive zones are then faced with a dilemma: either they let local real wages shrink to ease up the adjustment, or they accept growing unemployment figures. [...] This entails from Nation States further redistribution to the most vulnerable regions and populations, those most likely to bear the consequences of the aforementioned offshoring strategies, or of trade competition.* [Commissariat général du Plan 1993, p. 11 sq., personal translation]

In order to counter these distortionary effects, economic governance institutions were promoting three types of adjustment mechanisms: 1. Relief transfers targeted at manufacturing workers specifically hurt by import competition, in the spirit of the Kennedy-era *Trade Adjustment Assistance* (TAA); 2. Adjustment through out-migrations (facilitating *mobility* was one the main mottos of the OECD *Job Study* of 1994, and quickly became a leading policy objective in European countries); 3. Last but not least, adjustment through wages. In the early 1990s, economic institutions started to promote more decentralised wage negotiation mechanisms, on theoretical and empirical grounds alike; the theoretical justification being provided by *insider-outsider* models, and the evidence, by the catastrophic unemployment figures of Scandinavian countries with traditionally centralised negotiation systems (over 1990-1995, unemployment went up from 3 to 17% in Finland, from 2 to 10% in Sweden).

It is from these types of mechanisms that depressed regions were urged to expect redemption. After all, macro policies of the late 1980s had explicitly opted for vast employment transfers between regions, most notably at the expense of ancient industrial bastions (Britain being the most well known example).

However, in the age of the great neokeynesian models, it was hypothesised that adjustment through transfers and wages could be insufficient to bring back prosperity to the depressed regions. Policies which explicitly accepted large hikes in *between*-region inequality through industrial decline (typically, the British economic strategy of the early 1990s), could create cumulative processes of outmigrations and decline of local demand, transforming transitory local employment shocks into long-term *premia* or penalties to local job dynamics. One of the most quoted of the neokeynesian models of the time addressing that issue is [Blanchard and Katz 1992]; it is this model that we will adapt to our framework.

Towards an augmented Blanchard-Katz model

According to the original model, we'll be using both employment and income figures. Our main variables shall be:

- The unemployment rate among the adult population in region c, centred around the national average[12], denoted $u_{c,t}$. Its opposite is the employment rate among the adult population, $e_{c,t}$;
- The logged evolution of the total stock of employment (also centred) denoted $\Delta n_{c,t}$[13];
- The logged participation rate (also centred) denoted $p_{c,t}$; these three are drawn from the INSEE's Census;
- Data on log yearly wages are from the DADS, data about fiscal income, from the IRCOM database.

The model itself starts with a neoclassical labour demand curve; as usual, it is downward slopping (low local wages w_c make a region attractive and drive local employment n_c up). d is therefore the elasticity of labour demand, and z_c the intercept of the demand curve:

$$w_{c,t} = -d \times n_{c,t} + z_{c,t} \tag{6}$$

[12]It is the usual practice, in the original article, as in the subsequent literature, to use the raw rate of unemployment and not its evolution, though there is evidence from the Dickey-Fuller tests that the bulk of local processes might not be stationary; this was the case in the original article; it is also the case in our replication, whether we use the city or the employment zone level. Minimising four criteria for the choice of the order of the AR univariate model (Akaike, HQ, SC and FPE) would lead us to choose $p = 2$ for this variable.

[13]Formally, it is the log value of employment in one geographical zone at time t, minus the same logged value at time $t-1$, centred around the national evolution between $t-1$ and t. As in the original article, the Dickey-Fuller tests suggest that for the immense majority of cities or ZEs, logged employment is $I(1)$ while the evolution of logged employment is stationary. Minimising four criteria for the choice of the order of the AR univariate model (Akaike, HQ, SC and FPE) would lead us to choose $p = 2$ in most cases for this variable.

On the short term, local labour supply is assumed to be infinitely inelastic.

Then a time dimension is added. The labour supply line and the intercept of the labour demand line z_c are allowed to evolve over years. Both are modelled as random walks with a white noise ε, a drift α, and a relation to wages:

$$z_{c,t+1} - z_{c,t} = -a \times w_{c,t} + \alpha_c^d + \varepsilon_{c,t+1}^d \tag{7}$$
$$n_{c,t+1} - n_{c,t} = b \times w_{c,t} + \alpha_c^s + \varepsilon_{c,t+1}^s \tag{8}$$

The global dynamics are straightforward: a region with high local wages will attract interstate migrants but it will also discourage firms from coming: a and b go for the mobility of firms and workers.

Solving the model brings two difference equations, for wage and employment growth:

$$w_{c,t+1} = (1 - db - a)w_{c,t} + (\alpha_c^d - d \times \alpha_c^s) + (\varepsilon_{c,t+1}^d - d\varepsilon_{c,t+1}^d) \tag{9}$$
$$\Delta n_{c,t+1} = (1 - db - a)\Delta n_{c,t} + (b \times \alpha_c^d + a \times \alpha_c^s) + (b \times \varepsilon_{c,t+1}^d + \varepsilon_{c,t+1}^s - (1-a)\varepsilon_{c,t}^s) \tag{10}$$

From which it is easy to deduce steady-state values:

$$\overline{w}_c = \frac{1}{a + db} \times \alpha_c^d - \frac{d}{a + db} \times \alpha_c^s \tag{11}$$

The wage is set to remain stationary around a state-specific average determined by local institutions (the αs) and by national dynamics : the higher the inter-region mobility of workers b, the higher the specialisation of a region in one bundle of goods (higher a), the more important the local institutions in the determination of wages.

The mean employment growth rate, on the other hand, is non-stationary: local institutions have long-term diverging consequences (an innovation in local labour demand drives new migrants in, which stimulate local demand, creating new jobs, and so on ...):

$$\overline{\Delta n}_c = \frac{b}{a + db} \times \alpha_c^d + \frac{a}{a + db} \times \alpha_c^s \tag{12}$$

Local dynamics are not limited to employment: we must include workers transitioning to unemployment and those leaving the workforce, hence the introduction of the unemployment rate u and the workforce as a share of the total adult population l:

$$z_{c,t+1} - z_{c,t} = -a \times w_{c,t} + \alpha_c^d + \varepsilon_{c,t+1}^d \tag{13}$$
$$l_{c,t+1} - l_{c,t} = b \times w_{c,t} - g \times u_{c,t} + \alpha_c^s + \varepsilon_{c,t+1}^s \tag{14}$$
$$w_{c,t} = -d \times (l_{c,t} - u_{c,t}) + z_{c,t} \tag{15}$$
$$h \times w_{c,t} = -u_{c,t} \tag{16}$$

We rewrite the labour demand dynamics (7) with (13), and labour demand statics (6) with (15). Labour supply (14) is slightly modified: it is expressed in terms of workforce l_c and unemployment u_c (with $l_{c,t} - u_{c,t} = n_{c,t}$); the main novelty here is that high local unemployment can discourage migrants to come (for fear of higher local taxes or sluggish job market dynamics). Equality (16) is the simplest way to state that higher unemployment drives wages down (it is a simplified WS relation). All coefficients are within $]0, 1[$.

The original article does not explicitly solve this system, but we can easily derive from it two difference equations:

$$w_{c,t+1} = \left(1 - \frac{a + d(b + gh)}{1 + dh}\right) \times w_{c,t} + \alpha_c^d + \varepsilon_{c,t+1}^d - \frac{d}{1 + dh}(\alpha_c^s + \varepsilon_{c,t+1}^s) \tag{17}$$
$$\Delta l_{c,t+1} = \left(1 - \frac{a + d(b + gh)}{1 + dh}\right) \times \Delta l_{c,t} + \left(\frac{b + gh}{1 + dh}\right)(\alpha_c^d + \varepsilon_{c,t+1}^d) + \left(1 + \frac{a}{1 + dh}\right)(\alpha_c^s + \varepsilon_{c,t+1}^s) \tag{18}$$

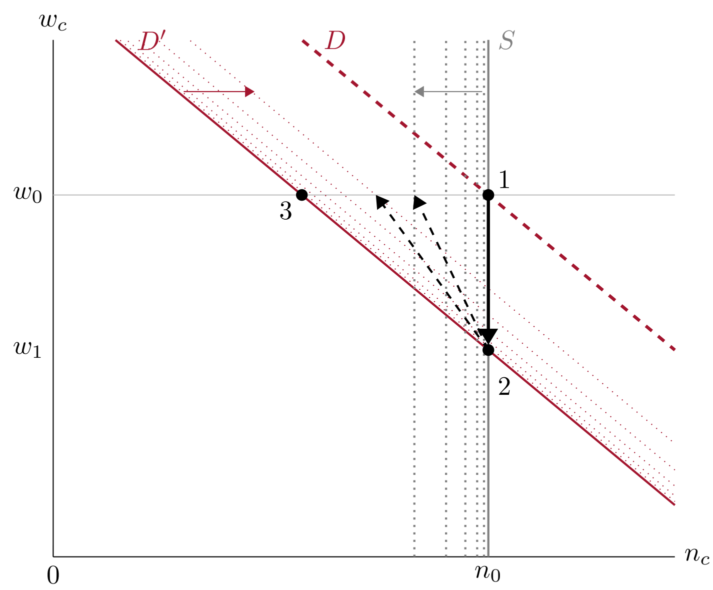
Now let us compute the marginal impact of an innovation of 1 in the demand error term ε^d; we denote this marginal effect with a tilde:

$$\tilde{w}_{c,t} = \left(1 - \frac{a + d(b + gh)}{1 + dh}\right)^t \xrightarrow[t \to \infty]{} 0$$

$$\tilde{l}_{c,t} = \frac{b + gh}{a + d(b + gh)}\left(1 - \left(1 - \frac{a + d(b + gh)}{1 + dh}\right)^t\right) \xrightarrow[t \to \infty]{} \frac{b + gh}{a + d(b + gh)}$$

A negative demand shock (i.e. a negative innovation in the error term ε_c^d) has a direct impact on the labour demand curve: its intercept z_c goes down: graphically (in figure 4, the purple curve D is shifted to the bottom, and the economy transitions from point 1 to point 2. On the short-run, the reaction to a decline in labour demand is a decline of wages. On the longer term, this decline has two opposite effects: faced with lower wages, workers migrate out of the region (S is shifted to the left), but conversely firms tend to migrate to that region to benefit from these lower wages (point D' is shifted to the right). Local economy will return to its pre-crisis wage level, but there will be a permanent negative impact on employment: the new equilibrium will be somewhere on the straight line between point 3 and point 1, the precise location depending on the parameters:

Figure 4: Predicted reaction to a negative local labour demand shock in an augmented Blanchard-Katz model



The impact of each parameter of the model on the amplitude and the persistence of the labour demand shock is summarised in the following table:

		Demand shock impact	
		Persistence of wage shift	Magnitude of employment shift
d	Elasticity of local labour demand	*non-monotonic*	−
a	Interstate mobility of firms	−	−
b	Interstate mobility of workers	−	+
h	Flexible labour market institutions	−	+
g	Reputational premium of a territory	−	+

Significance of the original framework in its policy context

The Blanchard-Katz model has become a widely used tool of policy evaluation; the ECB, most notably, has replicated it several times to gauge the differential impact of a macro shock on European regions [Beyer and Smets 2015b]. In these replications, the neokeynesian undertones of the original model are lost, while the policy recommendations of the original article were clear-cut, and relatively suggestive in their original context. To sum them up:

- If the main policy objective is to resorb short term shocks affecting regions, any form of flexibility goes (flexible labour market institutions / policies fostering between-region's migration);
- But if on the contrary the objective is to limit the long-term impact of a negative shock on employment, the main urge should be to stop the multiplicative effect on local demand (this is the neokeynesian dimension of the model). It implies that:
  - If there's a *migration adjustment*, it must come from firms migrating to the region to benefit from lower wages, not from workers leaving the zone (and depressing local demand further);
  - If there must be a *labour market adjustment*, it is preferable to have higher local unemployment than to let wages do down (because it would accelerate out-migrations of workers and hurt local demand).

In the context of that time, i.e. the great employment transitions of the 1980s in the U.S. and U.K., which had left many old industrial bastions crippled, the model must also be understood as a neokeynesian objection against the idea that workers' mobility and flexible labour markets alone could redeem regions left depressed by employment sags.

2.2 A confounder: inherited regional hysteresis

Factoring out the effect of the national cycle and of regional hysteresis

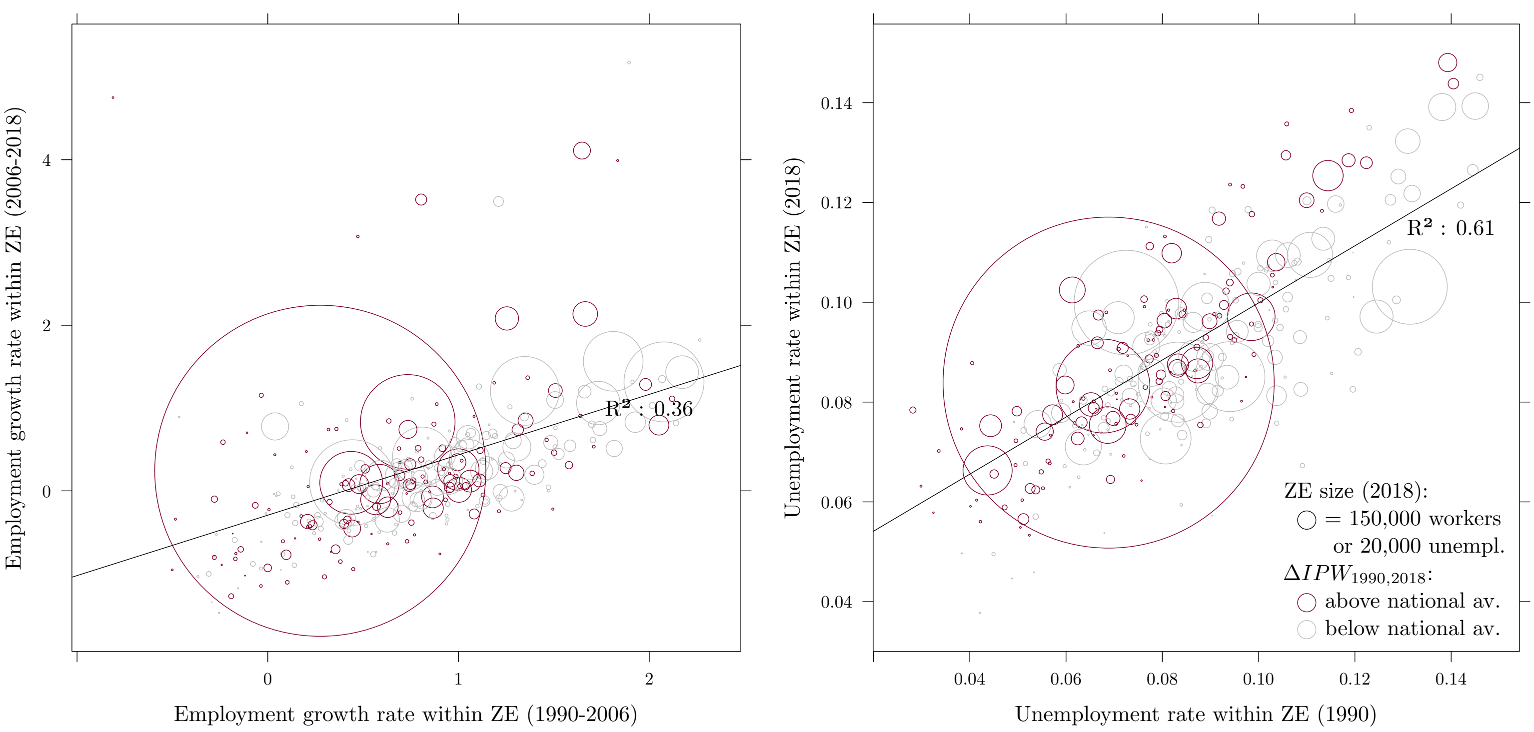
Industrial decline predates the era of trade globalisation. Over our period of study (1990-2018), 1.8 million manufacturing jobs were lost, and the share of manufacturing within total employment dropped by 9.4 pp to reach 13.3%. Yet over the two preceding decades (1968-1990), the corresponding figures are 700,000 jobs lost and a −6.4pp drop in the share. The old industrial bastions, which were known in the late 1960s for their low unemployment rates and frequent labour shortages [Salais 1969], were among those districts which were hurt the most by the crises of the 1970s (see the RHS panel of fig. 53).

There is however little indication that this early industrial decline might be a sizeable confounder to our empirical strategy. The correlation between employment growth rates over the decades before and after 1990 is but very loose (see LHS panel of fig. 53) At the left tail of the regression line, we find old industrial districts which have been hurt by the crisis of the late 20th c. and by the Great Recession alike (Lens-Hénin). At the right tail, on the contrary, we mostly find areas dependent on service economy (with an overrepresentation of the Parisian metropolis), which have grown relatively smoothly over the past 50 years. The fit is loose mainly because the fortunes of districts are strongly polarised across space. At the bottom of the panel, we find regions lying mainly on a Sedan-Nevers line, districts known to have been, in the late 19th c., at the frontier of early human capital accumulation, and early industrialisation; these districts have not been hurt by the crisis of the heavy industries in the 1970s, but have suffered much more during the Great Recession because they were still highly reliant on industrial employment. Conversely, at the top of the panel, we find exclusively areas from the North-West, noted historically for late human capital accumulation - late industrialisation ; these regions have resisted the Great Recession thanks to a employment mix made of agriculture and tourism.; in our robustness exercises (see notably tab. 43), we fail to detect any significant correlation between this early industrial decline and exposure to import competition after 1990; the raw correlation could even be, paradoxically, negative, i.e. the ancient industrial bastions of the North-East had lost so many manufacturing jobs over the 1970s and 1980s, that they entered the era of trade globalisation as relatively protected regions (for the ZE of Valenciennes, which had the second highest unemployment rate of the country in 1990 - 14.7% - we find a $\Delta IPW_{1990,2018}$ about \$2,000 per head below the national average).

Now coming to our time period of interest (1990-2018), employment figures draw a consistent picture of divergence between regions. As noted by [Davezies 2012], during the late 20th c., metropolitan regions enjoyed more dynamic labour markets, but this came at the expense of greater exposure to international competition and greater vulnerability to the phases of the global business cycles. On the contrary, the Great Recession saw the dawn of an era in which major cities are at the same time more dynamic but also more resistant to the cycle; as shown in the LHS of fig. 5, metropolitan ZEs are the clear winners of the post-crisis period in terms of employment growth; though they belong to the districts most exposed to trade competition, the implied negative marginal impact on industrial employment is largely offset by the growth of other sectors; trade competition exposure is a problem for middle-sized ZEs, not for metropolises.

Another significant feature of the 1990-2018 period is the strong persistence of unemployment rates across regions (see RHS panel of fig. 5). Regressing the 2018 on the 1990 rate yields a R-squared of 0.61 (0.86 for the post-crisis period 2009-2018 and even 0.96 over 2006-2009), to compare to the figures of previous periods (0.17 over 1968-1990, [Blanchard and Katz 1992] reporting a similar 0.21 figure for the U.S. over 1970-1990).

Figure 5: The persistence of unemployment and of employment growth within ZEs post-exposure

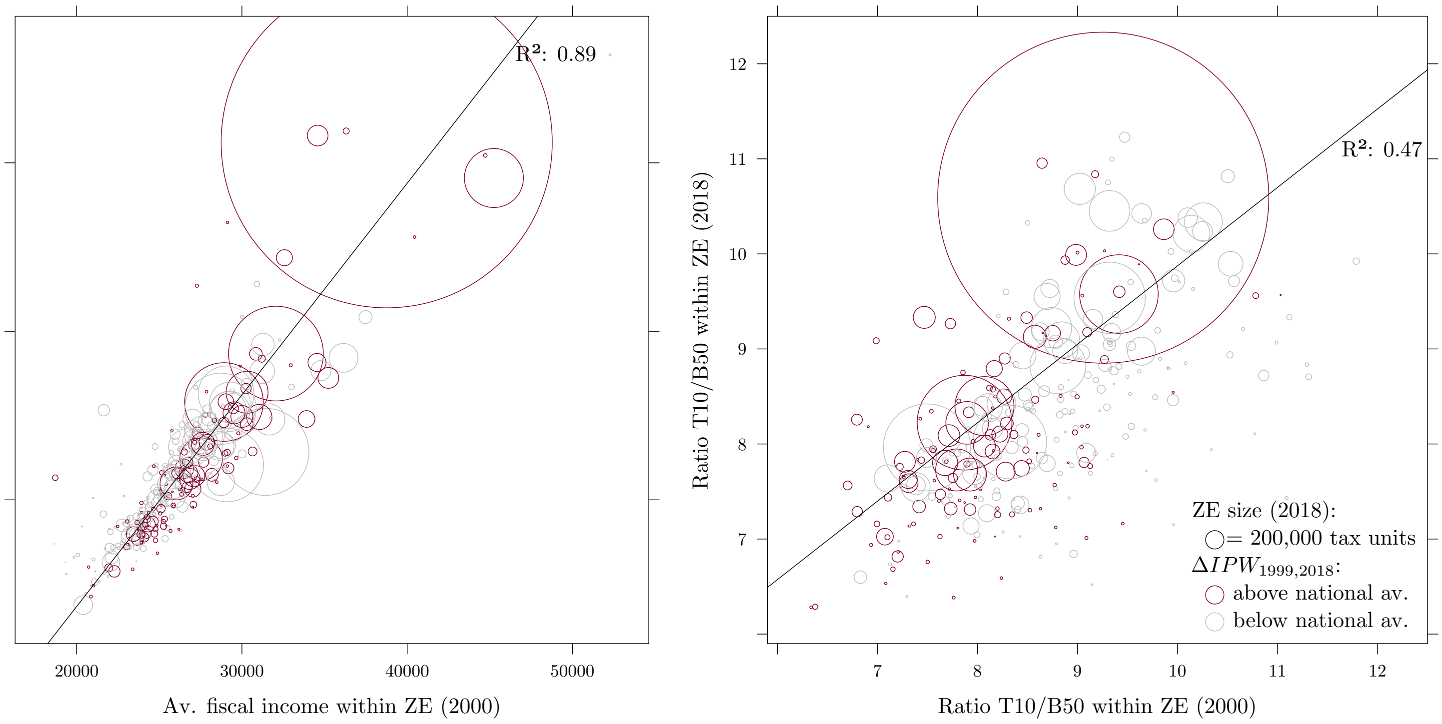


Note: The unit of interest is the *Zone d'emploi* (INSEE 2010 definition). Data are from the INSEE's Census. Reported statistics include the average growth rate of total employment within the ZE (LHS panel) and the unemployment rate (ILO-INSEE definition, RHS panel). Circle sizes provide the related population of the ZE (total number of employed workers, or of unemployed people). A red circle indicates that the average exposure to import competition within the ZE (expressed in 2022 kUSD per worker) is above the national average (and vice versa for grey circles). We plot the regression line of the variable on the y-axis on the variable of the x-axis, weighting by the related population of each ZE.

Consistent with the idea that trade exposure might act as a buffer of other polarising forces – *à la* [Acemoglu 2003] – if we restrict ourselves to the 129 most exposed ZEs (those for which $\Delta IPW_{1990,2018}$ is above the national average), we find, in all four panels of fig. 5, 6 and 53[14] regression lines with a significantly higher slope (with the implicit idea that growth in fast-growing zones is magnified by trade exposure, while laggard regions are further hurt by it).

[14] When it comes to average fiscal income, the persistence over time is almost perfect, with little differential features (districts with a more educated population are generally found above the regression line). When it comes to the ratio T10/B50, we find more persistence still, especially after the Great Recession (R-squared of 0.91 when we regress the values 2018 over those of 2006). The post-crisis era was one of polarised fortunes. For the old industrial districts, it does not seem like the Great Recession brought any change in local inequality (local public spending surely acting as a stabiliser). Service districts saw a slight rise. But more interesting : the industrial districts of the East which were among the most equal, experienced a sharp rise in local variance of income ; for the Western district which resisted the 2008 crisis, it is the very contrary; traditionally unequal, they experienced a sharp decline in the variance of income.

Figure 6: The persistence of the fiscal income and income inequality within ZEs (2000 *versus* 2018)



Note: The unit of interest is the *Zone d'emploi* (the INSEE's commuting zone, 2010 definition). Data are from the INSEE's Census and the IRCOM base. Reported statistics include the average fiscal income within ZE expressed in euros of 2022 (LHS panel) and the population-weighted average of the ratio T10/B50 of the *communes* belonging to each ZE (RHS panel). Circle sizes provide the related population of the ZE (total number of tax units in 2018). A red circle indicates that the average exposure to import competition within the ZE (expressed in 2022 kUSD per worker) is above the national average (and vice versa for grey circles). We plot the regression line of the variable on the y-axis on the variable of the x-axis, weighting by the related tax unit population of each ZE.

This divergence story is further illustrated when we attempt to replicate the well known graphs of convergence of [Barro and Sala-i-Martin 1992], plotting start-of-the-period average income per head within each region, versus the growth rate of income over the period. As we see in fig. 7, over ZEs protected from trade competition, we indeed find a convergence pattern *à la* Barro and Sala-i-Martin (i.e. poorer districts grow faster). Over exposed ZEs on the contrary, we find a significant divergence.

Another, more formal way, to illustrate that persistence of employment and income fortunes of regions is to fit the time series of our main variables for each commuting zone (ZE) into a simple autoregressive model. In almost every case, when the first-differenced series is indeed stationary, the usual criteria tests recommend an $ARIMA(2,1,0)$ model when the period of estimation is 2006-2018, and $ARIMA(3,1,0)$ when the period is 2000-2019. For the employment rate, the main specification writes:

$$\Delta n_{c,t} = \beta_{1,c} + \beta_{2,c}(L)\Delta n_{c,t-1} + \varepsilon_{i,t} \tag{19}$$

Average coefficients of the resulting estimation for each variable are reported in table 45.

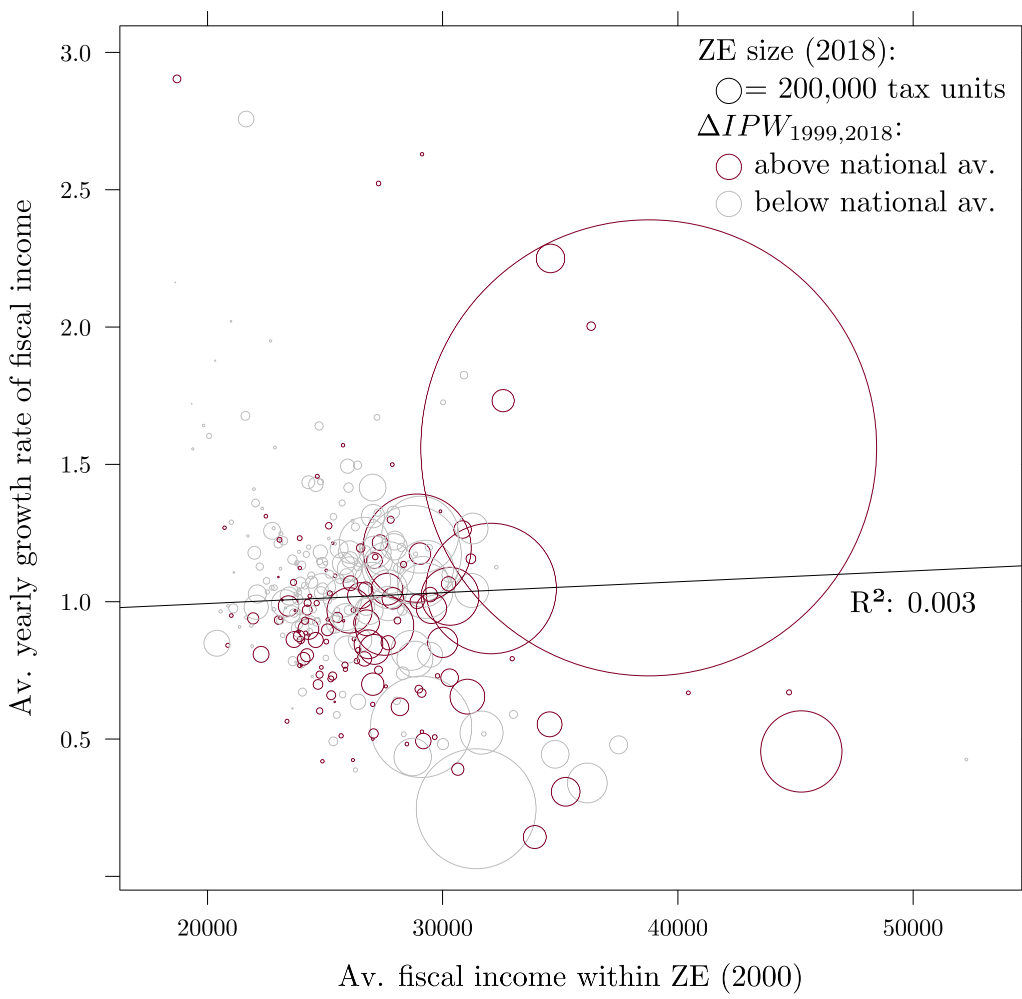
Estimating (19) at the ZE level yields results which are stunningly similar to the original article. In the Blanchard-Katz framework, estimated on 1950-1992 U.S. data, the heterogeneity of employment shocks across regions show no convergence over time; in a typical *Rust bell* state like Michigan, an innovation of 1 in ε will result in an implied response of 0.86 after 10 years and it will remain at that level over the long-run ; at the other extreme, it will be 3.15 for an agricultural state like Wyoming, and around 2 for the *Sun Belt* states. Across the French employment zones or ZEs, we find even more extreme values at the lower tail (0.48 for Châlons-en-Champagne) or at the upper tail, the most dynamic ZEs being found in the North-West or the Southern touristic regions (3.74 for Cannes). Strict comparison is not possible (especially since there are signs that responses are magnified when we focus on tinier territories), but the general conclusion of Blanchard & Katz − "The correct image of employment evolution is one of regions growing at different rates, with shocks having largely permanent effects" − can easily be applied to our framework.

Hysteresic regions exhibit significantly sharper responses to the shock

In the literature of the 1990s, we still find many objections against the Barro and Sala-i-Martin convergence paradigm, especially when applied to inter-regional dynamics in Northern economies. Some of these objections had the idea to transfer the hysteresis scheme of [Blanchard and Summers 1986] to regional inequalities. Such an hypothesis was discussed in European [Baddeley, Martin, and Tyler 1998] and American context [Bartik 1991] alike: hence the concept of regional employment hysteresis, i.e. the idea that past realisations of employment dynamics in a region are like "stored" as a permanent premium or penalty to local employment rates, or to local unemployment figures. Blanchard and Katz were sceptical about the original argument of [Bartik 1991] because they found no significant correlation, on U.S. postwar data, between the growth of employment and unemployment rates. Yet over the INSEE's data, it is straightforward to isolate such a significant correlation, as table 47 shows.

Our preferred strategy to isolate hysteresic regions, is to regress local employment growth rates on the national counterpart, in order to discriminate the local dynamics from the national cycle. This breakdown is a classical exercise in the European literature about regional unemployment dynamics, commonly known as the Brechling-Thirlwall breakdown [Brechling 1967, Thirlwall 1966]; it also serves a more trivial objective, i.e. many European regional time series of employment cannot be made stationary by simple differentiation; regressing regional values

Figure 7: No income convergence between ZEs over 2000-2018, mostly because of exposed ZEs



Note: The unit of interest is the *Zone d'emploi* (the INSEE's commuting zone, 2010 definition). Data are from the INSEE's Census and the IRCOM base. Reported statistics include the average fiscal income within ZE expressed in euros of 2022 and the average yearly growth rate of fiscal income over 2000-2018. Circle sizes provide the related population of the ZE (total number of tax units in 2018). A red circle indicates that the average exposure to import competition within the ZE (expressed in 2022 kUSD per worker) is above the national average (and vice versa for grey circles). We plot the regression line of the variable on the y-axis on the variable of the x-axis, weighting by the related tax unit population of each ZE.

on their national or EU counterpart, and storing the residuals as an independent variable is one simple way to retrieve a stationary time series.

The main specification relies on the logged employment stock, but this time not centred on the national average; we denote it N, the variable without subscript c being the national figure:

$$\Delta N_{c,t} = \alpha_c + \beta_c \Delta N_t + \eta_{c,t} \tag{20}$$

Evaluating (20) for each commuting zone (ZE) will allow us to retrieve three major results:

- The α_c of each region, which might be seen as the simplest way to empirically assess the αs of the model, i.e. the structural attractiveness of the region;
- The R^2 will be an index of the level of synchronisation between the local and the national employment dynamics;
- The coefficient β_c assesses the local reaction to a national shock.

Blanchard and Katz exclude the hypothesis that β_c might be significantly different depending on the phase of the cycle ; i.e. that some districts might have a pro-cyclical reaction in times of crisis, and a counter-cyclical one in times of take-up (and vice versa). Yet on our data, there is a very strong suspicion of it; that is why we will estimate (20) separately for the crisis phase (2007-2012) and the take-up phase (2013-2018).

In the original article, the main outcome of that exercise is straightforward: industrial states have a consistent profile of low-α, high-β, high-R^2; this profile is reversed for oil and agricultural states. Our results at the commuting zone level, exhibited in table 48, indicate a more nuanced picture. We could classify ZEs along two criteria:

- *Employment hysteresis* – We found the regional hysteresis we were searching for in the form of a strong positive correlation between the employment growth rate over 1968-1990 $\Delta N_{c,1968-1990}$ and the α_c estimated through (20), exactly the type of "stored" premium or penalty hypothesised by Bartik[15]. We'll then define a *hysteresis district* as a ZE in which $\Delta N_{c,1968-1990}$ and $\alpha_{c,2006-2012}$ are below their respective national average (a condition which is very near to $\Delta N_{c,1968-1990} < 0 \cap \alpha_{c,2006-2012} < 0$). This pertains to 45% of the ZEs, which accounted for 25% of national employment in 2018. This distinction might be seen as an index of resistance to the crises of the late 20th c., gauging how much local dynamics was fit for the earlier stages of European integration and the globalisation of employment dynamics;
- *Cyclicity* – We propose a simple criterion to discriminate over the cyclical component: a ZE is said to be pro-cyclical if the sum of the normalised elasticities is above 0;.

Table 7: Some descriptive statistics about subtypes of ZEs

	Δ Chin. import. expos.			Share	Automation	Offshorability	Fiscal income (2008-2018)			
	1990-1999	1999-2008	2008-2018	manuf. empl. (2008)	1999-2008	index (2008)	Initial av. inc.	Evol.	Ini. ratio T10/B50	Evol.
synchronised & Counter-cyclical ZEs	284	2120	723	15.3	1.61	0.15	38677.2	−1.3	8.02	−1.07
synchronised & Pro-cyclical ZEs	272	1719	685	12.4	1.67	0.11	41056.5	+4.8	8.88	−0.19
Hysteresic & Counter-cyclical ZEs	266	2389	753	19.7	1.91	0.23	32910.1	−3.9	7.59	−1.01
Hysteresic & Pro-cyclical ZEs	239	1991	645	18.1	2.03	0.08	33445.8	−3.5	7.78	−1.22

*p<0.1; **p<0.05; ***p<0.01

Note: The main variables have been described in 3. We report average values by subtypes of ZEs, computed using, as weights, the total individual population reported in the Census at the related date, or the related start-of-the period year.

To comment with more details about these categories:

- *Hysteresic zones* – With their normal-unemployment + low-employment profile, these districts fit perfectly well within the Blanchard-Katz framework. These are districts with high rates of out-migrations, low rates of participation and high shares of retired workers within the population; i.e. they have been hurt by the crises of the late 20th c., and the adjustment happened through population dynamics. They draw a consistent picture of a less urban, less educated, more working-class labour force. Wages in these regions grow more slowly, and are much more reactive to the cycle, but these remain egalitarian communities, with some of the lowest ratios T10/B50 of the country. But most important for our model: in these hysteresis districts, the elasticities of the employment growth rate to a national shock β_c are consistently different between the upward and the downward stages of the cycle:
  - *Hysteresic & Counter-cyclical zones* – When a hysteresis district has low elasticites, it often means that the elasticity is greater in times of crisis: i.e. the zone loses more on the downward stage of the cycle than it gains on the upward stage. This category pertains mainly to former bastions of the heavy industry that have been badly hurt by the crises the 1970s. The main consequence of employment hysteresis

[15] [Baddeley, Martin, and Tyler 1998] similarly, using UK local unemployment time series which are much more precise than ours, identify clear-cut structural breaks in local unemployment rates of the main Northern industrial regions of the country around the early 1980s.

is that local job structure has been shifted away from the national dynamics[16]; either job losses have been so massive in the late 20th c. that local evolution is now completely desynchronised from national dynamics (Saint-Etienne, Montluçon), or there is still some form of synchronisation, but at the cost of extreme vulnerability to recessions (ZEs on the Nevers-Sedan line or in rural Lorraine are among the rare instances of positive elasticity during the Great Recession & negative elasticity after 2013);

  – *Hysteresic & Pro-cyclical zones* – When a hysteresis district has high elasticities, it generally means that the elasticity is greater in times of recovery. These are mainly rural regions where the massive job losses of the 1970s-1980s were compensated by out-migrations, and which are now synchronised with the business cycle, but highly reliant on the different forms of European protectionism : agriculture, car industry (this is where we find the rural decentralised car industry centres of the 1960s, most notably Peugeot-Sochaux).

- *Non-hysteresic zones* – These ZEs do not have consistently lower unemployment rates, but it is where we find the highest employment and participation rates. Incomes tend to grow faster, unfettered by the employment cycle. If within hysteresis districts we found decorrelated relations between upward and downward elasticities of employment growth, within these regions, there's almost perfect equality between the two β_c. To sum it up, the transition to a globalised economy in the late 20th c. resulted there in massive job gains, with two different ways of integrating to the international division of labour:

  – *synchronised counter-cyclical zones* – These are the areas with no hysteresis and low elasticities; they are mainly metropolises, which are shielded from the cycle by the predominance of service economy and NTIC in their job mix. If there be winners of the so-called skilled-biased technical change, they are to be found among these areas;

  – *synchronised pro-cyclical zones* – These are the areas with no hysteresis and high elasticities; they tend to overreact to the business cycle, be it downward or upward. They are embedded into the globalised economy, but highly dependent on it; these include touristic resorts, but also productive regions protected by Dixit-Stiglitz differentiation (wine valleys, specialised industries ...). These are unequal regions (with local T10/B50 ratios 1 point over the national average, and declining much more slowly than elsewhere).

Can we isolate any connection between these hysteresic pattern inherited from the 1970s-1980s? If we replicate specification (3), introducing crossed variables between the explanatory and dummies for types of ZEs, we find that counter-cyclical are not significantly different from pro-cyclical ones (the coefficient on the crossed being 0.19, t=0.65); desynchronisation from the international business cycle does not paradoxically provide extra protection; in sheer descriptive statistic, counter-cyclical zones are even a bit more exposed. It is quite a different story for hysteresic ZEs; here we find a significantly lower crossed coefficient (-1.97, t=2.16, as opposed to a main one of -3.04, t=2.64); regions which inherited from the 1970-1980s consistently lower levels of employment (a lower α_c in specification 20) suffered more from the shock; unsurprisingly, the effect is driven by extractive industries and chemistry.

We must therefore heed to the fact that the Brechling-Thirwall breakdown of specification (20) applied to retrieve transformed, stationary time series, factors out part of the structural parameters of the employment impact, most importantly the hysteresic, stored employment premia theorised by Bartik we discussed above. The factored-out variables we'll now be using are the purely regional variations in the employment and income dynamics, independently of the national and global cycle (the impact of the elasticities β_c), and independently from the structural attractiveness of the territory inherited from past employment realisations (the a_c).

2.3 Predicted *versus* actual longer term effects

A panel VAR framework

The most straightforward empirical transcription of our model is a panel VAR setting with $\Delta n_{c,t}$ as the causal determinant:

- Formally, it implies that the evolution of the employment stock $\Delta n_{c,t}$ is allowed to affect current values of other variables (the employment and participation rates e and p), but not vice versa. At $t = 0$, there is a one standard deviation innovation in the error term of the driving variable, denoted $\eta^n_{c,t}$, which sets the VAR model in motion;

- From a theoretical standpoint, this setting is meant to embody the main assumption that the root cause of the persistence of labour employment fortunes and misfortunes in recent decades were differential labour demand shocks; i.e. local employment declined because the bundle of goods the region was producing had become obsolete, or because it was now possible to produce it elsewhere at a lower cost, or because production could easily be automated, etc.[17]

[16] This shifting has been a classical object of analysis for socialist economics in the late 20th c. France. Postalthusserians interpreted it through a revised version of the R. Vernon product lift cycle model [Vernon 1966], arguing that with the decline of heavy industries and the correlative destructuring of working-class political organising, these regions would become the dominated margins of the knowledge economy; i.e. innovation, R&D, and the early stages of production of a new item would be concentrated within the metropolises; later on, at the stage of mass production, employment would be decentralised to these peripheral regions, where producers would benefit from a local depoliticised & underpaid labour force [Castells 1972; Lipietz 1977]. The turn-of-the-millennium was in fact of quite a different nature: as emphasised by [Damette 1994], it was not a reconfiguration of the structure and of the connection between the dominant and exploited part of it, but on the very contrary, a transition from the "integrating domination" of the postwar fast-growth era (where inequalities between regions were considerably higher than they are today, but where laggard regions had an effective role in the production process) to what he calls a "forsaking domination" where between-region inequality is indeed far lower, but where laggard zones are utterly desynchronised from the cycle, reliant mainly on public employment, transfers and tourism.

[17] The authors argue that a convincing argument in favour of the labour demand shocks hypothesis is the strong negative correlation between local employment growth rates and local unemployment growth rates: if the proximate cause of regional labour misfortunes had been labour demand shocks, we should find in the data a *negative* correlation between employment growth rates and unemployment rates (i.e. employers hire massively to respond to demand shocks). If on the contrary, the proximate cause is the impact of labour supply shocks, that correlation should be *positive* (migrants arriving face some sort of "wait unemployment"). What the authors found on U.S. data is a negative (but not significant) correlation; on our data, we find an even more robust negative correlation, see table 47.

The main specification for this panel VAR equation writes:

$$\begin{pmatrix} \Delta n_{c,t} \\ e_{c,t} \\ p_{c,t} \end{pmatrix} = \begin{pmatrix} \gamma_{c,1,0} \\ \gamma_{c,2,0} \\ \gamma_{c,3,0} \end{pmatrix} + \begin{pmatrix} 0 & \gamma_{c,1,2} & \gamma_{c,1,3} & \gamma_{c,1,4} \\ \gamma_{c,2,1} & 0 & \gamma_{c,2,3} & \gamma_{c,2,4} \\ \gamma_{c,3,1} & 0 & \gamma_{c,3,3} & \gamma_{c,3,4} \end{pmatrix} (L) \times \begin{pmatrix} \Delta n_{c,t} \\ \Delta n_{c,t-1} \\ e_{c,t-1} \\ p_{c,t-1} \end{pmatrix} + \begin{pmatrix} \eta^n_{c,t} \\ \eta^e_{c,t} \\ \eta^p_{c,t} \end{pmatrix} \tag{21}$$

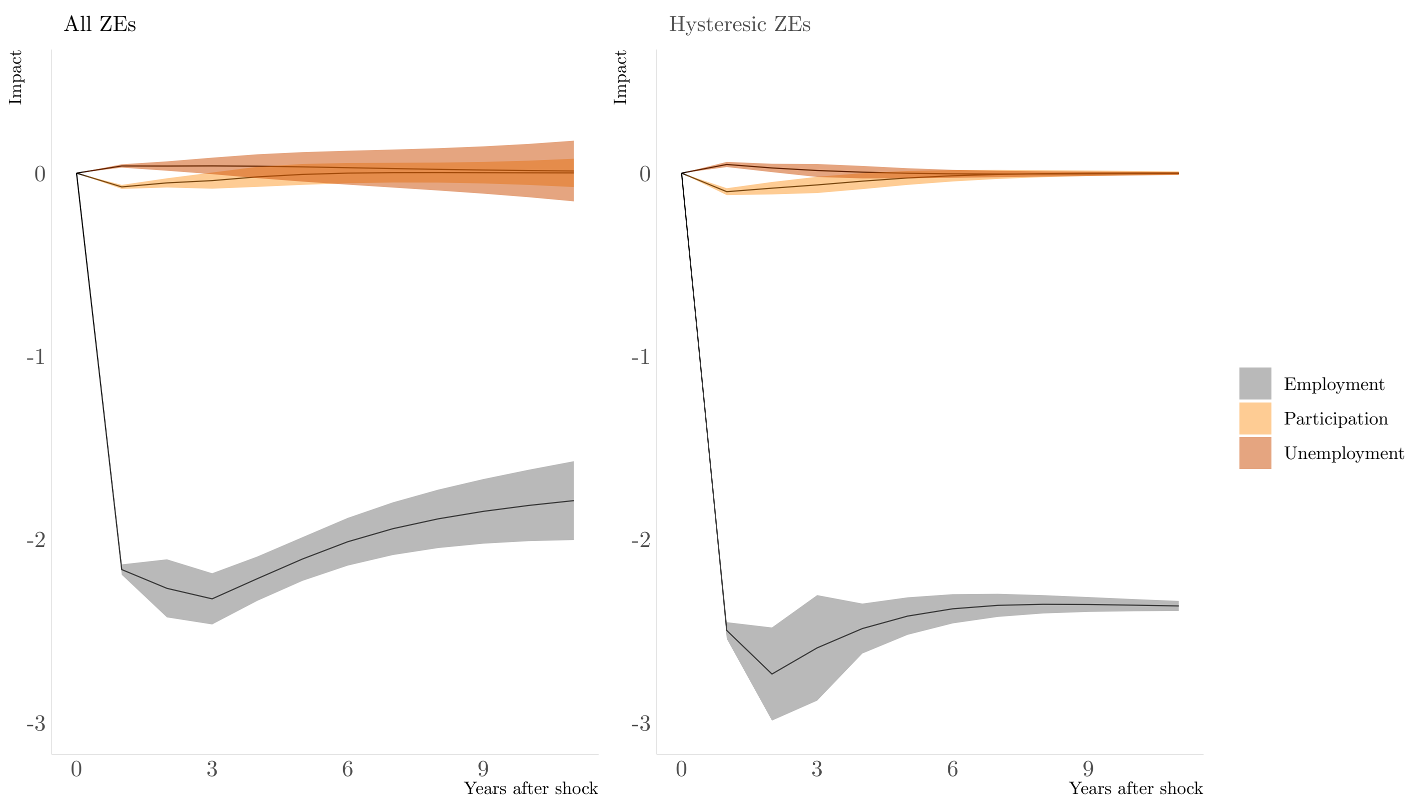
We allow for two lags, like the original model, and we apply some further adjustments to fit that model to the European context[18].

If the employment dynamics following a shock follows the predictions, we should see, following a negative labour demand shock, employment and income figures which remain permanently stuck at a contracted level.

Employment response

The main results of the estimation of (21) are reported in figure 8: from Δn_c and e_c, it is easy to deduce the dynamics of raw employment and unemployment, which are drawn in the figure:

Figure 8: OIRF estimates of the long-term impact of a general employment shock at the ZE-level using panel VAR specification (21)



Note: The unit of interest is the *Zone d'emploi* (ZE, INSEE 2010 definition), denoted by subscript c. Defining N, L, U, and P respectively as total employment, labour force, stock of unemployed people, and total adult population, we fit in the panel VAR model (21) three main variables: the year-to-year employment growth rate $\ln(N_{c,t}) - \ln(N_{c,t-1})$, the employment rate $\ln(E/L)$ and the participation rate $\ln(L/P)$. Data are drawn from the INSEE's Census, collected at the *commune* level and then aggregated at the level of the ZE. To ensure that each variable is stationary, we applied a Brechling-Thirlwall breakdown, regressing each time series of the local variable on its national counterpart, retrieving the residuals and interpreting them as the purely local component of employment dynamics. The Cholesky ordering of the OIRF setting is: employment growth rate - employment rate - participation. The impact of unemployment has been reconstructed as the opposite of the logged employment rate. Employment has been reconstructed cumulatively. An hysteresis zone has been defined herein above by two conditions: the employment growth rate between 1968 and 1990 has been below the national average, and for which the intercept of the Brechling-Thirwall breakdown is below the national weighted mean of all zones (i.e. a zone for which past job destructions have resulted in a permanent penalty to the employment growth rate in recent data). We report 95% conf. intervals.

In our model, a worker who loses its job has three options: migrating, exiting to unemployment, or exiting to inactivity. In the original article, the main conclusion is that, at the apex of the shock of -1SD, unemployment jumps at $+0.3$ SD, while participation declines at -0.05 SD, i.e. migrations account for 65% of the adjustment. In the existing European literature, generally focused on the region level [Beyer and Smets 2015a; Lesuisse 2020], it oscillates around 70 and 80%. In our setting, it is above 90%, which is not surprising since we use a spatial

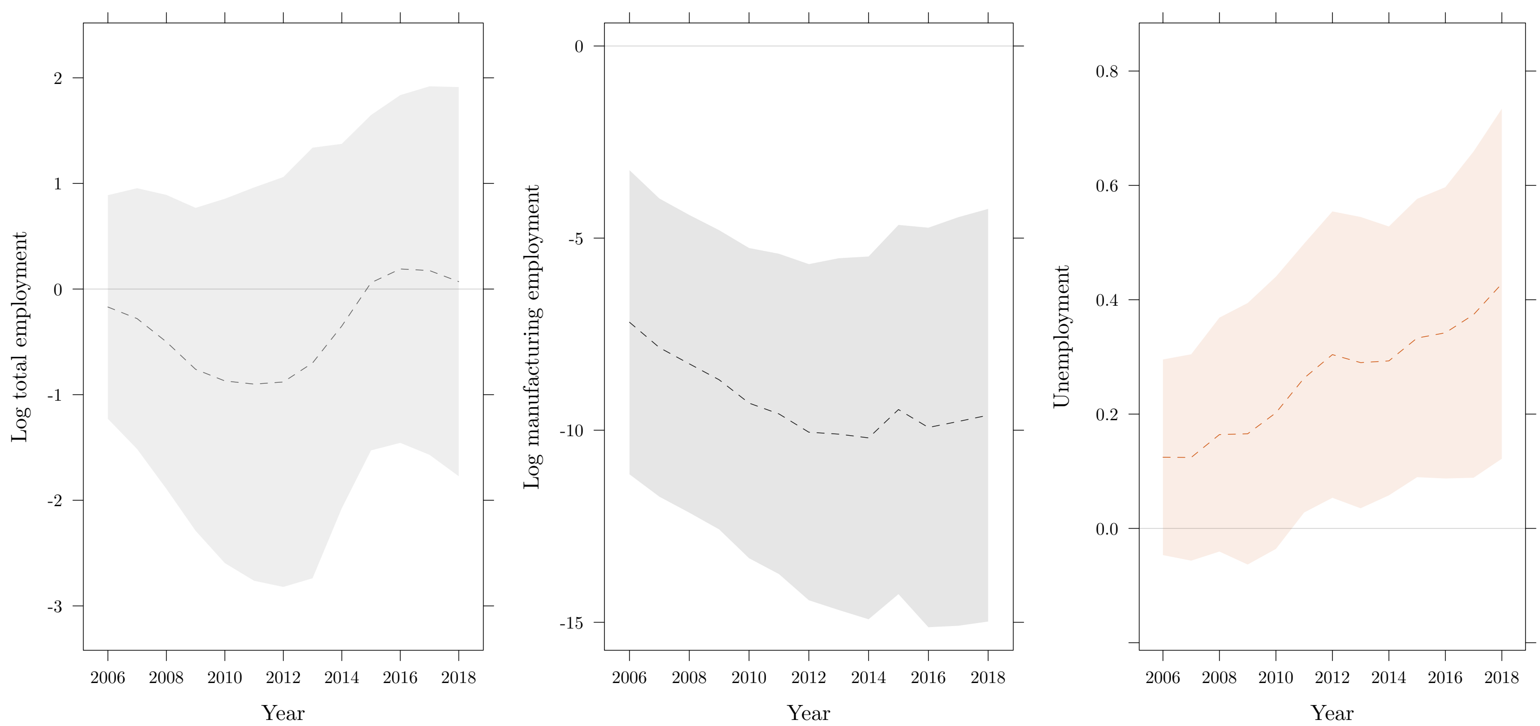
[18] Following the recommendations of the econometric literature [Baltagi, Griffin, and Xiong 2000], we estimate (21) by pooling all ZEs together, allowing for zone fixed effects computed through FOD (and not first-differences, in order to preserve our time interval which is still very short). As emphasised in the subsequent literature, attempts to replicate the Blanchard-Katz model in European context face a stationarity issue: the employment rate e and the participation rate p have a unit root in most regions when we use the logged variable minus the logged national mean value. The main correction used in the recent literature [Halleck-Vega and Elhorst 2016, Halleck-Vega and Elhorst 2017, Lesuisse 2020] focuses on the idea of cyclical sensitivity and common factors: i.e. if regional variables tend to oscillate in parallel to the national average, some regions might over or under-react to national shocks. Here it means applying to each variable the Brechling-Thirlwall breakdown that we applied herein above to N and u ; i.e. regressing the time series of the local variable on its national counterpart, and storing the residuals, which are then interpreted as the specific regional trend of the variable. We used the array of unit root tests of the *xtunitroot* function of STATA to check for stationarity; for every single variable obtained through the aforementioned adjustment, we reject the hypothesis that panels have a unit root with a 1% risk at least. For all the panel VAR estimates reported in this section, we have checked that our PVAR model is stable by computing the moduli of the companion matrix based on the estimated parameters. Applying the Granger causality test, we confirmed that Δn has a causal impact on the other variables (with a 5% risk) in every setting.

unit roughly 20 times thinner than most European studies, 60 times thinner than the U.S. states of the original article[19]. Even when we attempt breakdown by subtypes of ZEs, unemployment and inactivity account for 10% at most of the adjustment.

To summarise these estimates, we might say that at the level of the ZE, a one standard deviation shock to the employment stock, or equivalently a decline of the stock of -2.1 log points, should yield: a maximum negative impact on employment of -2.35 log points after three years, and -1.74 after ten years; $+0.04$ log points to the unemployment rate in the short run, the impact vanishing after 3 years; -0.08 log points to the participation rate in the short run, the impact vanishing after 4 years.

In figure 9, we compare these values with the actual ZE-level longer term effects of the China shock, estimating model 3 using the 1999-2008 exposure as the main explanatory, and the evolution of some major employment variables as the dependent, in the spirit of [Acemoglu, Autor, et al. 2016]. When the shock is fully realised, at the end of the decade[20], it brings a -7.85 log points decline to the manufacturing employment stock, and a -0.51 log points decline for to the entire stock. We see that the impact is fully realised approximately 12 years after the beginning of the exposure decade, a result consistent with structural [Caliendo, Dvorkin, and Parro 2015] and reduced form approaches [Autor, Dorn, and Hanson 2021] alike. The impact on industrial employment decline is considerable, with no sign of recovery in the longer term. The short-term reaction of total employment and unemployment are in the spirit of the Blanchard-Katz model, but their longer term fortunes are not[21]; total employment reverts to its pre-exposure values, while the marginal impact on unemployment continues to be felt 20 years after the start of the shock. It is like industrial employment is becoming an autonomous, desynchronised section of the economy, a king of sectoral hysteresis[22].

Figure 9: Long-term impact of the China shock (1999-2008 import exposure) on employment within ZEs



Note: The unit of interest is the ZE. Employment data are from the INSEE's Census. We estimate model (3) using the full vector of controls, taking as explanatory the import competition exposure index $\Delta IPW_{1999-2008}$ instrumented in the way described herein above, and as dependent, the evolution of the mentioned employment variable (log total employment, log manufacturing employment, unemployment rate) between 1999 and the year mentioned on the x-axis. Estimation is done for each year over 2006-2018, weighting observations by the total population of the ZE, and clustering S.E. at the INSEE superzones level. We report 95% conf. intervals.

Wage response

We now expand the panel VAR setting to wages, which are allowed to be influenced by the driving variable only:

$$\begin{pmatrix}\Delta n_{c,t}\\ w_{c,t}\end{pmatrix} = \begin{pmatrix}\gamma_{c,1,0}\\ \gamma_{c,2,0}\end{pmatrix} + \begin{pmatrix}0 & \gamma_{c,1,2} & \gamma_{c,1,3}\\ \gamma_{c,2,1} & 0 & \gamma_{c,2,3}\end{pmatrix}(L) \times \begin{pmatrix}\Delta n_{c,t}\\ \Delta n_{c,t-1}\\ w_{c,t-1}\end{pmatrix} + \begin{pmatrix}\eta^n_{c,t}\\ \eta^w_{c,t}\end{pmatrix} \tag{22}$$

It is difficult to detect any significant reaction when we estimate (22) at the ZE level. Estimated at the grand region level, on the contrary, we find a reaction which is very near the usual results of the literature (see fig. 11). We lack the corresponding DADS data to conduct a long-term analysis of the China shock for wages. We propose a decade-per-decade estimation for 2002-2008 and 2008-2018 in section 3.2.1.; when we attempt to estimate the impact of the 1999-2008 decade of exposure on the wage evolution over 2002-2018, it suggests that the negative impact found at the apex of the shock (at the end of the corresponding decade) vanishes after another decade.

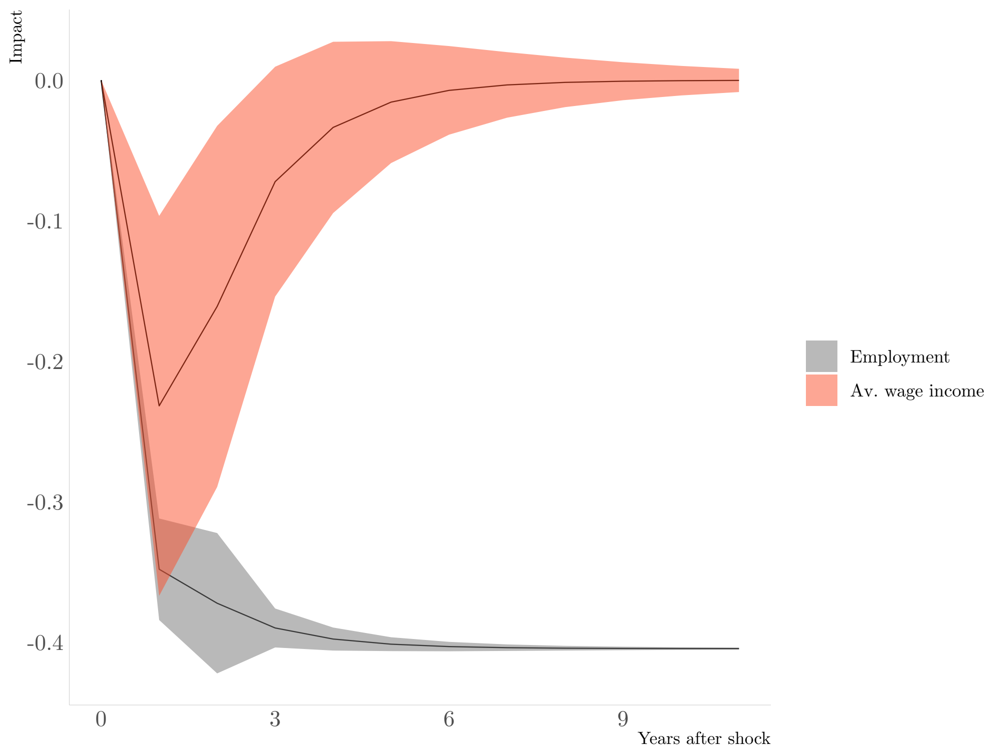
[19] There exists in American context studies at the MSA level [Bartik 1991]; the nearest European equivalent we were able to find is [Halleck-Vega and Elhorst 2016] for Netherlands' sub-regions.

[20] In most empirical calibrations inspired by [Autor, Dorn, and Hanson 2013], most importantly [Caliendo, Dvorkin, and Parro 2015], the full impact is deemed realised between 7 and 12 years after the start of the exposure period.

[21] It is mainly because we fail to find any sign of outmigration reactions following a trade shock. In American context, the standard correlation with the local demographic increase through migration is 0.9 [Turek 1985] and this reaction is generally deemed to sluggish by macro models [Davis, Fisher, and Veracierto 2021]; in French context, the general figure is around 0.4, 0.7 for the Paris metropolis (see figure 46).

[22] Local multipliers we get over the whole period 1990-2018, as we saw, are in the spirit of [Moretti 2010; Dijk 2016], but on the 1999-2008 interval, the reaction seems less pronounced.

Figure 10: OIRF estimates of the long-term impact of a general employment shock using panel VAR model (22)



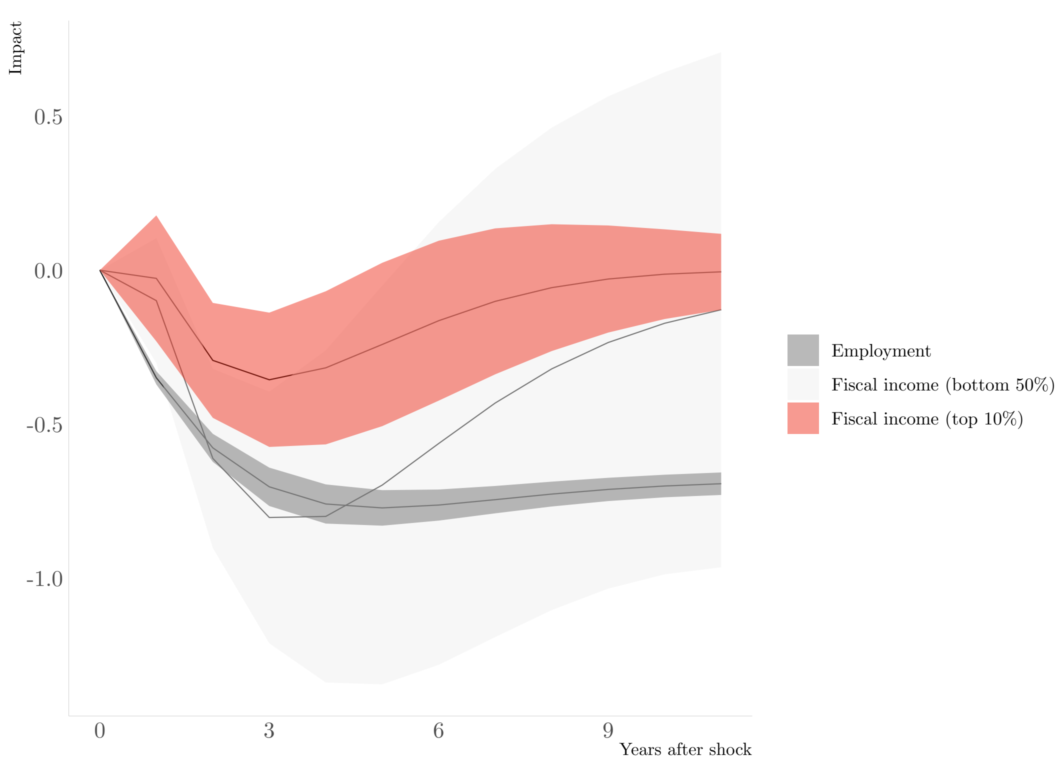
Note: The unit of interest is the *région* in the pre-2015 geography (i.e. at the time where there were 27 regions). We plot the average individual annual net wage reported within the 1/12 subsample of the DADS within each spatial zone. Period of estimation is 2009-2018. Other parameters are described in figure 8.

Income response

Wage data from the DADS allow for distinctions between SES statuses (the PCS scale of the INSEE). When we estimate (22) separately for each group, the reaction of wages among blue-collar and menial jobs seems more pronounced. Our idea then is to fit our panel VAR model framework, not with the av. fiscal income of the region, but concurrently the av. fiscal income of the 10% richest and 50% poorest tax units:

$$\begin{pmatrix} \Delta n_{c,t} \\ i^{t10}_{c,t} \\ i^{b50}_{c,t} \end{pmatrix} = \begin{pmatrix} \gamma_{c,1,0} \\ \gamma_{c,2,0} \\ \gamma_{c,3,0} \end{pmatrix} + \begin{pmatrix} 0 & \gamma_{c,1,2} & \gamma_{c,1,3} & \gamma_{c,1,4} \\ \gamma_{c,2,1} & 0 & \gamma_{c,2,3} & \gamma_{c,2,4} \\ \gamma_{c,3,1} & 0 & \gamma_{c,3,3} & \gamma_{c,3,4} \end{pmatrix} (L) \times \begin{pmatrix} \Delta n_{c,t} \\ \Delta n_{c,t-1} \\ i^{t10}_{c,t-1} \\ i^{b50}_{c,t-1} \end{pmatrix} + \begin{pmatrix} \eta^{n}_{c,t} \\ \eta^{it10}_{c,t} \\ \eta^{ib50}_{c,t} \end{pmatrix} \tag{23}$$

Figure 11: OIRF estimates of the long-term impact of a general employment shock using panel VAR model (23)



Note: The spatial units of interest is the *région* in the pre-2015 geography. Our dependent variables include the average fiscal income within each spatial unit and group (IRCOM series). Period of estimation is 2006-2018. Other parameters are described in figure 8.

The results of the estimation of (23) are plotted in figure 11. The estimated response to a 1SD shock on employment (-0.35 log points to the total stock) is a lingering negative impact on the total number of jobs in the region (-0.69 log points a decade after) and a significant negative shock on the fiscal income of both the top 10 and bottom 50%, at -0.34 log points after three years, and -0.77 log points respectively after three years, but the confidence interval comprise zero again four or five years after the shock itself[23].

Figure 12 compares these effects of a general shock to the long-term fiscal income effect of the China shock (1999-2008 decade of exposure) for all households.

Figure 12: Long-term impact of the China shock (1999-2008 import exposure) on mean fiscal income within ZEs

Note: The unit of interest is the ZE. Income data are from the IRCOM database. We estimate model (3) using the full vector of controls, taking as explanatory the import competition exposure index $\Delta IPW_{1999-2008}$ instrumented in the way described herein above, and as dependent, the evolution (in pp) of the average fiscal income reported in the zone between 1998 and the year mentioned on the x-axis. Estimation is done for each year, weighting observations by the total population of the ZE, and clustering S.E. at the INSEE superzones level. We report 90% confidence intervals.

In our section 3., we shall provide a much more detailed approach of the distributional dynamics of the shock, but in almost every setting, we'll be relying on a decade-by-decade approach, under the assumption that the exposure of one decade impacts the income evolution of the same decade. Yet fig. 12 suggests no sign of recovery in the longer term dynamics of fiscal income following an import shock. In a sense, fiscal income exhibits a reaction quite similar to the employment stock in the original Blanchard-Katz model, a pattern imputable, not much to purported local multiplicative effects through the income channel (since, as we'll see, redistribution deadens much of the income impact of the shock), but rather to the rise in local dependence to social transfers (which will be a pivotal aspect in the findings of the next section). It's primarily that increased dependence to redistribution which encapsulates, once the shock is realised, much of the impact of the multiple distortion mechanisms documented by New Keynesian models: multiplicative effects across labour markets [Bartik 1991], sluggish outmigrations, and regional hysteresis.

[23]Estimating (23) at the ZE level yields results which also indicate slightly better fortunes for the top incomes after a local shock, with a short-term impact (at year +3) on the local ratio T10/B50 in the order of magnitude of the ones reported above, but confidence intervals are too large to allow any definitive prediction on longer term distributional impacts.

3 Income and wage responses

3.1 *Between*-region impacts

3.1.1 Main estimates

The simplest way to gauge the impact of an import shock on between-inequalities is to evaluate specification (3) using the decadal rise in fiscal income within the geographical unit of interest as the dependent. This *between* approach is the only type of estimates found in [Autor, Dorn, and Hanson 2013] and in almost all of its replications. The results of the stacked decades' estimation[24] are reported in table 13. As far as comparison is possible, the estimates of [Autor, Dorn, and Hanson 2021] are almost similar to the results displayed there in column (2):

Table 8: Exposure to import competition and evolution of between regions inequalities

Dep. : Decadal change in the av. fiscal and disposable income (in pp)			
	Distribution of spatial units		
	Between départements	*Between ZEs*	*Between communes*
	(1)	(2)	(3)
Rise in import exposure:			
Panel A. IRCOM dataset − 1990-2018 − Fiscal income			
β_1	-4.34^{***}	-1.93^{**}	-1.41^{**}
S.E.	(1.08)	(0.82)	(0.59)
R^2	0.85	0.32	0.39
F-stat	44.6^{***}	12.6^{***}	1393^{***}
Obs.	282	912	71520
Panel B. − Filosofi dataset − 2012-2017 − Fiscal income			
β_1		-2.16^{*}	
S.E.		(1.13)	
R^2		0.44	
F-stat		6.9^{***}	
Obs.		304	
Panel C. − Filosofi dataset − 2012-2017 − Disposable income			
β_1		-1.41^{*}	
S.E.		(0.83)	
R^2		0.54	
F-stat		10.6^{***}	
Obs.		304	

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: Income data are from IRCOM and Filosofi databases. We report the estimation of the main coefficient of model (3), but the dependent variable is the evolution of the related average yearly fiscal or disposable income of the persons living within each geographical unit of interest. The main explanatory variable is the index ΔIPW, described herein above, which provides an estimation of the mean rise in Chinese imports (in value) per worker within each ZE. The instrument is the same ΔIPW, in which French trade data has been replaced by a control group of four countries (Japan, Germany, Spain, Switzerland) and all labour force variables are taken with a decadal lag. When the time period exceeds ten years, decades are stacked with the inclusion of a time dummy. Each specification includes the full vector of controls mentioned in table 3 ; when the model is estimated at the level of the *commune*, we ascribe to each city the explanatory and the instrument of the ZE to which it belongs; other controls, and the dependent, are city-specific. Observations are weighted by the start-of-the-period total tax units population. Standard errors are clustered at the level of the INSEE superzones.

Since metropolises and richer cities are overall more exposed, trade shocks have almost no impact on the income rankings of cities and regions. If we build a counterfactual scenario with $\Delta IPW_{1990-2018} = 0$, offering to each city a premium equal to the average exposure over 1990-2018 of the ZE it belongs to times the corresponding effects reported in table 13, replicating it on every issue of the IRCOM series, taking then the ratio top 10% versus bottom 50% of cities according to the average fiscal income of their inhabitants, we see that counterfactual values are barely distinguishable from the actual ones (see figure 13).

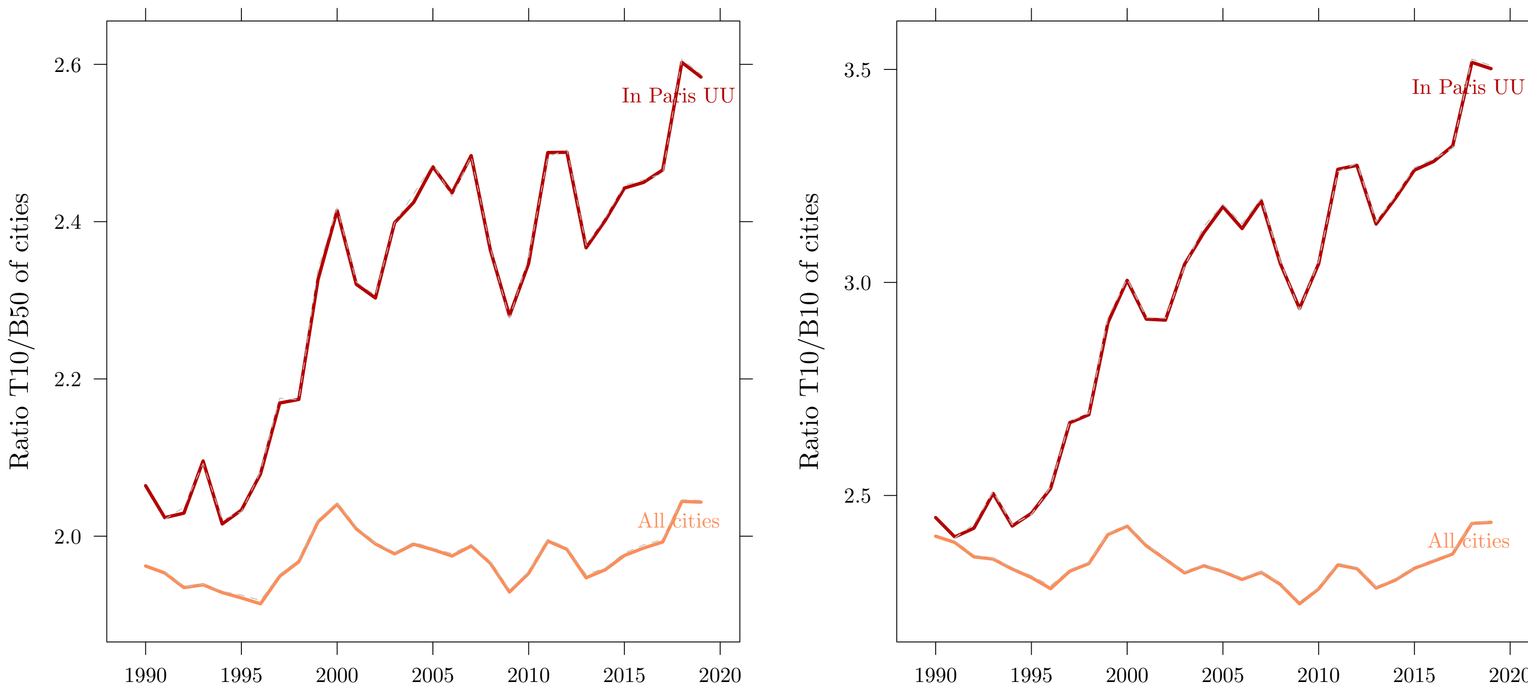
3.1.2 The divergence of the middle city

These marginal estimates indicate that more exposed regions tend to suffer from an income loss *ceteris paribus*, but they hardly illuminate the general picture about the distributional impact. Two main limitations stand in the way: 1. There are reasons to believe that these marginal effects might vary across types of regions ; 2. A consistent interpretation must juxtapose these marginal effects with the total rise in fiscal income per region, since, as emphasised by [Dorn and Levell 2021] a very same marginal decline could be barely felt in a fast-growing metropolis (which is likely to benefit from other channels of trade exposure), while it might have dreary consequences in a declining city. Figure 15 heeds to these two biases: 1. Relying on the IRCOM data series 1990-2018, we divide cities in weighted-deciles along three great variables: the average fiscal income of their inhabitants, their size (tax unit population) and their distance to the nearest of the 22 metropolises of the country. Over each subgroup, we reestimate marginal effects, multiplying them by the average decadal ΔIPW of the group, which provides us with

[24] Since in empirical tests [Autor, Dorn, and Hanson 2021] and econometrical calibrations [Galle, Rodríguez-Clare, and Yi 2022] alike, the impact of the trade shock is fully realised approximately seven years after the start of the exposure, it seems legitimate to estimate decade by decade (under the assumption that the shock of one decade impacts the income evolution of the contemporary decade, and not of the following one).

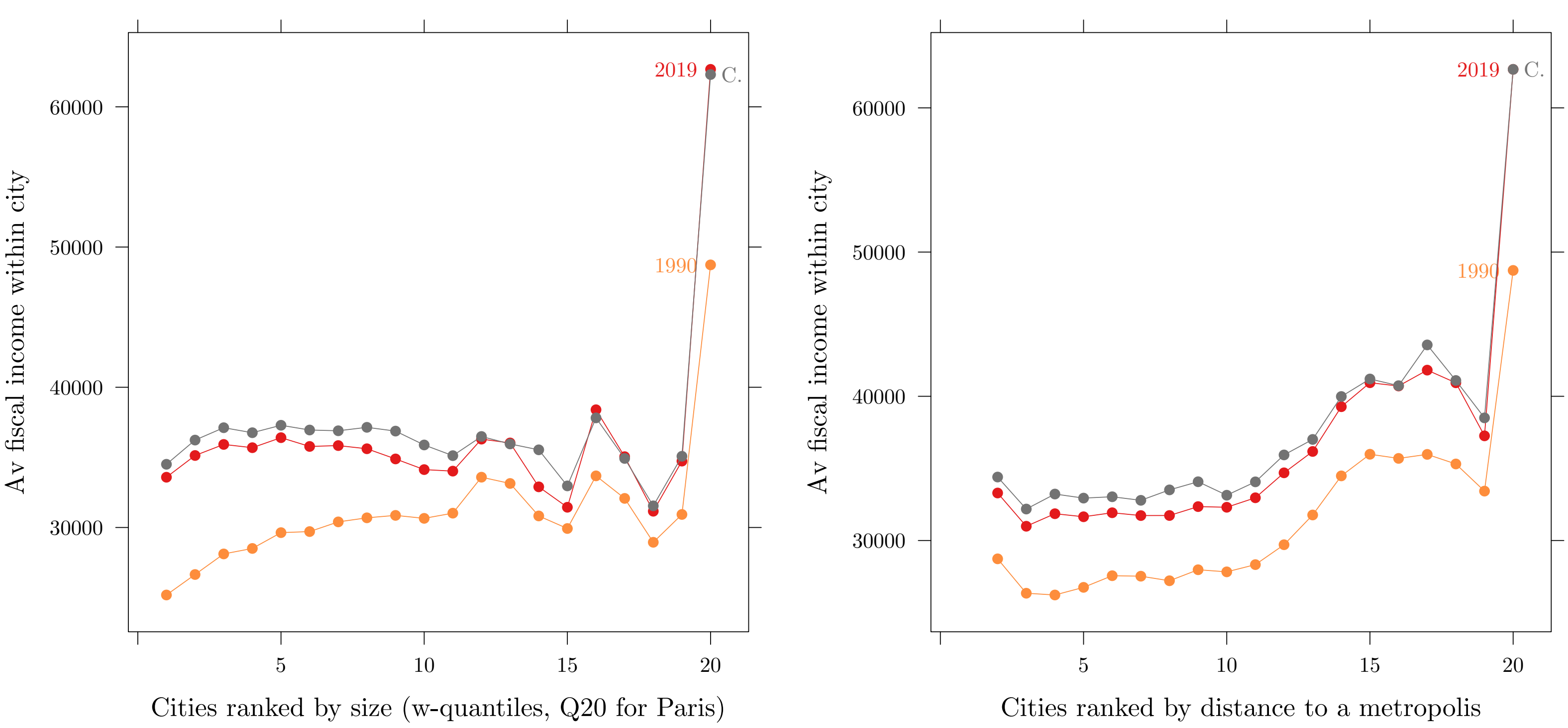
an estimate of the decadal income loss; 2. We then plot the actual mean decadal rise in fiscal income for each subgroup; summing with the estimated loss provides us with the counterfactual scenario estimates.

Figure 13: Counterfactual scenario with no exposure ($\Delta IPW_{1990-2018} = 0$) − Impact of between-cities inequality



Note: The unit of interest is the *commune*. Data are from the IRCOM database, over restriction 0 (i.e. over all *communes*, at the exception of those with fewer than 11 inhabitants). We consider each city as an individual which earns the average fiscal income of its inhabitants, and we compute the ratio of the average income of the 10% richest cities, over the 10% (or 50%) poorest; fractiles are computed with tax unit population weights, i.e. the 10% richest cities are not the 3600 richest *communes*, but the richest communes in which 10% of the national population lives. Through this computation, the *arrondissements* of Paris, Lyon and Marseille are taken as individual cities (the Western *arrondissements* of Paris always fall within our top 10%, while the Northeastern ones do not). See the annex for more details about the corrections applied to the original data of fiscal income. For the counterfactual scenario without trade exposure (plotted as a dashed grey line), we compute the marginal yearly equivalent of the decadal impacts reported in table providing to each city an income premium scaled by the exposure of the ZE it belongs to over 1990-2018.

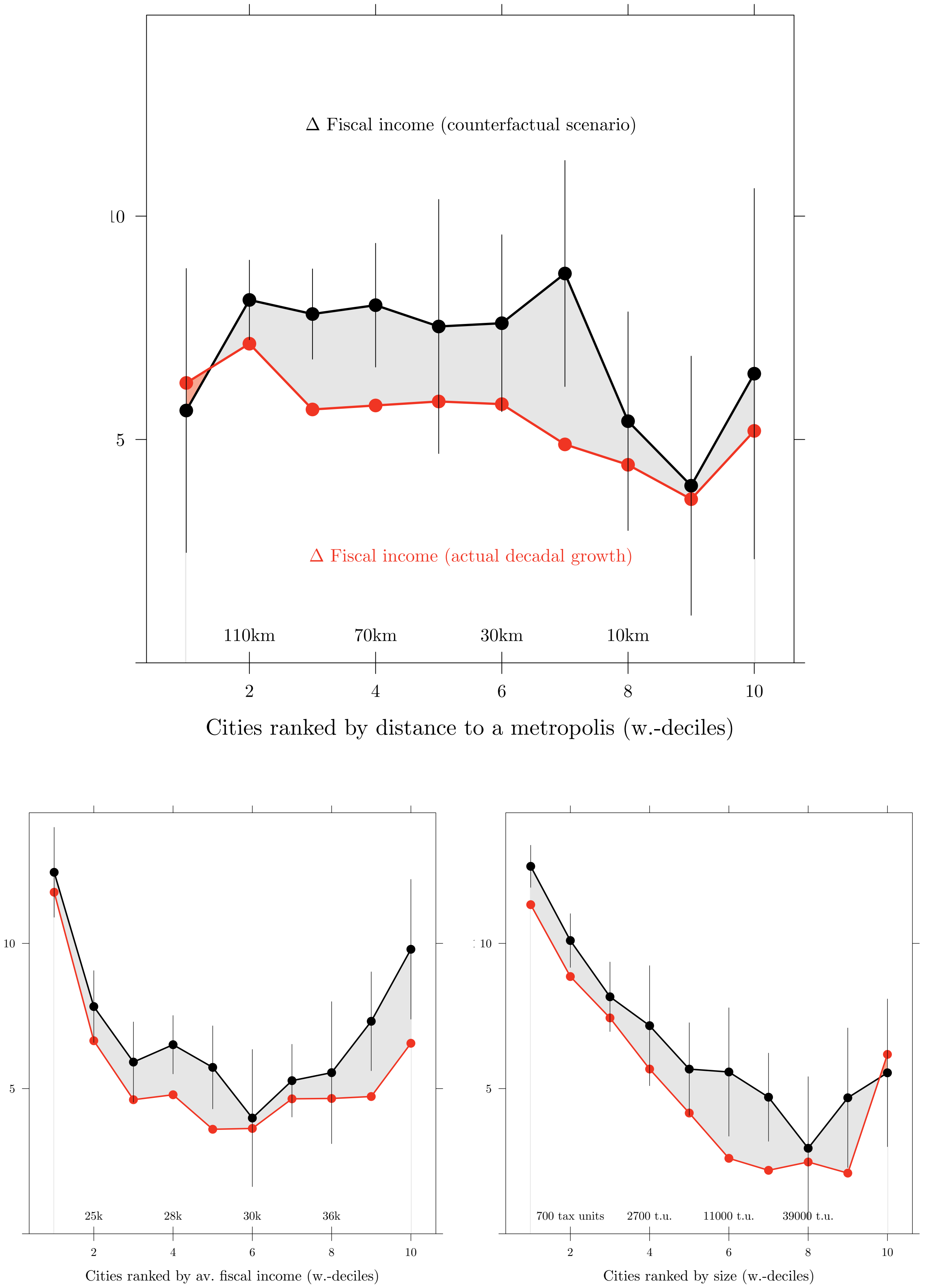
Figure 14: Counterfactual scenario with no exposure ($\Delta IPW_{1990-2018} = 0$) − Impact on the distribution of cities



Note: The unit of interest is the *commune*. Cities are ranked by 5% total-population-weighted quantiles of : 1. City size (tax unit pop.); 2. Haversine distance from the centroid of the city to the centroid of the nearest metropolis. Reported statistics is the average fiscal income within the commune, as reported in the IRCOM database (subsample R0, i.e. all cities with a tax unit pop. above 11). Actual values are reported in colour, counterfactual values with import exposure 1990-2018 set to zero, in grey.

As suspected, even though richer cities are slightly more exposed, when we plot the average rise in fiscal income over 1990-2018 over the distribution of cities along major variables (their av. fiscal income, their size, their distance to a metropolis), we systematically get a U-shaped curve; it is the average *commune* (in terms of income, of size, of proximity to bigger cities) which grows slower, the marginal impact of trade shocks being even more pronounced among these average cities.

Figure 15: Growth of av. fiscal income across types of cities − Actual growth *vs* Counterfactual scenario with no exposure ($\Delta IPW_{1990,2018} = 0$)



Note: Cities are divided by 10% total-tax-unit-population-weighted quantiles of the corresponding ranking variable. Specification 3 is estimated over each of these subgroups, for three decades (1990-1999, 1999-2008 and 2008-2018), using the full set of controls, and the decadal rise in fiscal income of the inhabitants of the city as the dependent. All growth rates are in pp. We ascribe to each city the ΔIPWs of the ZE to which it belongs; other controls are city-specific. We report the descriptive statistics average rise in fiscal income over the decade, plus the estimated loss of income with its 95% conf. interval. In all specifications, observations are weighted by the start-of-the-period population, and S.E. are clustered at the level of the INSEE superzones.

3.1.3 Winners and losers

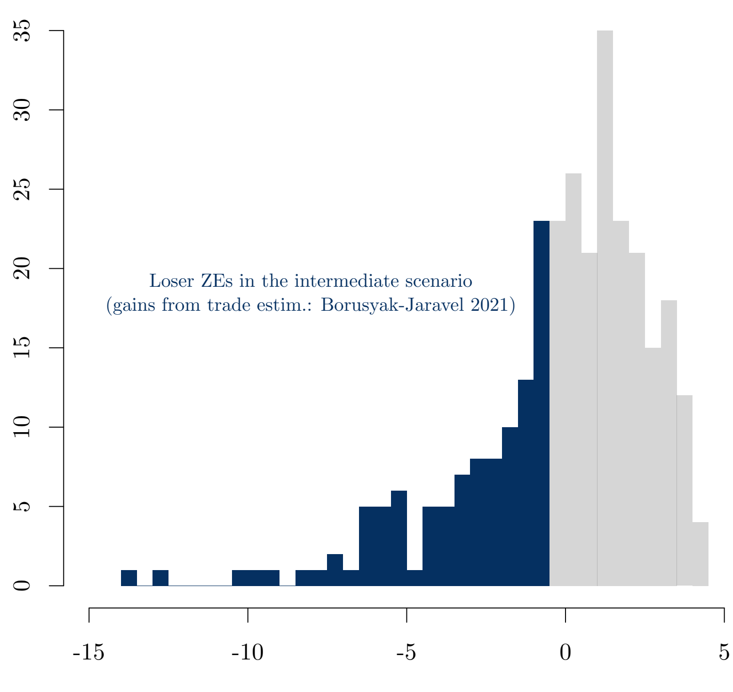
A reduced form approach cannot, *stricto sensu*, construct a counterfactual scenario in autarchy, in the spirit of the recent ones of [Adão, Carrillo, et al. 2020; P. Fajgelbaum and Khandelwal 2016]. In one of their companion articles, Autor, Dorn and Hanson [Autor, Dorn, and Hanson 2021] heed to that issue by juxtaposing their estimates of the negative income impact of import exposure, and some recent results about the aggregate gains from trade, relying on two types of evidence:

- Macro trade models which elaborate the relocation model of [Eaton and Kortum 2002] in order to gauge the aggregate welfare impact of the China shock. [Caliendo, Dvorkin, and Parro 2015], who calibrate their model with the ΔIPW of the original Autor-Dorn-Hanson article, and data about the U.S. in the first decade of the millennium, find a sharp negative impact on wages in the short-run due to local labour effects (with an average decline of -0.4 seven years after the shock), which is offset in the longer term by aggregate gains caused by increased competition, mainly the decline in prices and productivity gains; overall, 13 years after the shock, the average aggregate welfare gains is estimated at $+0.2$ log points (unweighted S.D. of 0.09). [Galle, Rodríguez-Clare, and Yi 2022], because they include a labour immobility factor in their model, find a similar positive impact ($+0.22$ over a 7 years period), but with a far greater variability between regions (unweighted S.D. of 0.31). As far as comparison is possible, [Borusyak and Jaravel 2021] seem to find slightly superior marginal gains of trade; they estimate that a 10% decline of trade costs with China results in a $+0.15$pp net welfare gain per worker; since, over 2000-2010, U.S.-China trade costs have declined by 27% (World Bank's ESCAP figures), the net gains would oscillate around $+0.4$pp;
- Micro evidence about the net gains of purchasing power due to trade competition. [Jaravel and Sager 2019], using estimates drawn from Autor-Dorn-Hanson papers, estimate that a 1% in import penetration within the CZ results in a decline of consumer price of about -1.4%; using a similar approach, [Dorn and Levell 2021] report estimates which are about half the size of Jaravel-Sager ones.

One of the main drawback of this approach lies in the fact that it focuses exclusively on inequality *between* regions, discarding the *within* region dimension. The implied bias for the estimation of the gains from trade through consumer prices is negligible according to the most recent micro analyses [Borusyak and Jaravel 2021], which tend to show that the share of goods imported from China within the consumption basket is relatively flat across deciles of the income distribution. For other channels (especially the impact of labour relocations), we might expect considerable differences between deciles; such differences are found even in the most optimistic micro approaches [Borusyak and Jaravel 2021].

Starting from these figures, [Autor, Dorn, and Hanson 2021] proceed with a relatively straightforward strategy; they estimate an extended version of model (3) using as main explanatory $\Delta IPW_{c,2001,2012}$, and as dependent, the variation of the average log personal income of all inhabitants of each CZ over 2001-2019. They retrieve the corresponding $\hat{\beta}_1$. They consider the distribution of the average predicted loss $\hat{\beta}_1 \times \Delta IPW_{c,2001,2012}$, centring it around its national weighted average. Regions, the net loss of which is below the purported decadal gains of trade discussed herinabove might be considered as the losers of trade liberalisation. Here, we emulate that strategy using the very same decades; we are forced to rely on models calibrated over U.S. data, the only framework relying on the Autor-Dorn-Hanson approach applied to European countries being [Caliendo, Dvorkin, and Parro 2015], which provides for France marginal gains of trade slightly superior to the U.S. estimates ($+0.23$).

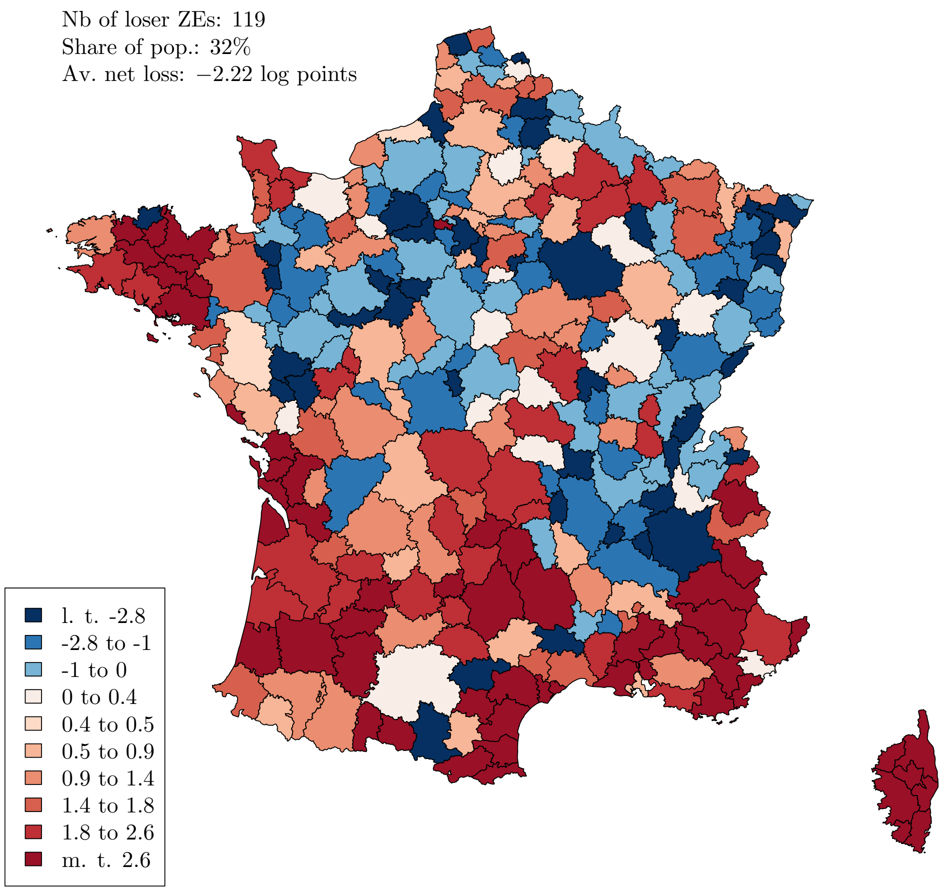
Figure 16: centred distribution of the income loss caused by a decade of import competition exposure



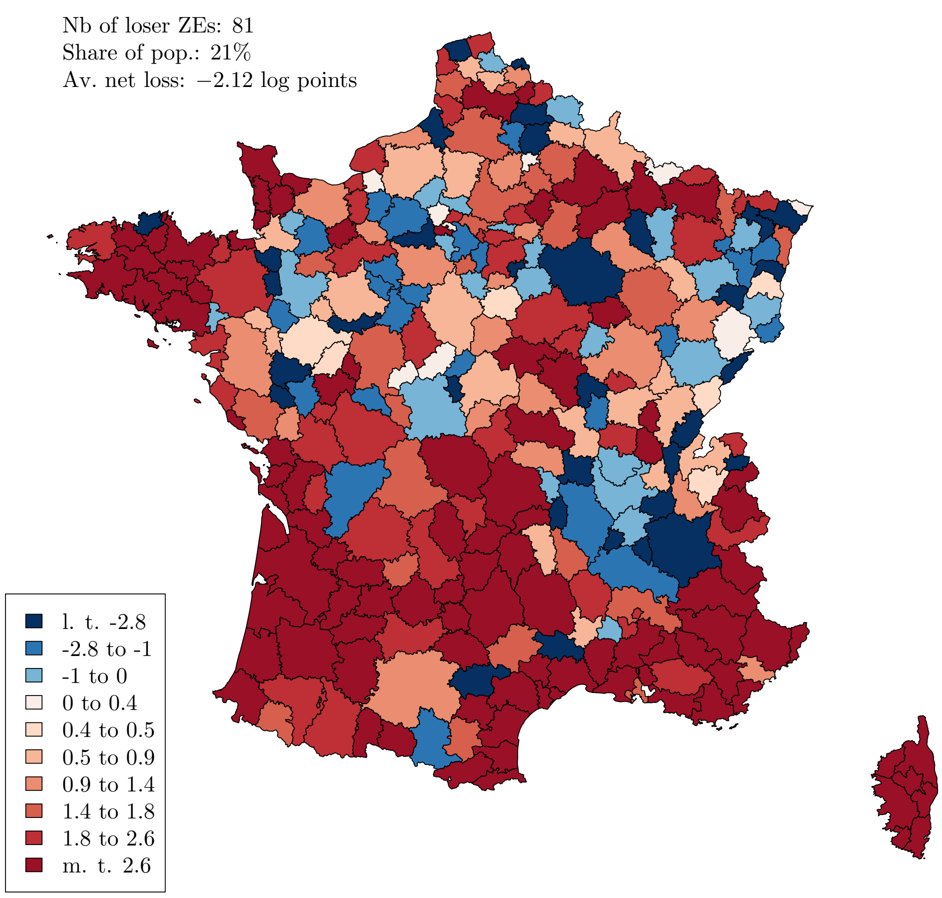
Note: The unit of interest is the *Zone d'emploi* (2010 INSEE definition). Income data are from the IRCOM base. We estimate model 3, using the import exposure index $\Delta IPW_{1999,2008}$ as the main explanatory variable (instrumented in the way described herein above), the full set of controls, and the variation in fiscal income (in log points) within the ZE over 1999-2018 as the dependent, weighting observations by the start-of-the-period population. We retrieve the corresponding coefficient $\hat{\beta}_1 = -2.27$ (t-stat: 1.95) and multiply it by the exposure of each ZE over 1999-2008, providing an estimate of the income loss caused by a decade of import exposure on local incomes. This histogram plots the distribution of that statistic, centred around the national weighted average. ZEs for which that statistics is below minus the aggregate gains from trade estimates of [Borusyak and Jaravel 2021] are displayed in blue.

Figure 17: Regions losing from import exposure in the sense of [Autor, Dorn, and Hanson 2021] using different estimates of gains of trade (marginal deviation of local incomes from national av. in log points)

(a) Aggregate gains for France estimated by [Caliendo, Dvorkin, and Parro 2015]



(b) Marginal gains from price effects estimated by [Jaravel and Sager 2019]



Note: The unit of interest is the *Zone d'emploi* (2010 INSEE definition). Income data are from the IRCOM. The reported statistics, described in figure 16, is the average loss in the average fiscal income of the ZE caused by a decade (1999-2008) of exposure to import imputed to competition from China, expressed in deviation from the national average (log points). Two estimates of gains of trade are added to that statistic; we plot the resulting values.

The $\hat{\beta}_1$ we retrieve[25] is −2.27. The weighted average of the corresponding estimated loss in av. fiscal income over 1999-2018 is −4.53 log points, an impact which, as far as comparison is possible, is about one half greater than the one found in [Autor, Dorn, and Hanson 2021]. Figure 16 plots the distribution of this statistic. We then differentiate with multiple estimates of gains of trade to detect the loser ZEs in the sense of Autor, Dorn and Hanson. Two polar scenarios based respectively on [Caliendo, Dvorkin, and Parro 2015] and [Jaravel and Sager 2019] are displayed as maps in 17, and an intermediate scenario based on [Borusyak and Jaravel 2021] in the histogram of figure 16. The variance of our estimated losses is almost twice the size of the Autor-Dorn-Hanson estimates (2.96 versus 1.22), not because of a difference in the marginal effect, but because exposure varies much more across ZEs than across U.S. commuting zones. As a result, the perimeter of the losing regions is much more robust to the choice of gains of trade estimates. In [Autor, Dorn, and Hanson 2013], 38% of the U.S. population lives in a losing CZ in the minimal scenario of based on Caliendo et alii; this figures drops to a tiny 7% in the maximal scenario based on Jaravel-Sager price estimates. The respective values in our setting are 32% and 22%.

3.2 *Within*-region impacts – Labour market

The differential impact through the wage distribution within each region, plotted in figure 20, is in the spirit of the UK estimates of [De Lyon and Pessoa 2021], of the US ones by [Chetverikov, Larsen, and Palmer 2016], of the French ones of [Malgouyres 2017a] and of [Autor, Dorn, and Hanson 2013] alike. On mean yearly wages reported in the DADS, the impact of a +$1000 rise in exposure is negative and 10% significant, at −6.27 log points (see table 9), a figure approximately one half greater than the ones found by [Autor, Dorn, and Hanson 2021] for the U.S. and by [Malgouyres 2017a] for 1995-2007 French data. As far as comparison is possible, when we focus on the within-region dimension, the difference in reaction to an import shock between the first and last quartiles of the wage distribution is one half wider than the one reported by [Chetverikov, Larsen, and Palmer 2016]. [Adão, Carrillo, et al. 2020] find a much more polarised reaction, but over the labour market of the developing country hardly comparable to our setting.

Table 9: Within-firm impact of the shock on wages, hours works, and part-time jobs

	*Dep. : Decadal change of corresponding variable (*département*-level)*									
	Period of estimation: 2002-2008 & 2008-2018									
Restriction	Gender		Age			Type of job			Firm size	
	Women	*Men*	*30−*	*31-50*	*51+*	*Blue-c.*	*Menial*	*Sup.*	*249−*	*250+*
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Rise in import exposure										
Panel A. Impact on the av. yearly wage income (log pts)										
−6.27*	−6.09**	−8.66	−17.3*	−6.61	−5.58	−6.99**	−8.56*	−9.19	−5.16*	−5.83
(3.66)	(2.84)	(6.01)	(9.25)	(4.05)	(6.75)	(3.11)	(4.82)	(6.84)	(2.91)	(4.59)
Panel B. Impact on the total nb of hours worked (log pts)										
−0.57	−2.04**	0.09	−5.44	−0.04	2.13	0.51	−3.94*	1.09	0.71	−4.09
(0.92)	(0.86)	(1.04)	(3.64)	(1.05)	(1.24)	(1.15)	(2.02)	(0.78)	(0.63)	(3.23)
Panel C. Impact on the share of part-time jobs										
2.05	3.02*	1.87	4.53	1.46	−0.58	0.96	1.59	1.51	1.04	5.21
(1.64)	(1.71)	(1.68)	(2.88)	(1.49)	(1.21)	(1.71)	(1.39)	(2.18)	(1.02)	(3.55)

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: The unit of observation is the *département* (oversee dép. and territories excluded). The dependent variable is the evolution of the related average yearly wage within the *département*, computed over the 1/12 subsample of the *Déclarations annuelles des données sociales* (DADS). The main explanatory variable is the index ΔIPW, described herein above, which provides an estimation of the mean rise in Chinese imports (in value) per worker within each ZE. The instrument is the same ΔIPW, in which French trade data has been replaced by a control group of four countries (Japan, Germany, Spain, Switzerland) and all labour force variables are taken with a decadal lag. We stack the two decades, including a time dummy. Each specification includes the full vector of controls. Observations are weighted by the start-of-the-decade total employment. Standard errors are clustered at the level of the INSEE superzones.

Nevertheless, some peculiarities of the French case must be heeded to:

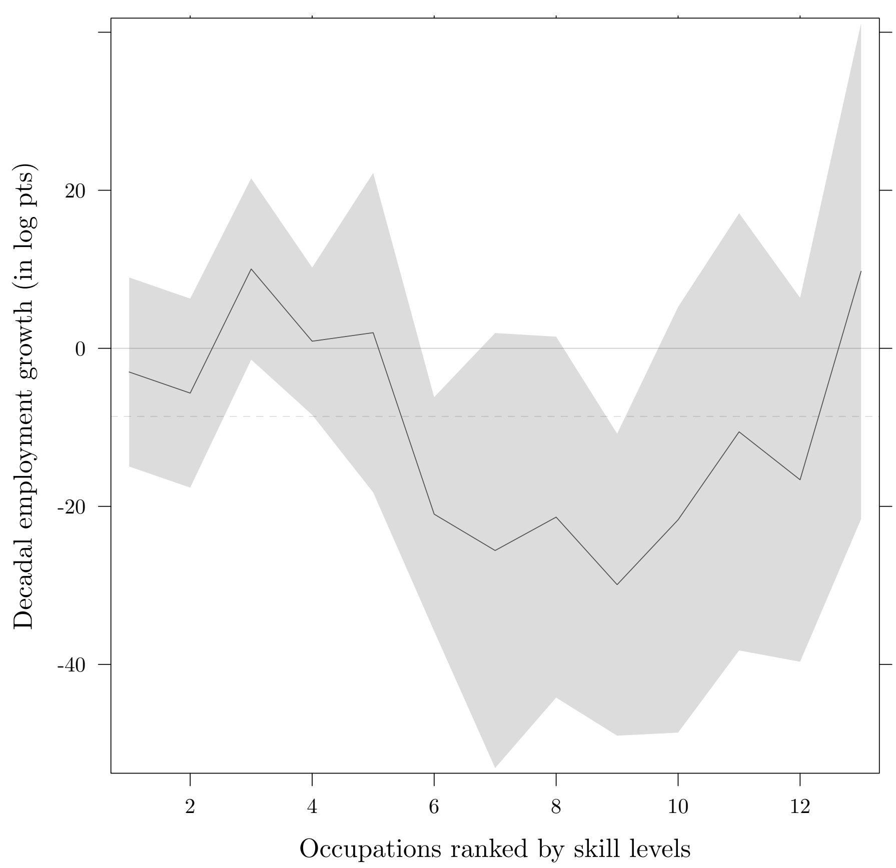
- *A seemingly negative impact on innovation within firms, even near the technological frontier* – The literature provides conflicting results when it comes to the issue of the innovation impact of the China shock. In the U.S. context, Autor and his coauthors argue that the shock had a significant depressing impact on innovation input (R&D spending) and output (patents) [Autor, Dorn, Hanson, Pisano, et al. 2016]; over European data on the contrary, [Bloom, Draca, and Van Reenen 2015] find that the Multifiber agreement bolstered patenting among more exposed firms. The usual argument to reconcile these findings involves the inverted-U-curve of [Aghion, Bloom, et al. 2005], i.e. a rise in competition on markets which used to be highly regulated shall have a positive impact on innovation, but exacerbated competition and high variance of the level of technological advancement might induce *laggard* firms to drop out of the innovation contest. Congruent with this setting, Aghion and his coauthors recently identified among French firms a general negative impact of the China shock on innovation output, an impact largely concentrated onto firms which have a prior productivity disadvantage and were lagging far behind the technological frontier [Aghion, Bergeaud, et al. 2021]. Actually, Aghion's findings draw a more disturbing picture: firms which are more exposed to Chinese import competition tend to invest less, compensating with: 1. A Ricardian focus on products for which France had a prior comparative advantage (an effect identified among all types of firms) ; 2. A rise of the Chamberlin differentiation, with an acceleration of the pace at which new products are introduced and old ones discontinued (among firms which are nearer to the technological frontier only). In a sense, the shock bolsters the demand for creative

[25] In this version of model 3, the explanatory is $\Delta IPW_{1999,2008}$, instrumented as described above, the dependent, the evolution of the average log fiscal income of the ZE between 1999 and 2018 (computed over the IRCOM base).

and innovative jobs at the very top of the wage and skill hierarchy (an impact clearly seen in figures 19 and 18), while those intermediate jobs which are usually needed to implement new innovations are needed less;

- *Downgrading polarisation in the non-exposed sector* – [**harrigantoubal2020polaris**] find in French context that trade shocks foster job polarisation within service firms with hikes in demand for either very low-skilled jobs (like retail workers) or very high-skilled ones (what they call *techies*), at the expense of lower-intermediate blue-collar jobs (this corresponds to the median worker in our fig. 19b, or to the longer vocational programs in fig. 18, the decline of which is primarily driven by non-manufacturing sectors);
- *Polarisation and skill-upgrading in the exposed sector* – Most of the European estimates suggest that the employment, wage, and skill impact of the China shock is straightforwardly regressive. Exposure to competition on final goods is doomed to hurt primarily low-skill-low-pay workers [Biscourp and Kramarz 2007; Dauth, Findeisen, and Suedekum 2021], while offshoring and the correlative rise of intermediate imports shall boost the skill intensity and the productivity of the remaining workforce, especially at the top of the wage distribution [Mion and Zhu 2013; Costa, Dhingra, and Machin 2019; **harrigantoubal2020polaris**]. This seems congruent with a general narrative where international competition forces unproductive firms out of the market, and fosters reallocation of labour towards more productive ones, in the spirit of the [Melitz 2003] model, or of models which interpret offshoring as a task-trading mechanism [Feenstra and Hanson 2001; G. Grossman and Rossi-Hansberg 2006]. However, as emphasised by [Malgouyres 2017a], applying an Autor-Dorn-Hanson framework to the employer-employee matched DADS datasets of the INSEE yields results which are not perfectly consistent with this general story: over DADS data for manufacturing firms, it's primarily upper-middle jobs that are destroyed (fig. 19a), while jobs at the two tails of the distribution are preserved, but both experience sharp wage cuts (fig. 19b). This result[26] is less surprising if we heed to the fact that Chinese imports have become more and more technology-intensive, to the point that, as emphasised by [Rodrik 2006] and [Schott 2008], China can hardly be used as a real-world equivalent of the low-income trade partner in dual HOS models. Our labour market findings are in fact relatively consistent with a setting *à la* Aghion in which firms lagging far behind the technological frontier rescind their innovation investments (hence the decline of these upper-middle jobs which are critical to the implementation of new innovations), while those firms which are still innovating hire only the most qualified creative jobs (the impact being clearly seen in fig. 19a and 18).

Figure 18: Import shocks and labour market response – Job polarisation (skills)

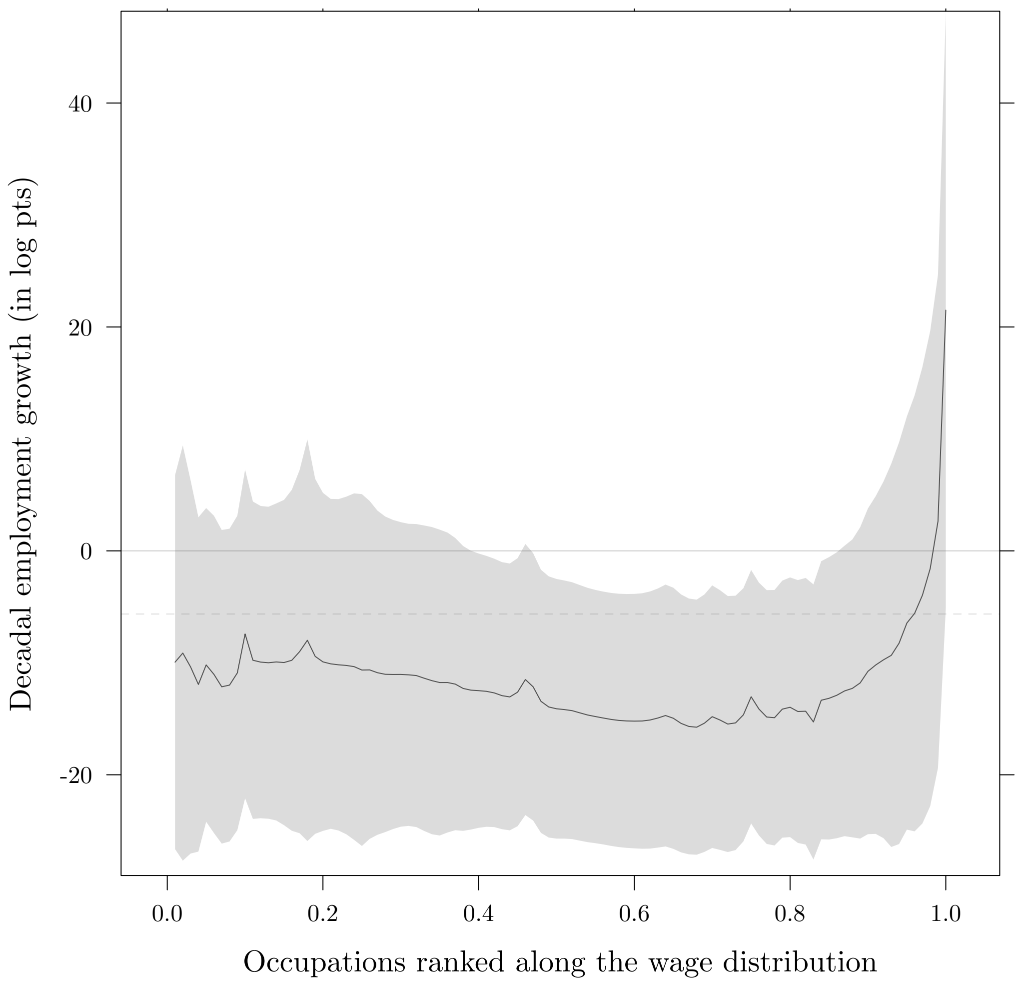


Note: The unit of interest is the *département*. Employment data are from the 1/12 microsample of the DADS. The main specification is still 3 with the full vector of controls and the Chinese imports' exposure index ΔIPW as the main explanatory, but this time the dependent is the decadal evolution (in log points) of the total stock of employment for a specific occupation. The decade estimated is 2008-2018, and we take the exposure index of 1999-2008 as the explanatory, with the corresponding instrumentation. Occupations are then ranked as follows: We group occupations using the PCS scale at the 4-digits level. We consider the individual diploma scale of the INSEE with slight modifications (1 - No schooling / 2 - No high school and no diploma / 3 - Some high school but no diploma / 4 - Middle-school diplomas (BDC-BEPC-CEP-DFEO) / 5 - Vocational education, short diplomas (CAP-BEP) / 6 - Vocational education, long diplomas (Bac-Tech-Bac-Pro) / 7 - High school diplomas (Bac) / 8 - Vocational higher education (BTS-DUT-DEUST) / 9 - Undergraduate short diplomas (DEUG-L1-L2) / 10 - Undergraduate long diplomas (L3-M1) / 11 - Graduate (M2) / 12 - Gradute "grande école" / 13 - PhD). Using the 2009 issue of the INSEE-ECMOSS dataset, we reconstruct the distribution of diplomas of the members of one specific occupation. We then compare that distribution to the nationwide distribution of diplomas of all workers, and we ascribe to that occupation the diploma which is most overrepresented compared to the national structure of diplomas for all workers. We estimate our main specification with, as dependent, the employment growth for the set of workers to which that specific diploma has been ascribed in each *département*, weighting observations by the start-of-the-decade log total employment, clustering standard errors at the level of the INSEE's superzones.

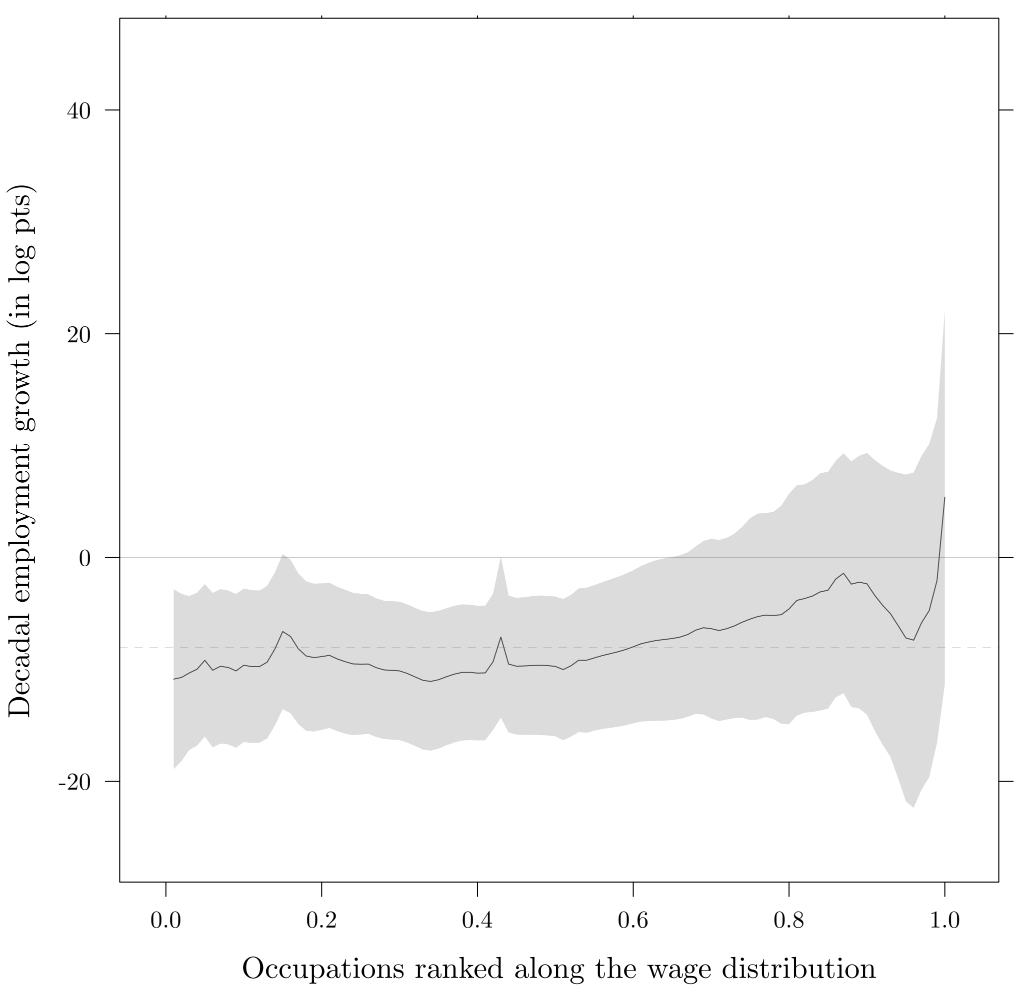
[26] These findings are congruent with the French estimates of [Malgouyres 2017a] and the Danish estimates of [**kellerutar2016polaris**], but also, in a sense, with the original Autor-Dorn-Hanson framework: in, [Autor, Dorn, and Hanson 2013], the sectoral pattern of the China shock is clear-cut: in the manufacturing sector, there are massive layoffs because of import exposure, but no decline of the average wage, while for non-manufacturing they find the reverse reaction (no employment decline, but a sharp negative impact on the wages of lower-paid service employees), i.e., in the exposed industrial sector, the adjustment is made through layoffs; in the non-exposed sectors on the contrary, the adjustment involves primarily a decline of wages; concurrently, there seems to be no impact of the China shock on the share of labour in the remuneration of factors [Autor, Dorn, Katz, et al. 2020], while it is the case French context according to [Aghion, Bergeaud, et al. 2021].

Figure 19: Import shocks and labour market response – Job polarisation (wages)

(a) Employment response by wage percentile – Manufacturing



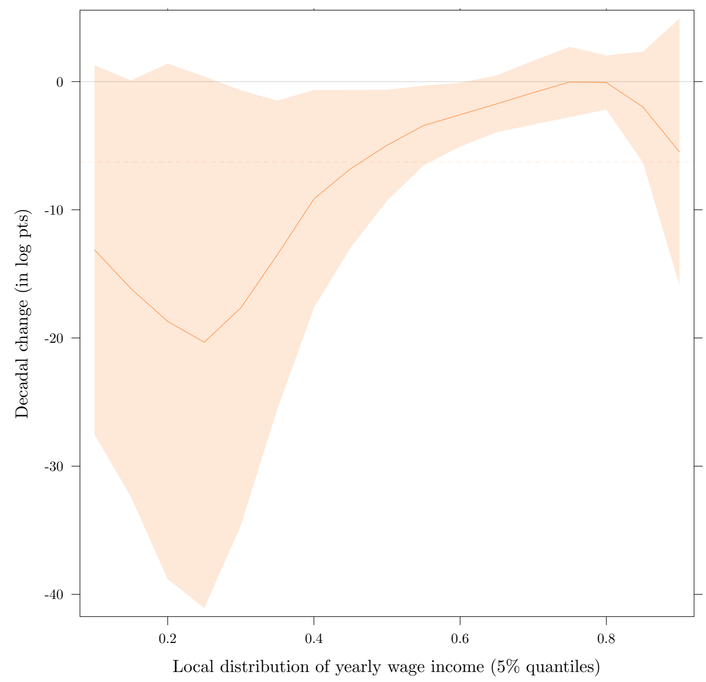
(b) Employment response by wage percentile – Non-manufacturing



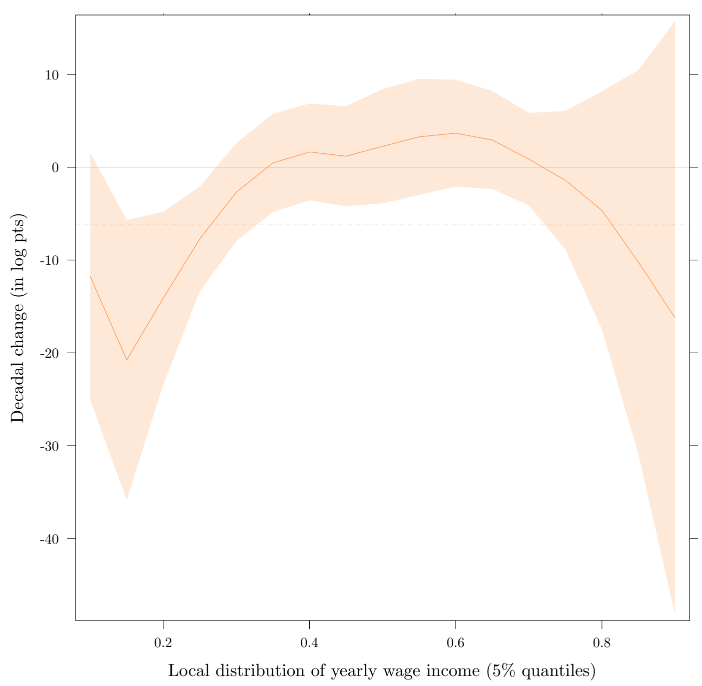
Note: The unit of interest is the *département*. Employment data are from the 1/12 microsample of the DADS. The main specification is still 3 with the full vector of controls and the Chinese imports' exposure index ΔIPW as the main explanatory, but this time the dependent is the decadal evolution (in log points) of the total stock of employment for a specific occupation. The decade estimated is 2008-2018, and we take the exposure index of 1999-2008 as the explanatory, with the corresponding instrumentation. Occupations are then ranked as follows: We estimate our main specific using as dependent the employment growth for each occupation of the PCS scale at the 3-digits level within each *départeent* (with a required minimum of 100 workers nationwide), weighting observations by the start-of-the-decade log occupation-specific employment, clustering standard errors at the level of the INSEE's superzones. The grouping method then applied is exactly similar to the one used by [Malgouyres 2017a] to construct its figure 7: we calculate how many workers in each occupation belong to the x-th percentile of the start-of-the-decade wage distribution. Then for each percentile, we construct a specific coefficient which is the weighted sum of the occupation-specific betas retrieved before. Standard errors are similarly reconstructed assuming independence between occupation-specific coefficients. The final coefficient plotted for each percentile is meant to provide an estimation of the job loss at a certain location of the start-of-the-decade wage distribution.

Figure 20: Import shocks and labour market response – Wages

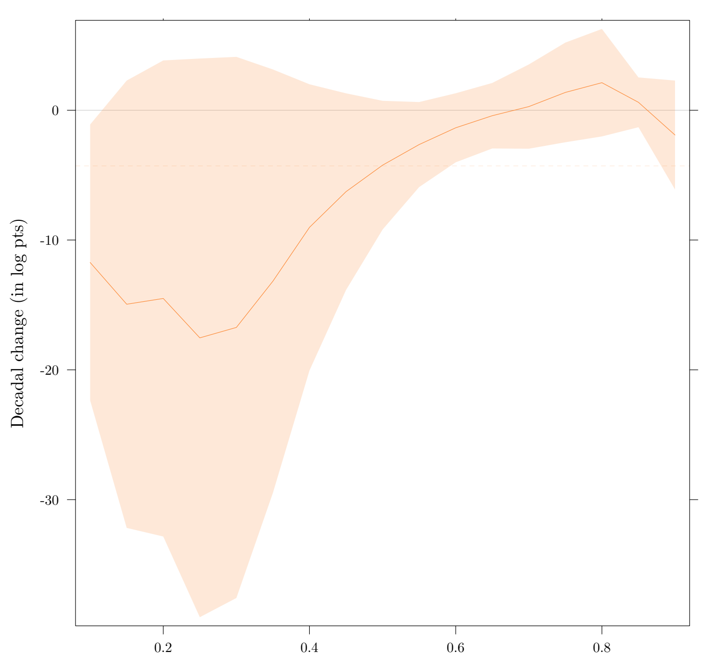
(a) Wage response – All sectors



(b) Wage response – Manufacturing



(c) Wage response – Non-manufacturing



Note: The unit of interest is the *département*. The main source is the 1/12 microsample of the DADS. The dependent variable is the evolution (in log points) or quantiles of the average yearly wage distribution of the *département*. The main explanatory variable is the index ΔIPW, described herein above, with the corresponding instrumentation. The period of estimation is 2002-2008 and 2008-2018; the two decades are stacked with the addition of a time dummy. All specifications include the full vector of controls. Observations are weighted by the start-of-the-decade total employment. Standard errors are clustered at the level of the INSEE superzones. The main line denotes the main coefficients, with the corresponding 95% conf. interval. A dashed line plots $\hat{\beta}_1$ when the dependent is the rise in log points of the av. yearly wage within the *département*, i.e. respec. -6.27 (t=1.71), -6.19 (t=1.07) and -4.29 (t=1.54).

3.3 *Within*-region impacts – Income

3.3.1 Regressive effects on pre-redistribution, but not on post-redistribution income

A specificity of French tax and income data is that they do not only provide averages for each region, but also detailed information about the distribution of income within each zone (at the level of the city since 1990, at the level of the *département* since the early 19th c.), most of the time in the form of a piece-wise function, from which income distribution interpolation methods pioneered by [Blanchet, Fournier, and Piketty 2017] can derive a continuous distribution with all its parameters.

It is then tempting to build a strategy with exploits the distinction of *between* versus *within*-region income dynamics, which is widely used at the international level [Chancel 2019a; Chancel and Piketty 2021; Bourguignon and Morrisson 2002] or to compare U.S. states [Bartik 1991] and EU countries [Blanchet, Chancel, and Gethin 2019; Blanchet, Chancel, and Gethin 2020].

When it comes to the estimation strategy for the second impact (the *within*-region distributional effect of the shock), there are very few items in the econometrical literature which address the specific problem we are faced with. In the absence of individual income data, the best approach that we know of is a group-level-treatment IV quantile regression, a framework which was still used informally in the early 2000s, most notably by [Angrist and Lang 2004], and which has been recently formalised by [Chetverikov, Larsen, and Palmer 2016]. It applies to experimental settings where the treatment is group-specific (whether that group is a school, a firm, a city, a region or else) and where endogeneity concerns involve the group, not the individual dimension. When micro data is missing, the first step of that estimation strategy consists merely in the retrieval of quantiles of the distribution of the dependent variable within each group of interest.

Using our ZEs (or another geographical unit) as groups, we can replicate specification (3), using now as the dependent variable, the evolution of each quantile of the *within*-distribution of income. The results of two group-level-treatment IV quantile regressions (at the city and ZE levels[27]) are provided in figure 21. We also have recourse to a strategy based on the shares of the total fiscal income of the region held by each *within*-decile, the outcome of which is plotted in figure 22.

Consistent with our wage estimates, we find that the impact on fiscal income is largely concentrated on the first three deciles of the *within*-region, a finding strikingly similar to the one of Chetverikov, Larsen, and Palmer 2016.

3.3.2 A rise in dependence to redistribution

Concurrently with [Autor, Dorn, and Hanson 2013] and with our disposable income findings, we identify a clear marginal impact of import exposure on the rise of the share of social transfers within the final income, illustrated in table 10, concentrated on minimum income. When we decompose that aggregate impact by decile of the *within*-ZE disposable income distribution, we find that the rise in shares is concentrated on deciles 2 and 3.

Table 10: Exposure to import competition and evolution of reliance on social transfers

Dep. : Decadal change in average share of social transfers in final income *Period of estimation: 2012-2017*						
		Types of transfers			*Restrictions*	
	All transfers	*Minimum inc.*	*Family allow.*	*Housing benefits*	*Top 10 ZEs*	*Bottom 50 ZEs*
	(1)	(2)	(3)	(4)	(5)	(6)
Rise in imports from China per worker:						
+Full vector of controls:	0.511**	0.411**	0.083	0.036	1.07***	0.469
	(0.24)	(0.19)	(0.05)	(0.04)	(0.21)	(0.37)
R^2	0.68	0.78	0.61	0.57	0.88	0.56
F-stat	18.2***	29.7***	13.2***	10.5***	6.1***	4.9***
Obs.	304	304	304	304	31	152

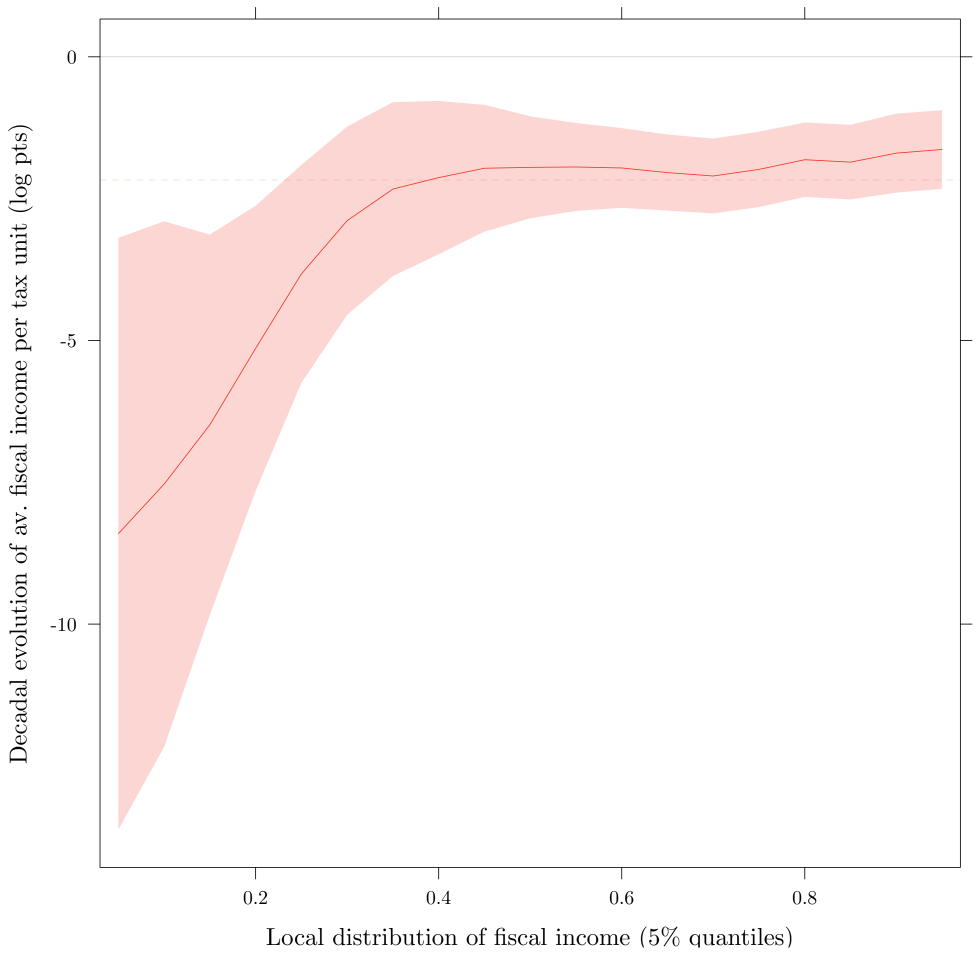
Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: The unit of observation is the ZE (*Zone d'emploi*, definition of 2010) . The dependent variable is the average evolution (in pp) of the share of each type of transfers within the final disposable (after-redistribution) income of a tax unit within the ZE of interest, as reported in the Filosofi database of the INSEE. The main explanatory variable is the index ΔIPW, described herein above, which provides an estimation of the mean rise in Chinese imports (in value) per worker within each ZE. The instrument is the same ΔIPW, in which French trade data has been replaced by a control group of four countries (Japan, Germany, Spain, Switzerland) and all labour force variables are taken with a decadal lag. All specifications include the full vector of controls (at the exception of offshorability and machine penetration indexes due to data limitation). Observations are weighted by the start-of-the-decade total population of households reported in the Filosofi database. Standard errors are clustered at the level of the INSEE superzones.

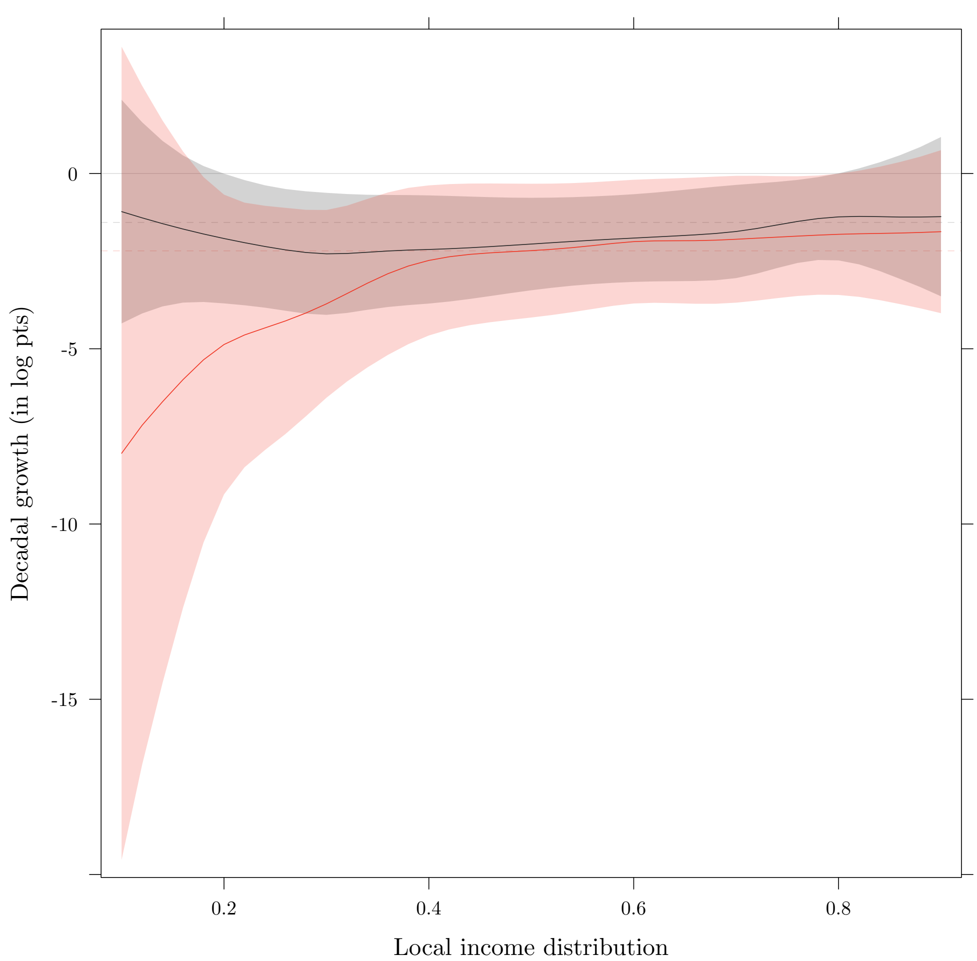
[27] In the estimation strategy based on the IRCOM income figures, the quantiles of each local distribution must be computed with the *gpinter* algorithm of [Blanchet, Fournier, and Piketty 2017]. To check for possible biases created by the interpolation, in the second panel of figure 21, we use non-interpolated quantiles, those which are provided in the Filosofi dataset. Note however that these quantiles are not directly computed by the INSEE over micro data; they are also reconstructed. See our annex A for more details.

Figure 21: Distributional impact of the shock *within* regions – Group-level-treatment IV quantile regressions using the Chetverikov-Larsen-Palmer estimator [Chetverikov, Larsen, and Palmer 2016]

(a) Impact on the distribution of fiscal income within cities-*communes*



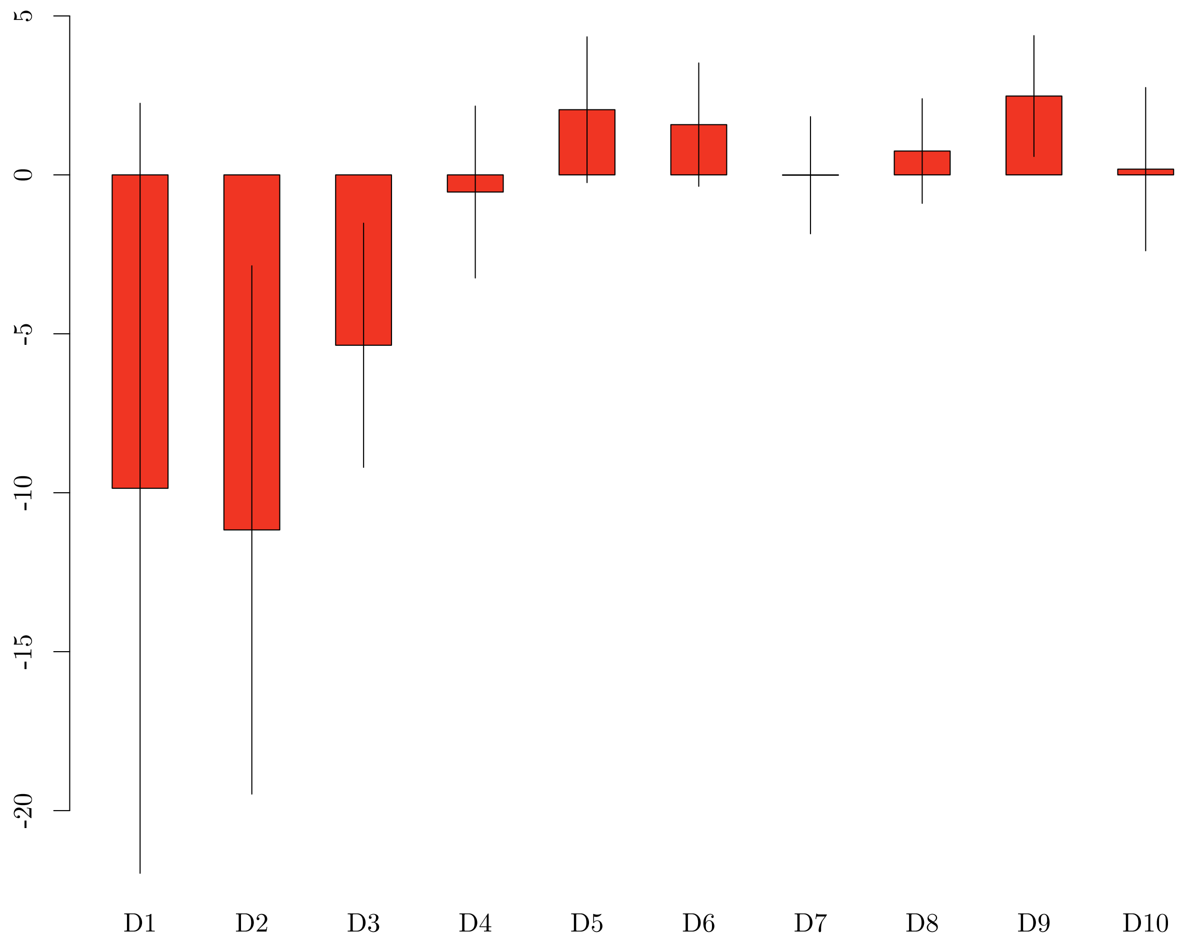
(b) Impact on the distribution of fiscal (*red*) and disposable (*grey*) income within commuting zones-*ZEs*



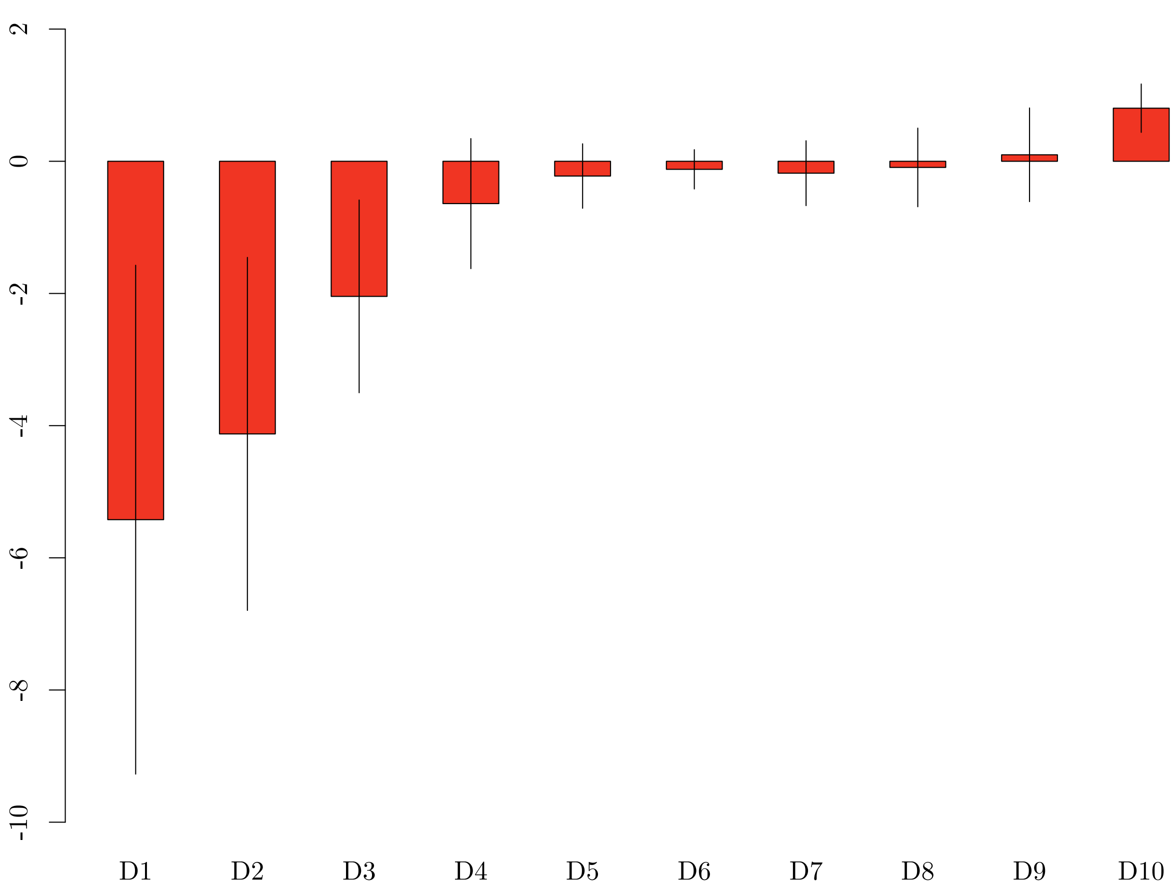
Note: The unit of interest is the city-*commune* and the commuting zone-ZE (*Zone d'emploi*, 2010 INSEE definition). The main source for income variables are respectively the IRCOM database (restriction 2) and the Filosofi database (full dataset). The specification is similar to 3, but this time the dependent variable is the decadal-equivalent evolution (in log points) of each quantile of the local distribution of fiscal (red) and disposable (grey) incomes. The main explanatory variable is the index ΔIPW, described herein above, which provides an estimation of the mean rise in Chinese imports (in value) per worker within each zone. The instrument is the same ΔIPW, in which French trade data has been replaced by a control group of four countries (Japan, Germany, Spain, Switzerland) and all labour force variables are taken with a decadal lag. The period of estimation is 2012-2017 for ZEs, 2001-2018 for cities. All specifications include the full vector of controls mentioned in table 3; for the city-level strategy, we ascribe to each city the indexes ΔIPW, the routine, offshorability, and machine penetration indexes of the ZE to which it belongs, other controls being city-specific. Observations are weighted by the start-of-the-decade total population of the zone reported in the IRCOM or Filosofi base. Standard errors are clustered at the level of the INSEE's superzones. The main line denotes the coefficient $\hat{\beta}_1$, with the corresponding 95% conf. interval; dashed line provide the corresponding $\hat{\beta}_1$ when the mean rise in fiscal income in the zone is the dependent, i.e. for the first panel -2.17 (t=5.38), and for the second one, respec. -2.13 (t=1.77) and -1.51 (t=1.64).

Figure 22: Distributional impact of the shock *within* regions – Shares estimates

(a) Decadal growth of the shares of total local fiscal income held by each *within*-decile of *départements* (in pp)

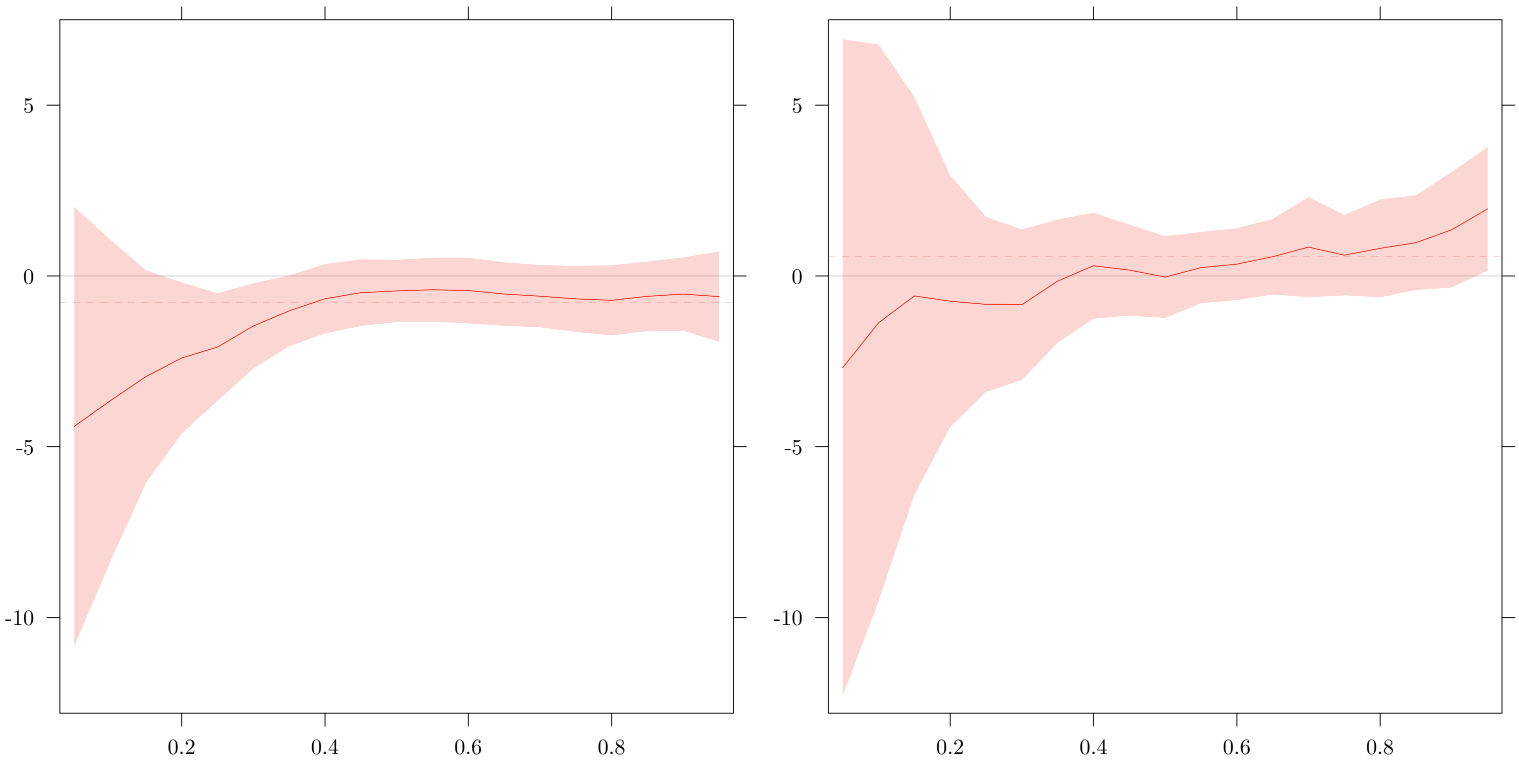


(b) Decadal growth of the share of total local fiscal income held by each *within*-decile of *communes* (in pp)



Note: The unit of interest is the *département* or the city-*commune*, the main source being the IRCOM database (restriction 0 for the latter case, restriction 2 for cities). The dependent variable is the evolution (in pp), for each within-region-decile of the fiscal income distribution, of its share of total regional income as a ratio of the initial share. The main explanatory variable is the index ΔIPW, described herein above, which provides an estimation of the mean rise in Chinese imports (in value) per worker within each *département*. The instrument is the same ΔIPW, in which French trade data has been replaced by a control group of four countries (Japan, Germany, Spain, Switzerland) and all labour force variables are taken with a decadal lag. The period of estimation is 2001-2018, i.e. two decades, and we estimate decade by decade (i.e. the ΔIPW of decade t is meant to have a causal impact on the income evolution of decade t only, not on the evolution in $t+1$). All specifications include the full vector of controls mentioned in table 43. Observations are weighted by the start-of-the-decade total population of the *département*. Standard errors are clustered at the level of the INSEE superzones. Bars denote the main coefficient, with the corresponding 95% conf. interval.

Figure 23: Isolating the role of offshoring in the fiscal income response

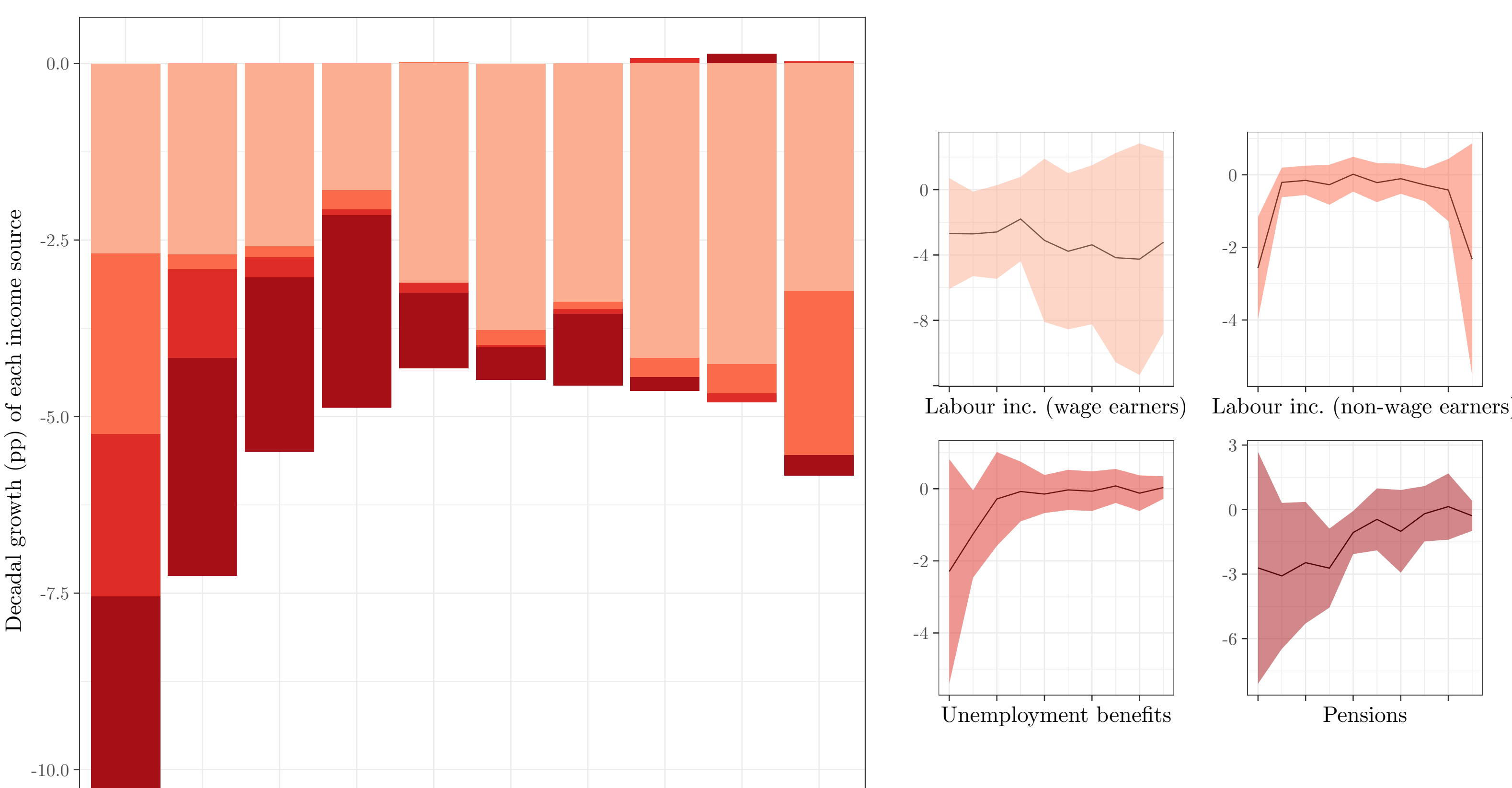


(a) Restriction to consumption goods

(b) Restriction to intermediate products

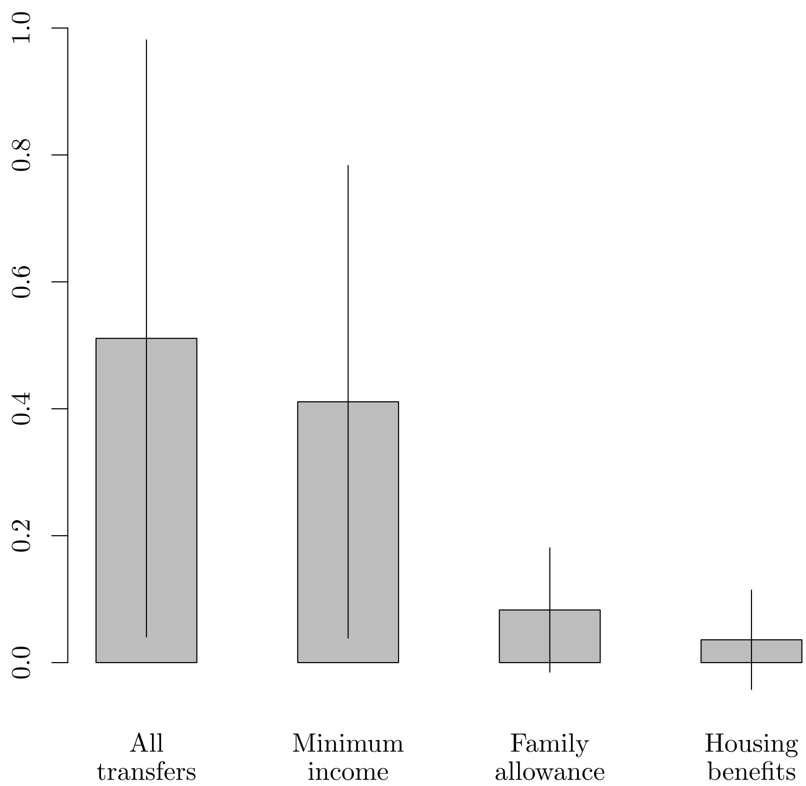
Note: These two figures are replicates of fig. 21a (the only difference being the removal of the offshoring index as a control variable), this time using as explanatory the exposure index ΔIPW described herein above and the corresponding instrument, but over a restriction to goods and products which are considered within the BEC-*Broad Economic Categories classification* as either intermediate or consumption goods. We use the concordance tables between the HS, STIC and BEC scales provided by the UN Statistics division: note that all products cannot be mapped from the latter to the former one.

Figure 24: breakdown of the fiscal income impact of the shock



Note: The unit of interest is the ZE. The main source is the Filosofi database (2nd-breakdown datasets). The dependent variables are the decadal evolution (in percentage points) of the four exclusive main sources of fiscal (pre-redistribution) income – labour incomes of wage earners or of independent workers, unemployment benefits, and pensions – for each household weighted by its number of *unités de consommation*. The main explanatory variable is the index ΔIPW, described herein above, which provides an estimation of the mean rise in Chinese imports (in value) per worker within each *département*. The instrument is the same ΔIPW, in which French trade data has been replaced by a control group of four countries (Japan, Germany, Spain, Switzerland) and all labour force variables are taken with a decadal lag. The period of estimation is 2012-2017. All specifications include the full vector of controls. Observations are weighted by the start-of-the-decade total household population of the ZE. Standard errors are clustered at the level of the INSEE superzones. Lines denote the main coefficient, with the corresponding 95% conf. interval.

Figure 25: Impact on the share of transfers in final income



Note: The unit of interest is the ZE. The main source is the Filosofi database. The dependent variables are the evolution of each type of transfer as a share of disposable income. The main explanatory variable is the index ΔIPW, described herein above, which provides an estimation of the mean rise in Chinese imports (in value) per worker within each *département*. The instrument is the same ΔIPW, in which French trade data has been replaced by a control group of four countries (Japan, Germany, Spain, Switzerland) and all labour force variables are taken with a decadal lag. The period of estimation is 2012-2017. All specifications include the full vector of controls (at the exception of offshorability and machine penetration indexes due to data limitation). Observations are weighted by the start-of-the-decade total population of the ZE. Standard errors are clustered at the level of the INSEE superzones. Bars denote the main coefficient, with the corresponding 95% conf. interval.

3.4 Summary and provisional conclusion

The main features of the income reaction plotted in figures 22 to 21, and simulated in **??**, critical for the interpretation of our last section, might be summarised as such:

- *A distributional impact less regressive than expected*, at least considering the few existing individual-level estimates [Adão, Carrillo, et al. 2020]; for the top 70%, the income impact is relatively flat, and remains sufficiently low so that, if we rely on usual estimates of gains of trade, these families might be surmised to be better-off once the trade shock is fully realised;

- *A sharp discontinuity of the fiscal income impact below the third decile* – The simulation of figure **??** indicates a flat and insensitive impact down till somewhere around the 3rd decile or first quarter, where the estimated loss brutally jumps down. As a consequence, the bottom 30% experience considerable fiscal income losses, up to −25pp over 1999-2008, which is far beyond any estimates of gains of trade. The implied decadal impact on the national ratio T10/B50 is +0.66 for the 1999-2008 decade, +0.16 for the 2008-2018 one. This finding is consistent with other estimates based on labour income [Chetverikov, Larsen, and Palmer 2016] or final income [Adão, Carrillo, et al. 2020] data[28]; concurrently, aggregate growth impact curves at the European level [Blanchet, Chancel, and Gethin 2019] display non-linear behaviours below the 3rd decile which might encapsulate some long-term effects of the shocks driven by international economic integration;

- *A quite different picture whether we focus on fiscal or on disposable income*; We had evidence, in European context, of redistributive policies being powerful enough to counter the depressing income effects of trade shocks, so that the final aggregate impact on personal income be nonsignificant [Utar 2018]. However, to our knowledge, the distributive dimension of this fiscal versus disposable difference had never been investigated before. It shows that, if at the aggregate level, redistribution reduces the total ratio T10/B50 by a rough one third margin [Chancel 2019b], in the case of trade shocks, it seems surprisingly efficient at suppressing almost the entirety of its inequality impact. However, there remains a margin of lower-middle-class households which fall within the hotspot of the employment and income impact of the import shock, but outside of the usual perimeter of many social transfers. More detailed individual data would be necessary to fully illuminate that point, but it is a crucial indication for the interpretation of the social and political side-effects mentioned hereinafter;

- *Little impact of the between-region divide* – If we replicate the Theil breakdown exercise of [Blanchet, Chancel, and Gethin 2019] at the nation's level, using inequalities of fiscal income *within* and *between* ZEs in 2012 over

[28] Studying the last decade, [Adão, Carrillo, et al. 2020] finds a decadal impact of the order of +1.7 to the ratio T10/B50 of income (in Ecuadorian contexts, where the initial ratio is twice greater than European figures).

the Filosofi database, we find that *between*-zones inequalities account for at most 6% of the total (the *between*-Theil index is 0.013, compared to an average *within*-Theil index of 0.203). In our Monte Carlo simulation at the ZE level, the first shock (the *between*-zones impact) has paradoxically a slight negative effect on the ratio T10/B50 (-0.006), consistent with the idea that more exposed regions are overall slightly richer. I.e., if regional variability is crucial to our identification strategy, in itself, inequality between regions mediates but a very marginal part of the income impact of shocks caused by international economic integration; we'll draw a very similar conclusion later about the political impact of such shocks;

- *Investigating the U.S. versus Europe divide* – We have seen that import exposure in French ZEs over 1990-2008 was very similar in magnitude to the ones of U.S. CZs. Yet both countries have followed quite different patterns of inequality dynamics in the last three decades. Actually, shocks caused by international economic integration might have put national social security schemes to test; European systems, as shown in our setting or in European equivalents [Utar 2018], have been surprisingly efficient at deadening the negative income impact of trade shocks, while we might doubt the ability of the U.S. system to do so. Besides, a common feature of the French, but also of the German case [Dauth, Findeisen, and Suedekum 2021] is that a large part of the most exposed regions are overall more export-oriented, more integrated and more dynamic, while in the articles of Autor, Dorn and Hanson, the general portrait of the most exposed zones draws a consistent picture of declining industrial bastions for which the import shock is the final straw;

- *Other channels* – To our final income estimates, we should add the distributional impact of other effects of trade exposure, most notably the export channel (which is slightly progressive in the most recent estimates of [Adão, Carrillo, et al. 2020]) and the expenditure channel (for which there is a heated debate to know whether [P. Fajgelbaum and Khandelwal 2016] or not [Borusyak and Jaravel 2021] lower income families' consumption patterns are more intensive in imported goods).

A distributional impact smoother than could be expected ; effects which are concentrated on the middle city and on the lower-middle-class worker; a rise of redistribution which counteracts the negative effects of the import shock, but fails to bring full compensation for families lying around the 3rd and 4th quantile of the national income distribution. It is quite obvious that such a structural income impact is doomed to spawn social and political reactions which will not be focused on distributional concerns; on the contrary, it seems that all is set to obfuscate the political consciousness of these issues. The discontinuity found in our estimation **??** seems to create a breach within the bottom 50%, with ominous social consequences on which we'll now be focusing.

4 Side effects on social polarisation

4.1 Introduction – Exploiting residential strategies to assess the rise of social intolerance

Towering the recent research on the social consequences of trade, [Autor, Dorn, and Hanson 2013] has spawned a very large empirical literature which draws a dreary picture of the longer term social consequences of the industrial decline spawned by import exposure: this wide array of investigations, reviewed by [Dorn and Levell 2021], ranges from unemployment hysteresis to the longer term impact on manufacturing workers' health or on the school performances of their children.

In our setting, as emphasised above, the most striking result of our income estimates is this sharp discontinuity in the fiscal income response found around the third decile, that rift created within the lower half of the income distribution. We hypothesised it could spawn within the middle-class further doubts about the legitimacy of redistributive policy, the social representation attached to the losing 30% crystallising around *Welfare queen*-like narratives. Another likely consequence involves the social behaviours of lower-middle-class families which lie just above the discontinuity, and for whom it might act as an incentive to embrace social strategies meant to preserve a distinction with derelict families of those bottom 30%; to sum it up, it is likely to bolster the rise, within middle-class families, of the triangular social consciousness we discussed above [Collovald and Schwartz 2006].

Such mechanisms have been well described in the context of a declining industrial bastion (Sochaux) by [Beaud and Pialoux 2003]; shrinking manufacturing employment does not only cripple unionism and political organising at the firm level; it creates a more profound rift within working families; in the aftermath of major layoffs of the late 1990s, there's a rise of social anomy within the families of lower-paid-lower-skilled workers; Beaud & Pialoux notice growing concerns, among households of more qualified blue-collars, about the declining social atmosphere of working-class suburbs, the families with sufficient financial leverage even starting to leave the district, which becomes more and more segregated throughout the 1990s and 2000s.

Within the economic literature, there is a simple model to formalise such cumulative processes: the tipping point model of Thomas Schelling [Schelling 1971]. It combines a spatial residential framework and a micro model where households of the dominant or privileged group have a distaste for residing near families of minority groups, with the intuition that, even if the disutility of the lower-group-proximity is very low, complete residential separation is the only Nash equilibrium of the setting[29]. A key feature is that the utility function of the members of the upper group is non-monotonous; it brutally jumps down when the share of the lower-group population in a district exceeds a certain level l^*, i.e. the upper members are indifferent to the presence of the other group till a certain level: subjectively and at a micro level, this level is a *threshold of tolerance*, but at the aggregate macro level, it becomes a *tipping point*, i.e. a massive flight reaction of the upper group. In this setting, one household moving away for a mixed district creates considerable negative externatilities for those remaining, generating a chain reaction which stops when the system reaches absolute segregation. I.e. even if mixity is a social optimum, and even if agents are aware of it, individual choices are set in a way that separation is the only Nash equilibrium.

To our knowledge, such a model has never been applied over the INSEE's data, while the dynamics of residential segregation is a widely debated issue within the French statistical literature [Préteceille 2009].

4.2 A Schelling model applied to the French context

A focus on tolerance thresholds towards minority people

One of the most well known empirical transcriptions of this model in the U.S. context is to be ascribed to [Card, Mas, and Rothstein 2008]. Their original intuition might be summarised as such. If tipping patterns do exist, we should be able to detect structural breaks in the variation of the share of the upper group within each district, when the share of the lower group evolves over time. However, a model which would use these two variations as dependent and explanatory variables would be faced with an obvious problem of colinearity. Hence the idea to compare the share of the lower population at the beginning of a decade, and the evolution of the upper population over that decade. We assume that over the decade, the upper group households will move to the most privileged districts which have the lowest shares of minority population.

Main specification

Formally, the explanatory variable is defined as the share of the lower group, in district i of region c at time $t-10$ (i.e. at the beginning of the decade):

$$l_{ic,t-10} = L_{ic,t-10}/N_{ic,t-10}$$

Where l and L go for the share and population of the lower group, and N the total population of the referenced district.

The dependent variable is the evolution of the share of the upper group, defined similarly as:

$$\Delta u_{ic,t} = (U_{ic,t} - U_{ic,t-10})/N_{ic,t-10}$$

For each region c, two specifically built algorithms (which are detailed in the annex) determine the tipping point, i.e. the initial share $l^*_{c,t-10}$ above which there is a brutal drop in the upper group population. The first algorithm (our *method 1*) is based on a structural break research device; the second one (our *method 2*) seeks a fixed point, modelising the variations of the dependent as a quartic polynominal and finding its roots. The points are identified over a randomly drawn 2/3 subsample of the data, and the whole model is estimated on the remaing 1/3. Variable

[29] A detailed review of these early models is provided by [Goffette-Nagot, P. Jensen, and Grauwin 2009]

δ is defined as the distance between the tipping point for the whole region and the lower group share in district i of that region:

$$\delta_{ic,t-10} = l_{ic,t-10} - l^*_{c,t-10}$$

The final model is specified as:

$$\Delta u_{ic,t} = p(\delta_{ic,t-10}) + d\mathbb{I}[\delta_{ic,t-10} > 0] + \tau_c + X_{ic,t-10}\beta + \varepsilon_{ic,t} \tag{24}$$

Where p() is fourth-order polynomial, τ_c a region-specific fixed effect, and X a matrix of controls. d is the coefficient of interest, which shall provide the magnitude and significance of the decline in the upper-group population over the tipping point.

The original article relies on data from the U.S. census. The district or *infra*-unit is the census tract, the *supra*-unit, the *Metropolitan Statistical Area*. In our setting, we'll use the INSEE's Census and the RFL-Filosofi database, with the district (IRIS) as the *infra*-unit, and the *Zone d'emploi* as the *supra*-unit[30].

Choice of the lower group

As emphasised by [Goffette-Nagot, P. Jensen, and Grauwin 2009], tipping models have been built to incorporate any definition of the privileged and underprivileged population, be it based on income, ethnicity, education level or another variable. In the original article, the authors focus on the white population, as opposed to the minority (African-American and Hispanic) population, over three decades between 1970 and 2000. They find consistent and significant tipping around their estimated zone-specific points. Typically, in Los Angeles 1990-2000, around a threshold of 15% minority population, there is a −7pp drop in the evolution of the white share. Over the decade 1970-1980, which just follows the banning of segregationist policies, tipping is happening for lower shares of minority population, and results in much more violent drops (with some extrema over −30pp).

Replicating that setting over the INSEE's data leaves us with a dual choice: we can pick either the migration status or the nationality. We shall use the latter one, which has the advantage to be stable over time (an individual might be granted French nationality over the period, while being born within or without French territory is a stable feature). We shall therefore define the majority group as the *natives* (people born on French metropolitan and oversee territory) and the minority group as the *migrants* (people born in a foreign country[31]). The census allows us to test other definitions of the upper and lower group, based on diploma, SES, or SES interacted with country of origin, but these alternative definitions perform but poorly (a wide array of categories have been tested, some major results being reported in our annex).

We develop in the corresponding annex a framework in which we find some evidence of tipping based on incomes, middle-class families reacting to a rise of the local population earning less than the national first decile of the income distribution, but we lack the data that would be necessary to ensure that these estimates are perfectly robust.

4.3 Tipping patterns

Estimated level

We present first our estimates of *origin*-based-tipping points. In table 11, we report the average value of the tipping thresholds estimated by method 1 (structural break) and method 2 (fixed point). Compared to the original article, we find very similar results, but our thresholds are generally a bit lower than Card and alii's recent values (which are respectively 14.5% and 13.9% for methods 1 and 2 over their last decade, 1990-2000).

Table 11: *Origin*-based-tipping - Estimated tipping points

	1999-2010		2010-217	
	Structural break (1)	Fixed point (2)	Structural break (3)	Fixed point (4)
Mean	10.39%	10.87%	7.78%	10.61%
SE	10.84	8.95	10.49	9.66
Without identified threshold	0	0	0	0
Correlations				
1999-2010 Structural break	1.00			
1999-2010 Fixed point	0.21	1.00		
2010-2017 Structural break	0.32	0.18	1.00	
2010-2017 Fixed point	0.17	0.22	0.19	1.00

Points are expressed in share of migrant pop. in district. Summary stats are unweighted.

[30] Contrary to exclusively urban *supra*-units like the AAV (*Aire d'attraction de ville*) or the UU (*Unité urbaine*), the ZE cover the entirety of the French territory; it is convenient for rural areas; for provincial metropolises, choosing the ZE or the agglomeration makes very little difference (it provides a definition of the metropolis which is generally wider than the UU but narrower than the AAV). One main difference between choosing the ZE or the UU-AAV lies in the treatment of the Paris metropolis; it is one huge AAV-UU, split in more than a dozen ZEs; such a division seems more natural, since a detailed analysis shows that tipping points are highly heterogeneous over the capital metropolis, with thresholds generally going down as we move from the center to the periphery. Yet as we show in our annex, the significance and magnitude of the estimations are not fundamentally altered by alternative choices.

[31] The INSEE has a narrower definition of what a *migrant* is, namely a person which is born in a foreign country, did not have the French nationality at birth, and has been living on the French territory for more than two years.

Estimated reaction around the threshold

Table 12 provides estimates for the tipping behaviour around *Origin*-based estimated thresholds. For each method, we provide the average drop in native population when the district-IRIS lies beyond the tipping point of the zone.

In the original article, Card and his coauthours found, for their last decade 1990-2000, a coefficient d of -7.1 and -9.3pp (for methods 1 and 2 respectively), meaning that the magnitude of the effect in our data is approximately between one third and one half of the American figures.

Table 12: *Origin*-based-tipping - Regression discontinuity model for change of native share around the tipping point

	Dependent var.: Change in native population in the district from $t-10$ to t					
	Method 1 - Structural break			Method 2 - Fixed point		
	Base	F.E.	Full	Base	F.E.	Full
	(1)	(2)	(3)	(4)	(5)	(6)
1999-2010 decade						
Beyond tipping point (coef. d)	$-.96$pp	-2.88pp***	-2.49pp***	-3.93pp***	-3.82pp***	-3.04pp***
SE	(1.11)	(0.81)	(0.74)	(1.34)	(0.91)	(0.78)
Observations	16271	16271	16271	16271	16271	16271
R^2	1.5%	18.8%	22.1%	1.3%	18.7%	21.9%
F-stat	16.63	3.91	4.6	8.97	4.12	4.78
2010-2017 decade						
Beyond tipping point (coef. d)	.83pp	-1.45pp**	-1.83pp***	$-.89$pp	-2.93pp***	-2.58pp***
SE	(0.69)	(0.54)	(0.64)	(0.71)	(0.72)	(0.66)
Observations	16281	16281	16281	16281	16281	16281
R^2	4.2%	7.3%	8.4%	4.2%	7.1%	8.5%
F-stat	14.61	5.85	3.44	7.82	5.69	3.32
Zone fixed effects		X	X		X	X
Controls			X			X

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: The unit of interest is the district (IRIS). Data are drawn from the INSEE's Census. We report the estimates of model 24, specifically the estimation of coefficient d. The dependent variable is the growth of the native population (defined as these persons who are born on French territory) within the district from t to $t-10$, as percentage of the base population at date $t-10$. The main explanatory variable is a dummy equal to one if the share of migrant population (those persons who are not born on French territory) is beyond the ZE-specific estimated tipping threshold. All specification also include a quartic polynomial in the deviation from the district's migrant share distance from the local tipping point, plus Z.E. fixed effects. The vector of controls, drawn from the INSEE's census, includes unemployment rate, share of working-class people, share of persons with no diploma, and share of vacant accommodation. Standard errors are clustered at the ZE level.

One important remark: we failed to replicate that setting over former issues of the INSEE's Census (1962, 1968, 1975, 1982, 1990). Actually, replication of the exact same specification is not possible before 1990. IRIS-level data were not provided before that time[32] and information about the country of origin are scarcely reported. For issues prior to 1990, we must rely on a very poor proxy, i.e. a strategy at the commune level[33], using nationality to define the upper and lower groups (French nationals / Foreigners). That strategy fails to identify tipping points: actually, the population of French citizens tends to rise between two issues of the Census in communes with the highest shares of foreigners. Using the share of persons repatriated from Algeria after the end of the war, or the share of Algerian Muslim population living on metropolitan territory (provided for the 1962 issue) leads to the very same result. In issues of the Census in which the share of SES categories can be matched over time, we find no tipping reaction from the upper and middle-class to the local shares of foreign population[34].

Yet we also fail to replicate our *origin*-based strategy on the 1990-1999 decade, while we have under hand everything we need to apply the very same specification. The algorithms estimate consistent tipping points, tipping reaction of the native population around these points is negative, but not significant, or significant at a 10% level in one specification only. This result is however consistent with data analyses [Préteceille 2009] and qualitative evidence [Beaud and Pialoux 2003] which suggest that ethnic segregation, after a slow decline during the 1980s, has been deterioriating since the late 1990s onward.

Graphical evidence

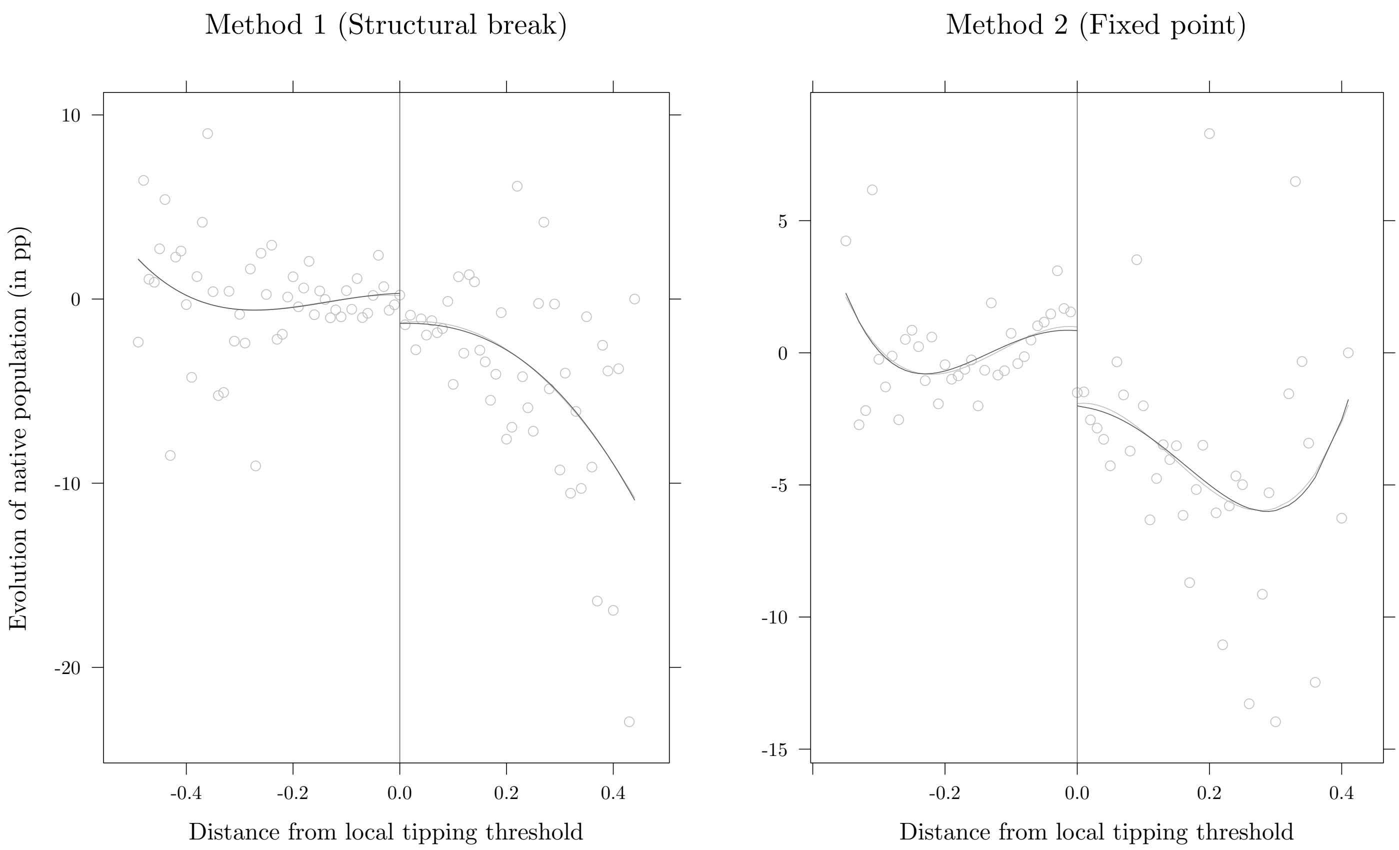
Figure 26 plots the results of the pooled analysis displayed in table 12 for methods 1 and 2. We put, on the x-axis, the share of migrant population (centred around the local tipping point), and on the y-axis, the evolution of the native population (centred around the zone's mean).

[32] A special subdivision of the communes, the *Ilôt*, did exist, but over these, information about nationality and country of origin are provided for one issue only - 1982 - and cannot therefore be matched over time.

[33] There, the commune is the *infra*-unit; as to the *supra*-unit, we tested different options: the 1994 ZEs, the 2010 ZEs, the *départements*; that choice does not change results fundamentally.

[34] Great care must be taken in interpreting these results: contrary to country of birth, nationality is not a stable individual characteristic and can change over the life course. The variations we observe might well be an upward trend in the naturalization process of foreigners.

Figure 26: *Origin*-based-tipping - Pooled estimation



Note: Results displayed come from the specification of table 12. x-axis is the share of migrant population within each district at the beginning of the decade minus the zone tipping point ($\delta_{ic,t-10}$). y-axis is the evolution of the native population as percentage of the base population at the beginning of the decade, centred around its mean ($\Delta u_{ic,t} - \overline{\Delta u_{ic,t}}$). Dots give averages in 1-percentage-point bins. We use two different fits: a 4th-order polynomial with an intercept shift at zero (light-grey) and a 3rd-order polynomial on the two subsamples right and left of zero.

Figure 27 gives an idea of what tipping looks like at a local level, taking the example of 4 zones. For this local approach, we do not center our variables: the x-axis reports the brute district migrant share, the y-axis, the brute evolution of native population share. We represent method 1 thresholds only.

The specific case of the Parisian metropolis is probably the most interesting one. The classical social geography of Paris is polarised over a East vs. West, bourgeois vs popular, axis. However, the analysis shall focus there, not on the extreme cases, but on these cities which lie just above or just below the threshold of tolerance.

These neighbourhoods share one pivotal common characteristics. Most of them are located adjacent to a priority district, but a district which does not belong to the worst cases, and is generally considered by the administration as "low-priority"[35]. Most of these special suburbs are found at the frontier between the East-West poles, around a North-South axis. On the South axis, we might mention Boissy-St-Léger, Longjumeau, Vigneux-sur-Seine/Montgeron, Dammarie-les-Lys/Le Mée-sur-Seine; that last example is paradigmatic : it is a relatively calm low-priority district, but provides the sharpest decline of native population among the whole data (-78.5pp over 1999-2010 in the IRIS immediately adjacent to the Fontainebleau forest, where l_{1999} was just 1.5pp below the threshold; the effect is surely driven by the Plaine-du-Lys QPV; the IRIS at the heart of that suburb exhibits a tipping reaction of -29.2pp, starting from a far superior initial migrant share of 22%)[36]. On the North axis, compelling examples include Franconville, Sannois, Tremblay-en-France, Asnières-sur-Seine[37], Mantes-la-Ville (the IRIS adjacent to Mantes-la-Jolie tips at -28.1pp over 1999-2010 for $l_{1999} = 9.8\%$ migrant population share). One important final remark. Almost every single aforementioned city is either a low-priority district, or a district which was not considered as a QPV before 2015 ; tipping behaviour seems to predate the official classification[38].

Robustness checks

One issue might imperil the zero conditional mean assumption in this framework. The variations of our dependent variable might be driven, not by departures or arrivals, but by some inner phenomenon: maybe the local native population rises because of demographic dynamics (extra births for instance).

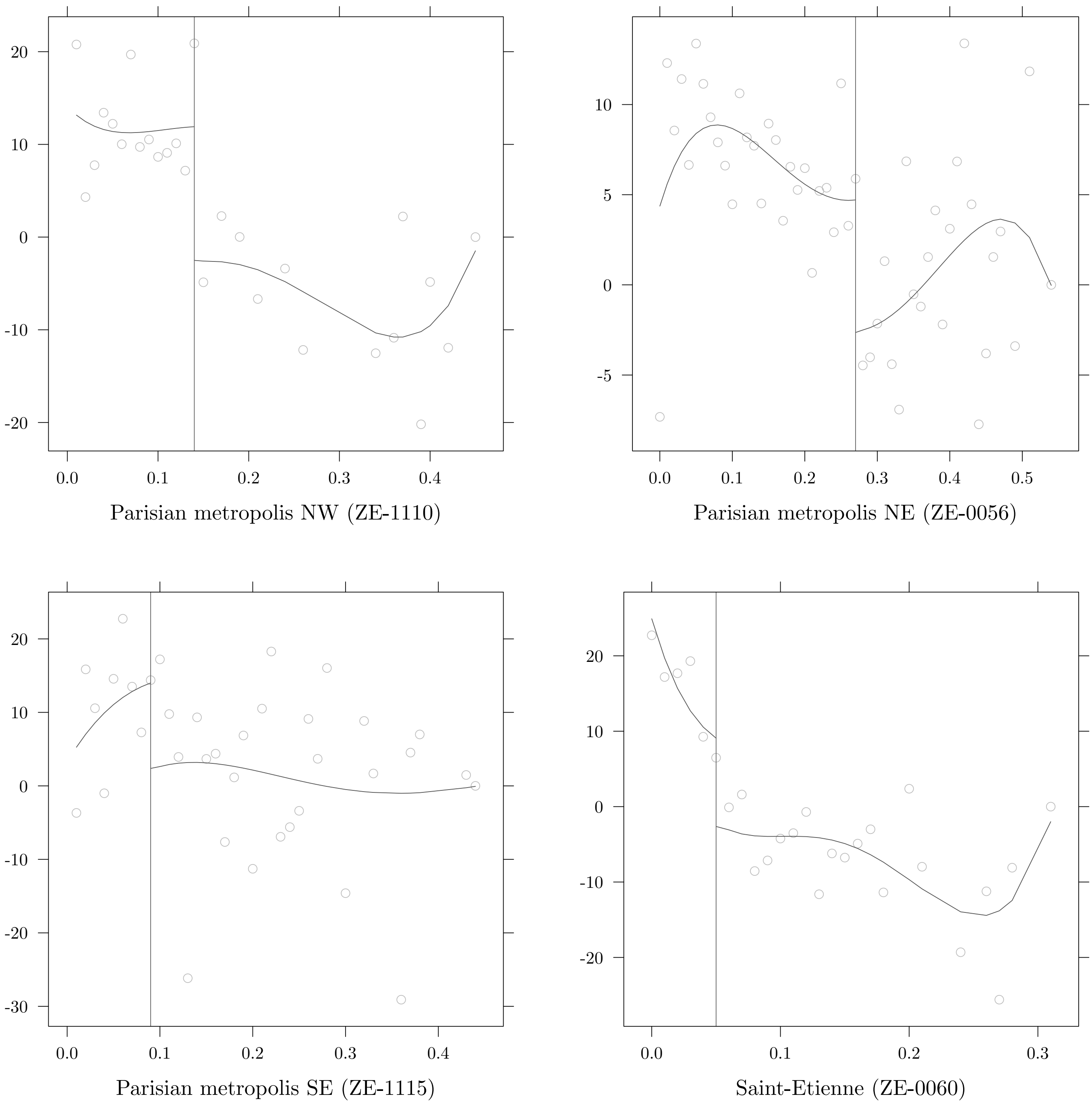
[35] In the technical jargon of the administration, such low-priority districts can be distinguished by the fact that they are not covered by the 2015 initiative known as the Nouveau programme national de renouvellement urbain (NPNRU), or by the fact that the NPNRU is considered of "regional" and not "national" interest.

[36] When it comes to *income*-based tipping, that commune provides also among the sharpest decline of the share of the middle 40% over 2000-2010 in the Paris zone ($\Delta u = -2.56pp$) and it lies just below the theoretical threshold for the area ($l_{2001} = 7.1\%$ vs. $l^*_{2000,Paris} = 7.2\%$)

[37] It is one of the most paradigmatic examples for the middle 40 vs bottom 10 analysis : $\Delta u_{2001,2011} = -1.8pp$ for $l_{2001} = 7.3\%$

[38] The same pattern can be identified in many provincial metropolises. In Toulouse, there is tipping West of the oldest and most well known QPV, Le Mirail, but the sharpest declines of the middle-class population are observed in recently created QPV units: Blagnac-Baradel, Colommiers-Val d'Aran, Toulouse-Soupetard. In Marseille, the sharpest drops of Δu are happening, not in the popular North, but in Southern (La Cravache - Le Trioulet) and Eastern priority districts (in M40vsB10 analysis, -16.2pp over 2001-2011 in the district just West of La Rouguière, where the initial l was relatively low, at 8.5%). In the zone of Saint-Etienne also, tipping is observed in peripheral cities relatively preserved till now: in Le Chambon-Feugerolles (in M40vsB10 analysis, -8.3pp, $l_{2001} = 10.2\%$) or at the heart of the ideal city planned by Le Corbusier at Firminy-Vert (in M40vsB10 analysis, -6.9pp, $l_{2001} = 12.6\%$).

Figure 27: *Origin*-based-tipping - Some illustrative zones



Note: x-axis is the share of migrant population within each district at the beginning of the decade ($l_{ic,t-10}$). y-axis is the evolution of the native population as percentage of the base native population at the beginning of the decade, ($\Delta u_{ic,t}$). Dots give averages in 1-percentage-point bins. We use two different fits: a 4th-order polynomial with an intercept shift at zero (light-blue) and a 3rd-order polynomial on the two subsamples right and left of zero. The tipping point estimated through method 1 is shown as a dashed grey line. Results come from four different zones (from top left to bottom right): ZE1110 (Parisian metropolis North-West / Mantes-la-Jolie), ZE0056 (Parisian metropolis North / Roissy), ZE1115 (Parisian metropolis South-East / Evry) and ZE0060 (Saint-Etienne metropolis).

To that objection we can provide both a theoretical and an empirical answer: 1. From a theoretical standpoint, we can show than the growth rates of each local upper group are independently distributed dependent of the size of the upper group at the beginning of the period[39]; 2. As to the empirical dimension, one solution is provided by the grand mobility dataset of the INSEE's Census, which allows us to know, for each recorded individual, in which commune that person lives at time t, and in which commune that person was living at time $t-5$ or $t-10$ (depending on the versions of the set). Over these datasets, it is possible to clearly isolate the flux of departures and arrivals from the inner dynamics. In the annex, we replicate our *origin*-based strategy on these mobility series; we find very similar tipping points, and coefficients d which are notably higher than the ones reported in table 12 (see table 57). If indeed there is a bias in our estimates because of the inner dynamics of the upper group, that bias leads us to underestimate the magnitude of d. Here's a quick overview of the other robustness checks provided in the corresponding annex:

- 1. Falsification exercises and placebos to check for biases in variables and selections:
  - 1.1. *Placebo tests on upper on lower groups* – We replicate our strategy on different alternative, closely

[39]This independence property is easily spoted on detailed data: to take the example of the *income*-based strategy, limiting ourselves to Paris, we see that destitute IRISes with very low values of $u_{i,2001}$ below 25% exhibit almost aleatory variations of the middle-class population (La Goutte d'Or III : $\Delta u_{i,2001,2011} = -1.11$ pp starting from $u_{i,2001} = 25.1\%$ over 2000-2010; La Courneuve - La Tour : +3.9pp starting from 20.1%). Conversely, the clearest tipping reactions are observed next to low-priority districts (with an extremum at -17.6pp in Asnières-sur-Seine, where in 2001, the deciles were almost mathematically drawn, with the bottom 10 making out 10.9% of the inhabitants, the middle 40, 39.93%).

related upper and lower groups. We find no consistent result on any other definition of the groups.

  - 1.2. *Flexible higher-order polynomials in controls* – We test the main specification with a quartic polynomial in each control variable, finding very similar estimates of d.
  - 1.3. *Alternative sources: Mobility series from the INSEE* – See above.

- 2. Robustness checks involving spatial interactions:
  - 2.1. *Alternative spatial levels* – There's a hotted debate in the spatial econometrics literature about a phenomenon known as MAUP (*Modifiable Areal Unit Problem*); i.e. in many empirical models, changing the spatial unit interest imperils the robustness of the estimates; coefficients which are sizeable and significant at the regional level might be nonsignificant at the county level. To check for such a risk, we tested different *supra*-units (UU, AAV) and different *infra*-units (the commune instead of the IRIS). Our estimates are robust to these changes;
  - 2.2. *Controlling for the proximity of a priority district* – See below tab. 13;
  - 2.3. *Controlling for spillover effects between districts* – See below tab. 13;
  - 2.4. *Checking spatial autocorrelation* – In the original article of Card and alii, there were suspicions of strong retroactive influence between districts (one upper group family leaves district i; families from neighbouring districts react with tipping; tipping from neighbours creates a retroactive effect in district i). We tested a Spatial Durbin model, very similar to the one reported in our annex H, to distinguish direct and retroactive effects on tipping; we found that these retroactive effects are most of the time nonsignificant: the decision to leave is influenced by the social situation in the immediate backyard only (a major difference with the American context).

4.4 Tipping patterns and import exposure – Towards an economic interpretation

Significance of tipping models in their original U.S. context

Residential segregation and tipping reactions should not be considered as a universal and timeless feature of hierarchical societies. As emphasised by [Cutler, Glaeser, and Vigdor 1999], in late 19th c. America, all indexes of spatial segregation were paradoxically lower than they are today, and even lower in the South than in the North [Kellogg 1977][40]; in a sense, the racial hierarchy was so stringently enforced at the time that no isolation or distinction was necessary for the dominant group. The real story of American segregation dawns in the 1910s. It is the half-century of the American Apartheid best described by Denton and Masey [Denton and Massey 1992]: flocks of black immigrants move from the South to the major Northern cities. As a reaction, white residents, afraid at the idea of being submerged, build a comprehensive complex of institutions to ensure that their neighbourhood will remain segregated, exclusively-white. Hence these ill-famed legal devices of that time like contracts with restrictive clauses prohibiting resale to black persons (*restrictive covenants*), quotas on mortgage lending to black residents (*redlining*), corrupt promoters threatening locals to sell to a black person in order to make extra-profits (the so-called *blockbusters*). These restrictions were stringently enforced : in 1950 Chicago, 80% of real estate transactions had a racial clause [Clark and Perlman 1948]. Quantitative evidence is striking : the 1950s-1960s are the all-time maxima of segregation indexes[41]. In 1970, the average black person was living in a tract with an average 68% Black share. Concurrently, and maybe paradoxically, in the academic fields, it was a time of great quantitative inventiveness, with the rise of the main segregation measures like the dissimilarity[42] and isolation[43] indexes [O. Duncan and B. Duncan 1955], but also a time of heated debate on whether or not segregation was a social evil in itself; in the Interwar period, Chicago School scholars had emphasised the fact that ghettos of the great metropolises were not necessarily places of social anomy [Whyte 1943] and could even act as springboards for economic opportunities [Halbwachs 1932], as a buffer for a faster integration within the nationwide labour market[44].

The Civil Rights policy of the Kennedy-Johnson administration put an end to this era of institutionalised segregation. Tipping models emerged by the same time, in the early 1970s; it was a paradoxical moment: the great Civil Rights Acts of 1964 and 1968 had curbed the tide of spatial segregation for the first time since a half century; yet the political coalition which had allowed the Democratic party to pass much of these policies fractured very early in the 1970s [Gethin, Martinez-Toledano, and Piketty 2021]; as a consequence, if the legal framework of the great Civil Rights Acts survived, many longer term provisions designed by the Johnson administration to foster residential desegregation were never applied; this is most notably the case of the 1968 Housing Act (one of the "three

[40] As shown by [Cutler, Glaeser, and Vigdor 1999], this will remain a permanent feature of Southern cities; throughout the 20th c., their mean dissimilarity index is systematically around 0.2 lower than the Northern average.

[41] In 1960, we get $D_c = 0.8$ and $I_c = 0.6$ [Cutler, Glaeser, and Vigdor 1999]

[42] The main segregation indexes of the American quantitative tradition were defined very early [O. Duncan and B. Duncan 1955], though attempts at axiomating their structure are very recent [Frankel and Volij 2011; Echenique and Fryer 2007]. The two most well known indexes are the dissimilarity and the isolation index. The dissimilarity index D_c is usually defined as:

$$D_c = \frac{1}{2}\sum_i \left| \frac{b_{c,i}}{B_c} - \frac{w_{c,i}}{W_c} \right|$$

Which averages the ratio of the different census tracts i within a city c, $b_{c,i}$ being the black population of the tract, $w_{c,i}$ the non-black population of the tract, B_c and W_c, the black and non-black population of the whole city. D_c is interpreted as the share of the district's population which should move out to ensure spatial homogeneity.

[43] Using the same notations as in previous footnote, the isolation index I_c is defined as:

$$I_c = \sum_i \left(\frac{b_{c,i}}{B_c} \times \frac{b_{c,i}}{b_{c,i} + w_{c,i}} \right)$$

The isolation accounts for the fact that, even in an ideal city where minority people are not segregated at all (i.e. $D_c = 0$), if that minority represents a very low percentage of the population, it is improbable for a majority person to meet a minority person.

[44] Recent research in economic history tends to inform that thesis. [Pérez 2019] reviews a wide range of primary sources and recent studies about Italians in the U.S., urging the fact that economic integration of migrants was generally smoother in the West than in the great ghettos of Northwestern cities. Consistently, when [Abramitzky, Boustan, and Eriksson 2014] apply their equation main specification to a restriction on urban centres, the initial wage premia become wage penalties; i.e. *ceteris paribus*, urban interstitial districts were not the best places to assimilate economically for a newcomer migrant.

most important pieces of legislation of [his] presidency" according to president Johnson himself) which required, in each new housing project, a 50% quota for public or moderate rents housing. That scheme was terminated by the Nixon administration, which replaced it with two types of measures, suffused with a *self-help* narrative, which would define much of the policy path of the next decades: 1. Voucher programs to help minority families leave impoverished districts (the direct ancestor of what *Moving to Opportunity* will be for the 1990s) ; 2. A dilution policy meant to dissolve ghettos by relocating minority people to a myriad of little public housing projects built among middle-class suburbs.

However elusive their results[45], these policy items were widely imitated in Northern America and in Europe, being in harmony with the intellectual atmosphere of the time: the primacy of individual-level approaches of inequality; the conjecture that, great political aggregates (nation, class, religion) being on the decline, social identity had become more volatile and self-defined, that the social destiny of a person now depended less on its background than on its environment; it was thought that manipulating this environment and the type of people with whom that person was interacting with on a daily basis could fundamentally change its social fortune. Critical and acritical transcriptions in social sciences spawned a surge of phenomenological and interactionnist approaches, the econometrical equivalent of which were the peer-effects literature spawned by the seminal model of [Manski 1993]. Within this vein, tipping models might be construed as a formal attempt to nuance the tenets of the time, with the paradoxical conclusion that, even a purely interactionnist framework, with rational and benevolent agents, can spawn a social structure as thoroughly segregated as it was at the time of the American Apartheid, a phenomenon against which interactionist policies are doomed to remain powerless. Very early, the peer-effects literature itself provided intuitions that network impacts where either spontaneous representations of social life unidentifiable within a cautious research program [Abdulkadiroglu, Angrist, and Pathak 2011] or that they were encapsulating more structural underlying forces [Card and Giuliano 2016]. Unsurprisingly, this pervasive intuition, common to other academic fields, was doomed to generate a dialectic reaction of equal magnitude in favour of structural and holistic approaches.

The French context – Fleeing districts rather than people

Our tipping estimates are consistent with recent findings about the evolution of segregation indexes within French cities. [Botton et al. 2020] conclude that, over the INSEE's Census data 1990-2015, the dissimilarity index of extra-European migrants is going *down* if we take the district (IRIS) as the infra-unit, but going *up* if we take the commune as the infra-unit. It seems to follow the pattern of an expanding ghetto mechanism : migrant population is rising in individual houses IRISs that lie just adjacent to a priority district (QPV, former ZUS); middle-class population seems to leave these districts for a commune farther away; that is why migrant segregation *within* the communes of French metropolises is declining, but the segregation *between* the communes is on the rise. Also consistent with the expanding ghetto hypothesis is the finding that, at the IRIS level, if the dissimilarity index of extra-European migrants is going *down*, while their isolation index is going *up* [Pan Ké Shon and Verdugo 2014]. This spatial concentration of segregation effects is a differential feature compared with the American estimates. It might suggest that native families are tipping, not necessarily to avoid special groups, but rather to avoid specific impoverished districts.

Besides, there is evidence that tipping reactions are not limited to the *origin* dimension. In our annex F, we manage to identify tipping thresholds based on *income*, with the families belonging to the first decile of the national fiscal income distribution as the lower-group. Since many districts with high concentrations of non-native population are plagued with an accumulation of social and economic disadvantages [Algan, Hemet, and Laitin 2016], we could fear that our main specification might be unable to disentangle the origin-based tipping reaction from other driving variables. It is all the more true if we heed to the well-identified feature of the French case that the social perception of these phenomena has crystallised over the official State-sponsored proto-*Affirmative Action* labels: the ZUS-QPV labels for cities, the ZFU label for firms, and the ZEP-REP labels for primary education. If the impact of the two latter ones is discussed [Lafourcade and Mayneris 2017], school labels are known to trigger sizeable tipping reactions; [Beffy and Davezies 2013] estimate that, when a school becomes eligible for such a label, over 50% white-collar children and more than 80% of children of teachers leave instantly, whatever the initial social conditions in the school or in the surrounding district.

Hence the idea to check whether or not the tipping reactions we identify are concentrated around the most impoverished suburbs of the country as defined by official district-level labels, formerly the list of the *Zones urbaines sensibles* (ZUS) and since 2015, the new list of the *Quartiers prioritaires de la politique de la ville* (QPV).

In order to do so, we adapt one of the robustness checks of the original article[46]. Using interacted dummies, we shall see whether our main tipping coefficient is significantly different depending on the distance to the nearest priority district (with three intervals: less than 1km, between 1 and 3, more than 3). The results of that exercise for *origin*-based tipping points are displayed in table 13, those for *income*-based tipping points in table 56. These tables might be read as such: 1. The main effect provides an idea of the tipping reaction in the immediate vicinity of poor suburbs; 2. Line 3 provides an idea of that reaction in urban context; 3. Line 5 gives an estimate for a rural or semi-rural context. In column (7) and (8), we also test for the existence of spillover effects between districts[47]:

[45] As to the flaws of *Moving to Opportunity*-like projects, we might quote [Ludwig, G. Duncan, and Hirschfield 2001] or [Chyn 2018] ; for the failure of relocation policies, [Oreopoulos 2003].

[46] Card, Mas & Rothstein hypothesised that tipping might be triggered by some missing variable, one of them being the proximity of a ghetto or minority-dominant district. Over their 1990-2000 decade, it seems that tipping more than 2 miles away from a ghetto is not 5% significant. They also gauge the impact of distance from the CBD to isolate the impact of the great postwar white flight from the center of U.S. cities, but that did not seem relevant for our setting.

[47] In the original article, Card and his coauthors found a very large negative coefficient on this interacted variable Beyond TP $\times$ None of neighbours with $l > l^*$, of far greater magnitude that the main effect (-32pp vs -3pp for the 1980-1990). I.e. when a district located in a preserved or privileged neighbourhood tips, tipping reaction is much more violent. The authors interpret that result as evidence for the existence of considerable spillover effects between districts. The share of lower-group population in the district of residence might actually be only a proxy of the real variable driving departures; it seems like to some white families, there's a huge preference, not only

Table 13: Tipping reaction by distance from nearest priority district (*origin*-base tipping)

	Dist. to ZUS		Dist. to QPV		Dist. to new QPV		By nearby spillovers	
	1999-2010	2010-2017	1999-2010	2010-2017	1999-2010	2010-2017	1999-2010	2010-2017
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Main effect: beyond tipping point	-7.03^{***}	-2.97^{***}	-6.65^{***}	-3.51^{***}	-4.96^{***}	-2.82^{***}	-3.23^{***}	-1.84^{***}
	(0.75)	(0.65)	(0.71)	(0.63)	(0.85)	(0.63)	(0.55)	(0.45)
Interacted: Beyond TP × Nearest priority district is 1-3 km away	2.33^{***}	0.32	2.19^{***}	1.29^{**}	1.25	0.81		
	(0.65)	(0.56)	(0.59)	(0.51)	(0.75)	(0.51)		
Total tipping effect	-4.69^{***}	-2.64^{***}	-4.47^{***}	-2.22^{***}	-3.71^{***}	-2.02^{*}		
	(0.65)	(0.52)	(0.65)	(0.51)	(0.64)	(0.54)		
Interacted: Beyond TP × Nearest priority district is >3 km away	5.61^{***}	1.93^{***}	6.03^{***}	2.76^{***}	2.54^{***}	1.36^{**}		
	(0.67)	(0.61)	(0.64)	(0.67)	(0.76)	(0.56)		
Total tipping effect	-1.43^{**}	-1.04^{*}	-0.62	-0.76	-2.42^{***}	-1.45^{***}		
	(0.58)	(0.52)	(0.61)	(0.57)	(0.58)	(0.48)		
Interacted: Beyond TP × None of neighbours with $l > l^*$							2.56^{**}	1.32
							(1.13)	(1.41)
Total tipping effect							-0.66	-0.53
							(1.21)	(1.44)

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: The unit of interest is the district (IRIS). Data are drawn from the INSEE's Census. We report the estimates of model 24, specifically the estimation of coefficient d. The dependent variable is the growth of the native population (defined as these persons who are born on French territory) within the district from t to $t-10$, as percentage of the base population at date $t-10$. The main explanatory variable is a dummy equal to one if the share of migrant population (those persons who are not born on French territory) is beyond the ZE-specific estimated tipping threshold. All specification also include a quartic polynomial in the deviation from the district's migrant share distance from the local tipping point, plus Z.E. fixed effects. The vector of controls, drawn from the INSEE's census, includes unemployment rate, share of working-class people, share of persons with no diploma, and share of vacant accommodation. The specification is still 24, now fully interacted with the indicated tract characteristic. We report results using the tipping points estimated with the fixed point method. In column 7 and 8, *neighbours* are the four closest IRISes (computed in terms of distance from centroid to centroid). Observations are weighted by total Census population, and clustered at the level of the region (pre-2015 geography).

In the original article, all estimates of d were robust to an interaction with the distance to the nearest ghetto, i.e. racial tipping was nationwide in the late 20th c. United States, not concentrated on urban areas or around ghettos. Here on the contrary, we often fail to identify significant tipping reactions more than 3km away from a priority district; it seems that the aggregate effect we identified might be considerably driven by a priority-district-stigma[48]. However, we find some indications of a slow transition to the nationwide framework characteristic of the U.S. case: in the 2000s, tipping seems highly concentrated around poor suburbs ; away from urban centres, we even find nonsignificant *positive* values of d for the *origin*-based strategy. The early 2010s display quite a different picture: tipping reactions have now become more systematic, uniformly distributed over the territory.

The interpretation of this concentration around QPVs is more challenging. One line of argument that the tipping framework itself can discard is the idea that isolation would be driven by non-natives themselves, a thesis [Cutler, Glaeser, and Vigdor 1999] label the *port of entry theory* which, in the U.S. context, owes much to the theses of the Chicago School mentioned above.

Its main basis in contemporary research is that paradoxical finding of the literature that, even if higher segregation within a city [Cutler and Glaeser 1997] or equivalently higher diversity [Hémet and Malgouyres 2018] might hurt the labour prospects of minority workers, these workers seem to draw sizeable benefits from local hiring networks [Bayer, Ross, and Topa 2008], which are known to be racially stratified (i.e. neighbours are more likely to help each other find a job if they are of the same race [Hellerstein, McInerney, and Neumark 2011]) and relatively efficient to help non-native workers to find a job in a labour market with hiring discrimination[49] [Dustmann, Glitz, et al. 2016[50]; Aslund, Hensvik, and Skans 2014; Giuliano, Levine, and Leonard 2009[51]; Bandiera, Barankay, and

for non-mixed districts, but also for non-mixed districts with non-mixed neighbours. That peculiarity is absent in our replication. The coefficient on the interacted variable Beyond TP × None of neighbours with $l > l^*$ is positive-significant, and total tipping reaction in districts with preserved neighbourhood, generally not significant. I.e., in the French context, the decision to leave is driven by the situation in the immediate vicinity, not by the situation of neighbouring districts, and at the margin, in the most preserved districts, spillover effects between IRISes might even be positive (middle-class family accept to remain in a district which has tipped, because neighbouring districts are still segregated (it is important to heed to the fact that the definition of what a preserved neighbourhood means varies considerably between zones.In provincial context, it generally means that all neighbouring districts have $l < 0.05$; within areas of the Paris metropolis, it is rather $l < 0.15$.).

[48] We must be extremely careful when comparing these results with the original article. Card and his coauthors defined a *ghetto* as a district with more than 60% minority population. If we use that definition over the INSEE's data, even with a far lower ceiling of 30% migrant share, distance to the nearest ghetto has no significant impact on the tipping behaviour. We find significant impacts when we use the distance to a priority district, and it is important to keep in mind that the QPV list is a very wide definition of what a poor suburb is (there are almost no districts in Paris or Lyon which are more than 3km away from a QPV). It is quite possible that our estimates simply capture the fact that tipping is exclusive to urban areas.

[49] It is an important feature of the works of [Aslund, Hensvik, and Skans 2014] or [Brown, Setren, and Topa 2012] that they seem to confirm the theory of statistical discrimination : for low-skilled jobs, in which a C.V. does not provide much information about the quality of a person, an extra source of information, coming from a referral for instance, might overcome the prejudice of managers against some types of persons.

[50] [Dustmann, Glitz, et al. 2016] use a German dataset about newly hired employees ; for each employee, the share of workers of the same ethnicity that this person is used as a proxy to determine whether or not that person was hired externally, or through an informal job search ethnic-based network. The authors estimate that employees hired through these informal channels earn higher wages, are less likely to leave the firm, but experience a slower rise in wage.

[51] [Aslund, Hensvik, and Skans 2014] use a Swedish linked employer-employee database to show that non-native managers significantly

Rasul 2009[52]]. Even though some evidence suggests that these informal racial networks might lock-up minority workers in low-skilled, bad paying jobs [Hellerstein, Kutzbach, and Neumark 2014], where pay raises are slower [Dustmann, Glitz, et al. 2016][53], these local networks might be a second-best for many non-native people. Yet on existing French data, it is difficult to substantiate such a thesis; in sheer descriptive statistics, we find no evidence that non-native families belonging to different SES groups tend to conglomerate together; the few significant correlation we found over Census data point to the opposite effect.

A more compelling interpretation involves the residential strategies of native middle-class families determined to avoid the reputational penalties attached to these districts, penalties which seems to have a strong impact on labour market outcomes. We allude to this special vein of the *Spatial Mismatch* literature [Gobillon, Selod, and Zenou 2005] which investigates the reputational effect of the place of residence indicated in a CV; in French context, there's evidence that indicating an address in a priority districts on one's CV is a considerable detrimental feature in a candidacy [Bunel, L'Horty, and Petit 2016], even once controlled for the ethnic origin of the applicant [Duguet, Leandri, et al. 2010]; more surprising, it seems like the burden of these penalties is borne by natives rather than by non-natives [Duguet, Gray, et al. 2020], i.e., employers who have overcome the racial prejudice will likely easily overcome the place-of-residence one, while prejudiced employers might overreact to the mention of a QPV in the CV of a native person.

The French context – Tipping reactions more correlated with the conservative than with the FN vote

Another feature which vindicates against an identitarian interpretation of our findings is the fact that, as opposed to the original article, we fail to correlate the level of tipping thresholds with indexes of attitude towards race, especially here, with the FN-RN vote. Actually, it is much easier to isolate a correlation with the centre-right than with the far-right vote, a finding consistent with recent European estimates of the effect of migrations on local political outcomes [Barone, D'Ignazio, et al. 2016; Dustmann, Vasiljeva, and Damm 2019].

In the original article, Card and his coauthors hypothesised that hostility towards contact between races would push the tipping point down. They find a significant correlation between the level of the threshold in each MSA and a specifically built Race attitude index. To take two extreme examples, in a city exemplifying the South like Memphis, the tipping point is 7 points lower than in a liberal city like San Diego.

We take a similar approach in the model of table 14, trying to find some correlation between the level of the tipping point and voting behaviour. We aggregated series from the CDSP-Sciences Po about city-level vote shares, to obtain vote shares at the zone level for each presidential election. We create five political aggregates (from far-right to far-left) and use the whole left-wing vote share as the base category. Empirically, the best predictors are the results of the election which happened next to the beginning of the corresponding decade (2002 for the 1999-2010 decade for instance).

Results suggest that tipping behaviour is polarised over the Left-Right axis. We find a strong correlation between the level of the tipping point and the conservative and centre-right vote, which consistently pushes the threshold down. Conversely, when estimated alone, the centre-left and far-left vote shares significantly push that threshold up[54]. The magnitude of the effect is sizeable : for the 2002 election, one extra percentage point for conservative candidates, as compared to the left base, is estimated to drive the tipping point 1.22pp down[55]. The negative marginal impact for the conservative total is not primarily driven by the Gaullist coalition: when we estimate the centre-right and Gaullist parties separately, we often find slightly superior coefficients for the latter ones (even in 2017, the coefficient on the LREM vote is negative, but nonsignificant).

More surprising, we fail to identify a significant impact of the Le Pen vote on tipping behaviour, even when we try a first-difference specification (to test the impact of the evolution of the FN-RN vote on the change in the level of the threshold). This might be explained by the heterogeneity of the far-right vote: we find very low tipping points in the old FN bastions of the South-East (3.9pp for Orange), while more recently conquered areas can exhibit extremely high thresholds (20.3pp for Béziers)[56].

Consistent with this interpretation, it is possible to obtain a significant negative far-right-coefficient in some specifications of table 14 with the addition of regional fixed effects, with the general picture that tipping points tend to be structurally lower (by a 10pp margin) in Provence and Nord regions (in which many old bastion cities of the party are found), but relatively high in other regions of the East where the FN scored recent victories.

recruit more frequently non-native workers ; however, when native managers recruit within a pool of former co-workers, the bias against non-native employees disappears, indicating that the effect is driven, not by taste-based discrimination, but by the fact that non-natives do not have access to sufficiently large informational networks. [Giuliano, Levine, and Leonard 2009] make a similar point on a US dataset.

[52] [Bandiera, Barankay, and Rasul 2009] show in a field experiment that when managers are paid with fixed wages, they will disproportionately recruit persons from their ethnicity, and that these matchs are generally sub-optimal for the firm. Performance pay for managers reduces that bias.

[53] These findings however are not always replicated in the literature : [Brown, Setren, and Topa 2012] is a good counter-example (in their setting, racial networks allow people to find jobs faster, and to be hired at a higher pay).

[54] Estimated alone, the vote shares of L. Jospin in 2002 and F. Hollande in 2012 are among the strongest predictors which are positive and 5% significant

[55] Thresholds are indeed considerably lower in relatively affluent areas which are centuries-old bastions of conservative parties (2.7pp for Les Herbiers, 5.1pp for Rambouillet).

[56] Note however that the correlation might indicate a reverse causality: there is a suspicion, from the descriptive data, that areas with the clearest hikes in the far-right vote between 2002 and 2012 were characterised, over 1999-2010, by low tipping thresholds (+6.6pp in the Le Pen vote share between 2002 and 2012 for Béthune, where the tipping threshold is estimated at 1.1pp over the decade). Besides, in unreported results based on the *Mobilité* datasets of the INSEE, we found some evidence that a rise in the native population coming from a district that has *tipped* over the decade is associated with a rise in the FN-RN vote.

Table 14: Determinants of the *origin*-based tipping point

	Dep. var.: Level of the tipping threshold							
Time period	Decade 1999-2010				Decade 2010-2017			
Election year	*pres. 2002 rd1*		*pres. 2007 rd1*		*pres. 2012 rd1*		*pres. 2017 rd1*	
Method for tipping point est.	*Str. br.*	*Fix. pt*	*Str. br.*	*Fix. pt*	*Str. br.*	*Fix. pt*	*Str. br.*	*Fix. pt*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Vote share in corresponding election								
Far-right	−0.07	−0.04	−0.01	−0.14	0.38	0.21	0.05	−0.002
	(0.16)	(0.22)	(0.29)	(0.32)	(0.25)	(0.24)	(0.13)	(0.08)
Right + centre-right	−1.22**	−0.18	−0.74**	−0.05	−0.58*	−0.31**	−0.46**	−0.17**
	(0.43)	(0.21)	(0.32)	(0.19)	(0.32)	(0.15)	(0.17)	(0.08)
Centre-left + far-left as base								
Controls	X	X	X	X	X	X	X	X
Obs.	304	304	304	304	304	304	304	304
R^2	0.18	0.14	0.09	0.15	0.23	0.05	0.31	0.05
F-stat	11.1***	8.3***	4.9***	8.4***	14.3***	2.7**	21.7***	2.7**

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: The unit of observation is the ZE (*Zone d'emploi*, INSEE 2010 def.). The dependent variable is the level of the tipping threshold of natives, reacting to non-native pop. shares, estimated with both methods in tables 12, over the data of the Census, for the decade indicated in the first row. The main explanatory variables are the vote shares in the presidential election of interest. Political groups are defined in annex table 58. Controls are taken from the Census and include: the share of residents in the IRIS who live in a single-family home, the share of those who own their home, the share of those who live in a public housing unit (HLM), and the average number of persons per room in the IRIS. Coefficients on the control variables are always nonsignificant. Observations are weighted by the total Census population, clustered at the level of the INSEE's superzones.

The French context – Consistently lower tipping thresholds in import-competition-exposed ZEs

Having discarded identity-based interpretations of the evolution of tipping thresholds, we can now assess the impact of industrial decline (and most specifically there, that specific decline which is driven by trade shocks) on tipping reactions. This hypothesis is tested in a 2SLS specification similar to model (3), the results of which are reported in table 15; rise in import exposure seems to indeed push the tipping threshold down.

Table 15: Simple model for the location of the *origin*-based tipping point

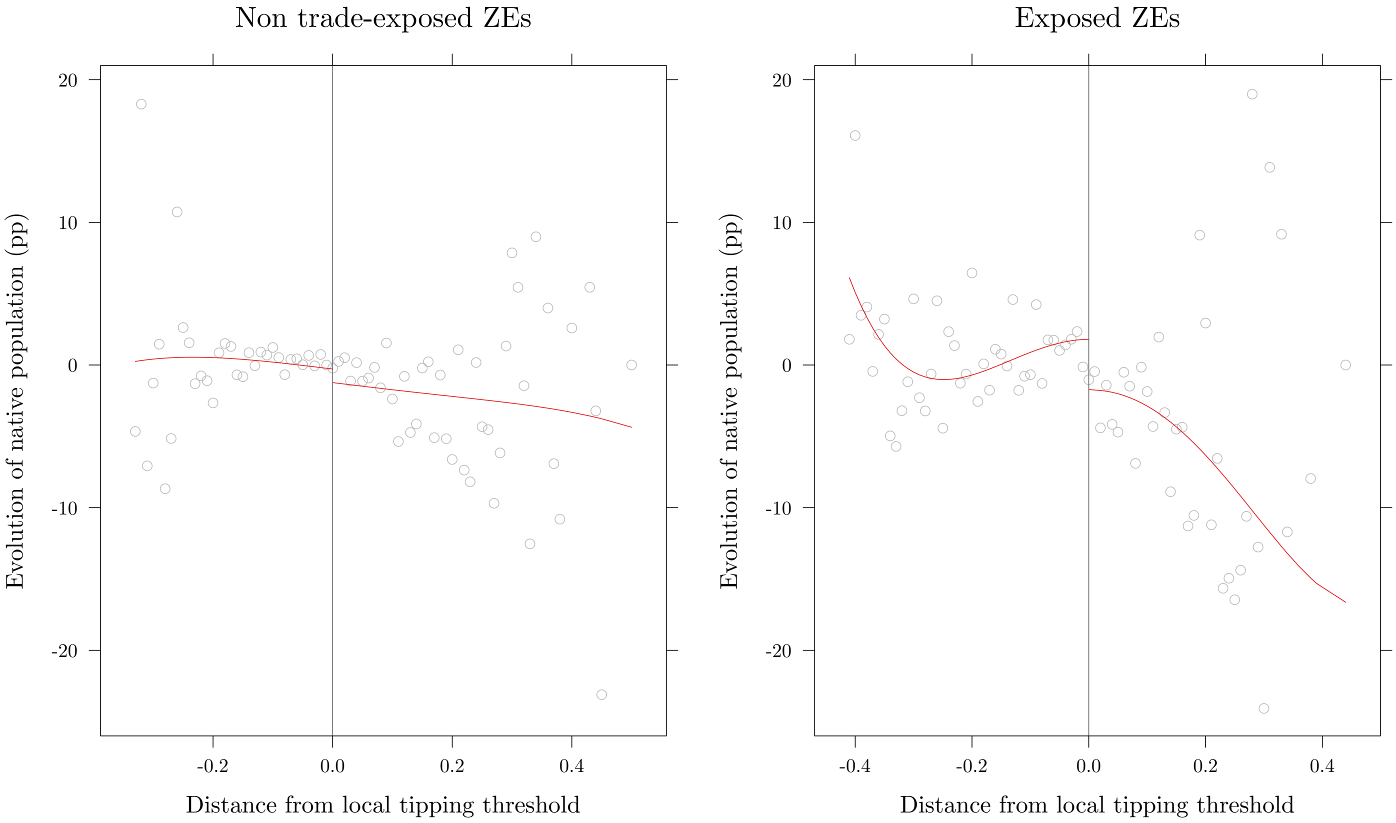
	Dep. var.: Evolution of the tipping threshold (1999-2010 versus 2010-2017)			
	Origin-based T.P.			
	(1)	(2)	(3)	(4)
Rise in import compet. exposure				
Panel A. 1999-2008 exposure				
	−2.21**	−4.36**		
	(1.01)	(1.21)		
Panel B. 2008-2018 exposure				
			−1.69	−1.93
			(5.42)	(9.31)
Controls		X		X
Obs.	304	304	304	304
R^2	0.18	0.29	0.02	0.31
F-stat	3.1***	3.9***	2.6*	3.9***

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: The unit of observation is the ZE (*Zone d'emploi*). The dependent variable is the difference between the tipping thresholds of each ZE for the second decade (2010-2017) and for the first decade (1999-2010) as estimated with the fixed point method, the estimation of which is reported in table 12. The main explanatory variable is the index ΔIPW, described herein above, which provides an estimation of the mean rise in Chinese imports per worker within each ZE. The instrument is the same ΔIPW, in which French trade data has been replaced by a control group of four countries (Japan, Germany, Spain, Switzerland) and all labour force variables are taken with a decadal lag. The model is otherwise similar to 3, with or without the full set of controls of the original specification. Observations are weighted by the total population reported in the 1999 Census. Standard errors reported in parentheses are clustered at the level of the INSEE's superzone.

Not only is the tipping point lower in more exposed regions, but the tipping reaction itself seems more pronounced. If we distinguish the winning and losing ZEs from trade competition exposure in the sense of the intermediate scenario of our third section above, and reestimate our main tipping model over a newly drawn sample for the first decade (1999-2010), we get a nonsignificant d among winning ZEs (-0.03, $t = 0.07$) versus -4.11 ($t = 5.87$) for losing ZEs. Figure 28 plots the results of that exercise.

Figure 28: Tipping reaction around the tolerance threshold depending on the type of ZEs



Note: The unit of interest is the district (IRIS). Data are from the INSEE's Census. The tipping threshold specific to each *Zone d'emploi* estimated through the structural break method is displayed as a vertical grey line. x-axis gives the share of migrant population within each IRIS at the beginning of the decade ($l_{ic,t-10}$) minus the tipping point of the ZE to which it belongs. y-axis is the evolution of the native population as percentage of the base native population at the beginning of the decade, ($\Delta u_{ic,t}$). Dots give averages in 1-percentage-point bins. We plot in red a 3rd-order polynomial on the two subsamples right and left of zero. We estimated the main tipping model separately on the pool of ZEs deemed winners and losers of trade competition exposure as defined in section 3 herein above (intermediate scenario based on gains of trade estimates by [Borusyak and Jaravel 2021]). Observations are weighted by the start-of-the-decade total population. S.E. are clustered at the INSEE superzones' level.

That such tipping behaviours are correlated with industrial decline caused by exposure to import competition is a novel finding, though it is perfectly consistent with the political effects of trade competition exposure general found in recent literature, an issue we will now be addressing.

5 Side effects on political polarisation

5.1 Elements of the political economy of trade

There's a growing body of literature connecting the rise of trade globalisation to political polarisation. Most notably, Autor, Dorn and Hanson expanded their original framework to the political dimension of the China shock [Autor, Dorn, Hanson, and Majlesi 2016b], arguing that "populists" congressional candidates (i.e. progressive democrats and MAGA republicans alike) fared better *ceteris paribus* in districts most exposed to Chinese import competition. The most provocative result of this set of working papers is a widely commented counterfactual scenario in which, had import exposure in the early decades of the 21th c. be one half lower, Democrats could have won the 2016 presidential and congressional race [Autor, Dorn, Hanson, and Majlesi 2016b]. If there exists many equivalent estimates in diverse national contexts (a wide range of results is reported in the literature review of [Rodrik 2021]), their significance is far from transparent. Most of the time, specifications are fed with vote shares changes, finding positive marginal impacts on the populist (i.e. far-left and far-right) vote, interpreting them through the outcome of political questionnaires about people's attitude towards trade, immigrants, democracy, etc., the (unsurprising) result of that exercise being most of the time a restatement of the old Lipsetian argument [Lipset 1959]; i.e., the import shock would bolster the illiberal, authoritarian ethos of working-class families most hurt by trade liberalisation, which would explain populist vote hikes.

Independently of their inner foibles (emphasised most notably by [Bourdieu 1979; Bourdieu 1980]), these approaches exemplify the pervasive influence of the early 1990s' academic consensus [Wood 2018] on research outside of economics and on nonacademic representations alike, with the widespread assumption that international economic integration had but marginally contributed to the rise of inequalities in the North, that political reactions emphasising the role of globalisation where therefore misled, and should be analyzed through a cultural, identitarian or psychological lens.

With the setting of [Autor, Dorn, and Hanson 2013], on the contrary, we closely follow each step of the import competition shock − its employment impact in the manufacturing sector, the multiplicative effect on services, the lingering unemployment reaction, the variability across groups and regions of the consequences over pre and post-redistribution income − and we are able to gauge which variable mediates the final political outcome [Dippel et al. 2017].

In this section, we'll therefore investigate first evidence about the usual political impact of a local economic shock in French context, comparing them then with the impact of trade competition-driven employment shock.

5.2 The traditional political response to an economic shock in French context

In French context, traditional political economy statics assign poorer and more equal communities to the Left, a result that can be extracted from the most recent and the most ancient data alike (see table 16). Properly construing that result implies a clear distinction between the post-1945 societies − where the income structure crystallizes into conflicts over redistribution [Piketty 2019] in a context of declining class identification and higher social mobility [Piketty 1995] − and the old socio-fiscal architecture: nonexistent redistribution, taxation schemes relying on indirect contributions [Bouvier 1973; Bouvier 1978; Piketty 2001] highly unpopular among republicans and socialists [Delalande and Spire 2010]. [Corbin 1975] reports that among the very poor (but also very equal) communities of Limousin, socialist activities refrained themselves from exposing their redistributive platform, knowing it would be highly unpopular among poor farmers, and rather played on pre-industrial political feelings, especially grievances against the Church and the local nobility. The political economy statics identified there relied much more on this type of representation than on the class divide, and was therefore easily transferable to non-republican or non-socialist alternatives (liberalism, jacobinism, bonapartism) [Marx 1852] for which similar correlations can be isolated.

Table 16: Raw correlation between some income variables and the left-wing vote shares in the 1849 parliamentary election

	Dep. var.: Vote share of the left-wing candidates					
	Dep. var.: Vote share of the left-wing candidates					
	(1)	(2)	(3)	(4)	(5)	(6)
Impact of a 1SD change						
Av. income within the département	-1.99	-2.61^{*}			-6.25^{**}	-7.01^{**}
	(1.46)	(1.44)			(3.15)	(3.25)
Ratio T10/B50 within the département			-0.77	-1.39	2.88	2.93
			(1.47)	(1.46)	(3.16)	(3.24)
Controls		X		X		X

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: The unit of observation is the *département* (1848 geography). General data is drawn from the *Statistique de la France* report published by the statistical division of the *Ministère des travaux publics*. Electoral data is from [Salmon 2001]. The dependent variable is the vote share of candidates which are self-reported as democrats or socialists. The explanatory variables are all derived from tax data about the *contribution personnelle et mobilière*, the receipts of which are provided as a piece-wise function for each *département* (here, for the fiscal year 1835). Means and ratios are then obtained through interpolation with *gpinter* [Blanchet, Fournier, and Piketty 2017]. Controls, when included, comprise only an index of religiosity drawn from the Boulard map [Boulard 1982; Le Bras 1931]. There are 84 observations, weighted by the number of tax units.

Arguably, there are signs, over ancient income data, of the emergence of a political consciousness of distributional issues, especially around landmarks in the history of working-class movement (we purposely focused on the 1849 and 1936 parliamentary elections in tables 16 and 17 resp.). Inequality drives the left-wing vote, the marginal effects being concentrated in more industrial and more urban areas[57]. In the 1936 election, the lead of the Left in marginally more unequal regions is entirely driven by the communist vote, which is stronger in industrial bastions of the Parisian metropolis and of the far North, two of the most unequal regions of the country at that time.

Table 17: Raw correlation between some income variables of 1929 and the left-wing vote in parliamentary elections

	Dep. var.: Vote share of the left-wing coalition in parliamentary election								
Explan. var.	All incomes		Wage income		Wealth				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Impact of a 1SD change									
Panel A. Parliamentary elections of 1924 − Cartel des gauches coalition									
−Average within the *département*	−0.52	−2.22	−0.44	−2.19	−0.23	−1.51			−0.15
	(1.07)	(1.89)	(1.01)	(1.79)	(0.88)	(1.32)			(2.33)
−Ratio T10/B50 within *dép.*							−1.12	−2.42	−2.25
							(1.11)	(1.79)	(3.19)
Panel B. Parliamentary elections of 1936 − Front populaire coalition									
−Average within the *département*	1.52*	−2.29	1.38*	−2.49*	1.08	−1.56			−5.78***
	(0.84)	(1.49)	(0.79)	(1.41)	(0.68)	(1.02)			(1.76)
−Ratio T10/B50 within *dép.*							1.89**	0.39	7.04***
							(0.86)	(1.43)	(2.43)
Controls		X		X		X		X	X

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: The unit of observation is the *département* (1929 geography). General data is drawn from the *Annuaire statistique de la France* published by the statistical division of the *Présidence du Conseil*, issues [Présidence du Conseil 1930] to [Présidence du Conseil 1938]. Electoral data is from [Lachappelle 1924;Lachappelle 1936]. The dependent variable is the vote share for the entirety of reported left-wing coalition. The explanatory variables are: for columns (1) and (2), the total tax levy of the income tax (IGR) within the *département* divided by its total population (over our datasets, we are not provided with the number of taxpaying persons); for columns (3) and (4), the same statistics, but for the special *cédulaire* tax on wages; for columns (5) and (6), the average inheritance reported within each *département*, and in columns (7) and (8), the ratio of the average top 10% inheritances over the average bottom 50% ones; this last statistics was interpolated with *gpinter* [Blanchet, Fournier, and Piketty 2017], from the piece-wise distributions provided for each *département*. Controls, when included, comprise the share of rural population, the share of blue-collar within local population, and an index of religiosity drawn from the Boulard map [Boulard 1982; Le Bras 1931]. There are 85 observations, weighted by the number of votes cast in the corresponding election.

Table 18: Simple model for the political impact of the Great Depression

	Dep. var.: Evo. of vote shares (1924-1936) in pp	
	ΔSocialist vote	ΔCommunist vote
	(1)	(2)
Impact of a 1SD change		
Δ*Personal income*$_{1929,1935}$	−2.05	−1.19*
	(1.61)	(0.67)
Δ*Personal wage income*$_{1929,1935}$	−0.36	−0.06
	(1.57)	(0.66)
Δ*Personal wealth*$_{1929,1935}$	−0.62	−0.57
	(1.17)	(0.48)
Δ*Ratio T10/B50 of wealth*$_{1929,1935}$	−0.79	0.47
	(0.99)	(0.41)
Δ*Ratio T10/B10 of wealth*$_{1929,1935}$	−2.06*	1.12**
	(1.23)	(0.51)

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

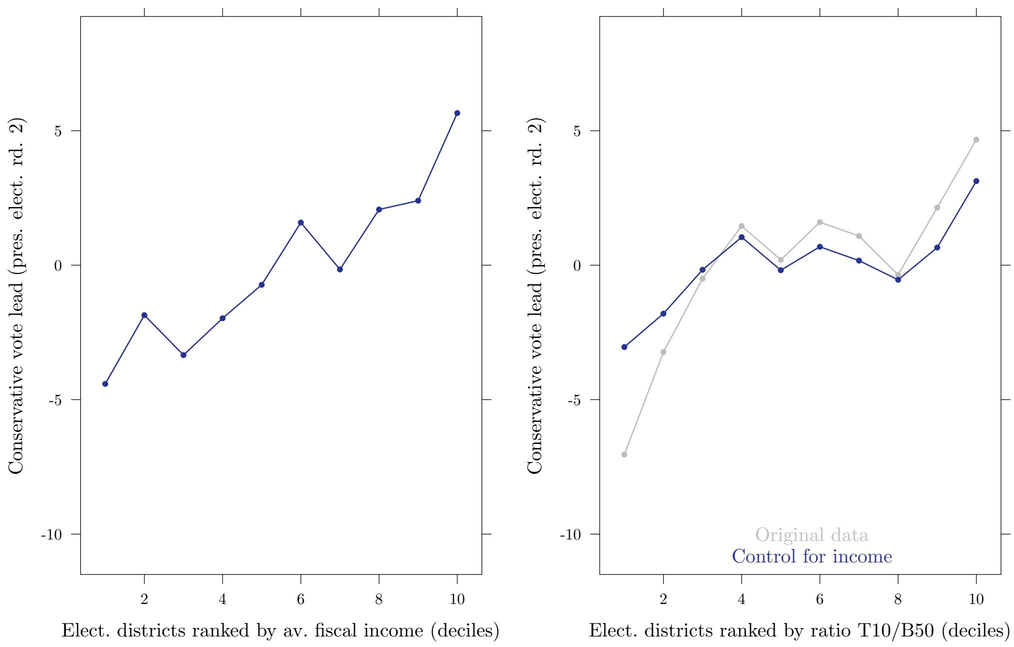
Note: The unit of observation is the *département* (1929 geography). General data is drawn from the *Annuaire statistique de la France* published by the statistical division of the *Présidence du Conseil*, issues [Présidence du Conseil 1930] to [Présidence du Conseil 1938]. Electoral data is from [Lachappelle 1924;Lachappelle 1936]. The dependent variable is the evolution of vote shares for the SFIO and SFIC parties between the 1924 and 1936 parliamentary elections. The explanatory variables are the evolution of the explanatory described in table 17 over 1929-1935; controls are start-of-the-period values of the ones mentioned in table 17. There are 85 observations, weighted by the number of vote cast in the 1936 election.

In the late 20th c., the decline of rural communities and early dechristianization pushed conflicts over redistribution on the forefront of political debate. The olf Left *statics* (poorer and more equal community lean left) and the old Left *dynamics* (a community turns to the Left when faced with declining incomes or rising inequalities) crystallised (see fig. 29, 30 and 31), with some looming fissures however: 1. Very early, the FN vote borrowed to the left its statics and dynamics alike: it boomed in times of economic hardship and could claim predominance among

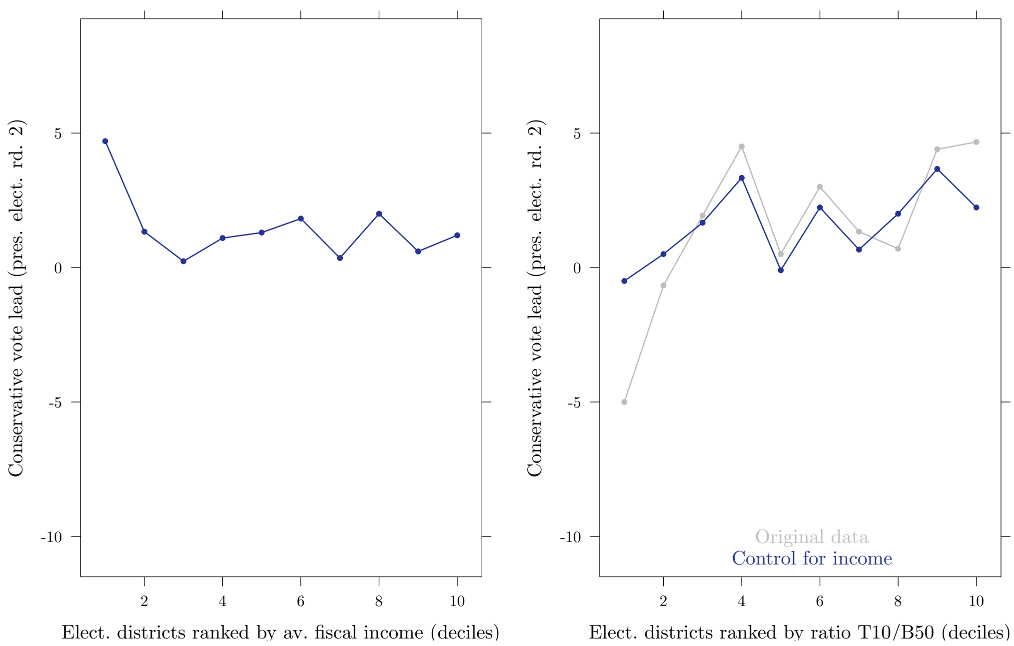
[57] Indeed, if during the fordist era, the predominance of manufacturing employment in a region was associated with lower wage and income inequalities, prior to the war, it was the very contrary.

Figure 29: Income, income inequality, and the left-right cleavage at the regional level (I) − Descriptive statistics

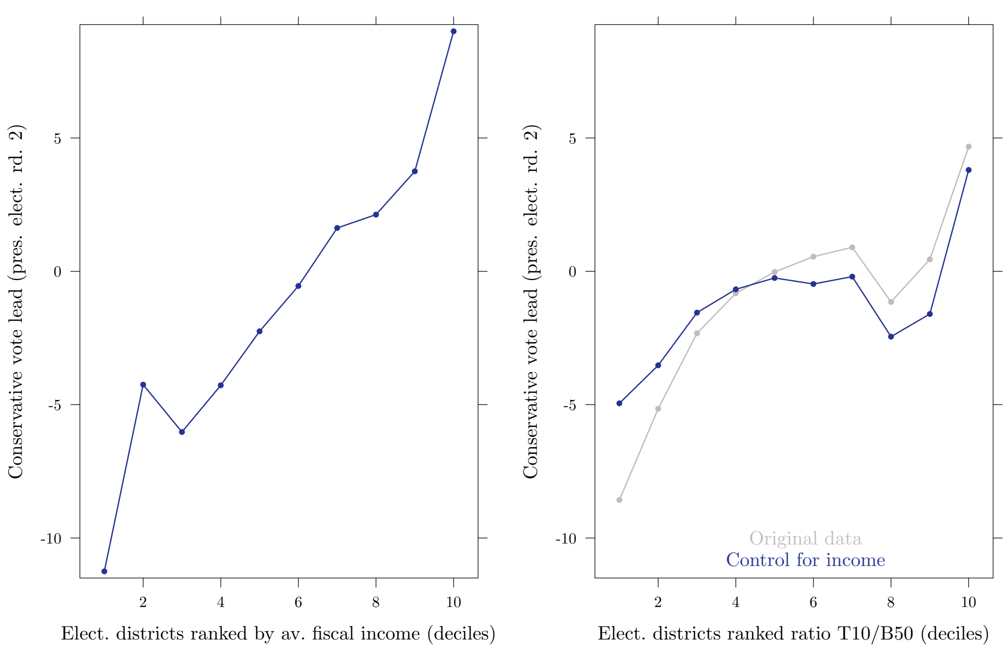
(a) All presidential elections 1965-2012 (1969 and 2002 excl.)



(b) 1965-1981



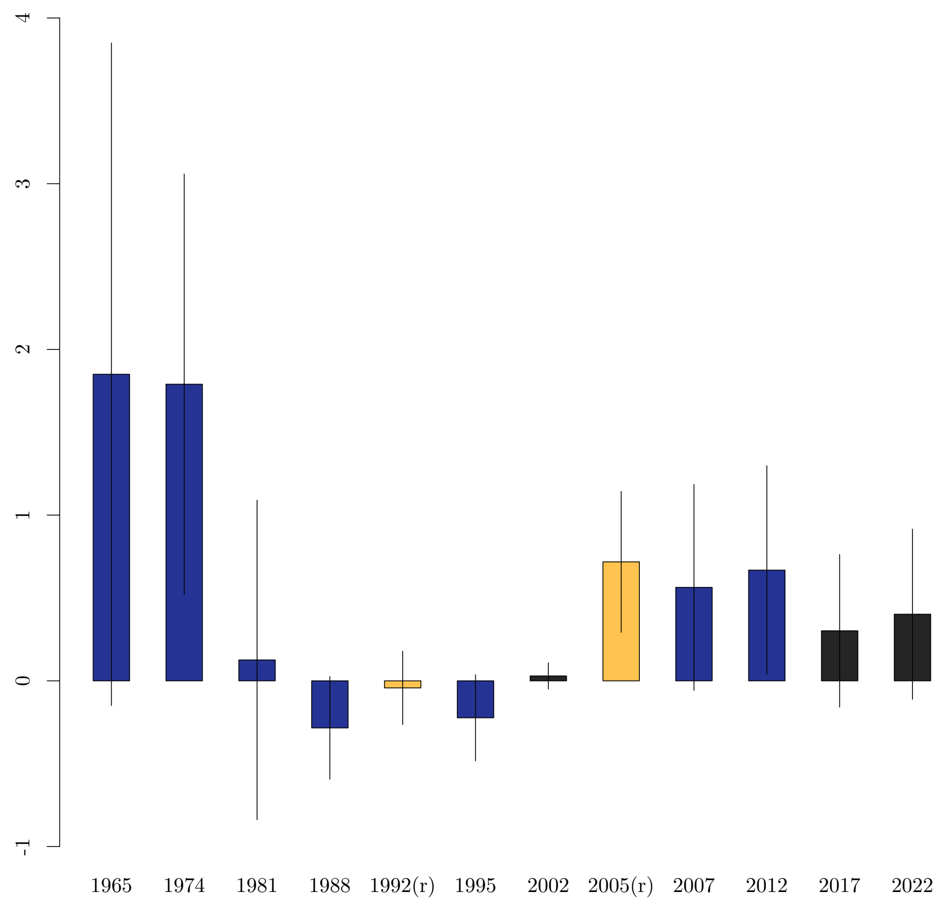
(c) 1988-2012



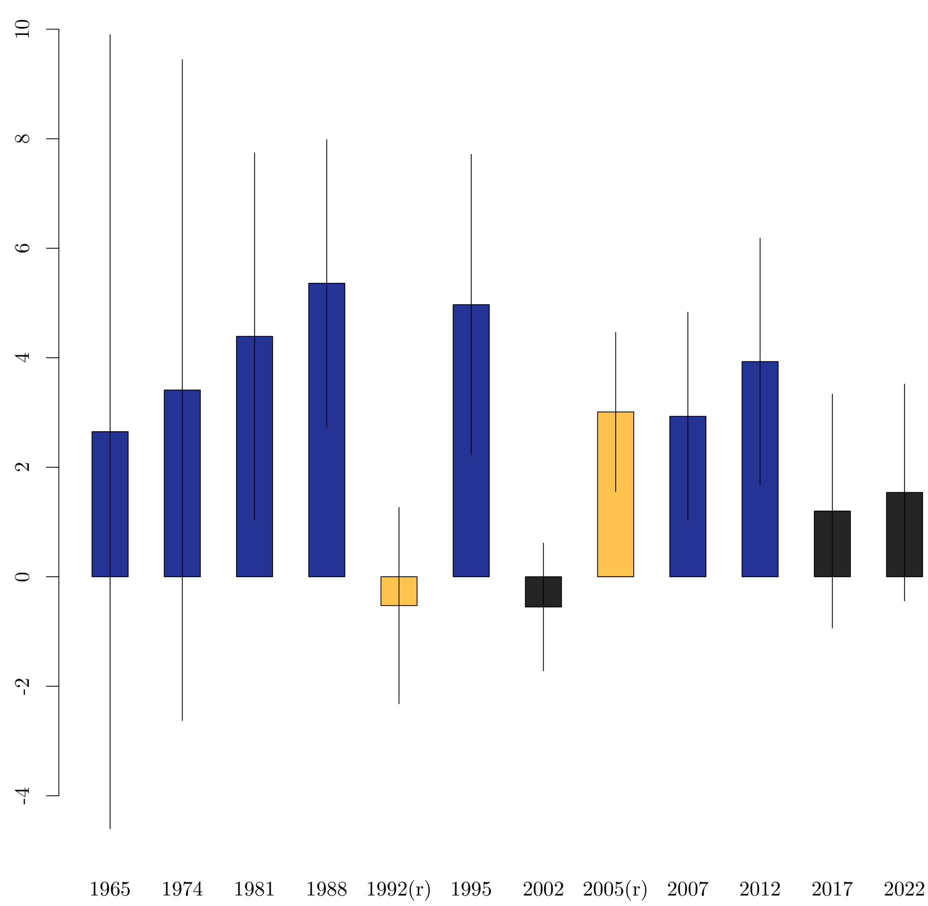
Note: Electoral results are taken from the datasets of the CDSP-Sciences Po. Income variables are from the IRCOM database, or interpolated from the INSEE's Census (see our annex A). On the x-axis, we sort electoral districts (*circonscriptions législatives*) in deciles according to: 1. The average fiscal income of their inhabiting tax units; 2. The population-weighted average of the ratio T10/B50 within the *communes* of the electoral district. We use the political geography at the time of the election; we follow the concordance table *communes-circonscriptions électorales* of the CDSP. For the ratio T10/B50, we provide the original vote lead, and the vote lead once the impact of income has been factored out.

Figure 30: Income, income inequality, and the left-right cleavage at the regional level (II) − Marginal effects

(a) Marginal effect of the av. fiscal income of the electoral district over the right-wing vote lead in the 2nd rounds of presid. elec.



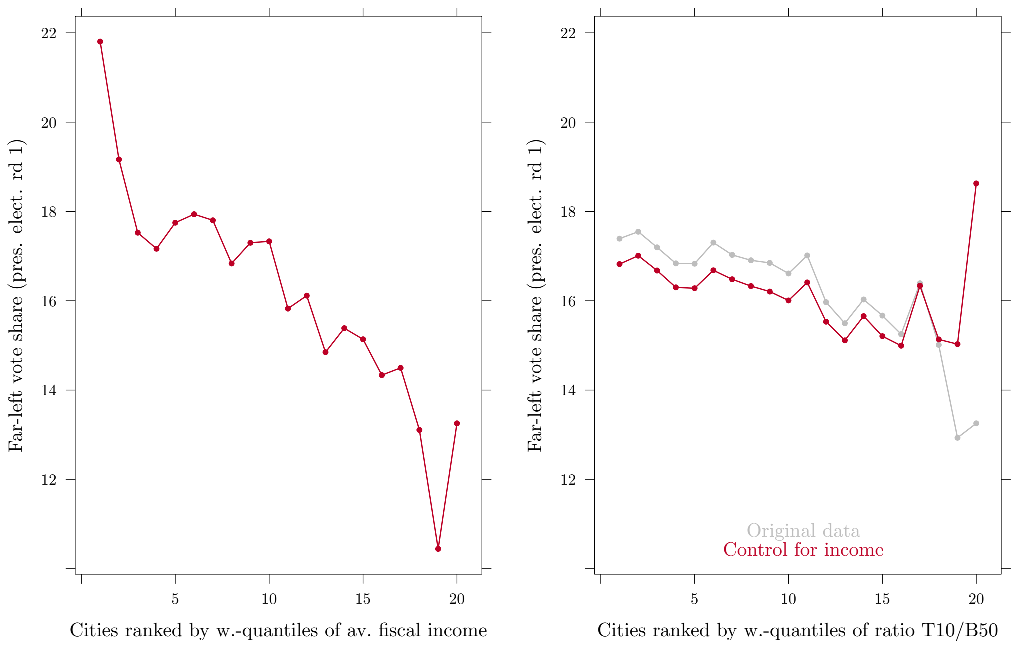
(b) Marginal effect of the ratio T10/B50 of fiscal income within the electoral district



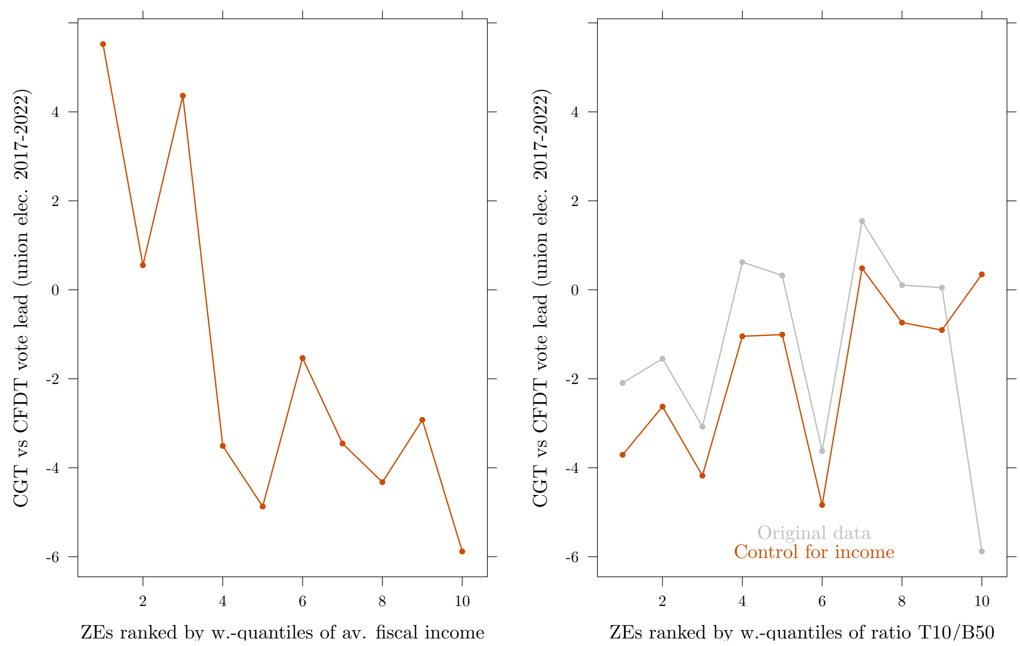
Note: The unit of interest is the electoral district (*circonscriptions législatives*). We use the political geography at the time of the election; we follow the concordance table *communes-circonscriptions électorales* of the CDSP-Sciences Po. Electoral results are taken from the datasets of the CDSP. Income variables are from the IRCOM database, or, for elections prior to 1995, interpolated from the INSEE's Census (see our annex A). We regress the vote lead (in pp) of the conservative candidate in the second round of the presidential election (or the vote lead of the *Yes* to the two referendums about Europeans treatises, or the vote lead of the non-FN/RN candidate), on the average fiscal income of the district (expressed in thousands euros of 2021), the population-weighted average of the ratio T10/B50 of the *communes* of the district, plus a set of socio-demographic controls (the unemployment rate, the share of rural population, the share of blue-collar population, and an index of religiosity, i.e. the share of the population of the district which lives in a parish with high churchgoing rates as defined by [Boulard 1982; Le Bras 1931]).

Figure 31: Some indexes of polarisation (I)

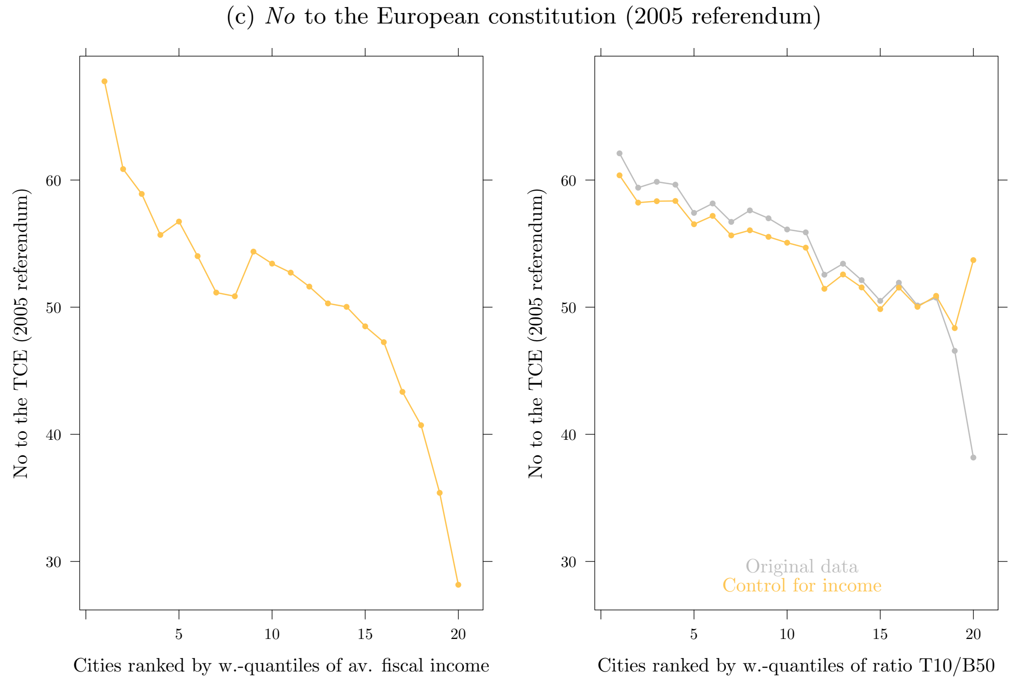
(a) The far-left vote (2002 to 2017 pres. elect.)



(b) Radical vs Moderate unionism (vote lead of the CGT over the CFDT, union elect. 2017-2022)



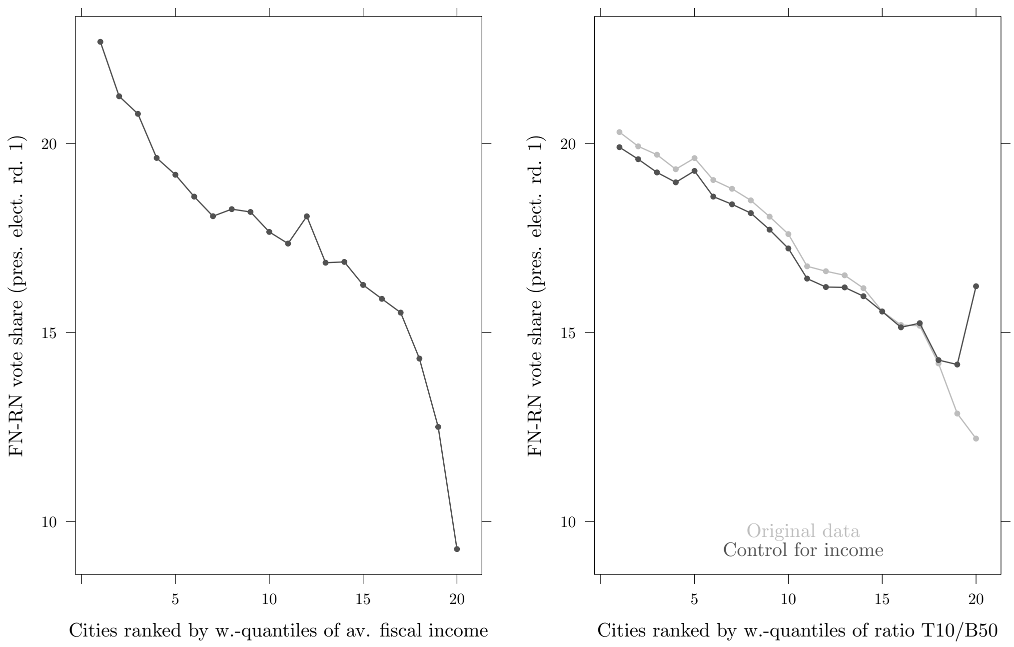
(c) *No* to the European constitution (2005 referendum)



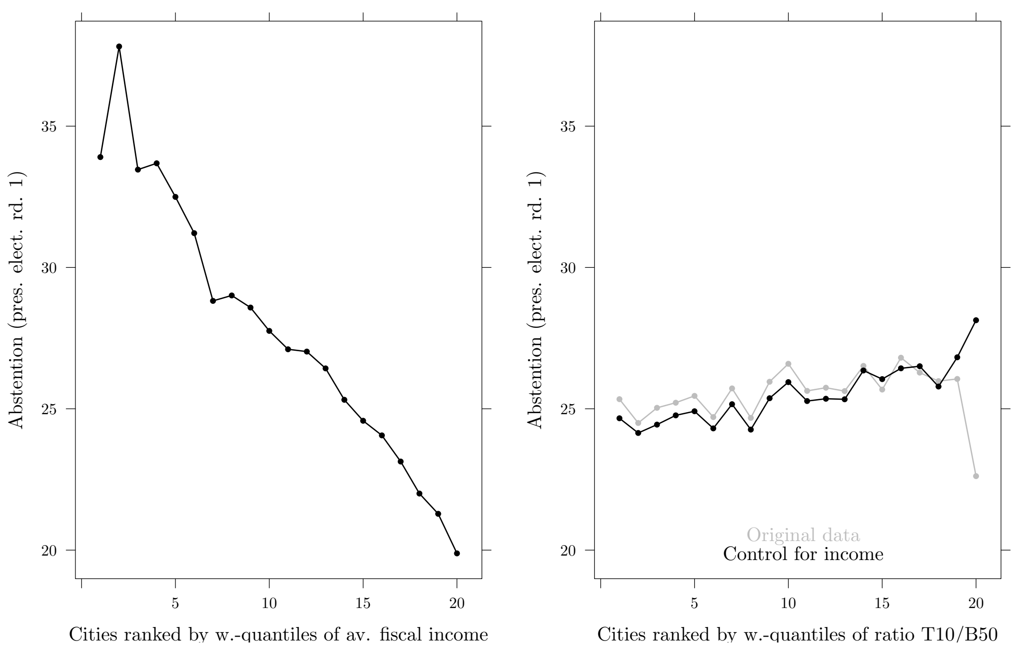
Note: Electoral results are taken from the datasets of the CDSP-Sciences Po, except for the union representation elections, for which data crom from the *Ministère du Travail*. Income variables are from the IRCOM database, or interpolated from the INSEE's Census (see our annex A). On the x-axis, we sort cities (*communes*) or *zones d'emploi* in deciles according to: 1. The average fiscal income of their inhabiting tax units; 2. The population-weighted average of the ratio T10/B50 within the *communes* of the electoral district. We use the political geography at the time of the election; we follow the concordance table *communes-circonscriptions électorales* of the CDSP. For the ratio T10/B50, we provide the original vote lead, and the vote lead once the impact of income has been factored out. For all plots, quantile 20 gives the value for Paris (either the *commune* or the corresponding ZE1101).

Figure 32: Some indexes of polarisation (II)

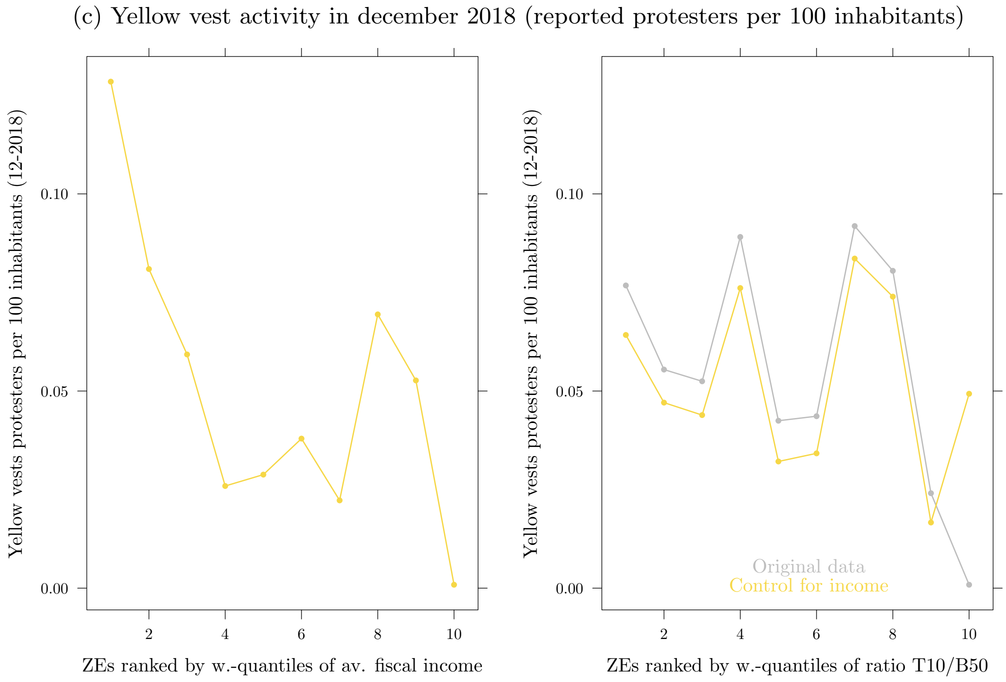
(a) Average far-right (FN-RN) vote in 1st rd. of the presidential elect. (2002 to 2022)



(b) Abstention (2022 presidential elect. rd. 1)



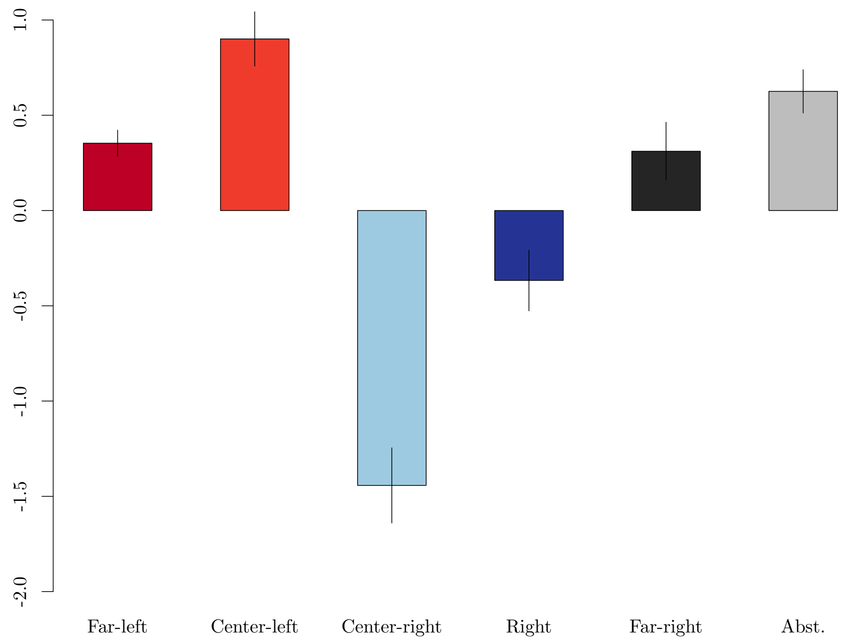
(c) Yellow vest activity in december 2018 (reported protesters per 100 inhabitants)



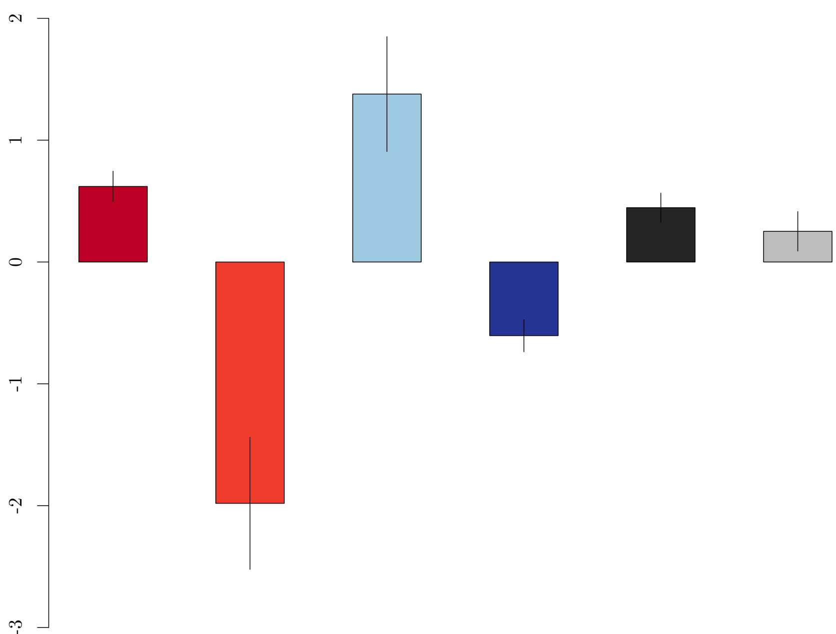
Note: Electoral results are taken from the datasets of the CDSP-Sciences Po. Yellow vests activity data courtesy of Daniel Cohen (raw dataset used in [Algan, Beasley, et al. 2019]). Income variables are from the IRCOM database, or interpolated from the INSEE's Census (see our annex A). On the x-axis, we sort electoral districts (*circonscriptions législatives*), cities (*communes*) or *zones d'emploi* in deciles according to: 1. The average fiscal income of their inhabiting tax units; 2. The population-weighted average of the ratio T10/B50 within the *communes* of the electoral district. We use the political geography at the time of the election; we follow the concordance table *communes-circonscriptions électorales* of the CDSP. For the ratio T10/B50, we provide the original vote lead, and the vote lead once the impact of income has been factored out. For all plots, quantile 20 gives the value for Paris (either the *commune* or the corresponding ZE1101).

Figure 33: FD-panel estimates (2002-2017) of the impact of economic shocks on vote shares (in presidential elect. round1)

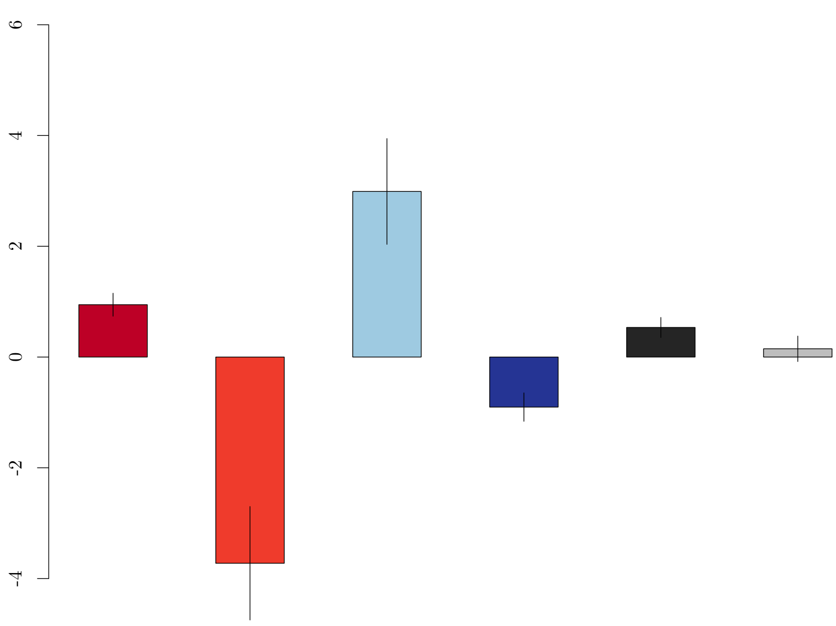
(a) Marginal impact of an unemployment shock



(b) Marginal impact of a negative shock on income



(c) Marginal impact of a rise of the ratio T10/B50



Note: Electoral data are drawn from the CDSP datasets, income data from the IRCOM base (restr. 2), controls from the INSEE's census. The unit of interest is the *commune*. Our main specification is a first-difference panel regressing vote shares in first round of presidential election, on the average fiscal income (in euros of 2022), the ratio T10/B50 of the *commune*, the unemployment rate, log employment, plus a set of controls (density, share of retired persons, share of blue-collar in pop. and share of people commuting on everyday basis for their job). Four elections are estimated (2002, 2007, 2012 and 2017). S.E. are clustered at the *département* level; we report the 95% conf. inter.

poorer communities, though statistically these marginal effects did not exhibit the robustness of their left-wing equivalents (see table 60); 2. The full unravelling of the economic significance of the centre-right (*orleanist*) vote was slow in the making, liberal themes suffusing much of centrist political life and the leading centre-right party, the UDF, developing an almost anti-system rhetoric; yet the rise of the LREM coalition saw a reversion to more classical right-wing statics and dynamics; 3. With an electorate still highly class-polarised late in the Mitterrand era, the socialist party failed to benefit from the take-up of the late 1990s [Lefebvre and Sawicki 2006] and was the clear marginal loser of the economic shocks of the early 21th c. (see figure 33); in the second rounds of the presidential elections, the old Left *statics* and *dynamics* was still operating (see table 59); in the longer term however, our findings are consistent with individual data which show that the left dominance among lower-income workers eroded from the mid-1990s onward and had virtually vanished by 2017 [Gethin, Martinez-Toledano, and Piketty 2021].

All in all, over the first two decades of the 21st c., an economic shock (a rise of local unemployment, a decline of personal income, a rise in within-region inequality) is associated with a significant increase in support for left-leaning options, or for options which are seen as alternatives to the incumbent parties (the centrist UDF, or the FN-RN and its allies), the effect being more robust across specifications for the latter channel.

These are features to be kept in mind when we'll compare the impact of a general economic shock to the impact of trade competition-driven one.

5.3 The specific impact of an import-competition-driven shock (1995-2022)

Empirical specification

To assess the impact of an employment shock specifically driven by the rise of import competition from China, we update specification (3) using as explanatory is the rise in import exposure per worker ΔIPW for the first decade of the millennium (in our setting, 1999-2008), and the dependent, the evolution of vote shares for specific parties between a pre-exposure event (here, the 1995 presidential election), and the elections which happened over and after the exposure period (here, presidential elections from 2002 to 2017). We built five non-exclusive political aggregates (detailed in the corresponding annex) to gauge the political fortunes of each side.

Main results

The main coefficients $\hat{\beta_1}$ for each political aggregate and each election is plotted in figure 34 for the first rounds of presidential elections, and in figure 35 for the second rounds. Figures are built in order to ensure comparability with [Autor, Dorn, Hanson, and Majlesi 2016b].

The results are clear-cut; trade-induced economic shocks do not trigger a polarisation reaction like in [Autor, Dorn, Hanson, and Majlesi 2016a], but rather a general shift to right-wing options, moderate and radical alike. Prior to 2017, when the left-right cleavage was still the dominant polarising force, a trade shock had a clear negative impact on the PS vote, in the first and second rounds alike, and tended to favour conservative options, from moderate liberals of the UDF to the far-right FN vote. After 2017, parties which emerged from the downfall of the PS inherited its negative marginal impacts, while the FN-RN vote benefited from the decline of the grand Gaullist party.

Our results are at variance with the existing literature. Up till now, the expected political impact of an import shock was either: 1. Increased support for the far-right [Colantone and Stanig 2017; Malgouyres 2017a; Dippel et al. 2017]; 2. Alternatively, increased support for populist parties of both sides [Autor, Dorn, Hanson, and Majlesi 2016b; Autor, Dorn, Hanson, and Majlesi 2016a; Barone and Kreuter 2021]; 3. More generally, a shift towards a more nationalistic or identitarian political atmosphere [Ballard-Rosa, A. Jensen, and Scheve 2022; Cerrato, Ferrara, and Ruggieri 2018; Steiner, Harms, et al. 2020]. To our knowledge, such significant positive impacts for all right-wing parties (moderate and pro-European ones included) coupled with massive negative impacts for all left-wing options (globalist and anti-globalist alike) are a home exception, with no equivalent in the extensive literature review of the subject by [Rodrik 2021].

The magnitude of the effect for conservative and far-right candidates is not trivial. In a simple counterfactual scenario with $\Delta IPW_{1999,2008}$ 60% below its actual value, Mrs Le Pen fails to qualify for the second round of the presidential election in 2017. Conversely, with $\Delta IPW_{1999,2008}$ 75% above its actual value, the PS candidate loses the second round of 2012. The only existing French replication of [Autor, Dorn, Hanson, and Majlesi 2016a] to our knowledge, the one of [Malgouyres 2017b], tested only the FN vote, finding marginal impacts slightly inferior to ours.

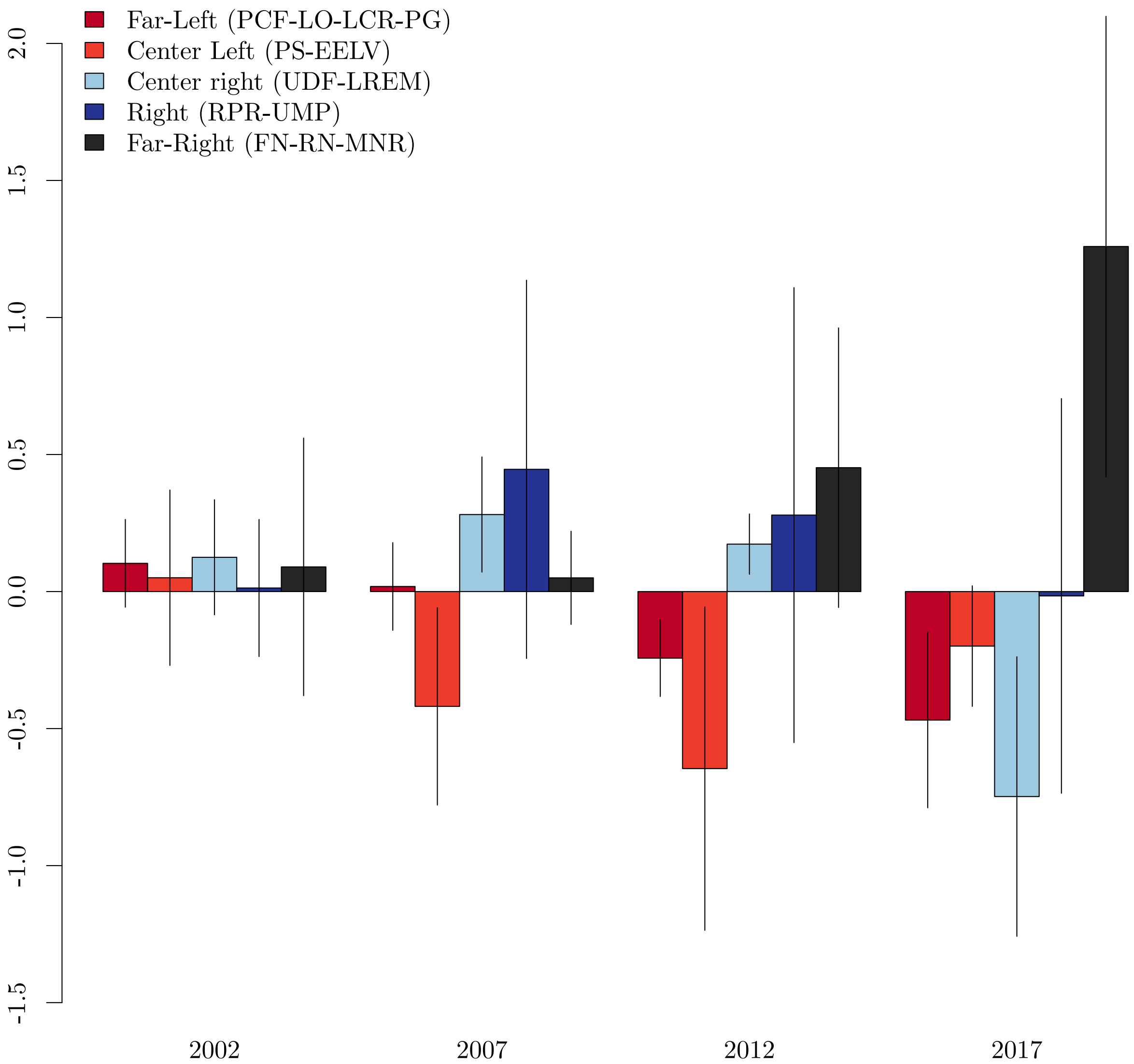
In the U.S. context, import exposure has been linked to increased voter turnout, to sharp rises in campaign donations for radical candidates of both sides, but also to changes in news network viewership (at the expense of liberal cable news channels like CNN and MSNBC) [Autor, Dorn, Hanson, and Majlesi 2016b]. In French context, we fail to find a similar impact on abstention (see figure 36a), but we find some evidence of a negative marginal impact on political options which are commonly associated with centrist or centre-left politics, for instance a decline of the support to centrist unions (see figure 36b) or to European federalism (see 36c).

Robustness checks and discussion of the findings

The interpretative features highlighted by [Autor, Dorn, Hanson, and Majlesi 2016a] are well fitted for the polarisation reaction they identify in the U.S. context, but are difficult to transfer to a European setting. Here are some of their major theses:

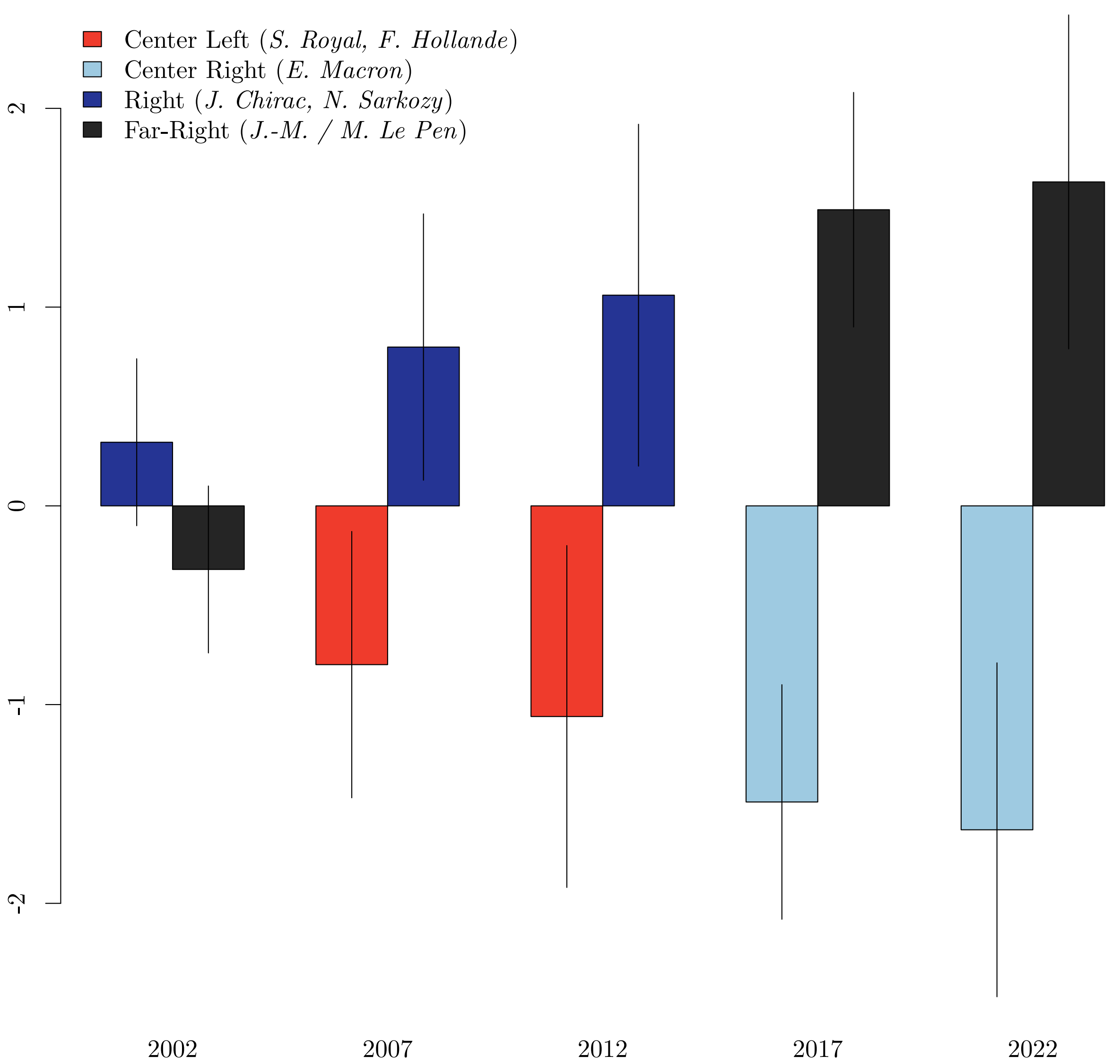
- *Protectionist reactions* – In U.S. [Feigenbaum and Hall 2015] and European contexts alike [Mayda and Rodrik 2005; Davenport, Dorn, and Levell 2021], a rise in import competition exposure is associated with an increased support to protectionist policies, an increase which is pervasive across groups and political

Figure 34: Political impact of a rise in import competition exposure within the ZE (I) − First rounds of the presidential elections



Note: The unit of interest is the *Zone d'emploi* (ZE, INSEE def. of 2010). Electoral data are drawn from the CDSP datasets, trade data from the Comtrade base, other variables from the INSEE's Census. We estimate model (3) for the sole decade 1999-2008, with the full vector of controls mentioned in table 2. The main explanatory variable is still the rise in exposure to imports from China per ZE over the decade (expressed in thousands USD per worker) but the dependent variable is now the evolution of the vote share of five political aggregates (Far-left, Centre-left, Centre-right, Right, Far-right) between the first round of the 1995 presidential election, and the first round of the presidential election mentioned on the x-axis. Political aggregates are defined in the corresponding annex. Observations are weighted by the number of votes cast (*blancs-nuls* excluded). Standard errors are clustered at the level of the INSEE's superzones; we report 95% conf. intervals.

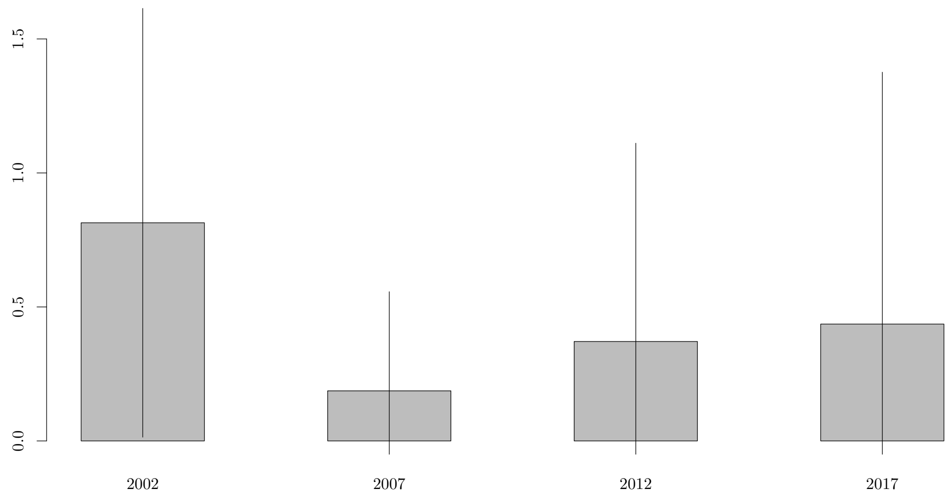
Figure 35: Political impact of a rise in import competition exposure within the ZE (II) − Second rounds of the presidential elections



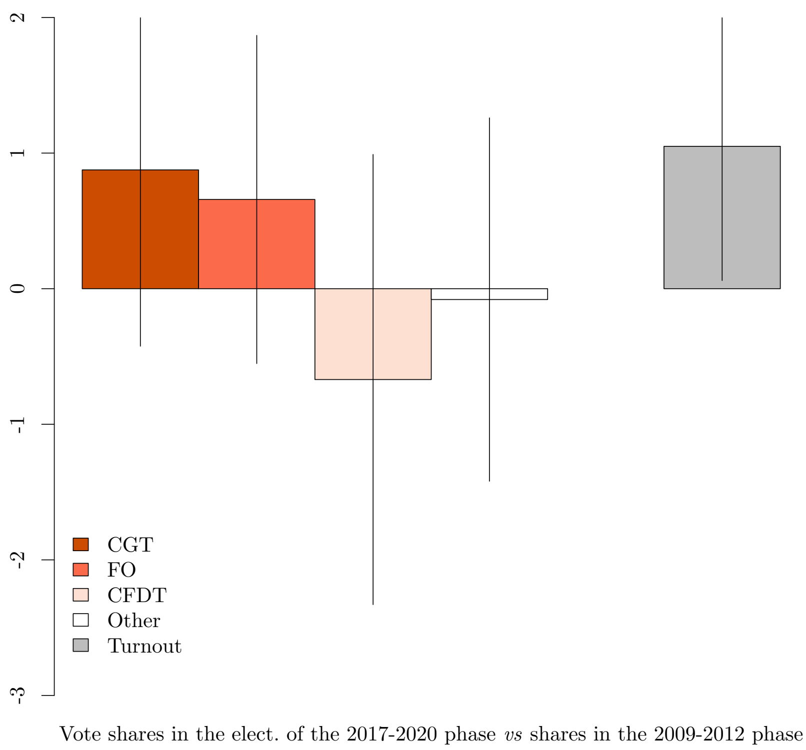
Note: The unit of interest is the *Zone d'emploi* (ZE, INSEE def. of 2010). Electoral data are drawn from the CDSP datasets, trade data from the Comtrade base, other variables from the INSEE's Census. We estimate model (3) for the sole decade 1999-2008, with the full vector of controls mentioned in table 2. The main explanatory variable is still the rise in exposure to imports from China per ZE over the decade (expressed in thousands USD per worker) but the dependent variable is now the evolution of the vote shares, defined as such: 1. For the 2022 and 2017 presidential elections, the vote share of M. Le Pen in the 2nd round, minus the vote share of J.-M. Le Pen in the 2nd of the 2002 election (and vice versa for E. Macron 2017/2022-rd2 minus J. Chirac 2002-rd2); 2. For the 2012 and 2017 election, the vote share of N. Sarkozy in the corresponding 2nd round, minus the vote share of J. Chirac in the 2nd round of 1995 (and vice versa F. Hollande 2012-rd2 and S. Royal 2007-rd2 minus L. Jospin 1995-rd2); 3. For the 2002 election, the vote share of J.-M. Le Pen in the 2nd round, minus the vote share of all far-right forces (FN+MNR) in the 1st round of the 1995 presidential election (and vice versa, the vote share of J. Chirac in the 2nd round of 2002, minus the vote share of all non-far-right forces in the 1st round of the 1995 election). Observations are weighted by the number of votes cast (*blancs-nuls* excluded). The absolute start-of-the-period vote shares which are differentiated in the computation of changes in vote shares are included as a control variable. Standard errors are clustered at the level of the INSEE's superzones; we report 95% conf. intervals.

Figure 36: Political impact of a rise in import competition exposure within the ZE (III)

(a) Impact on abstention (first rounds of presidential elections)



(b) Impact on professional elections



(c) Some other tests

Note: The unit of interest is the ZE. Electoral data are drawn from the CDSP datasets (and from the *Ministère du Travail* for professional elections), trade data from the Comtrade base, other variables from the INSEE's Census. We estimate model (3) for the sole decade 1999-2008, with the full vector of controls mentioned in table 2. The main explanatory variable is still the rise in exposure to imports from China per ZE over the decade (expressed in thousands USD per worker) but the dependent variable is now : A. The evolution of abstention (as a percentage of registered voters) between the 1st round of the 1995 presidential election, and the 1st of the presidential election mentioned on the x-axis; B. The evolution of the vote shares of the three main unions (CGT, FO, CFDT) and the vote shares of all other options between the set of professional elections organised over 2009-2012, and the set of elections organised over 2017-2022; we also test for the turnout of the professional elections; note that because some candidates might be endorsed by several unions, the sum of vote shares can exceed the unity; C. The evolution of the vote share in support of new EEC-EU treatises between the 1992 and 2005 referendums; the Yellow vests activity (expressed in numbers of reported protesters per 10k inhabitants). Extra controls include respectively: A. Abstention in the 1st round of the 1995 presidential election; B. The vote share of the corresponding union over the 2009-2012 period, plus the whole left-wing vote share in the 1st round of the 2007 presidential election; C.1. The far-right vote in the first round of the 1995 presidential election; C.2. The centre-right vote share in the first round of the 2017 election. Observations are weighted by: A. The number of registered voters; B. Total Census population; C. The number of expressed votes. Standard errors are clustered at the level of the INSEE's superzones; we report 95% conf. intervals for panel (a) and (b), 10% risk ones for panel (c).

preferences; yet if we trust our estimates, this does not transcribe into a systematic support for protectionist parties or platforms. Our findings, as many prior results[58] do not fit into the well known narrative about the decline of the left-right cleavage and the correlative rise of a new *globalist versus anti-globalist* political divide encapsulating all issues (trade, State regulation, entrepreneurship, but also attitudes towards minorities or international cooperation). In our setting, trade shocks are much more reactivating the left-right divide that they are blurring it;

- *Decline of class consciousness and realignment over identity* − Trade shocks not only decompose the local industrial structure; they also tend to shift the focus away from class-oriented conflicts, favoring identification to groups defined by nationality or identity, and not by income or worker status. There is a socialist version of this line of argument [Bourdieu 1979], which puts the emphasis on the political organising channel, i.e., on the decline of unions and political traditions caused by massive layoffs in manufacturing [Beaud and Pialoux 2003]. There is also a liberal version, focused on group identification [G. Grossman and Helpman 2018; Gennaioli and Tabellini 2019]. Widely used to construe the Brexit vote shares [Colantone and Stanig 2017; Colantone and Stanig 2018], this line of reasoning can however hardly explain why import shocks seem to bolster left political organising at the margin[59] or outside[60] the main left parties;

- *Competition over resources of declining redistribution schemes* − In our setting as in [Autor, Dorn, and Hanson 2013], trade shocks make poor families more dependent on social transfer. From this, Autor, Dorn and Hanson hypothesise a reaction in the spirit of [Alesina, Baqir, and Easterly 1999]; each ethnic group will turn to the party which he judges most likely to divert local transfers away from the concurrent groups, hence an increase of the white support to Republicans and of the minority support to Democrats, a constant of the last half-century [Gethin, Martinez-Toledano, and Piketty 2021], particularly in the South [Kuziemko and Washington 2018]. One of the main findings meant to substantiate this hypothesis was that, when [Autor, Dorn, Hanson, and Majlesi 2016b] plot the equivalent of our figure 34 over a restriction to predominantly non-Hispanic-white districts, they find marginal gains for radical republicans only, while when they focus on predominantly non-white districts, progressive democrats seem to be the main winners of a trade-induced shock. However, as emphasised by [Malgouyres 2017b], such a conclusion is extremely difficult to replicate in French context; in order to obtain a positive, 5% significant marginal effect of trade exposure on the far-left vote in the model plotted in figure 34 (2017 election first round), we should drop almost 95% of our observations, focusing on those ZEs with the highest shares of non-native population (above 16%). Even on that very restricted set, we indeed find a positive marginal impact of +1.12 for the far-left vote, but we also find a 1% significant marginal impact of +0.44 for Mrs Le Pen (while [Autor, Dorn, Hanson, and Majlesi 2016b] get a negative marginal impact on the republican vote in predominantly non-white districts). Arguably, there's some evidence in UK [Colantone and Stanig 2018] and US [Cerrato, Ferrara, and Ruggieri 2018] contexts alike that further imports exposure is associated with a polarisation of local communities around the ethnicity cleavage. At first glance, this might be consistent with the findings of our section 4. However, the many differences between our results and the ones of [Card, Mas, and Rothstein 2008], most notably those emphasised in table 13 − i.e., the idea that tipping households are trying to avoid certain districts rather than certain ethnic groups − cautions against hurried interpretations[61];

- *Decline of the local provision of public goods and public services* − There is consistent evidence in U.S. and European contexts that trade shocks lead to a decline of local real estate prices, hence a decline of local tax levies, and a decline of the quality of local public goods' provision [Feler and Senses 2017]. In French context, this line of reasoning has been widely used to construe the most recent political crises, especially the rise of the FN-RN or the *Yellow vests* movement [Algan, Beasley, et al. 2019]. This argument is however, as emphasised by [Davezies 2021; Dherbécourt and Deschard 2019], hard to substantiate over existing data. As shown in table 64, using a wide range of INSEE indexes, we fail to replicate the two steps of the argument, i.e.: 1. We do not detect an impact of import exposure on the local provision of public services; 2. Nor do we manage to identify a significant connection between the decline of local public services and the FN-RN vote.

Actually, among these diverse interpretations, it seems like the great economic channels have been relatively overlooked; the literature found significant marginal impacts of import exposure on manufacturing employment, unemployment, incomes, dependence to social transfers... but this literature seems to have little to say about their indirect impact on the political outcomes. A solution is ventured by [Dippel et al. 2017] in the form of a specific method which consists in a second-stage regressing vote shares on a mediating variable (for instance, the rise of unemployment, or the decline of incomes) itself instrumented by the home ΔIPW, using the control group's ΔIPW as a conditioning variable. With this strategy, they are able to fully explain the marginal positive impact of the China shock on Germany's FDP vote with economic mechanisms only. We tried to replicate their setting on the model plotted in figure 34; in our results, the evolution of personal fiscal income at the ZE's level, the rise of the total stock of unemployed people, and the increased dependence on social transfers, are the only significant mediating variables, with which we are able to explain respectively 15.3, 13.1 and 12.4% of the +1.3pp marginal impact in favour of Mrs Le Pen (2017, first round) plotted in figure 34. The connection with our section 3 is then straightforward:

- As emphasised above, the fiscal income impact of an import shock plotted in figure 21 is relatively flat across the income distribution down to somewhere around the third decile, where it brutally jumps down, creating a rift within popular classes, opposing a lower middle-class which remains employed, and those who lose their job and become increasingly dependent on social transfers. From this, the connexion is straightforward with what the specialists of the FN-RN vote have called the *triangular consciousness* of the FN voters [Collovald and Schwartz 2006], i.e. the pervasive feeling that they are caught between the winners of globalisation on the one hand, and a *lumpenproletariat* threatening to absorb them on the other. The fear of falling behind

[58] Most notably those evolving radical democrats in [Autor, Dorn, Hanson, and Majlesi 2016a].
[59] See figure 7 in [Autor, Dorn, Hanson, and Majlesi 2016a].
[60] See our figure 36.
[61] In some of our specifications, we found signs that tipping migrants might bolster the FN vote in their place of arrival, but this result was not robust enough across models and sources to be reported.

the sharp discontinuities identified there might nourish injunctions to embrace a more conservative lifestyle [Cartier et al. 2008; Lechien and Siblot 2019], typical the residential strategies we identified in our section 4;

- The fiscal impact of the shock is piece-wise, while the redistribution curves plotted by [Blanchet, Chancel, and Gethin 2019] are almost linear between deciles 2 and 9, meaning that somewhere (here, around above the 3rd quantile), there could be a group of people who lie within the hot-spot of the shock, but are not compensated enough by transfers; in an fig. 25, the sharpest negative response of disposable income is found precisely at centile 29. It is not difficult to picture how any variation of the anti-redistribution narrative might work on this group. The very heated debate about the inactivity and poverty traps created by the French redistributive architecture [Anne, L'horty, and Dollé 2002; Gomel and Méda 2014] might have overlooked the fact that voter's perceptions of the legitimacy of redistribution might be based, not much on its aggregate effect, but on its differential impact when a community is faced with a specific shock.

However, if this income-based interpretation seems relevant for the conservative and far-right vote, it hardly illuminates other findings of this section. Arguably, the identification strategy of [Dippel et al. 2017], as well as many other approaches we tested, work very well to explain the FN-RN vote shares, but perform very poorly for other parties[62].

Indeed, one of the most striking features of our findings are these large negative marginal impacts suffered by all left-wing options. As emphasised above, it is without equivalent within Western democracies[63]. It is like, from the late 1990s onward, the centre-left socialist party has been identified by voters as the party of globalisation, to which any other political option was preferred when a community was faced with a trade shock. This stigma set in very early and was inherited by parties which emerged from the decline of the PS. Its effect has never faltered since; changes in platforms or rhetoric have left it almost unaltered. When the Gaullist coalition embraced a more globalist and pro-trade discourse, it still benefited from that stigma; conversely, dissident socialist parties, even when they developed some form of protectionist rhetoric, inherited that stigma all the same.

An objectivist analysis would linger over the irrationality of import-competition exposed voters who paradoxically turn to pro-trade and anti-redistribution parties. Yet, as emphasised by [Dippel et al. 2017], what might be at stake there are not party agendas, but more broadly the *Weltanschauung* on which a social and political coalition relies. In the early 1960s, caught between Communists and Gaullists, unfit for a political sphere in which the legitimacy gained in the *Résistance* was the main political capital, many non-communist left parties awaited the fading of the national consciousness as the *Aufhebung* that would cast away a political structure in which they could not compete. Rants against general De Gaulle's lack of support to European integration, or against the *chauvinism* of a *revisionnist* PCF, became one of the trademarks of centre-left discourse[64]. While the parties which had acquired monopoly over the resistant *ethos* promoted the values of national unity, political voluntarism, in short, the values of the identity with the Self, the non-communist Left developed a *pluralistic* theory of politics[65], which transfigurated its exclusion from the dual monopolistic structure of dominant parties into performative predictions about the rise of a new social order defined by a *multiplication of normative producers*[66], an *indetermination of power structures*[67], in short, a rhapsodic world governed by exogeneous forces against which Gaullist technocrats and Communist activits would be powerless, a world in which, as opposed to wartime, there would be *no simple or straightforward answers*[68]. When social-democrats reached power, amidst the rise of the Great Moderation, these tenets were easily transcribed into an argument about international integration being *written in the stars*[69], governed by *unfathomable, irresistible laws*[70], against the consequences of which, *all, in vain, had been tried*[71], with no other option left than *walking down the path of reason*[72]. To voters faced with the aftermath of the shock, it was the discourse of the estrangement to the Self, a discourse which, once its utility had been exhausted and its role fulfilled was very simply and straightforwardly *aufgehoben.*

[62] If we try to replicate figure 33 over a restriction to cities lying in the most exposed ZEs, or if we include variables that cross exposure and some major explanatory, we indeed find that in most exposed districts, the income and inequality channels which nourishes the FN-RN vote are reinforced: i.e., a decline of income or a rise in inequality within the city tends to bolster the far-right vote, and it is even more the case when the city is more exposed to import competition. However, we find very similar bolstering mechanisms for the left-wing vote, the marginal effects being much more robust across specifications. Replicating the strategy of [Dippel et al. 2017] for left-wing vote shares yields results which suggest a decline in the unemployment and income channels which usually bolster local left-wing support, but the marginal impacts fall short of significance. Both approaches allow us to explain to a large extent the marginal negative impact on the LREM vote, but not the gains experienced by traditional conservatives.

[63] [Autor, Dorn, Hanson, and Majlesi 2016b] find such strong negative impacts for the moderate democratic vote, but not for more progressive democrats; similarly, [Dippel et al. 2017] report negative coefficients for the Green party vote shares, not for the SDP and Die Linke.

[64] It is relatively ironic to see that an algorithm trained at detecting *populist* rhetoric on present-day political platforms [Docquier et al. 2022], when it is provided with political leaflets of French parties of the 1960s, tend to conclude that not only the Communist party, but also the Gaullist union, are populist options ; conversely, the socialist party, even at the time it was allied to the PCF, is never interpreted as a populist party, because its platform always contain some form of support for European federalism and international cooperation in general.

[65] It is at least under this derogatory term that it was, at the time, derided by Althusserian political science [Poulantzas 1968]

[66] P. Ricoeur, *La pluralité des instances de justice* in *Le Juste*, 1995

[67] C. Lefort, *L'invention démocratique*, 1981

[68] P. Ricoeur, "Le risque", *L'Unité française*, 1941

[69] J. Delors, *En sortir ou pas*, 1985

[70] Y. Montand, *Vive la crise*, 22-02-1984

[71] F. Mitterand, *L'Heure de vérité*, 25-10-1993

[72] P. Ricoeur, *Le Monde*, 10-12-1995

6 Conclusion

There are few topics in empirical economics which might better exemplify the shortcomings of an approach focused on the aggregate macro effects of shocks than the issue of trade integration. Arguably, it is very easy to extract from this thesis the main line of argument of the pervasive consensus of the 1990s on trade and inequality reviewed by [Wood 2018]:

- In French context, trade shocks involving a developing partner like China have but insensitive effects on *between*-region inequality; there is a horizontal redistribution between more and less exposed zones (with a marginal commuting-zone-level impact of −2.16pp), but more exposed ones are overall a bit more affluent, making the distributional impact across regions neutral;
- The impact of the *within* structure of disposable income is almost flat, and moderate enough (decadal zone-level marginal impact of −1.41pp) so that there is little doubt that other channels of the trade shock leave a very large majority of households better off once the total impact of trade openness is realised;
- Like in many European equivalent settings [Dauth, Findeisen, and Suedekum 2021], we find evidence that the most socially vulnerable regions (declining heavy-industry bastions of North-East, or deep rural districts) are relatively protected from such import shocks, and that on the contrary the most exposed regions might benefit from exposure to international trade dynamics through other channels; we find no evidence of import exposure hurting the provision of public goods, or the quality of public services, or more generally the attractivity of a region (since we even isolate in-migration fluxes to the most exposed zones);
- If we except a three times lower marginal impact on manufacturing employment, almost every single other consequence connected with the rise of Chinese imports (fiscal income regressive impact, social and political polarisation) is impossible to replicate for data about developed trade partners like Germany.

Yet the most interesting set of results involves the much commented specific nature of Chinese exports [Rodrik 2006; Schott 2008], their becoming increasingly technology-intensive over the last two decades, and the fact that Chinese competition hardly fits into a traditional North-South HOS framework:

- Over the INSEE's data, the negative impact of the China shock over manufacturing employment (a decadal marginal effect of −4.02pp at the commuting zone-*ZE* level) is maximal around the 7th decile of the wage distribution, and for workers with undergraduate degrees, a finding at variance the popular image of Chinese imports hurting mainly declining light industries: in a sectoral approach, textiles, but also steelworks, microelectronics and computers, are the main drivers of the effect. These results are consistent with a job polarisation story [Mion and Zhu 2013; Malgouyres 2017a], but also with the idea that the China shock might drive little firms lagging behind the technological frontier out of the innovation competition [Aghion, Bergeaud, et al. 2021];
- The sharpest negative wage response (overall marginal impact of −6.27log points) is found around the 2nd decile of the wage distribution in both the exposed and non-exposed sectors, driven mainly by female workers faced with a rise of part-time work and a decline in hours worked. That wage impact is however negative for almost everyone along the whole distribution; we then fail to isolate statistically significant wage premia for top jobs, though we find a sizeable impact of the shock on the demand for the most productive workers (in exposed and non-exposed firms alike);
- If the impact of the shock on post-redistribution income is neutral, the effect on pre-redistribution income is highly regressive, with an overall decadal marginal response of −2.16pp, with a sharp discontinuous effect around the 3rd decile (the marginal impact being estimated at −8.01pp at decile 1);
- The neutral impact on disposable income comes at the expense of a marked rise of the share of transfers within the final income (+0.51pp), these families which lose the most in terms of disposable income being found at the 3rd decile, within the hotspot of the fiscal income impact but outside of the scope of many transfer schemes;

We hypothesise that discussions about the existence of inactivity traps in European Welfare architectures − a debate particularly heated in French context [Anne, L'horty, and Dollé 2002; Gomel and Méda 2014] − might have overlooked the dynamic dimension: the perceived legitimacy of a redistributive system might depend less on the absolute level of transfers at time t, than on the way this system reacts to a particularly salient shock. In our setting, if we trust our estimates of section 5, apprehensions about the rise of a new rift within democracies opposing the winners and losers of international economic integration seems to be a performative more than a descriptive discourse; we find no evidence of trade shocks having so massive and intense income effects that they would create an autonomous social group having its own consciousness; it even seems to be the contrary, with the pre-redistribution impact being so closely concentrated on the earliest deciles of the income distribution, that it might concurrently foster trans-class identification, with a rise of anti-redistribution sentiments among the lower-middle-class. In the late 1990s or early 2000s, in the absence of detailed approaches to the income response, policymakers, caught between an academic consensus stressing the distributional innocuousness of trade on the one hand, and heated political reactions to international economic integration on the other hand, were left but with the dual option of discarding the academic consensus to embrace the significance of the political response [Sapir 2011; Klein and Pettis 2021] or of embracing the consensus and discarding political grievances as a mere identitarian reaction. The correct interpretation might be more trivial. In our setting, evidence suggests that international economic integration was relatively well accepted, but that it induced the rise, within the middle-class, of a social consciousness more averse to redistribution in general and to groups perceived as benefiting from it.

References

Abdulkadiroglu, Atila, Joshua Angrist, and Parag Pathak (2011). "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools". In: *Econometrica* 82.

Abramitzky, Ran, Leah Boustan, and Katherine Eriksson (2014). "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration". In: *Journal of Political Economy* 122.3, pp. 467–506.

Acemoglu, Daron (2003). "Patterns of Skill Premia". In: *The Review of Economic Studies* 70.2, pp. 199–230.

Acemoglu, Daron, David Autor, et al. (2016). "Import Competition and the Great US Employment Sag of the 2000s". In: *Journal of Labor Economics* 34.S1, S141–S198.

Acemoglu, Daron and Pascual Restrepo (2020). "Robots and Jobs: Evidence from US Labor Markets". In: *Journal of Political Economy* 128.6, pp. 2188–2244.

Adão, Rodrigo, Costas Arkolakis, and Federico Esposito (2019). *General equilibrium effects in space: Theory and measurement.* Tech. rep. National Bureau of Economic Research.

Adão, Rodrigo, Paul Carrillo, et al. (2020). *International Trade and Earnings Inequality: A New Factor Content Approach.* NBER Working Papers 28263. National Bureau of Economic Research, Inc.

Adão, Rodrigo, Michal Kolesár, and Eduardo Morales (2019). "Shift-Share Designs: Theory and Inference". In: *The Quarterly Journal of Economics* 134.4, pp. 1949–2010.

Aghion, Philippe, Antonin Bergeaud, et al. (2021). *Opposing firm-level responses to the China shock: horizontal competition versus vertical relationships?* Tech. rep. National Bureau of Economic Research.

Aghion, Philippe, Nicholas Bloom, et al. (2005). "Competition and Innovation: an Inverted-U Relationship". In: *The Quarterly Journal of Economics* 120.2, pp. 701–728.

Alesina, Alberto, Reza Baqir, and William Easterly (1999). "Public Goods and Ethnic Divisions". In: *The Quarterly Journal of Economics* 114.4, pp. 1243–1284.

Algan, Yann, Elizabeth Beasley, et al. (2019). *Qui sont les Gilets jaunes et leurs soutiens ?* Tech. rep. 2019-03. OBE-CEPREMAP-CEVIPOF.

Algan, Yann, Camille Hemet, and David D. Laitin (2016). "The Social Effects of Ethnic Diversity at the Local Level: A Natural Experiment with Exogenous Residential Allocation". In: *Journal of Political Economy* 124.3, pp. 696–733.

Alvaredo, Facundo, Denis Cogneau, and Thomas Piketty (2021). "Income inequality under colonial rule. Evidence from French Algeria, Cameroon, Tunisia, and Vietnam and comparisons with British colonies 1920–1960". In: *Journal of Development Economics* 152, p. 102680.

Angrist, Joshua and Kevin Lang (2004). "Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program". In: *American Economic Review* 94.5, pp. 1613–1634.

Angrist, Joshua and Jorn-Steffen Pischke (2010). "The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics". In: *Journal of economic perspectives* 24.2, pp. 3–30.

Anne, Denis, Yannick L'horty, and Michel Dollé (2002). "Transferts sociaux locaux et retour à l'emploi ; suivi d'un commentaire de Michel Dollé, et d'une réponse au commentaire, de Denis Anne et Yannick L'Horty". In: *Economie Et Statistique* 357, pp. 49–78.

Aslund, Olof, Lena Hensvik, and Oskar Nordström Skans (2014). "Seeking Similarity: How Immigrants and Natives Manage in the Labor Market". In: *Journal of Labor Economics* 32.3, pp. 405–441.

Autor, David and David Dorn (2013). "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market". In: *American Economic Review* 103.5, pp. 1553–1597.

Autor, David, David Dorn, and Gordon Hanson (2013). "The China Syndrome: Local Labor Market Effects of Import Competition in the United States". In: *American Economic Review* 103.6, pp. 2121–68.

— (2016). *The China Shock: Learning from Labor Market Adjustment to Large Changes in Trade.* CEPR Discussion Papers. C.E.P.R. Discussion Papers.

— (2021). *On the Persistence of the China Shock.* NBER Working Papers. National Bureau of Economic Research, Inc.

Autor, David, David Dorn, Gordon Hanson, and Kaveh Majlesi (2016a). *A Note on the Effect of Rising Trade Exposure on the 2016 Presidential Election.* MPRA Paper 112889. University Library of Munich, Germany.

— (2016b). "Importing Political Polarization? The Electoral Consequences of Rising Trade Exposure". In: *American Economic Review* 110.

Autor, David, David Dorn, Gordon Hanson, Gary Pisano, et al. (2016). *Foreign Competition and Domestic Innovation: Evidence from U.S. Patents.* Tech. rep. C.E.P.R. Discussion Papers.

Autor, David, David Dorn, Gordon Hanson, and Jae Song (2014). " Trade Adjustment: Worker-Level Evidence". In: *The Quarterly Journal of Economics* 129.4, pp. 1799–1860.

Autor, David, David Dorn, Lawrence Katz, et al. (2020). "The fall of the labor share and the rise of superstar firms". In: *The Quarterly Journal of Economics* 135.2, pp. 645–709.

Autor, David and Lawrence Katz (1999). "Chapter 26 - Changes in the Wage Structure and Earnings Inequality". In: *Handbook of Labor Economics.* Ed. by Orley C. Ashenfelter and David Card. Vol. 3. Elsevier, pp. 1463–1555.

Autor, David, Lawrence Katz, and Melissa Kearney (2006). "The Polarization of the U.S. Labor Market". In: *American Economic Review* 96.2, pp. 189–194.

Autor, David, Frank Levy, and Richard Murnane (2003). "The Skill Content of Recent Technological Change: An Empirical Exploration". In: *The Quarterly Journal of Economics* 118.4, pp. 1279–1333.

Baddeley, Michelle, Ronald Martin, and Peter Tyler (1998). "Transitory shock or structural shift? The impact of the early 1980s recession on British regional unemployment". In: *Applied Economics* 30.1, pp. 19–30.

Ballard-Rosa, Cameron, Amalie Jensen, and Kenneth Scheve (2022). "Economic decline, social identity, and authoritarian values in the United States". In: *International Studies Quarterly* 66.1, sqab027.

Baltagi, Badi, James M. Griffin, and Weiwen Xiong (2000). "To Pool Or Not To Pool: Homogeneous Versus Hetergeneous Estimations Applied to Cigarette Demand". In: *The Review of Economics and Statistics* 82.1, pp. 117–126.

Banderier, Gaston (1970). *Les Revenus des ménages en 1965.* Vol. 34. Collections de l'INSEE, série M.

Banderier, Gaston and Pierre Ghigliazza (1974). *Les Revenus des ménages en 1970.* Vol. 40. Collections de l'INSEE, série M.

Bandiera, Oriana, Imran Barankay, and Imran Rasul (2009). "Social Connections and Incentives in the Workplace: Evidence From Personnel Data". In: *Econometrica* 77, pp. 1047–1094.

Barone, Guglielmo, Alessio D'Ignazio, et al. (2016). "Mr. Rossi, Mr. Hu and politics. The role of immigration in shaping natives' voting behavior". In: *Journal of Public Economics* 136.C, pp. 1–13.

Barone, Guglielmo and Helena Kreuter (2021). "Low-wage import competition and populist backlash: The case of Italy". In: *European Journal of Political Economy* 67, p. 101970.

Barro, Robert and Xavier Sala-i-Martin (1990). *Economic Growth and Convergence across The United States.* NBER Working Papers 3419. National Bureau of Economic Research, Inc.

— (1992). "Convergence". In: *Journal of Political Economy* 100.2, pp. 223–51.

Bartik, Timothy J. (1991). *Who Benefits from State and Local Economic Development Policies?* Books from Upjohn Press. W.E. Upjohn Institute for Employment Research.

Bayer, Patrick, Stephen L. Ross, and Giorgio Topa (2008). "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes". In: *Journal of Political Economy* 116.6, pp. 1150–1196.

Bayet, Alain and Martine Julhès (April 1996). "Séries longues sur les salaires". In: *INSEE-Résultats* 457.105.

Beaud, Stéphane and Michel Pialoux (2003). *Violences urbaines, violence sociale. Genèse des nouvelles classes dangereuses.* Fayard.

Beffy, Magali and Laurent Davezies (2013). "Has the "Ambition Success Networks" Educational Program Achieved its Ambition?" In: *Annals of economics and statistics*, pp. 271–293.

Beyer, Robert and Frank Smets (2015a). "Labour market adjustments and migration in Europe and the United States: How different?" In: *Economic Policy* 30, pp. 643–682.

— (2015b). *Labour market adjustments in Europe and the US: How different?* Working Paper Series 1767. European Central Bank.

Biscourp, Pierre and Francis Kramarz (2007). "Employment, skill structure and international trade: Firm-level evidence for France". In: *Journal of International Economics* 72.1, pp. 22–51.

Bivens, Josh (2007). *Globalization, American Wages and Inequality: Past, Present, and Future.* Tech. rep. Economic Policy Institute Working Paper 279.

Blanchard, Olivier and Lawrence Katz (1992). "Regional Evolutions". In: *Brookings Papers on Economic Activity* 23.1, pp. 1–76.

Blanchard, Olivier and Lawrence H.l Summers (1986). "Hysteresis and the European Unemployment Problem". In: *NBER Macroeconomics Annual* 1, pp. 15–78.

Blanchet, Thomas, Lucas Chancel, and Amory Gethin (2019). "How Unequal is Europe? Evidence from Distributional National Accounts, 1980-2017". In: *WID Working paper* 2019/3.

— (2020). "Why is Europe More Equal than the United States?" In: *WID Working paper* 2020/19.

Blanchet, Thomas, Juliette Fournier, and Thomas Piketty (2017). "Generalized Pareto Curves: Theory and Applications". In: *WID working paper.*

Bloom, Nicholas, Mirko Draca, and John Van Reenen (2015). "Trade Induced Technical Change? The Impact of Chinese Imports on Innovation, IT and Productivity". In: *The Review of Economic Studies* 83.1, pp. 87–117.

Borjas, George, Richard Friedman, and Lawrence Katz (1997). "How Much Do Immigration and Trade Affect Labor Market Outcomes?" In: *Brookings Papers on Economic Activity* 28.1, pp. 1–90.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel (2021). "Quasi-Experimental Shift-Share Research Designs". In: *The Review of Economic Studies* 89.1, pp. 181–213.

Borusyak, Kirill and Xavier Jaravel (2021). *The Distributional Effects of Trade: Theory and Evidence from the United States.* NBER Working Papers 28957. National Bureau of Economic Research, Inc.

Botton, Hugo et al. (2020). *L'évolution de la ségrégation résidentielle en France : 1990-2015.* Document de travail 2020-09. France Stratégie.

Boulard, Fernand, ed. (1982). *Matériaux pour l'histoire religieuse du peuple français XIXe-XXe siècles.* EHESS-Presses de la FNSP-CNRS.

Bourdieu, Pierre (1979). *La Distinction. Critique sociale du jugement.* Minuit.

— (1980). *Le Sens pratique.* Minuit.

Bourguignon, François and Christian Morrisson (2002). "Inequality Among World Citizens: 1820-1992". In: *American Economic Review* 92.4, pp. 727–744.

Bouvier, Jean (1973). "Le système fiscal français du dix-neuvième siècle: étude critique d'un immobilisme". In: *Deux siècles de fiscalité française, dix-neuvième, vingtième siècle: histoire, économie, politique.* Ed. by Jean Bouvier, Robert Schnerb, and Jacques Wolff. Paris-La Haye: Mouton.

— (1978). "Histoire financière et problèmes d'analyse des dépenses publiques". In: *Annales ESC* 33.2, pp. 207–215.

Brechling, Frank (1967). " Trends and cycles in British unemployment". In: *Oxford Economic Papers* 19.1, pp. 1–21.

Brown, Meta, Elizabeth Setren, and Giorgio Topa (2012). "Do Informal Referrals Lead to Better Matches? Evidence from a Firm's Employee Referral System". In: *Journal of Labor Economics* 34.

Bunel, Mathieu, Yannick L'Horty, and Pascale Petit (2016). "Discrimination based on place of residence and access to employment". In: *Urban Studies* 53.2, pp. 267–286.

Burstein, Ariel and Jonathan Vogel (2017). "International trade, technology, and the skill premium". In: *Journal of Political Economy* 125.5, pp. 1356–1412.

Caliendo, Lorenzo, Maximiliano Dvorkin, and Fernando Parro (2015). *Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock.* Working Papers. Federal Reserve Bank of St. Louis.

Campagne, Nathalie, Didier Contencin, and Christelle Roineau (1996). *Les revenus fiscaux des ménages en 1990.* Vol. 453. INSEE-Résultats.

Canceill, Geneviève, Antoine Chastand, and Olivier Choquet (1981). *Données statistiques sur les familles.* Vol. 86. Collections de l'INSEE, série M.

Card, David and Laura Giuliano (2016). "Can Tracking Raise the Test Scores of High-Ability Minority Students?" In: *American Economic Review* 106.10, pp. 2783–2816.

Card, David, Alexandre Mas, and Jesse Rothstein (2008). "Tipping and the Dynamics of Segregation". In: *The Quarterly Journal of Economics* 123.1, pp. 177–218.

Cartier, Marie et al. (2008). *La France des petits-moyens.* La Découverte.

Castells, Manuel (1972). *La question urbaine.* Paris: Maspero.

Cerrato, Andrea, Federico Maria Ferrara, and Francesco Ruggieri (2018). “Why Does Import Competition Favor Republicans? Localized Trade Shocks, Voting Behavior, and Scapegoating in the U.S.” In: *SSRN Electronic Journal*.

Chancel, Lucas (2019a). “Global inequality: a matter of class or of nationality?” In: *WID Working paper* 2019/3.

— (2019b). “Ten facts about income inequality in advanced economies”. In: *WID Working paper* 2019/3.

Chancel, Lucas and Thomas Piketty (2021). *Global Income Inequality, 1820-2020: The Persistence and Mutation of Extreme Inequality*. PSE Working Papers.

Chetverikov, Denis, Bradley Larsen, and Christopher Palmer (2016). “IV Quantile Regression for Group-Level Treatments, With an Application to the Distributional Effects of Trade”. In: *Econometrica* 84, pp. 809–833.

Chyn, Eric (2018). “Moved to Opportunity: The Long-Run Effects of Public Housing Demolition on Children”. In: *American Economic Review* 108, pp. 3028–3056.

Citino, Luca and Andrea Linarello (2021). “The impact of Chinese import competition on Italian manufacturing”. In: *Review of International Economics*.

Clark, Thomas C. and Philip B. Perlman (1948). *Prejudice and Property : An Historic Brief against Racial Covenants*. Public Affairs Press.

Cogneau, Denis, Yannick Dupraz, and Sandrine Mesplé-Somps (2018). “African states and development in historical perspective: Colonial public finances in British and French West”.

Colantone, Italo and Piero Stanig (2017). *The Trade Origins of Economic Nationalism: Import Competition and Voting Behavior in Western Europe*. BAFFI CAREFIN Working Papers 1749.

— (2018). “Global Competition and Brexit”. In: *American Political Science Review* 112, pp. 1–18.

Collovald, Annie and Olivier Schwartz (2006). “Haut, bas, fragile : sociologies du populaire”. In: *Vacarme* 6.

Commissariat général du Plan (1993). *Monde-Europe, repères et orientations pour les Français, 1993-1997*. Dunod-La Documentation française.

Corbin, Alain (1975). *Archaïsme et modernité en Limousin au dix-neuvième siècle (1845-1880)*. Paris: Marcel Rivière et Cie.

Costa, Rui, Swati Dhingra, and Stephen Machin (2019). *Trade and Worker Deskilling*. Tech. rep. 25919. National Bureau of Economic Research.

Cutler, David M. and Edward L. Glaeser (1997). “Are Ghettos Good or Bad?” In: *The Quarterly Journal of Economics* 112.3, pp. 827–872.

Cutler, David M., Edward L. Glaeser, and Jacob L. Vigdor (1999). “The Rise and Decline of the American Ghetto”. In: *Journal of Political Economy* 107.3, pp. 455–506.

Damette, Félix (1994). *La France en villes*. Paris: La Documentation française.

Dauth, Wolfgang, Sebastian Findeisen, and Jens Suedekum (2021). “Adjusting to Globalization in Germany”. In: *Journal of Labor Economics* 39.1, pp. 263–302.

Davenport, Alex, David Dorn, and Peter Levell (2021). “Import Competition and Public Attitudes Towards Trade”. In: *CEPR Discussion Paper No. DP16339*.

Davezies, Laurent (2012). *La crise qui vient. La nouvelle fracture territoriale*. Paris: Le Seuil.

— (2021). *L’État a toujours soutenu ses territoires*. Paris: Le Seuil.

Davis, Morris A., Jonas D.M. Fisher, and Marcelo Veracierto (2021). “Migration and urban economic dynamics”. In: *Journal of Economic Dynamics and Control* 133, p. 104234.

De Lyon, Josh and Joao Paulo Pessoa (2021). “Worker and firm responses to trade shocks: The UK-China case”. In: *European Economic Review* 133, p. 103678.

Delalande, Nicolas and Alexis Spire (2010). *Histoire sociale de l’impôt*. La Découverte.

Denton, Nancy A. and Douglas S. Massey (1992). *American Apartheid, Segregation and the Making of the Underclass*. Harvard University Press.

Dherbécourt, Clément and Flore Deschard (2019). *La répartition territoriale des emplois publics*. Document de travail 2019/3. France Stratégie.

Dijk, Jasper Jacob van (2016). “Local employment multipliers in U.S. cities”. In: *Journal of Economic Geography* 17.2, pp. 465–487.

Dippel, Christian et al. (2017). *Instrumental Variables and Causal Mechanisms: Unpacking The Effect of Trade on Workers and Voters*. Tech. rep. 23209. National Bureau of Economic Research.

Docquier, Frederic et al. (2022). *Globalization and Populism: The Last Sixty Years*. working paper not yet circulated.

Dorn, David and Peter Levell (2021). *Trade and Inequality in Europe and the US*. IZA Discussion Papers 14914. Institute of Labor Economics (IZA).

Duguet, Emmanuel, David Gray, et al. (2020). “Labour market effects of urban riots: An experimental assessment”. In: *Papers in Regional Science* 99.3, pp. 787–806.

Duguet, Emmanuel, Noam Leandri, et al. (2010). “Are Young French Jobseekers of Ethnic Immigrant Origin Discriminated against? A Controlled Experiment in the Paris Area”. In: *Annals of Economics and Statistics* 99/100, pp. 187–215. (Visited on 09/07/2022).

Duncan, Otis and Beverly Duncan (1955). “A Methodological Analysis of Segregation Indexes”. In: *American Sociological Review* 20.2, pp. 210–217.

Dustmann, Christian, Albrecht Glitz, et al. (2016). “Referral-based Job Search Networks”. In: *Review of Economic Studies* 83.2, pp. 514–546.

Dustmann, Christian, Kristine Vasiljeva, and Anna Damm (2019). “Refugee Migration and Electoral Outcomes”. In: *Review of Economic Studies* 86.5, pp. 2035–2091.

Eaton, Jonathan and Samuel Kortum (2002). “Technology, Geography, and Trade”. In: *Econometrica* 70, pp. 1741–1779.

Echenique, Federico and Jr. Fryer Roland G. (2007). “A Measure of Segregation Based on Social Interactions”. In: *The Quarterly Journal of Economics* 122.2, pp. 441–485.

Edwards, Lawrence and Robert Z. Lawrence (2010). *US Trade and Wages: The Misleading Implications of Conventional Trade Theory*. NBER Working Papers. National Bureau of Economic Research, Inc.

Elhorst, J.Paul (2010). “Applied Spatial Econometrics: Raising the Bar”. In: *Spatial Economic Analysis* 5.1, pp. 9–28.

Fajgelbaum, Pablo and Amit Khandelwal (2016). “Measuring the Unequal Gains from Trade”. In: *The Quarterly Journal of Economics* 131.3, pp. 1113–1180.

Fajgelbaum, Pablo D et al. (2019). “The Return to Protectionism*”. In: *The Quarterly Journal of Economics* 135.1, pp. 1–55.

Feenstra, Robert and Gordon Hanson (2001). *Global Production Sharing and Rising Inequality: A Survey of Trade and Wages*. NBER Working Papers 8372. National Bureau of Economic Research, Inc.

Feigenbaum, James and Andrew Hall (2015). “How Legislators Respond To Localized Economic Shocks”. In: *Journal of Politics* 77.4, pp. 1012–1030.

Feler, Leo and Mine Senses (2017). “Trade Shocks and the Provision of Local Public Goods”. In: *American Economic Journal: Economic Policy* 9.4, pp. 101–143.

Fichte, Johannes G. (1799). *Der geschlossene Handelsstaat*. Fichtes sämmtliche Werke. Herausgegeben von I. H. Fichte, Band 1-8, Berlin: Veit Comp., 1845/1846.

Firpo, Sergio, Nicole M. Fortin, and Thomas Lemieux (2011). *Occupational Tasks and Changes in the Wage Structure*. IZA Discussion Papers. Institute of Labor Economics (IZA).

Frankel, David M. and Oscar Volij (2011). “Measuring school segregation”. In: *Journal of Economic Theory* 146.1, pp. 1–38.

Galle, Simon, Andrés Rodríguez-Clare, and Moises Yi (2022). “Slicing the Pie: Quantifying the Aggregate and Distributional Effects of Trade”. In: *The Review of Economic Studies*.

Garbinti, Bertrand, Jonathan Goupille-Lebret, and Thomas Piketty (2018). “Income inequality in France, 1900–2014: Evidence from Distributional National Accounts (DINA)”. In: *Journal of Public Economics* 162, pp. 63–77.

Gennaioli, Nicola and Guido Tabellini (2019). *Identity, Beliefs, and Political Conflict*. CESifo Working Paper Series 7707. CESifo.

Gethin, Amory, Clara Martinez-Toledano, and Thomas Piketty (2021). *Clivages politiques et inégalités sociales: une étude de 50 démocraties (1948-2020)*. Paris: Seuil.

Giuliano, Laura, David Levine, and Jonathan Leonard (2009). “Manager Race and the Race of New Hires”. In: *Journal of Labor Economics* 27.

Gobillon, Laurent, Harris Selod, and Yves Zenou (2005). “The Mechanisms of Spatial Mismatch”. In: *Urban Studies* 44.

Goffette-Nagot, Florence, Pablo Jensen, and Sebastian Grauwin (2009). “Dynamic models of residential segregation: Brief review, analytical resolution and study of the introduction of coordination”. In: *HAL, Post-Print* 96.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift (2020). “Bartik Instruments: What, When, Why, and How”. In: *American Economic Review* 110.8, pp. 2586–2624.

Gomel, Bernard and Dominique Méda (2014). “Le RSA, innovation ou réforme technocratique ?” In: *Revue Française de Socio-Economie* 1.13.

Grossman, Gene and Elhanan Helpman (2018). *Identity Politics and Trade Policy*. Tech. rep. 25348.

Grossman, Gene and Esteban Rossi-Hansberg (2006). *Trading Tasks: A Simple Theory of Offshoring*. NBER Working Papers 12721. National Bureau of Economic Research, Inc.

Grossman, Gene M., Elhanan Helpman, and Philipp Kircher (2017). “Matching, Sorting, and the Distributional Effects of International Trade”. In: *Journal of Political Economy* 125.1, pp. 224–264.

Hakobyan, Shushanik and John McLaren (2016). “Looking for Local Labor Market Effects of NAFTA”. In: *The Review of Economics and Statistics* 98.4, pp. 728–741.

Halbwachs, Maurice (1932). “Chicago, expérience ethnique”. In: *Annales* 4.13, pp. 11–49.

Halleck-Vega, Solmaria and J.Paul Elhorst (2016). “A regional unemployment model simultaneously accounting for serial dynamics, spatial dependence and common factors”. In: *Regional Science and Urban Economics* 60.C, pp. 85–95.

— (2017). “Regional labour force participation across the European Union: a time–space recursive modelling approach with endogenous regressors”. In: *Spatial Economic Analysis* 12.2-3, pp. 138–160.

Hansen, Bruce E. (2000). “Sample Splitting and Threshold Estimation”. In: *Econometrica* 68, pp. 575–603.

Hellerstein, Judith K., Mark J. Kutzbach, and David Neumark (2014). “Do labor market networks have an important spatial dimension?” In: *Journal of Urban Economics* 79.C, pp. 39–58.

Hellerstein, Judith K., Melissa McInerney, and David Neumark (2011). “Neighbors and Coworkers: The Importance of Residential Labor Market Networks”. In: *Journal of Labor Economics* 29.4, pp. 659–695.

Helpman, Elhanan et al. (2016). “Trade and Inequality: From Theory to Estimation”. In: *The Review of Economic Studies* 84.1, pp. 357–405.

Hémet, Camille and Clément Malgouyres (2018). “Diversity and Employment Prospects: Neighbors Matter!” In: *Journal of Human Resources* 53.3, pp. 825–858.

Hufbauer, Gary Clyde et al. (1993). “NAFTA: An assessment”. In.

Hummels, David et al. (2014). “The wage effects of offshoring: Evidence from Danish matched worker-firm data”. In: *American Economic Review* 104.6, pp. 1597–1629.

INSEE (1983). *Annuaire statistique 1984*.

— (1986). *Annuaire statistique 1987*.

Jaravel, Xavier and Erick Sager (2019). *What are the price effects of trade? Evidence from the US and implications for quantitative trade models*. CEP Discussion Papers. Centre for Economic Performance, LSE.

Kaldor, Nicholas (1940). “A Model of the Trade Cycle”. In: *The Economic Journal* 50.197, pp. 78–92. (Visited on 08/27/2022).

Katz, Lawrence (1992). “Understanding Recent Changes in the Wage Structure”. In: *NBER Reporter* 4.

Kaufmann, Daniel and Tobias Renkin (2018). *Manufacturing Prices and Employment after the Swiss Franc Shock*. en. Tech. rep. Zurich.

Kellogg, John (1977). “Black Urban Clusters in the Postbellum South”. In: *Geographical Review* 67.3, pp. 310–321.

Klein, Matthew C. and Michael Pettis (2021). *Trade Wars Are Class Wars*. Yale University Press.

Kovak, Brian K. (2013). “Regional Effects of Trade Reform: What Is the Correct Measure of Liberalization?” In: *American Economic Review* 103, pp. 1960–76.

Krugman, Paul (1993). “The Uncomfortable Truth about NAFTA: It’s Foreign Policy, Stupid”. In: *Foreign Affairs* 72, p. 13.

Krugman, Paul, Richard N. Cooper, and T. N. Srinivasan (1995). “Growing World Trade: Causes and Consequences”. In: *Brookings Papers on Economic Activity* 1, pp. 327–377. (Visited on 08/27/2022).

Kuziemko, Ilyana and Ebonya Washington (2018). “Why Did the Democrats Lose the South? Bringing New Data to an Old Debate”. In: *The American Economic Review* 108.10, pp. 2830–2867.

Lachappelle, Georges (1924). *Elections législatives du 11 mai 1924*. G. Roustan.

— (1936). *Elections législatives, 26 avril et 3 mai 1936*. Le Temps.

Lafourcade, Miren and Florian Mayneris (2017). *En Finir avec les ghettos urbains ?* Vol. 44. CEPREMAP-Editions rue d’Ulm-Presses de l’Ecole normale supérieure.

Lawrence, Robert and Matthew J. Slaughter (1993). “International Trade and American Wages in the 1980s: Giant Sucking Sound or Small Hiccup?” In: *Brookings Papers on Economic Activity* 24.2 Microeconomics, pp. 161–226.

Le Bras, Gabriel (1931). “Pour un examen détaillé et pour une explication historique de l’état du catholicisme dans les diverses régions de France”. In: *Revue d’histoire de l’Église de France* 17.

Lechien, Marie-Hélène and Yasmine Siblot (2019). “Eux/nous/ils ? Sociabilités et contacts sociaux en milieu populaire”. In: *Sociologie* 10.1, pp. 1–15.

Lefebvre, Rémi and Frédéric Sawicki (2006). *La société des socialiste : le P.S. aujourd’hui*. Editions du Croquant.

Lesage, James and R. Pace (2014). “The Biggest Myth in Spatial Econometrics”. In: *Econometrics* 2.

Lesage, James and Robert K. Pace, eds. (2009). *Introduction to spatial econometrics*. Chapman and Hall/CRC.

Lesuisse, Pierre (2020). *Employment, unemployment, participation and migration: a regional study*. Post-Print. HAL.

Lipietz, Alain (1977). *Le Capital et son espace*. Paris: Maspero.

Lipset, Seymour Martin (1959). “Democracy and Working-Class Authoritarianism”. In: *American Sociological Review* 24.4, pp. 482–501.

Loonis, Vincent and Marie-Pierre de Bellefon, eds. (2018). *Manuel d’analyse spatiale - Handbook of spatial analysis*. Vol. 131. INSEE Méthodes.

Lopez-Acevedo, Gladys and Raymond Robertson (2012). *Sewing success?: Employment, wages, and poverty following the end of the multi-fibre arrangement*. World Bank Publications.

Louf, Rémi and Marc Barthelemy (2016). “Patterns of residential segregation”. In: *PloS one* 11.6, e0157476.

Ludwig, Jens, Greg Duncan, and Paul Hirschfield (2001). “Urban Poverty And Juvenile Crime: Evidence From A Randomized Housing-Mobility Experiment”. In: *The Quarterly Journal of Economics* 116, pp. 655–679.

Malgouyres, Clément (2017a). “The Impact Of Chinese Import Competition On The Local Structure Of Employment And Wages: Evidence From France”. In: *Journal of Regional Science* 57.3, pp. 411–441.

— (2017b). *Trade Shocks and Far-Right Voting: Evidence from French Presidential Elections*. RSCAS Working Papers 2017/21. European University Institute.

Manski, Charles F. (1993). “Identification of Endogenous Social Effects: The Reflection Problem”. In: *The Review of Economic Studies* 60.3, pp. 531–542.

Margalit, Yotam (2011). “Costly Jobs: Trade-related Layoffs, Government Compensation, and Voting in U.S. Elections”. In: *American Political Science Review* 105, pp. 166–188.

Marx, Karl (1852). *Der achtzehnte Brumaire des Louis Bonaparte,*

Mayda, Anna Maria and Dani Rodrik (2005). “Why are some people (and countries) more protectionist than others?” In: *European Economic Review* 49.6, pp. 1393–1430.

Melitz, Marc J. (2003). “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity”. In: *Econometrica* 71.6, pp. 1695–1725. (Visited on 07/01/2022).

Melo, Rodrigo Azuero, TD Rodriguez, and D Zarruk (2016). *Distance and travel time between two points from google maps*.

Ministère chargé du logement (2020). *La loi solidarité et renouvellement urbains (SRU) - Bilan triennal 2017-2019*.

Ministère des travaux publics (1837). *Statistique de la France*. Paris: Imprimerie royale.

Mion, Giordano and Linke Zhu (2013). “Import competition from and offshoring to China: A curse or blessing for firms?” In: *Journal of International Economics* 89.1, pp. 202–215.

Moretti, Enrico (2010). “Local Multipliers”. In: *American Economic Review* 100.2, pp. 373–77.

Novokmet, Filip, Thomas Piketty, and Gabriel Zucman (2018). “From Soviets to oligarchs: inequality and property in Russia 1905-2016”. In: *The Journal of Economic Inequality* 16.2, pp. 189–223.

OECD (2017). *Bilateral Trade Database by Industry and End-use (BTDIxE)*. Tech. rep.

Oreopoulos, Philip (2003). “The Long-Run Consequences Of Living In A Poor Neighborhood”. In: *The Quarterly Journal of Economics* 118, pp. 1533–1575.

Pan Ké Shon, Jean-Louis and Gregory Verdugo (2014). “Ségrégation et incorporation des immigrés en France : ampleur et intensité entre 1968 et 2007”. In: *Revue française de sociologie* 55, pp. 245–283.

Pérez, Santiago (2019). *Southern (American) Hospitality: Italians in Argentina and the US during the Age of Mass Migration*. NBER Working Papers 26127. National Bureau of Economic Research, Inc.

Piketty, Thomas (1995). “Social Mobility and Redistributive Politics”. In: *The Quarterly Journal of Economics* 110.3, pp. 551–584.

— (2001). *Les Hauts revenus en France au 20e siècle. Inégalités et redistribution, 1901-1998*. Paris: Grasset.

— (2011). “On the Long-Run Evolution of Inheritance: France 1820–2050”. In: *The Quarterly Journal of Economics* 126.3, pp. 1071–1131.

— (2013). *Le Capital au vingt-et-unième siècle*. Paris: Le Seuil.

— (2019). *Capital et idéologie*. Paris: Le Seuil.

Piketty, Thomas, Gilles Postel-Vinay, and Jean-Laurent Rosenthal (2006). “Wealth Concentration in a Developing Economy: Paris and France, 1807–1994”. In: *American Economic Review* 96.1, pp. 236–256.

— (2018). *The End of the Rentiers: Paris 1842-1957*. Working Papers. WID.

Poulantzas, Nicos (1968). *Pouvoir politique et classes sociales*. Maspero.

Présidence du Conseil (1930). *Annuaire statistique de la France 1930*. Direction de la statistique générale et de la documentation. Paris.

— (1937). *Annuaire statistique de la France 1937*. Direction de la statistique générale et de la documentation. Paris.

Présidence du Conseil (1938). *Annuaire statistique de la France 1938.* Direction de la statistique générale et de la documentation. Paris.

Préteceille, Edmond (2009). "La ségrégation ethno-raciale a-t-elle augmenté dans la métropole parisienne ?" In: *Revue française de sociologie* 50, pp. 489–519.

Rodrik, Dani (1997). *Has Globalization Gone Too Far?* Peterson Institute for International Economics.

— (2006). "What's So Special about China's Exports?" In: *China & World Economy* 14.5, pp. 1–19.

— (2021). "Why Does Globalization Fuel Populism? Economics, Culture, and the Rise of Right-Wing Populism". In: *Annual Review of Economics* 13, pp. 133–70.

Salais, Robert (1969). "L'emploi industriel a de nouveau augmenté de 1968 à 1969". In: *Economie et Statistique* 7.1, pp. 17–26.

Salmon, Frédéric (2001). *Atlas électoral de la France (1848-2001).* Paris: Le Seuil.

Samuelson, Paul A. (2004). "Where Ricardo and Mill Rebut and Confirm Arguments of Mainstream Economists Supporting Globalization". In: *Journal of Economic Perspectives* 18.3, pp. 135–146.

Sapir, Jacques (2011). *La Démondialisation.* Le Seuil.

Schelling, Thomas C. (1971). "Dynamic models of segregation". In: *The Journal of Mathematical Sociology* 1.2, pp. 143–186.

Schott, Peter K. (2008). "The relative sophistication of Chinese exports ['Manufacturing Earnings and Compensation in China']". In: *Economic Policy* 23.53, pp. 6–49.

Seys, Baudouin (1996). "L'évolution sociale de la population active". In: *Insee Première* 434.

Slaughter, Matthew J. (1998). *What Are the Results of Product-Price Studies and What Can We Learn From Their Differences?* NBER Working Papers. National Bureau of Economic Research, Inc.

Steiner, Nils D, Philipp Harms, et al. (2020). "Local trade shocks and the nationalist backlash in political attitudes: panel data evidence from Great Britain". In: *Gutenberg School of Management and Economics.*

Taffin, Claude and Thierry Debrand (2005). "Les facteurs structurels et conjoncturels de la mobilité résidentielle depuis 20 ans". In: *Économie et Statistique* 381, pp. 125–146.

Thirlwall, Anthony P. (1966). "Regional Unemployment As A Cyclical Phenomenon1". In: *Scottish Journal of Political Economy* 13.2, pp. 205–219.

Topalova, Petia (2010). "Factor Immobility and Regional Impacts of Trade Liberalization: Evidence on Poverty from India". In: *American Economic Journal: Applied Economics* 2, pp. 1–41.

Turek, Joseph H. (1985). "The Northeast in a National Context: Background Trends in Population, Income and Employment". In: *Economic Prospects for the Northeast.* Ed. by Harry W. Richardson and Joseph H. Turek. Philadelphia: Temple University Press.

Utar, Hale (2018). "Workers beneath the Floodgates: Low-Wage Import Competition and Workers' Adjustment". In: *The Review of Economics and Statistics* 100.4, pp. 631–647.

Vernon, Raymond (1966). "International Investment and International Trade in the Product Cycle". In: *The Quarterly Journal of Economics* 80.2, pp. 190–207.

Vignal, Cécile (2005). "Les espaces familiaux à l'épreuve de la délocalisation de l'emploi: Ancrages et mobilités de salariés de l'industrie". In: *Espaces et sociétés* 120-121.

Whyte, William F. (1943). *Street Corner Society: The Social Structure of an Italian Slum.* University of Chicago Press.

Wolf, Nikolaus (2009). "Was Germany ever united? Evidence from intra-and international trade, 1885–1933". In: *The Journal of Economic History* 69.3, pp. 846–881.

Wood, Adrian (2018). "The 1990s trade and wages debate in retrospect". In: *The World Economy* 41.4, pp. 975–999.

Wozniak, Abigail (2010). "Are College Graduates More Responsive to Distant Labor Market Opportunities?" In: *Journal of Human Resources* 45.4, pp. 944–970.

A Income data

RC-IRCOM database (*Revenus communaux*)

Overview

Before 1979, a person whose final income was not taxable was not mandated by the French administration to fill a tax declaration. The law of the 1st of July, 1979 paved the way for the principle of universal declaration[73], allowing the administration to gather data about low-income households. At that time, the share of non-taxpaying population had reached an all-time minimum of 35%.

Since 1990, the administration has been consolidating comprehensive data which provide, for each level of governance - statewide, regional (*région, département*) and local (*commune*) - the number of households which have filled a declaration, and several tax indexes. These are the RC (*Revenus communaux*) files, which provide the following variables[74]:

Table 19: Variables provided by the RC-IRCOM datasets

Index	Content	1990-1993	1994-1995	1996-1999	2001-2015	2016-2020
FN	Number of households	X	X	X	X	X
FS	Cumulated taxable income of all households		X	X	X	X
FM	Mean net taxable income per household	X		X		
IS	Cumulated net tax revenue		X	X	X	X
IM	Mean net tax revenue per household			X		
FIN	Number of taxpaying households		X	X	X	X
FIS	Cumulated income of taxpaying households		X	X	X	X
FIM	Mean income of taxpaying households					
FNN	Number of non-taxpaying households			X		
FNS	Cumulated income of non-taxpayers			X		
TSN	Number of wages (all households)				X	X
TSS	Cumulated wages				X	X
RPN	Number of pensions (all households)				X	X
RPS	Cumulated pensions				X	X

From 1990 to 1999, the data was published by the INSEE, each document considering the incomes of year N. From 1990 to 1998, the franc is the currency. From 2001 to 2020 on the contrary, the very same series was issued by the tax administration, formerly the DGI, and since 2008, the DGFiP. In accordance with the tax system, the data from year N cover the incomes of year $N-1$[75] These datasets are publicly available for the fiscal years 2004 to 2020 on the dedicated website of the DGFiP. For the 1990-2015 period, we use the cleansed data provided by the statistical service ADISP. For 2016-2020, our R code directly downloads and cleans the data on the DGFiP dedicated webpage.

Bracket breakdowns

In the series of the DGFiP (that is, since 2001), the administration provides, for approximately 15% of the most populated communes, a detailed breakdown of the variables through a set of brackets of disposable income (see figure 20).

Table 20: Income brackets of the RC series (in current euros)

	2001	**2002-2003**	**2004-2006**	**2007-2010**	**2010-2020**
1	$\leq$ 7623	$\leq$ 9000	$\leq$ 7500	$\leq$ 9400	$\leq$ 10000
2	7624 - 10673	9001 - 12000	7501 - 9000	9401 - 11250	10001 - 12000
3	10673 - 15246	12001 - 19000	9001 - 10500	11251 - 13150	12001 - 15000
4	15246 - 22869	19001 - 31000	10501 - 12000	13151 - 15000	15001 - 20000
5	22869 - 38113	31001 - 78000	12001 - 13150	15001 - 16900	20001 - 30000
6	$\geq$ 38114	$\geq$ 78001	13151 - 15000	16901 - 18750	30001 - 50000
7			15001 - 19000	18751 - 23750	50001 - 100000
8			19001 - 23000	23751 - 28750	$\geq$ 100001
9			23001 - 31000	28751 -38750	
10			31001 - 39000	38751 - 48750	
11			39001 - 78000	48751 - 97500	
12			$\geq$ 78001	$\geq$ 97501	

For the RC files which are partially publicly available, that reporting is constrained by the standards of statistical privacy of the INSEE :

[73] Code général des impôts, art. 170, al. 1

[74] Not mentionned here is the CIMR variable: in 2019, when the tax administration transitioned to tax withholding (*prélèvement à la source*), a special tax credit, the CIMR (*Crédit d'impôt pour la modernisation du recouvrement*) was created. Consequently, for the tax year 2019, variable IS provides the net tax revenue without the CIMR, variable ISCIMR the same tax revenue once the CIMR is applied, and variable CIMR the global amount of the tax credit, such that : $\mathrm{IS}_i - \mathrm{CIMR}_i = \mathrm{ISCIMR}_i$

[75] There is consequently no break in the series ; the INSEE dataset for year 1999 computes the incomes of year 1999, the DGFiP dataset for year 2001, the incomes of year 2000.

Table 21: INSEE privacy rules applied to the IRCOM 1990-1999 and Filosofi datasets

Range	**Rule**
Fewer than 50 households or 100 persons	Absolute privacy (data publication forbidden)
Between 50-1000 households or 100-2000 persons	Limited authorisation (publication of the mean and median of variables)
More than 1000 households or 2000 persons	Bracket breakdown allowed for sub-populations above 11 households minimum

As opposed to the INSEE, the tax administration has less stringent reporting rules, most importantly, a minimum threshold of 11, not 50, households.

Restrictions

These rules of secrecy define the preliminary restriction on the RC dataset:

- Restriction zero (R0) : All cities for which the number of tax units and total fiscal income is provided. Excluded communes are those for which the administration has fewer than 11 declarations (or for which one person makes out 85% or more the whole fiscal income); they account for a microscopic part of the national population, far less than 0.1%;

Over the RC-IRCOM files, we apply three restrictions to the raw sample R0:

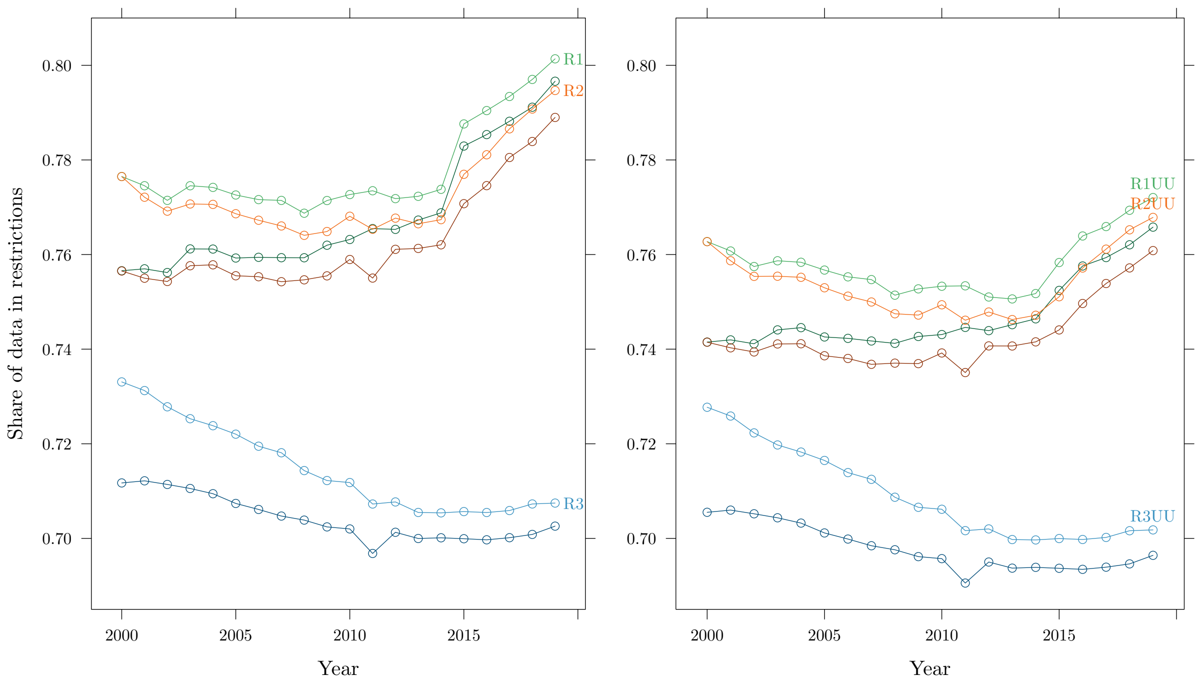
- 1^{st} restr. : Communes for which we have the average income plus data on at least one subbracket (R1 fluctuates over years, but never contains more than 18% of the communes);
- 2^{nd} rest. : Communes over which provided data are sufficient to apply the *gpinter* interpolating algorithm built by [Blanchet, Fournier, and Piketty 2017]. Before we proceed with the interpolating, we want to exclude these observations for which data reported is too poor to reach accurate estimates. If we want to use the full version of the algorithm, which requires information, not only about fractiles and quantiles, but also about averages within each bracket, the interpolation algorithm must be provided with a continuous set starting from one of the tails of the distribution. For instance, for year 2001-2003, when there are 6 brackets, acceptable profiles include first two, first three, first four, first five, full, last five, last four, last three and last two brackets. It seems natural to exclude profiles with too little information. Then, for datasets with 6 or 8 brackets, we require a minimum of three intervals ; for datasets with 12, we require a minimum of 6;
- 3^{rd} rest. : Communes for which full interpolation with *gpinter* is possible over each year over 2000-2019.

Table 22: Number of communes within R0, R1 and R2 of the RC-IRCOM dataset

Year	ADISP index	Nb. obs.	No data	R1	R2	All brack.	Last miss.	2 last miss.	3 last miss.	First miss.	First 2 miss.	First 3 miss.
1990	lil-0276	36607	2333									
1991	lil-0276	36607	2317									
1992	lil-0276	36607	2310									
1993	lil-0276	36607	2158									
1994	lil-0277	36602	2078									
1995	lil-0278	36603	1940									
1996	lil-0279	36564	1992									
1997	lil-0280	36599	1794									
1998	lil-0193	36607	1560									
1999	lil-0281	36597	1542									
2001	lil-0566	36598	131	4642	4639	4632	0	7	0	0	0	0
2002	lil-0567	36601	130	4641	4591	3186	9	1396	0	0	0	0
2003	lil-0568	36000	118	4641	4603	3405	4	1193	1	0	0	0
2004	lil-0569	36603	114	4828	4728	3740	2	750	236	0	0	0
2005	lil-0570	36605	107	4864	4767	3800	6	724	237	0	0	0
2006	lil-0571	36606	108	4864	4760	3818	5	738	199	0	0	0
2007	lil-0572	36604	115	4905	4791	3897	6	704	184	0	0	0
2008	lil-0573	36604	104	4945	4814	4104	1	578	131	0	0	0
2009	lil-0574	36605	108	4978	4850	4188	6	537	119	0	0	0
2010	lil-0630	36605	110	5092	4923	4217	4	581	121	0	0	0
2011	lil-0751	36605	106	5141	5022	4214	1	807	0	0	0	0
2012	lil-0913	36594	115	5186	5079	4413	2	662	0	0	2	0
2013	lil-0939	36594	129	5223	5106	4538	2	566	0	0	0	0
2014	lil-0998	36587	124	5287	5125	4518	4	603	0	0	0	0
2015	lil-1102	36578	119	5331	5145	4576	3	565	0	1	0	0
2016	ircom2016	36603	100	5804	5460	4372	0	647	0	0	414	27
2017	ircom2017	35554	105	5688	5399	4395	0	525	0	0	448	31
2018	ircom2018	35426	98	5755	5542	4851	0	529	2	89	12	0
2019	ircom2019	35267	98	5816	5616	5145	0	431	0	0	32	8
2020	ircom2020	35007	91	5923	5708	5183	0	455	0	0	53	17

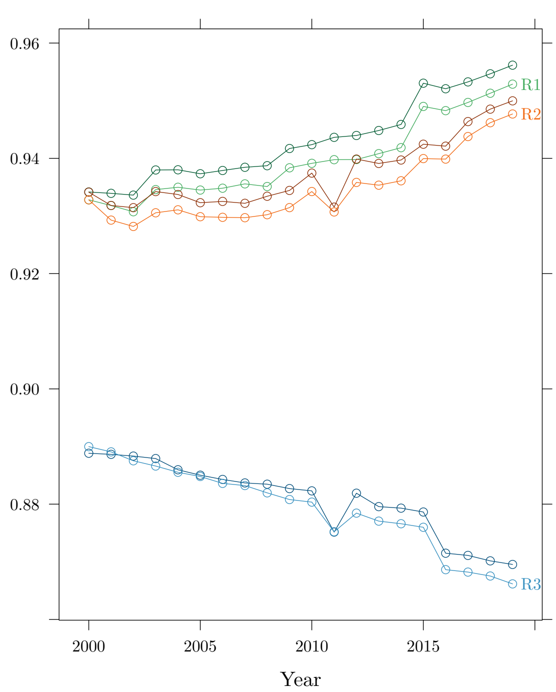
The 3^{rd} restriction includes 3859 communes ; the 2^{nd} one fluctuates around 4200 communes, that is about 15% of the total number of cities. In spite of the implicit bias, we'll still be using the RC files, but restricted to these cities which are part of what the INSEE calls a urban unit (UU)[76]. That restriction can be seen as exemplifying urban life and economic dynamics. With such a choice, as illustrated in figure 37, we lose almost no data on R3. On R2, we lose not more than 1.5% of the national population.

Figure 37: Share of household population (dark shades) and fiscal income (light s.) within restrictions (as % of R0)



In our UU restrictions, the selection bias is considerably reduced. We retain on R2 at least 93.3% of the national population living on a UU, and 93.2% of the fiscal income[77].

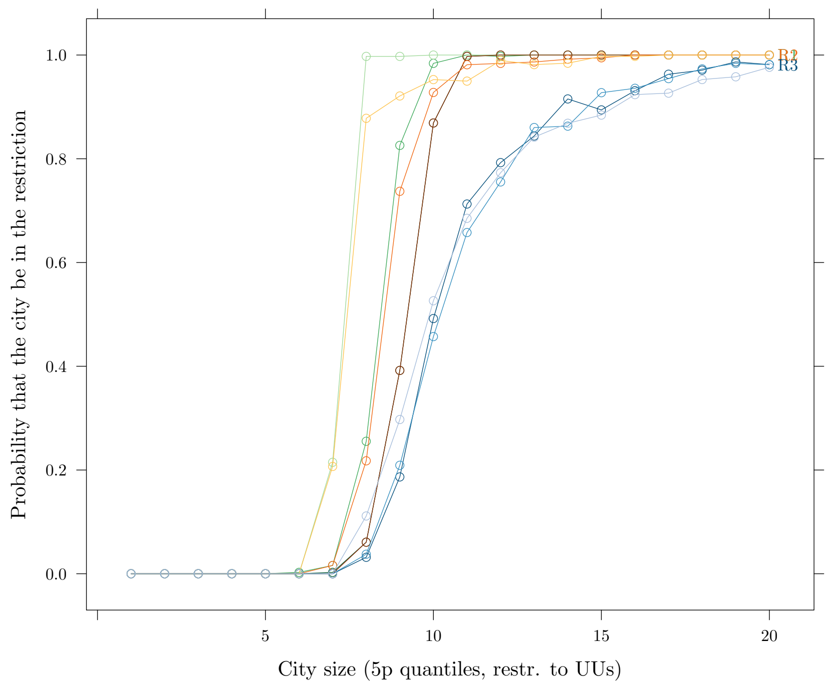
Figure 38: Shares of data within UU restrictions (as % of R0-UU)



For cities which do not lie in a UU, the probability to be even in the 1^{st} restriction is almost null (respectively 1.1, 1.7 and 3.2% for 2000, 2010 and 2019). That probability jumps brutally to 0.9 or more after a threshold, which, depending on restrictions and years, fluctuates around 1500 tax units in the city. With its extra rules on time consistency and merging, the 3^rd restriction tends to focus on larger cities.

[76]Defined as a homogeneous urban zone of 2000 inhabitants at least with no gap between constructions of more than 200 meters.
[77]Compared to the national total, cities in our UU restrictions are richer; that is why, on a R0 basis, shares of income within restrictions are always above shares of population. On a R0-UU basis, it is the opposite, simply because littler, out-of-restriction UUs are often homogeneous middle-class suburban centres which tend to have slightly higher average incomes.

Figure 39: City size and probability to be in a restriction (2000, 2010 and 2019, lighter shades for recent years)



When we have recourse only to data from the three years (most notably, 2000, 2008 and 2018), we apply a less stringent restriction. This relaxed restriction concatenating data for 2000, 2010 and 2019 is defined as R4 and contains around 4100 communes.

Among our different restrictions, we were forced to remove more than 85% of French communes. An interesting approach might be to import from the INSEE general socio-economic data about French communes, to draw a relatively subtle portrait of the communes which happen to end inside or outside our restriction.

Table 23: Socio-economic portrait of communes within and outside of the 3rd restriction

Variable	**Weighted mean (out of 3rd restr.)**	**Weighted mean (within 3rd restr.)**
Growth of the population (2013-2018)	0.29	0.447
... of which share due to natural increase	0.199	0.386
... of which share due to migrations	0.093	0.06
Share of 1-person household (2018)	31.44	38.084
Share of foreigners (2018)	4.38	8.62
Share of immigrants (2018)	6.3	11.67
Share of retired people (2018)	28.35	25.82
Unemployment level (2018)	10.63	15.27
Share of insecure jobs (2018)	14.48	16.56
Share of services in employment (2018)	67.11	80.87
Share of working-class (INSEE-CSPs workers + employees)	28.77	27.97
Participation to the 2017 presidential election	82.6	76.51
M. Le Pen vote at the 2017 presidential election (1^{st} rnd)	25.9	20.2

Unsurprisingly, communes picked by the tax administration are major unites located not far from a metropolis. Hence a population which is overall younger, which higher share of immigrants and foreigners, and a steadier growth of the population. Note however that migration explains a higher share of the population's growth out of the restriction (surely because of people leaving city centres). As to economic variables, activity inside our restriction is based mainly on services, hence a lower share of working-class persons, but also far higher shares of insecure menial jobs, and a higher level of unemployment. Over each variable, standard deviation is larger within the restriction.

There is no clear connection between the size of the commune and the average taxable income of its inhabitants (regressing the former on the latter gives a R^2 which, over years, is never above 1%. As a consequence, the proportion of population and fiscal income which lies inside our 3^{rd} restriction is not fundamentally altered between 2000 and 2019, ranging between 68 and 73%.

Particulars

City districts ("arrondissements") are usually provided for the three major cities (Paris, Lyon, Marseille), but data is missing for the year 1996 (for which we only have data for the whole city). In almost all specifications and figures, our default option is to merge *arrondissements* together.

Mayotte had to be excluded from the restrictions, since data are missing for 2005-2011.

A concern is raised by the multiple mergers between communes that happened over that period. Over 2000-2019, according to the geographical COG system of the INSEE, 3769 such events were recorded. We used the web scraping tools of R to pick on the COG webpage of the INSEE the codes of all communes for which there was a merger, an absorption, or another event, comparing it with our restricted dataset. We identified 265 pertinent events. We then apply the following rule : if there is no year-to-year rise of the taxpaying population of more than 5% over the whole period, we keep the observation. If on the contrary such a rise occurs, we delete it. Table ?? indicates the corresponding changes applied:

Table 24: Merged cities in the RC-IRCOM dataset

Year	Code	Event	Data	Share	Rule
2018	22050	Léhon (22123) merges into Dinan (22050)	No missing data	0 %	Merged
2016	35176	Messac (35176) absorbs Guipry (35129) to become Guipry-Messac	No data missing	0%	Merged
2019	44003	Ancenis (44003) and Saint-Géréon (44160) merge into Ancenis-Saint-Géréon (44003)	No missing data	0 %	Merged
2019	50025	Saint-Martin-des-Champs (50516) merges into Avranches (50025)	No missing data	0 %	Merged
2010	59183	Saint-Pol-sur-Mer (59540) Fort-Mardyck (59248) and Mardyck (59380) merge into Dunkerque (59183)	59380 missing	4.2 %	Merged
2017	74010	Annecy-le-Vieux (74011) Cran-Gevrier (74093) Meythet (74182) Pringy (74217) and Seynod (74268) merge into Annecy (74010)	No missing data	0 %	Merged
2019	78158	Le Chesnay (78158) and Rocquencourt (78524) merge into Chesnay-Rocquencourt (78158)	No missing data	0 %	Merged
2019	78551	Fourqueux (78251) merges into Saint-Germain-en-Laye (78551)	No missing data	0 %	Merged
2019	85194	Château d'Olonne (85060) and Olonne-sur-Mer (85166) merge into Les Sables d'Olonne (85194)	No missing data	0 %	Merged
2019	91228	Courcouronnes (91182) and Evry (91228) merge into Evry-Courcouronnes (91228)	No missing data	0 %	Merged

Filosofi database (*Dispositif sur les revenus localisés fiscaux et sociaux*)

The Filosofi database from the INSEE (formerly labeled, before 2011, RFL-RDL) provides, since the civil year 2000, detailed analysis of the income dynamics at a territorial level. Its main differences with the RC database might be summarised as such:

- Information is provided for infra-city level, with a special territorial division known as the IRIS. It was built by the INSEE to become the French equivalent of the US census tract. All communes with more than 10000 inhabitants, and most communes between 5000 and 10000, were split into little districts, with an average of 2500 inhabitants. Each remaining commune was artificially transformed into an IRIS;
- Income dynamics are provided for the household, the individual, and the *unité de consommation* (the first adult of the household is counted as one UC, any other person above 14 y.o. as 0.5, any person below 14 as 0.3). In this thesis, we systematically use the household as the pertinent level;
- The two main income statistics provided are the disposable income (*revenu disponible*) and the recorded income (*revenu déclaré*). The recorded income is simply the taxable income as reported in the ERFS database (see below), the INSEE applying some corrections, most notably to take into account capital gains. Disposable income is the post-tax post-redistribution living standard of an household; formally, it is the sum of all sources of income of the household (labour income, capital income, but also transfers including unemployment benefits and pensions) minus the burden of the income tax, the SSCs, and the *CSG-CRDS*;
- Over the FILOSOFI dataset, R1 = R2, i.e. there is no partial reporting; information is provided either for every single decile, or for none of them. In 2001 for instance, the whole dataset covers 22 million households or 56 million persons, deciles being provided for IRISes which cover 74% of the whole household population, and 73% of the individual population. Fiscal information is extremely limited for the missing communes (the household mean and median are rarely provided), and since the reporting rules of the INSEE are more stringent, data covers a more limited share of household population;
- The main drawback of that dataset lies in the fact that bracket averages are not provided ; interpolation with *gpinter* must rely only on the value of deciles for each IRIS, making the estimation much more imprecise, imprecision which is doomed to be magnified by the thinness of the geographical subdivision.

For the decade 2000-2010, our restriction (interpolation is possible for both years) is limited to 14331 IRISs (one third of the total) covering 70.5% of the individual population.

Table 25: Share of the household population within restrictions, depending on the unit of interest

	R1	R1-AAV	R1-Metropolis AAV	R1-UU	Paris-Lyon-Marseille AAVs
FILOSOFI 2001 (2^{nd} res.)	74% of 22m.	79% of 21m.	89% of 11m.	92% of 18m.	92% of 6m.
RC 2001 (2^{nd} res.)	76% of 32m.	80% of 31m.	90% of 14.9m.	93% 26m.	95% of 8.4m.

A strategy to infer inequality variables from the INSEE's Census

There's something frustrating in the idea of limiting our analysis to the last two decades only because most of the datasets do not provide the local structure of income prior to the 1990s. One possible strategy is to predict the level of inequality, proxied with the T10/B50 ratio, using the sole data from the Census. Contrary to the U.S. Census, the INSEE never recorded individual wages or earnings. We tried different reconstruction strategies before choosing our final option.

One possible strategy relies on the average income of each SES status. A large dataset on wages, the DADS (*Déclaration annuelle des données sociales*), initiated in 1950, provides us with the average wage for different types of jobs (blue-collar, menial, etc.). Aggregated data exist in the historical archives of the INSEE [Bayet and Julhès April 1996]. It is then possible to impute to each SES category his average income, and to compute the ratio of the top 10% versus the bottom 50% using this structure. We tested this method on the most recent Censuses (1999,2010,2017) and it actually yields extremely poor results, with a correlation between predicted and actual values below 2%.

Another option would be to fit on recent data a linear model which predicts the ratio T10/B50 from usual Census variables like the demographic structure, the SES shares, etc. We tested a simply strategy relying on the INSEE's CPS scale. When estimated on three recent issues of the Census (1999,2010,2017), the correlation between these predicted values and the one extracted from the IRCOM dataset with *gpinter*, over R3, oscillates around 60%. However, projecting it over past values yields estimates which are clearly at variance with national data available for these past decades ; most importantly, the national ratio is largely overestimated, since even the inclusion of ratio variables fails to take into account the evolution of inequalities across SES statuses (the average wage of a white-collar workers was 2.84 times the average wage of a menial worker in 1950, compared with 2.71 in 1994) [Seys 1996].

The best strategy we were able to work out relies on a special series of the INSEE called the ERF (*Enquête revenus fiscaux*). Since 1962, it draws from administrative data a large sample of taxpayers to analyse their socio-economic and demographic characteristics. The most helpful piece of data are tables which report, for each SES status, the distribution of the group's population along the income scale (in the form of generally 8 to 10 quantiles of income, and detailed information about how many tax units of each SES lies within these quantiles). If we assume that each SES within each zone has the same distribution of income than the nationwide SES, we can interpolate each city's income distribution.

Such a test is possible for recent data, for which we can compare our interpolation and the real values provided by the tax administration (the IRCOM base). Let us take the most recent data (i.e. the Census 2018, the ERF 2018, and the IRCOM 218). The raw correlations displayed in table 26 are encouraging:

Table 26: Comparing interpolated and real values (Census 2018 / ERF 2018 / IRCOM 2018) I

Variable	Comparison method	Outcome
Average income of the top 10% within the city	City by city raw correlation	71.7%
Average income of the bottom 50% within the city	City by city raw correlation	79.7%
Ratio T10/B50 within the city	City by city raw correlation	39.1%
Average income within the electoral district	District by district raw correlation	88.1%
Average income of the top 10% within the electoral district	District by district raw correlation	81.5%
Average income of the bottom 50% within the district	District by district raw correlation	85.4%
Ratio T10/B50 within the district	District by district raw correlation	61.9%
Average within-city Ratio T10/B50 within the district	District by district raw correlation	58.5%

There's almost perfect linear fit between the incomes interpolated and the real ones, but the variance of interpolated values is abnormally low; i.e. we are able to reconstruct the ranking of cities or districts along our preferred variable; we obtain very consistent national averages, but the reconstructed distribution of cities or district has abnormaly tight tails. When then proceed with a back-of-the-envelope calculation, which consists in normalising the variable interpolated from the Census, adding to it the national mean provided by [Garbinti, Goupille-Lebret, and Piketty 2018]. Besides, since incomes for Paris are specifically reported in almost every issue of the ERF, we are at liberty to treat Paris as the virtual top fractile (generally the top 3%) of the distribution, which provides us with a proxy of the corresponding standard deviation of the national distribution, provided that the dominance of Paris within the national structure of income has remained relatively unaltered over the last semi-century [Piketty, Postel-Vinay, and Rosenthal 2006; Piketty 2011 Piketty, Postel-Vinay, and Rosenthal 2018].

Table 27: Comparing interpolated and real values (Census 2018 / ERF 2018 / IRCOM 2018) II

Variable	Interpolated	Real
Average adult income (nationwide level in euros)	38,249	37,757
Average income of the top 10% (nationwide level in euros)	111,803	120.855
Average income of the bottom 50% (nationwide level in euros)	19,112	17,227

We rely on the successive versions of the ERF over 1965-1990. One must heed to the fact that data is much more detailed for the earlier issues of the study. Besides, in the 1960s, the SES scale provided a very good proxy of the income distribution; farmworkers were found in deciles 1 and 2, farmers around decile 3, blue-collars around deciles 4 and 5, employees around decile 5 and 6, and so on.

Table 28: Distribution of taxable income per socio-economic status in 1965 (ERF 1965, reconstructed from [Banderier 1970])

Code	CPS name	Taxable income (in thousands of French Francs)									
		3 or less	3 to 6.5	6.5 to 10	10 to 15	15 to 20	20 to 30	30 to 60	60 to 100	100+	Total
0	Farmers	42.9	30	12.5	7.7	3	2.1	1.4	0.4	0	100
1	Farmworkers	14.9	37.9	26.3	12.3	5.4	2.8	0.4	0	0	100
2	Self-employed	4	8.7	12.9	18.4	14.6	17.5	17	4.7	2.2	100
	Bosses & Big shopkeepers	2.1	4.5	5.4	6.6	7.5	20.2	39.5	12.4	1.8	100
	Craftsmen & Small shopkeepers	4.3	9.4	14.2	20.4	15.8	17	13.2	3.4	2.3	100
3	White-collar	0.2	0.4	1.1	3.5	7.9	24.8	45.7	11.8	4.6	100
	Professional services	0.2	0.4	1	3.4	7.5	17.1	39.2	16.7	14.6	100
	Professors, engineers & managers	0.2	0.4	1.1	3.5	8.1	26.3	46.9	10.8	2.7	100
4	Intermediate	0.6	1.6	5.7	21	23.3	29	16.8	1.6	0.4	100
	Public services	0.8	1.9	5.8	25.6	28.3	23.9	12.9	0.7	0.1	100
	Technicians & grey-collar	0.4	1.4	5.5	18.3	20.5	31.9	19.1	2.1	0.8	100
5	Employees-Menial jobs	3.3	8.1	21.4	29.9	18	14.9	4.3	0.1	0	100
6	Blue-collar	2.9	11.4	25.4	31.4	16.1	10.6	2.2	0	0	100
	Skilled workers	1.8	9.2	18.9	34.9	19.1	12.6	3.1	0	0	100
	Unskilled workers	3.9	13.1	30.3	28.7	13.5	9	1.5	0	0	100
7	*Maids & Domestics*	22.6	44.1	16.7	8.2	3.4	2.8	2.1	0.1	0	100
8	Others										
9	Out of the workforce	29.1	28.6	15.5	11.9	6.1	5.5	2.8	0.4	0.1	100
	Retired farmers	51.6	30.2	7.6	4.3	2	1.9	1.7	0.4	0.3	100
	Retired self-employed	29.9	27.5	14.8	9.9	5.6	5.4	4.7	1.4	0.9	100
	Retired civil servants	12.8	19.8	23.7	18.9	8.4	7.4	5.7	1.7	1.6	100
	Retired wage earners	23.4	30.3	17.3	10.3	5.7	5.1	4.4	2.1	1.4	100
	Others	44.1	29.5	10.2	7.1	3.6	3.2	1.9	0.3	0.1	100
	National distr. fractile	13.5	29.5	46.3	66.2	78.8	90.8	98.2	99.5	100	

Note: [Banderier 1970] provides, for 7 out of 9 main socioeconomic statuses of the INSEE's CPS scale, the distribution of taxable income (8 pairs of fractiles-quantiles or 9 brackets). For the remaining 2 main SES and for all sub-statuses of the scale, he only provides the two quartiles and the median (I italicised the categories for which we are only provided with this limited data). In order to build a consistent dataset, when 3 quantiles only are provided, we interpolate the whole distribution to deduce the 8 quantiles of the main SES. Data from italicised categories are therefore a reconstruction. For sub-statuses, we corrected the results of the interpolation to make it fit the raw data.

Table 29: Distribution of taxable income per socio-economic status in 1970 (ERF 1970, reconstructed from [Banderier and Ghigliazza 1974])

Code	CPS name	Taxable income (in thousands of French Francs)									
		3 or less	3 to 6.5	6.5 to 10	10 to 15	15 to 20	20 to 30	30 to 60	60 to 100	100+	Total
0	Farmers	21.1	26.3	16.1	14.3	8.1	7.1	5.6	1	0.4	100
1	Farmworkers	2.2	14	29.2	27.8	12.2	10.6	4	0	0	100
2	Self-employed	2.1	3.8	6.2	11.9	12.3	20.5	28	9.2	6	100
	Bosses & Big shopkeepers	0.7	1.2	1.9	0.8	0.6	1.2	34.2	33	26.5	100
	Craftsmen & Small shopkeepers	2.3	4	6.5	13.6	14.2	23.7	27.5	5.3	3	100
3	White-collar	0	0.3	0.4	1.2	2.3	11	50.9	26.2	7.7	100
	Professional services	0.3	0.4	0.5	1	1.6	6.6	29.1	37.3	23.3	100
	Professors, engineers & managers	0.1	0.3	0.4	1.2	2.2	10.5	54.1	24.6	6.5	100
4	Intermediate	0.1	0.6	1.2	5.9	12.5	32.3	41.8	4.7	0.9	100
	Public services	0.4	0.8	1.3	5	17.4	38.6	32.3	3.3	0.8	100
	Technicians & grey-collar	0.3	0.7	1.2	4.8	4.6	29.9	52.7	4.7	1	100
5	Employees-Menial jobs	1.3	3.1	6.2	20.8	19.9	27.3	19.5	1.7	0.2	100
6	Blue-collar	0.9	3.6	10.3	24.8	21.7	25.6	12.8	0.3	0	100
	Skilled workers	0.9	3.1	9.1	22.4	23.8	26.8	13.4	0.3	0.1	100
	Unskilled workers	1.3	8.8	15.9	29.7	20	15.4	8.6	0.3	0.1	100
7	*Maids & Domestics*	8.4	9.8	16.6	29.2	13.2	10.8	8.1	2.2	1.7	100
8	Others										
9	Out of the workforce	13.1	23	17.5	17.4	10.3	9.9	7.5	1	0.3	100
	Retired farmers	29	35.6	13.9	8.1	3.8	3.6	3.3	1.5	1.2	100
	Retired self-employed	16.9	21.2	16.2	16.3	7.5	7.4	7.4	4.2	2.9	100
	Retired civil servants	7.9	9.2	10.9	22	18.3	14.6	11.3	3.2	2.6	100
	Retired wage earners	14.9	20.3	18.7	19.5	8.3	7.6	6.4	2.2	2.1	100
	Others	21.7	28.3	15.9	11.9	6.5	6.3	6.1	1.9	1.4	100
	National distr. fractile	6	16.6	27.4	44.5	58.8	78.1	95.8	98.9	100	

Note: [Banderier and Ghigliazza 1974] provide, for 7 out of 9 main socioeconomic statuses of the INSEE's CPS scale, the distribution of taxable income (8 pairs of fractiles-quantiles or 9 brackets). For the remaining 2 main SES and for all sub-statuses of the scale, he only provides the two quartiles and the median (I italicised the categories for which we are only provided with this limited data). In order to build a consistent dataset, when 3 quantiles only are provided, we interpolate the whole distribution to deduce the 8 quantiles of the main SES. Data from italicised categories are therefore a reconstruction. For sub-statuses, we corrected the results of the interpolation to make it fit the raw data.

Figure 40: Distribution of taxable income per socio-economic status in 1975 (ERF 1975, quoted from [Canceill, Chastand, and Choquet 1981])

Catégorie socioprofessionnelle du chef de famille	Inférieur à 15 000 F	15 000 à 20 000 F	20 000 à 30 000 F	30 000 à 40 000 F	40 000 à 55 000 F	55 000 à 70 000 F	70 000 à 100 000 F	100 000 à 200 000 F	Supérieur à 200 000 F
Agriculteur exploitant.	34,0	15,2	23,1	13,8	8,0	3,1	1,9	ε	ε
Salarié agricole	6,2	12,2	31,0	24,3	16,7	5,9	3,7	0	0
Artisan, petit commerçant, patron pêcheur............	5,7	6,0	16,4	19,4	22,8	12,7	10,1	5,8	1,1
Industriel, gros commerçant...........	3,8	1,5	4,4	6,6	12,0	8,8	24,1	31,0	7,8
Profession libérale...	2,5	2,2	6,6	5,0	11,9	12,9	20,1	31,7	7,1
Cadre supérieur.....	ε	ε	1,2	3,1	11,6	19,0	32,1	28,3	3,5
Cadre moyen.........	ε	ε	4,4	14,3	32,3	24,7	18,3	4,3	ε
Employé............	1,7	2,6	14,6	24,5	31,9	14,5	8,5	1,5	ε
Ouvrier qualifié......	1,3	2,7	16,5	30,7	31,9	11,1	5,2	ε	ε
Ouvrier non qualifié.	3,5	5,4	23,0	28,8	27,7	8,6	2,3	ε	0
Inactif..............	8,5	24,8	26,1	16,4	12,4	6,0	3,8	1,8	ε
Ensemble....	**5,3**	**8,3**	**16,8**	**19,9**	**23,1**	**12,1**	**9,2**	**4,4**	ε

NOTE. — ε = pourcentage inférieur à 1 %.

Champ : Familles. Cf. annexe III.
Source : Enquête « Revenus fiscaux des ménages de 1975 ».

Figure 41: Distribution of taxable income per socio-economic status in 1979 (ERF 1979, quoted from [INSEE 1983])

Tranches de revenus (en francs)	Agriculteurs exploitants	Salariés agricoles	Professions indépendantes	Cadres supérieurs	Cadres moyens	Employés	Ouvriers	Inactifs
Moins de 15.000 (y compris déficits).....	18,2	6,7	3,1	1,4	1,2	2,6	3,4	9,0
De 15.000 à moins de 25.000...........	18,1	19,1	4,7	0,4	1,2	4,6	5,2	18,7
De 25.000 à moins de 35.000...........	17,3	23,9	7,9	0,7	2,7	11,7	14,8	21,8
De 35.000 à moins de 45.000...........	10,5	18,6	9,2	1,6	6,3	17,8	16,9	12,9
De 45.000 à moins de 60.000...........	12,6	12,4	12,6	4,7	18,0	18,9	21,8	13,5
De 60.000 à moins de 80.000...........	10,2	7,8	16,1	11,3	21,4	19,5	20,6	10,5
De 80.000 à moins de 100.000..........	4,9	7,3	10,8	13,8	20,5	12,8	10,4	5,7
De 100.000 à moins de 150.000.........	4,8	3,8	14,5	34,0	22,9	9,7	6,3	5,8
De 150.000 à moins de 200.000.........	1,7	0,4	8,2	17,0	4,1	1,9	0,5	1,3
200.000 et plus..........................	1,7	0,0	12,9	15,1	1,7	0,5	0,1	0,8
Total..............................	100	100	100	100	100	100	100	100

Figure 42: Distribution of taxable income per socio-economic status in 1984 (ERF 1984, quoted from [INSEE 1986])

Tranches de revenus (en France)	Agriculteurs exploitants	Professions indépendantes	Cadres supérieurs	Professions intermédiaires	Employés	Ouvriers	Inactifs
Moins de 15.000 (y compris déficits)	9,6	4,7	0,9	0,8	1,7	1,9	4,0
De 15.000 à moins de 25.000..................................	5,3	2,0	0,5	0,7	2,0	1,4	1,9
De 25.000 à moins de 35.000..................................	8,2	2,2	0,5	0,7	2,7	2,5	14,3
De 35.000 à moins de 45.000..................................	10,0	3,6	0,8	1,0	4,1	4,7	8,9
De 45.000 à moins de 60.000..................................	13,5	7,2	1,3	3,6	11,5	14,1	19,1
De 60.000 à moins de 80.000..................................	14,2	9,1	1,9	9,0	20,0	19,2	15,0
De 80.000 à moins de 100.000.................................	10,8	8,8	3,3	11,2	14,6	16,5	10,9
De 100.000 à moins de 150.000	14,6	19,2	18,3	30,7	26,3	28,2	15,5
De 150.000 à moins de 200.000	5,7	12,5	20,1	23,5	11,5	8,8	5,6
200.000 et plus...	8,1	30,7	52,4	18,8	5,6	2,7	4,8
Total...	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**

Figure 43: SES tax. inc.str. (ERF 1990, [Campagne, Contencin, and Roineau 1996])

	< 0	40 000	60 000	80 000	100 000	125 000	150 000	200 000	300 000	500 000	>>>>>>	ENSEMBLE
AGRICULTEUR EXPLOITANT	18 257	67 348	64 406	61 301	65 227	62 156	51 217	74 285	54 902	29 672	18 197	566 968
ARTISAN, COMMERCANT, CHEF ENTR	16 241	78 645	63 825	88 344	88 243	135 415	115 513	174 411	217 229	164 300	99 269	1 241 435
PROFESSION NON COMMERCIALE	1 389	7 392	5 105	6 663	5 513	14 588	12 300	31 921	45 014	82 353	93 432	305 670
CADRE SUPERIEUR	4 990	13 264	23 887	22 613	45 977	77 465	91 340	257 243	574 678	481 580	124 983	1 718 020
PROFESSION INTERMEDIAIRE	.	51 180	62 684	102 710	221 415	361 864	370 079	690 722	751 525	157 347	13 330	2 782 856
EMPLOYE	61	160 513	202 645	329 643	388 955	332 972	242 773	408 978	204 826	41 531	3 667	2 316 564
OUVRIER QUALIFIE	.	116 353	191 905	411 060	492 696	534 151	519 987	693 926	329 539	43 814	3 093	3 336 524
OUV NON QUALIFIE OU AGRICOLE	.	158 168	189 774	323 004	231 086	227 226	174 583	163 322	71 896	9 251	1 632	1 549 942
INACTIF	2 614	1 264 031	1 167 108	1 158 136	1 015 968	1 050 779	687 027	839 771	645 698	206 310	58 079	8 095 521
ENSEMBLE	43 552	1 916 894	1 971 339	2 503 474	2 555 080	2 796 616	2 264 819	3 334 579	2 895 307	1 216 158	415 682	21 913 500

Two types of robustness checks can be applied to test for the accuracy of our estimates:1. We can check if the reconstructed set fits the national structure of fiscal income provided by the World Inequality Database (which relies mostly on [Garbinti, Goupille-Lebret, and Piketty 2018]; 2. For the years 1990 to 1999, we can use the IRCOM set and compare the individual mean per city predicted by the interpolation, and the real values provided by the tax administration; 3. Finally, for every single year after 1999, we can check if our estimates fit the structure of income of each city or IRIS interpolated from the IRCOM or RFL sets.

Actually, it is important to heed to the fact that perfect comparison between these sources is impossible: 1. *Issues with the base unit* – Data from the Census are computed at the household level, data from the DGFiP, at the tax unit level, and the ERF has his own definition, nearer to the tax unit one. In 1990 for instance, there were 21.5 million households in the Census, 14.1 million tax units according to the DGFiP, and 14.4 million tax units in the ERF; 2. *Taxable vs. fiscal income* – The ERF (just as the IRCOM set) is expressed in taxable, not in fiscal income. It creates considerable inconsistencies, especially when it comes to the upper tail of the distribution. For instance, in 1990, according to [Campagne, Contencin, and Roineau 1996], the actual total amount of capital gains recorded by the national compatibility is 4.54 times the level estimated by the ERF; the ratio is 1.47 if we consider all types of income. We applied ratios from [Garbinti, Goupille-Lebret, and Piketty 2018] which should normally eliminate this bias ; yet we are forced to apply the same ratio within each city, which is all but realistic; 3. *Inconsistencies in the ERF* – The ERF is known to generally slightly underestimate the total taxable income (by about 4% in 1990) especially because the families at the very top of the distribution are underrepresented in the sample used for the estimations ; actually, it is both tails which are underrepresented (see table 3.7 in [Campagne, Contencin, and Roineau 1996]), implying that our interpolation with the ERF will consistently *underestimate* the incomes of the top 10% and *overestimate* the incomes of the bottom 50%.

To take the example of our first attempt, the Census year 1962, once the correction ratios from [Garbinti, Goupille-Lebret, and Piketty 2018] are applied, the ratio of the national structure of income is estimated to be 10.69 ; the weighted mean of the within-city ratio is 10.75 ; this is very close to the actual national values provided by [Garbinti, Goupille-Lebret, and Piketty 2018], namely 10.32 for 1962 and 10.89 for 1965.

Table 30: Comparing interpol. ratios T10/B50 and [Garbinti, Goupille-Lebret, and Piketty 2018]

		National ratio T10/B50	
Census year	ERF year	*Predicted*	*Real*
1962	1965	10.54	10.32
1968	1970		9.64
1975	1970	9.29	8.67
	1975	5.44	
1982	1975	6.65	7.01
	1979	6.02	7.01
	1984	7.005	7.01
1990	1984	7.52	8.03
	1990	6.69	8.03

Table 31: Final choice of estimation strategy

Election	Census year	ERF year
1965 (pres.)	1962	1965
1969 (pres.)	1975	1970
1974 (pres.)	1975	1970
1981 (pres.)	1982	1984
1988 (pres.)	1990	1990

Between 1990 and 1999, we can compare the values obtained from the Census+ERF interpolation and the values obtained from the IRCOM interpolation. For these years however, the IRCOM sets provide nothing more than the average taxable within each city, over a subsample which is large, but narrower that the full Census (for 1990: 35990 vs. 36594 communes, 12.7 million taxable units versus 14.4 in the ERF). There's a non-linear, almost logarithmic relation between the values predicted from the Census and the one provided by the IRCOM set; we tend to underestimate the average income in richer cities, and to overestimate it in poorer cities, making the *between* T10/B50 ratio far lower (1.45 vs. 2.19). The correlation between both values remains however relatively strong, above 55%.

Table 32: Comparing the interpolated values of 1990 (Census 1990 / ERF 1990) with the WID values

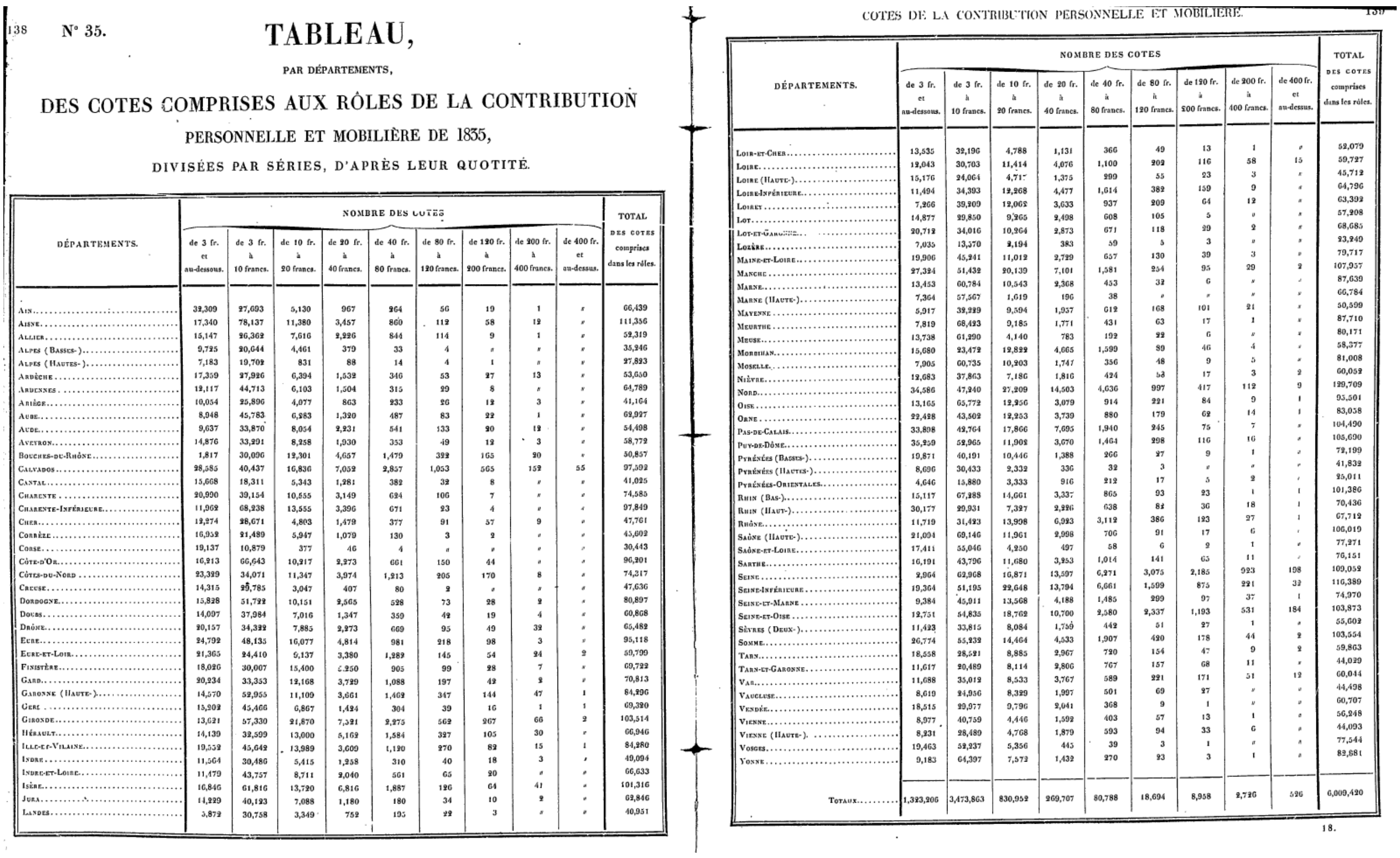
	Civil year 1990	
Variable	*Predicted*	*Real*
Average fiscal income for all households within the city (in 2021 euros)	31,088	31,170
Average fiscal income of the bottom 50% (in 2021 euros)	14,437	12,929
Average fiscal income of the top 10% (in 2021 euros)	96,644	103,783
Ratio T10/B50	6.69	8.03

Table 33: Comparing the interpolated values of 1990 (Census 1990 / ERF 1990) with the IRCOM set

	Set used for interpolation	
Variable	*Census+ERF*	*IRCOM*
Average fiscal income in the 10% richest cities (in 2021 euros)	50,719	58,136
Average fiscal income in the 50% poorest cities (in 2021 euros)	34.897	26,559
Ratio T10/B50 (*between* cities)	1.45	2.19

Pre-1945 Census data

Figure 44: Distribution of tax levies of the *Contribution personnelle et mobilière* in 1835 [Ministère des travaux publics 1837, p. 138]


138 N° 35.

TABLEAU,

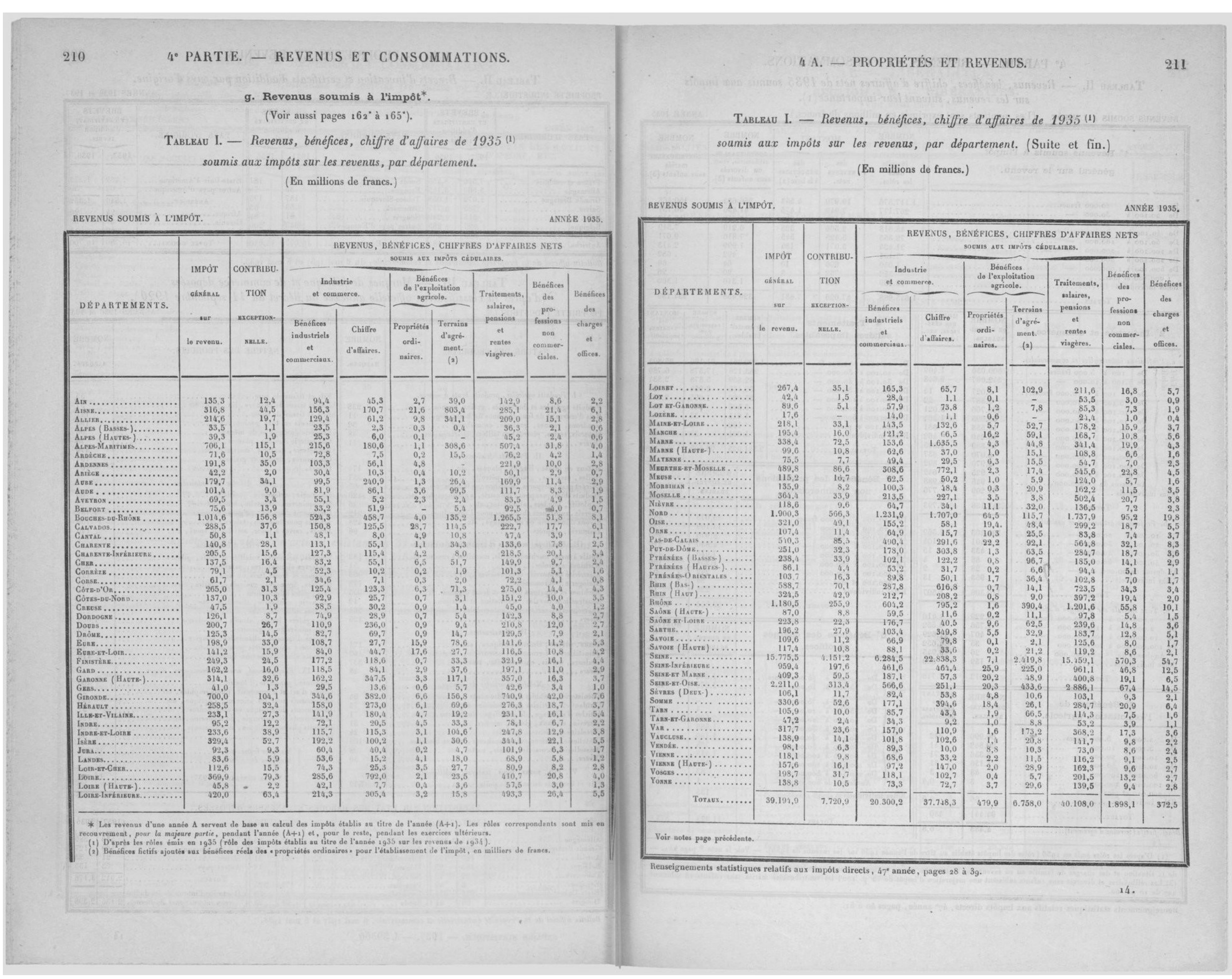
PAR DÉPARTEMENTS,

DES COTES COMPRISES AUX RÔLES DE LA CONTRIBUTION PERSONNELLE ET MOBILIÈRE DE 1835,

DIVISÉES PAR SÉRIES, D'APRÈS LEUR QUOTITÉ.

COTES DE LA CONTRIBUTION PERSONNELLE ET MOBILIÈRE. 139

Figure 45: Global amount of tax levies per *département* for some major contributions [Présidence du Conseil 1937]


210 4e PARTIE. — REVENUS ET CONSOMMATIONS.

g. Revenus soumis à l'impôt*.

(Voir aussi pages 162 à 165).

TABLEAU I. — *Revenus, bénéfices, chiffre d'affaires de 1935* (1) *soumis aux impôts sur les revenus, par département.*

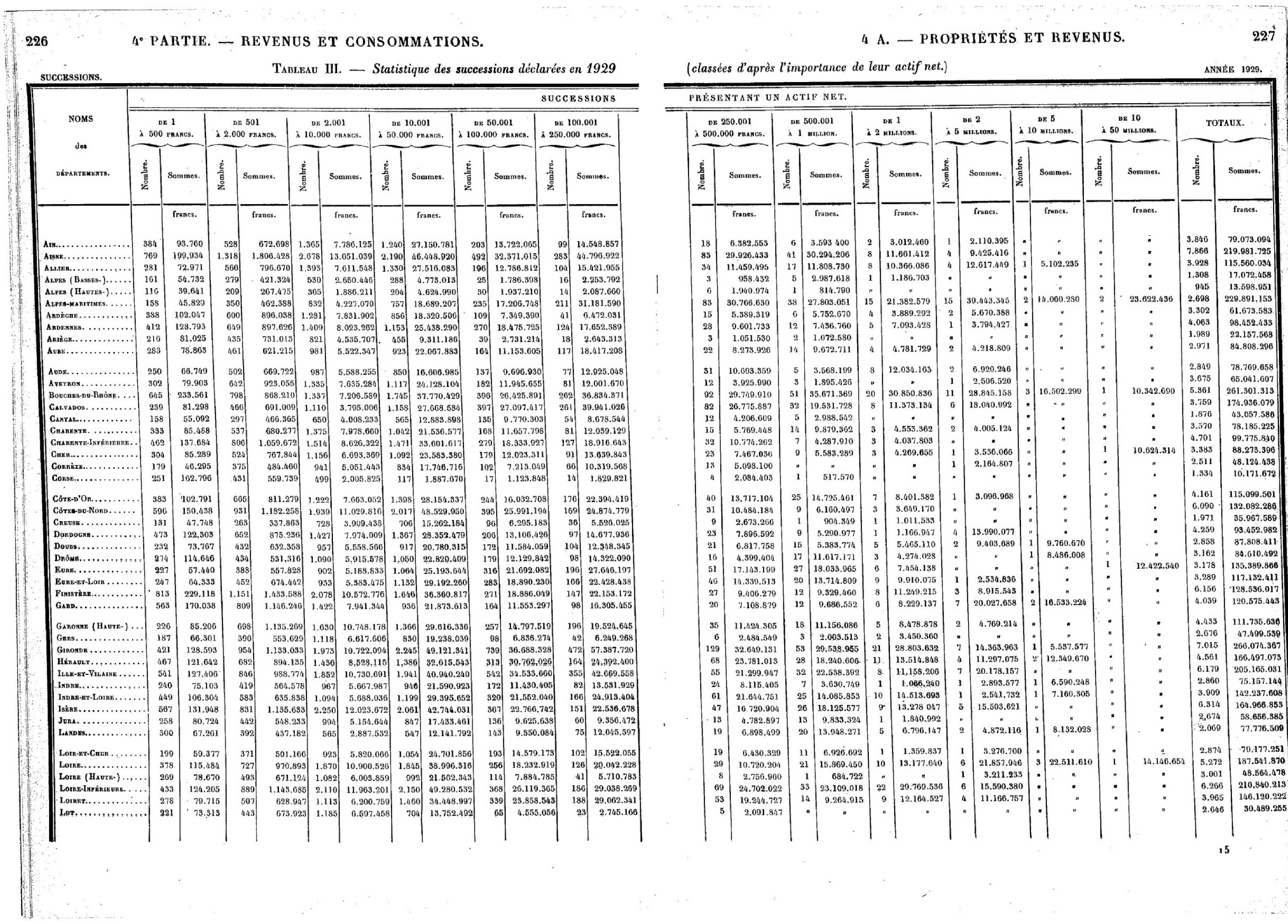
(En millions de francs.)

4 A. — PROPRIÉTÉS ET REVENUS. 211

TABLEAU I. — *Revenus, bénéfices, chiffre d'affaires de 1935* (1) *soumis aux impôts sur les revenus, par département.* (Suite et fin.)

(En millions de francs.)

Figure 46: Distribution of inheritances recorded in 1929 by the tax administration (I) [Présidence du Conseil 1930]



226 4e PARTIE. — REVENUS ET CONSOMMATIONS.

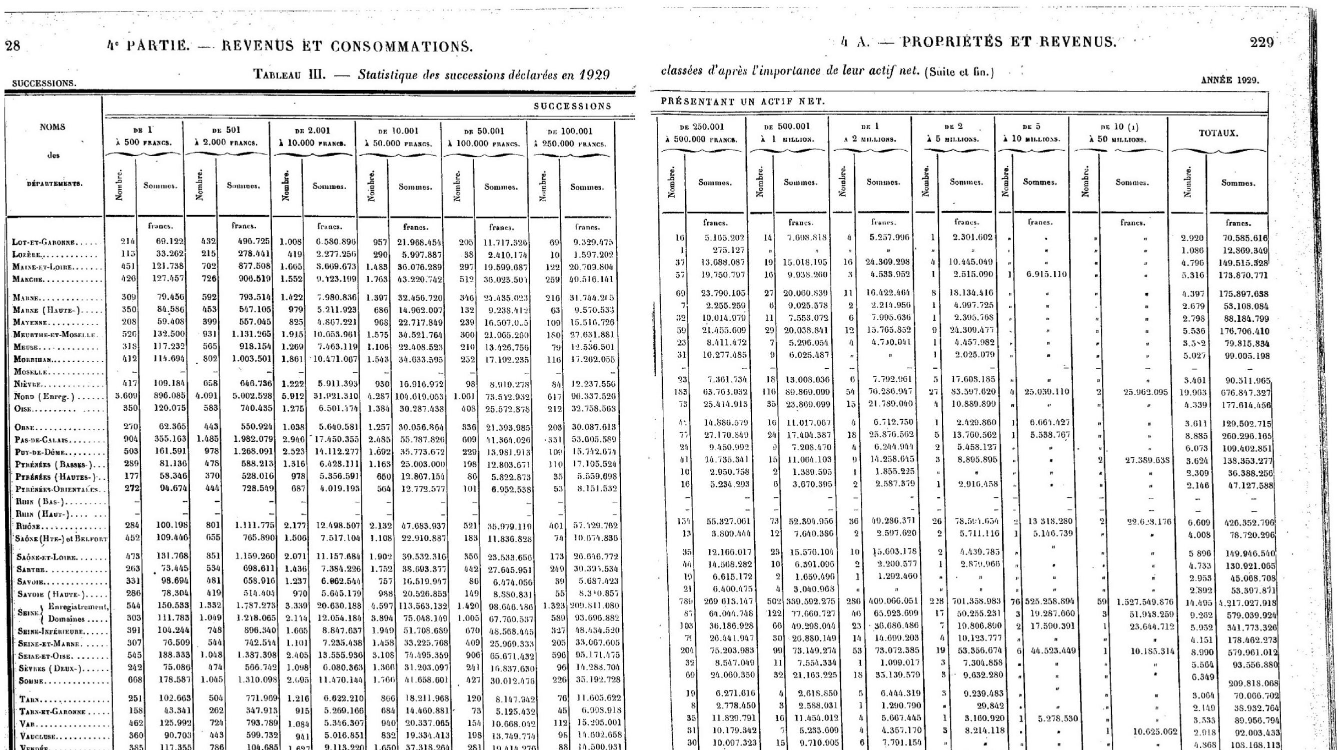
TABLEAU III. — *Statistique des successions déclarées en 1929*

4 A. — PROPRIÉTÉS ET REVENUS. 227

(classées d'après l'importance de leur actif net.)

ANNÉE 1929.

Figure 47: Distribution of inheritances recorded in 1929 by the tax administration (II) [Présidence du Conseil 1930]



228 4e PARTIE. — REVENUS ET CONSOMMATIONS.

TABLEAU III. — *Statistique des successions déclarées en 1929*

4 A. — PROPRIÉTÉS ET REVENUS. 229

classées d'après l'importance de leur actif net. (Suite et fin.)

ANNÉE 1929.

Bulletin de statistique et de législation comparée, juillet 1930.

B Corrections to income data

Following the standards of [Garbinti, Goupille-Lebret, and Piketty 2018], we're considering not the taxable income, but the fiscal income, i.e. the income before official pre-tax deductions (most importantly, the 20% extra tax allowance for wage earners which was scrapped in 2005). For the year-to-year ratios applied, we strictly follow the suggestions of Garbinti-Goupille-Lebret-Piketty ; their annex table TD4-section B, which recommends the following multipliers depending on taxable income fractiles:

Table 34: From taxable to fiscal income: upgrade rates by taxable income fractiles (for tax allowances and deductions)

Year	P0-90	P90-95	P95-99	P99-99.5	P99.5-99.9	P99.9-99.99	P99.99-100
1962-2005	1.43	1.43	1.39	1.33	1.25	1.19	1.11
After 2005	1.18	1.18	1.18	1.15	1.11	1.11	1.11

Applying that last rule is not trivial for the RC files, where bracket breakdown is not always available, and where bracket breaks do not follow precisely matched quantiles of the income distribution. We then follow that simple pattern:

- *Within* inequalities approach − To update information about the distribution of incomes within each city or zone, we first run the *gpinter* interpolation on available information on taxable income, and apply weighted upgrade rates afterwards on the average incomes of interest (after 2005, 1.18 for the bottom 50% and a weighted coefficient of 1.13 for the top 10% and 1.12 for the top 1% ; before 2005, respectively 1.43, 1.36 and 1.25); these rates are computed based on the national structure of incomes; it is one of the main biases of this reconstruction strategy: we assume that the national 10% richest living in the poorest *départements* have marginal capital gains comparable to the richest Parisian families, which is all but realistic[78];
- *Between* inequalities approach − When we are concerned only with city or district-level averages, we compute a weighted average upgrade rate based on the national structure of incomes (that average rate oscillates around 1.36 over 1962-1999, stabilising around 1.41 over 2000-2005, and around 1.17 after 2005). As to upgrade rates to account for capital gains at the top of the distribution, we also compute, for the top decile, a mean ratio based on the national structure of incomes (it is 1.058 in 1990, 1.045 in 2001, 1.036 in 2008 and 1.011 in 2014). Since over the R0 restrictions, we are most of the time not provided with within inequalities data, we are forced to apply the same ratio to all cities; for larger zones (like the *département* in the IRCOM base, or the *ZE* in the Filosofi base) we can account for some differences: i.e., we compute the share of the local population which belongs to the top 10%, and apply two different upgrade rates for these top 10% and the bottom 90%.

[78] That is why computation on RC files tend to slightly overestimate inequality in little cities, and underestimate it in larger metropolis

C Complements about income estimates

City size and average fiscal income within the city

We want to check whether city size is a significant predictor of income inequality (with, as proxy, the T10/B50 index, i.e. the ratio of the average fiscal income of the top 10% in the income distribution, over the same average, for the bottom 50%). Our main explanatory variables are the city size and the size of the UU (*unité urbaine* as defined by the INSEE) to which the city belongs. Our sample is drawn from the RC database, to which we apply our 3^{rd} restriction rule, plus a restriction to cities belonging to a UU. With such corrections applied, our sample consists in 3877 cities, or 23.3 million households, 89% of the population living in a UU or 71% of the national population (2000 levels). We gather the relevant control variables from the census of the INSEE, and in order to preserve consistency over time, we restrict ourselves to three years : 2000, 2010 and 2019.

We rely on the usual tests to choose between pooled OLS, fixed effects, & random effects. We start from a very simple model, with our dependent plus our two explanatory variables. With the Breusch-Pagan Lagrange multiplier test, we reject the null (p-value $= 2.2 \times 10^{-16}$). RE effects is to be preferred to pooled OLS. With the Fisher test to choose between FE and pooled OLS, we reject the null (p-value $= 2.2 \times 10^{-16}$). We can therefore definitively rule out the pooled OLS option. With the Hausman test to choose between FE and RE, we reject the null (p-value $= 7.5 \times 10^{-13}$). We therefore opt for a panel fixed effects model:

$$\text{RatioT10B50}_{c,u,t} = \beta_1(\text{Citysize}_{c,u,t} - \overline{\text{Citysize}}_{c,u}) + \beta_2(\text{UUsize}_{c,u,t} - \overline{\text{UUsize}}_{c,u}) + (X_{c,u,t} - \overline{X}_{u,c})\gamma + \varepsilon_{c,u,t} - \overline{\varepsilon}_{u,c} \quad (25)$$

Where $\text{RatioT10B50}_{c,u}$ is the aforementioned T10/B50 ratio of fiscal income, $\text{Citysize}_{c,u}$ is the city size computed as the number of tax units in commune c belonging to UU u; the population of the UU is likely computed as $\text{UUsize}_{c,u}$. Overlines indicate variable averaged over time. Finally, we add a vector of controls $X_{c,u}$ which includes population density, unemployment, share of working-class people (CSP− as defined by the INSEE) share of foreigners, share of people with no diploma, and share of vacant accommodation within each city. Standard errors are clustered at the city level.

Results are displayed in Table 35. In all but one of our specifications, coefficients on $Citysize$ and $UUsize$ are positive, 1% significant, and of considerable magnitude, whether they are estimated separately or combined. In our preferred specification (9), an extra 100.000 households within a city (or equivalently, 250.000 more inhabitants) should result in a rise of the T10/B50 ratio by 4.59 ; a similar rise within the urban unit brings an extra jump of 0.75 to that ratio. An interesting feature is that, in the combined specification (6), without controls, the coefficient on $Citysize$ is not significant ; it becomes so when we include socio-economic controls, especially the share of working-class people and the share of foreigners (the coefficient on both of these two variables being positive, 1% significant). We might interpret that as such: In descriptive stats, once controlled for individual city effects, larger cities are not more unequal, but become so when they belong to a large metropolis; but once controlled for SES shares, larger cities are found to be more unequal, and to become even more so when they belong to a large metropolis. Within large metropolis, selection effects and the polarisation of employment induces a concentration of workers at the tails of the income distribution, but also a greater spatial segregation. High-income and low-income families tend to gather in separated districts, and inequality within each of these districts remains artificially low. That artifact is removed once we introduce SES controls.

Table 35: Ratio T10/B50 vs city size and urban unit size

	OLS (cross-section, 2000 data)			Panel (fixed effects)					
				Dependent variable : ratio T10/B50					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
City size	3.5e−5***		3.3e−5***	2.93e−5***		−8.6e−7	5.28e−5***		4.59e−5***
	(4.6)		(4.7)	(3.12)		(0.08)	(3.11)		(3.85)
UU size		1.2e−7***	5.8e−8**		8.41e−7***	8.44e−7***		7.91e−7***	7.48e−7***
		(9.9)	(3.05)		(8.86)	(8.15)		(6.37)	(6.06)
Controls							X	X	X
Observations	3877	3877	3877	11,631	11,631	11,631	11,631	11,631	11,631
R^2	0.003	0.006	0.004	0.011	0.011	0.011	0.12	0.12	0.12
F-stat	123.7***	23.5***	64.6***	8.96***	88.26***	44.13***	139.58***	144.33***	128.93***

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

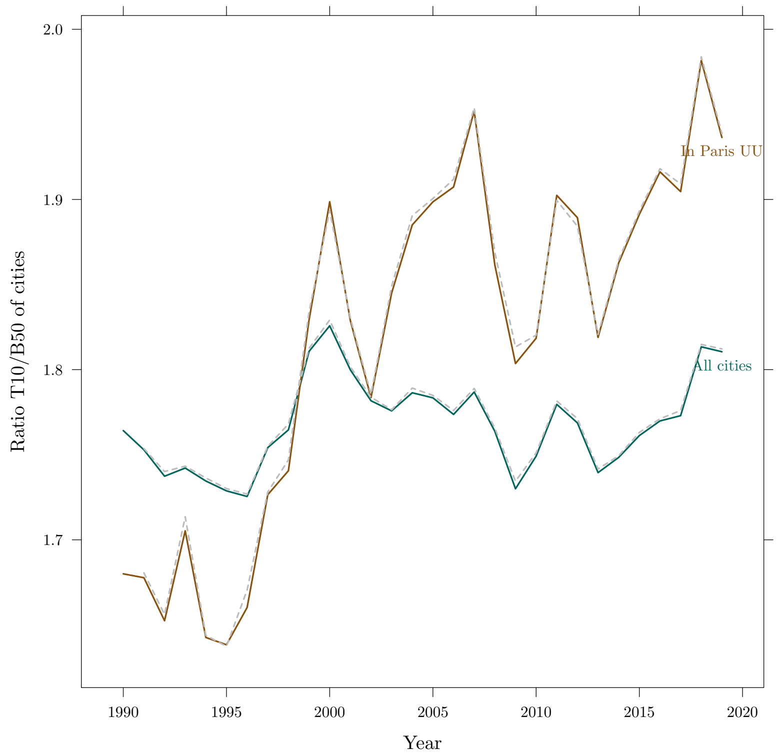
Note: The unit observation is the city (*commune*) at one of the three years 2000, 2010 & 2019. For the three main cities which are divided into subdistricts (*arrondissement*), each subdistrict provides one observation. Dependent variable is the ratio of the average fiscal income of the richest 10% within the city, over the same average, for the bottom 50%. Explanatory variables include the city and urban unit (*unité urbaine*) size (number of tax units). Controls include population density, unemployment, share of working-class people (CSP− as defined by the INSEE) share of foreigners, share of people with no diploma, and share of vacant accommodation within each city. Standard errors are clustered at the city level.

Inequalities *between* and *within* cities

To compute ratios of between-inequality, we create for each city an ideal Quetelet-like "average man" which earns the mean fiscal income of the commune, and we take ratios on R0 over the IRCOM base. In the following graphs, we also plot the counterfactual scenario of section 3 with $\Delta IPW_{1990-2018} = 0$, as a dashed grey line. By way of comparison, we computed between-inequality ratios over U.S. census tracts using series from the Census about the mean income of households in the past 12 months (code S1902). Unsurprisingly, the U.S. figures are clearly higher (the ratio T10/B50 lying around 3.3 for all Census tracts), and even more so when we focus on some major Metropolitan statistical areas like New York (3.8) or Miami (4.3). We also plot the ratio T10/B50 within cities against major ranking variables, with the corresponding counterfactual.

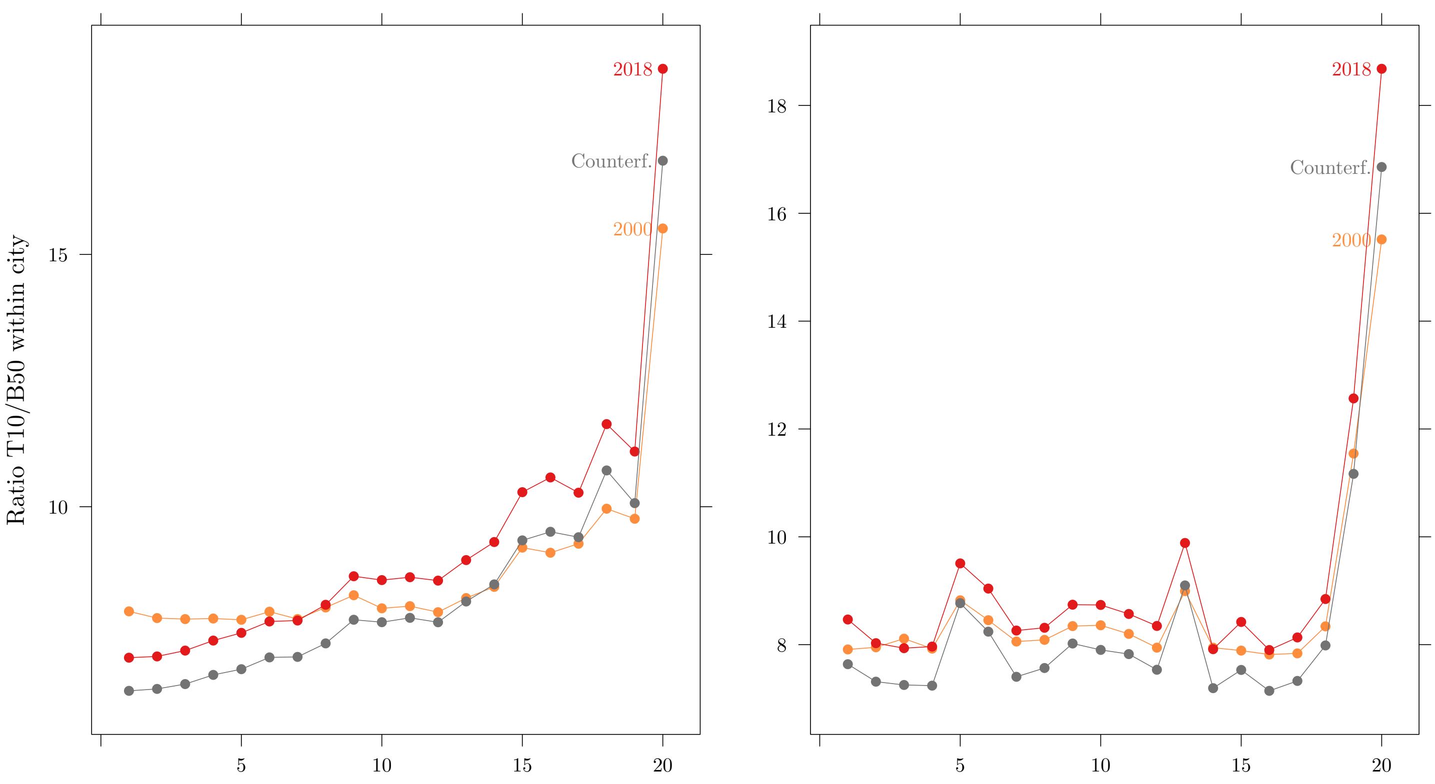
Figure 48: Counterfactual no-exposure scenario and inequality within and between cities : Complements

(a) T50/B50 ratio of communes



Note: See fig. for details.

(b) Ciy size (left) and av. fiscal income within city (right) *vs* Ratio T10/B50 (real fig./ counterf. scen. with $\Delta IPW_{1999,2018} = 0$)



Note: The unit of inter. is the *commune*. Cities are ranked by 5% total-population-weighted quantiles of : 1. City size (tax unit pop.); 2. Average fiscal income of their inhabitants. Reported statistics is the within-city ratio of the av. fiscal income of decile 10 of the income distribution, over the av. income of deciles 1 to 5, computed over the IRCOM databse (subsample R2, intersection of 2001, 2009 and 2019 fiscal year issues). The counterfactual scenario is derived from an estimation of model 3 over the first two decades of the 21st c., using ΔIPW as the explanatory (instrumented as described above), the full vector of controls, and as dependent, the evolution of the share of the city's total income held by the bottom 50% and the top 10% respectively. We retrieve a main coefficient of resp. -3.03 and 0.78; we ascribe to each group of each city a share premium proportionate to the exposure of the ZE to which the city belongs, times the opposite of the main coefficient.

Distance to a metropolis and income inequalities within the city

We picked the GPS coordinates of each city, and computed the distance to the nearest one among the 15 French metropolis, with two definition of distance: 1. Direct distance, *as the crow flies* (computed with the haversine formula); 2. Driving distance, computed using Google Distance Matrix API and the *gmapdistance* function of R [Melo, Rodriguez, and Zarruk 2016]. Computation has been made with a uniform departure time (Thursday, Jan. 27th, 7.30am GMT). Inequality curves plotting, over R2, Distance to a metropolis vs. Inequality exhibit a clear L shape pattern. Plotting (over R0) Distance vs. Average income results in a similar U-shape curve.

Figure 49: Direct distance to a metropolis (weighted centiles) *versus* Average fiscal income

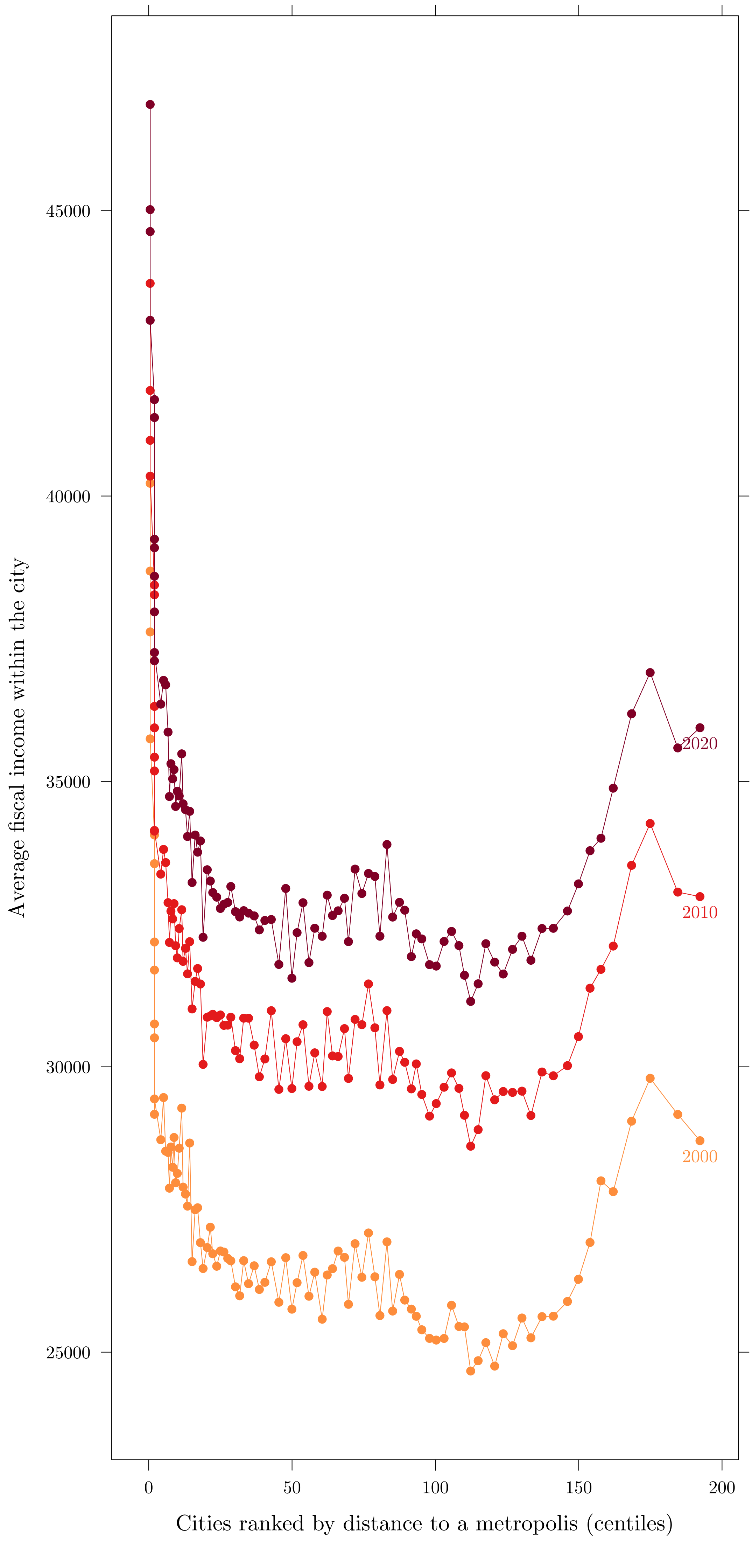
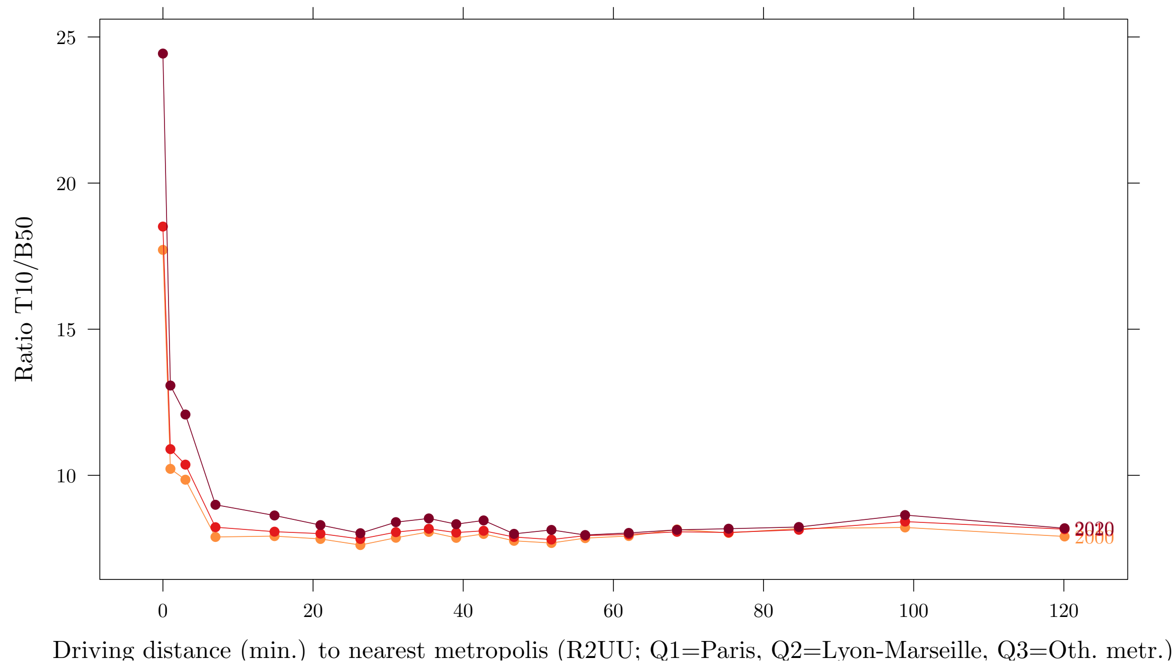
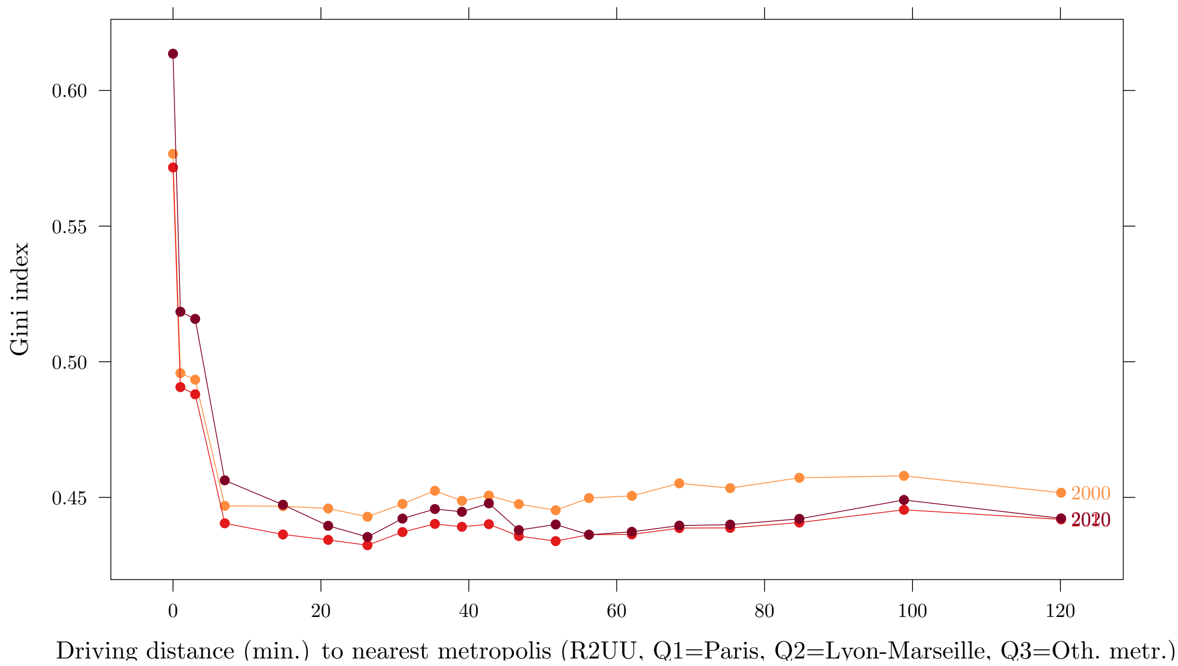


Figure 50: Cities ranked by pop.-weighted quantiles of driving to nearest metropolis *versus* major inequality indexes

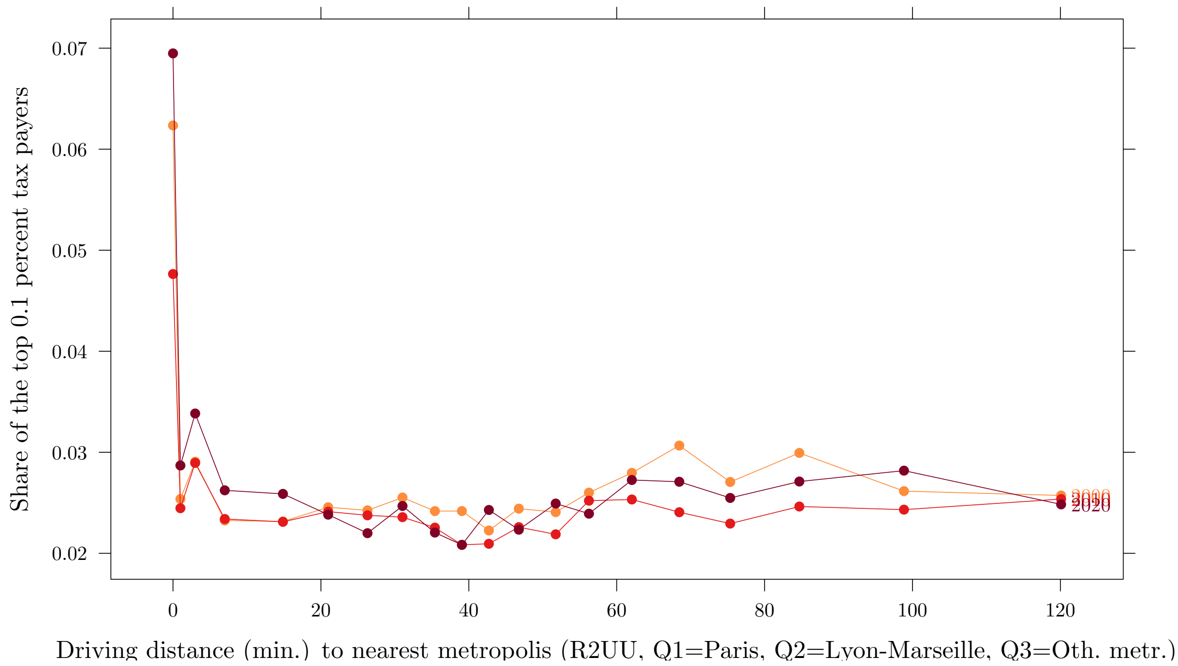
(a) *Versus* the ratio T10/B50 within the city



(b) *Versus* the Gini index



(c) *Versus* the share of the top 0.1 percent



D Complements to the shift-share strategy

Alternative dependent and explanatory var.

Some results of section 1 are replicated with: 1. An altern. dep. var: the evo. of the ratio of total manuf. employ. over the total 15-64 y.o. labour force of the ZE; 2. An altern. exp. var: ΔIPW, this time normalised, not by total employment, but by manufacturing employment (in this comput., ΔIPW is about 3 times larger).

Table 36: Exposure to imports from China and change in manufacturing employment at the ZE level

	Dep. : Decadal change in total manufacturing employ. per pop. (in pp) 1990-2018					
	(1)	(2)	(3)	(4)	(5)	(6)
Rise in imports from China per worker	-0.284^{***}	-0.305^{***}	-0.268^{***}	-0.192^{***}	-0.258^{***}	-0.152^{***}
over the decade (in 2022 kUSD)	(0.04)	(0.05)	(0.05)	(0.04)	(0.05)	(0.04)
Extra controls:						
Share of employ. in manufacturing		-0.22^{***}	-0.24^{***}	-0.25^{***}	-0.31^{**}	-0.31^{***}
		(0.03)	(0.05)	(0.04)	(0.11)	(0.09)
Share of women in lab. force				0.06		0.12^{***}
				(0.05)		(0.04)
Share of foreign-born in pop.				-0.05		-0.04
				(0.05)		(0.04)
Share of higher educ. in pop.				-0.08^{**}		-0.07^{*}
				(0.03)		(0.04)
Share of insecure jobs				0.09		0.08
				(0.12)		(0.15)
Share of routine jobs					0.79^{*}	0.61^{*}
					(0.42)	(0.36)
Offshorability of manuf. jobs					-0.05	-0.11
					(0.06)	(0.07)
Penetration of robots					-0.17	-0.02
					(0.21)	(0.21)
Regional dummies			X	X	X	X
R^2	0.03	0.29	0.34	0.36	0.36	0.38
F-stat	6.9^{***}	99.6^{***}	17.8^{***}	16.8^{***}	17.2^{***}	16.6^{***}
First stage: Instrumenting by the rise	0.71^{***}	0.72^{***}	0.71^{***}	0.69^{***}	0.62^{***}	0.61^{***}
in imports to a group of control countries	(0.03)	(0.04)	(0.04)	(0.05)	(0.04)	(0.05)
R^2	0.91	0.92	0.91	0.91	0.92	0.92
F-stat	2838^{***}	2147^{***}	334.9^{***}	291.2^{***}	341.1^{***}	302.1^{***}
Obs.	912	912	912	912	912	912

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: This is a replicate of table 3, withΔIPW now normalised by the manufacturing job stock.

Table 37: Exposure to import competition and evolution of wages

	Dep. : Decadal change in the average yearly wage (expressed in 2022 euros) within the département *(in pp)* *Period of estimation: 2002-2008 & 2008-2018*					
	All *départements*			Restrictions		
	All wages	*Manufacturing sector*	*Non-manuf. sectors*	*Top 10 dép.*	*Middle 40 dép.*	*Bottom 50 dép.*
	(1)	(2)	(3)	(4)	(5)	(6)
Rise in imports from China per worker:						
+ Full vector of controls:	-0.579^{***}	-0.721^{***}	-0.498^{**}	-1.48^{***}	-0.248^{***}	-0.245^{**}
	(0.17)	(0.22)	(0.21)	(0.11)	(0.07)	(0.102)
R^2	0.88	0.44	0.91	0.75	0.98	0.95
F-stat	89.12^{***}	8.4^{***}	108.3^{***}	2.3^{*}	95.6^{***}	62.7^{***}
Obs.	192	192	192	20	74	96

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: The unit of observation is the *département*. Dep. var. is the evo. of the av. y. wage, computed over the 1/12 subsample of the DADS. Main exp. var. is the index ΔIPW, described herein above, instrumented as described above. Two decades are stacked tog. with a time dummy, including the full vector of controls (sectoral ones excluded) plus the st.-of-dec. mean wage. Obs. are weight. by tot. pop. and SEs clust. at ZEAT level.

Table 38: Exposure to imports from China and evolution of occupations within each ZE

	Dep. : Decadal change in population log counts or shares of total adult population							
	Population evolution			Population breakdown				
	Total change (1)	*Natural increase* (2)	*Migration increase* (3)	In labour force (4)	*Employ. in services* (5)	*Unemploy.* (6)	Retired (7)	Other inactivity (8)
Panel A. Change in log counts								
Rise in imports from China per worker:								
No controls:	-0.08	-0.25^{***}	0.16	-0.04	-0.28	1.29^{***}	1.15^{***}	-1.37^{**}
	(0.19)	(0.04)	(0.21)	(0.21)	(0.26)	(0.44)	(0.42)	(0.61)
Full vector of controls:	0.43	-0.21^{**}	0.46^{*}	0.31	0.13	1.83^{***}	1.52^{***}	-0.42
	(0.26)	(0.08)	(0.27)	(0.22)	(0.21)	(0.48)	(0.55)	(0.72)
Panel B. Change in shares of adult pop.								
Rise in imports from China per worker:								
No controls:				0.01	-0.05	0.08^{***}	0.27^{***}	-0.17^{**}
				(0.11)	(0.06)	(0.015)	(0.06)	(0.07)
Full vector of controls:				-0.08	-0.11	0.07^{***}	0.24^{***}	-0.11
				(0.14)	(0.08)	(0.015)	(0.08)	(0.09)
Obs.	608	608	608	608	608	608	608	608

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: The unit of observation is the ZE (*Zone d'emploi*, definition of 2010). The dependent variable is the evolution of the related population, either in log counts, or expressed in shares of the total adult population (i.e. all individuals aged 15 y.o. or more) of the ZE. Exp. var. ΔIPW and corr. instr. have been described herinabove. We stack two decades (1990-1990 and 1999-2008) and include a time dummy for the second decade. Observations are weighted by the start-of-the-decade total population of the ZE. Standard errors are clustered at the level of the INSEE superzones.

Table 39: Exposure to import competition and evolution of reliance on social transfers

	Dep. : Decadal change in average share of social transfers in final income *Period of estimation: 2012-2019*					
	All ZEs				*Restrictions*	
	All transfers (1)	*Family allow.* (2)	*Housing benefit* (3)	*Minimum income* (4)	*Top 10 ZEs* (5)	*Bottom 50 ZEs* (6)
Rise in imports from China per worker:						
+Full vector of controls:	0.081^{**}	0.018	0.013^{**}	0.052^{**}	0.108^{***}	0.024
	(0.032)	(0.012)	(0.006)	(0.019)	(0.02)	(0.06)
R^2	0.68	0.58	0.56	0.78	0.86	0.81
F-stat	18.9^{***}	12.3^{***}	11.5^{***}	31.8^{***}	6.1^{***}	7.1^{***}
Obs.	304	304	304	304	31	152

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: The unit of observation is the ZE (*Zone d'emploi*, definition of 2010) or the *département*. The dependent variable is the average evolution (in pp) of the share of each type of transfers within the final disposable (after-redistribution) income of a tax unit within the ZE of interest, as reported in the Filosofi database of the INSEE. Exp. var. ΔIPW and corr. instr. have been described herinabove. All specifications include the full vector of controls (at the exception of offshorability and machine penetration indexes due to data limitation). Observations are weighted by the start-of-the-decade total population of the ZE. Standard errors are clustered at the level of the INSEE superzones.

Table 40: Exposure to import competition and evolution of incomes (*between*-inequality)

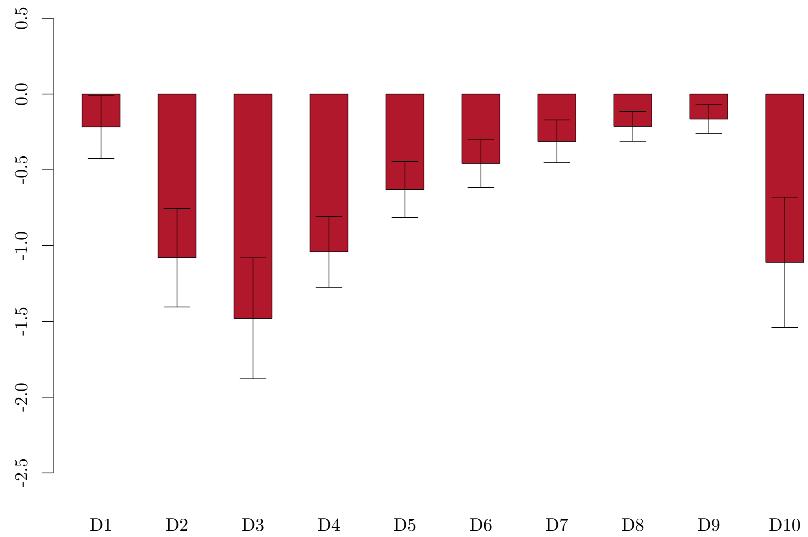
	Dep. : Decadal change in average yearly fiscal income (expressed in 2022 euros) within the spatial unit of interest (in pp) *Period of estimation: 2000-2008 and 2008-2018*								
	Fiscal income						... of which labour income		
	Département-level			*ZE*-level			*ZE*-level		
	All dép. (1)	*Top 10 dép.* (2)	*Bottom 50 dép.* (3)	All ZEs (4)	*Top 10 ZEs* (5)	*Bottom 50 ZEs* (6)	All ZEs (7)	*Top 10 ZEs* (8)	*Bottom 50 ZEs* (9)
Rise in imports from China per worker:									
+ Full vector of controls:	-0.68^{***}	-0.71^{*}	-0.35^{*}	-0.94^{***}	-1.61^{**}	-0.24	-0.47^{***}	-0.78^{**}	-0.15
	(0.17)	(0.36)	(0.21)	(0.25)	(0.74)	(0.27)	(0.14)	(0.39)	(0.09)
R^2	0.82	0.95	0.93	0.48	0.86	0.49	0.54	0.86	0.38
F-stat	24.1^{***}	11.2^{***}	35.1^{***}	16.6^{***}	15.8^{***}	9.1^{***}	21.1^{***}	16.7^{***}	5.9^{***}
Obs.	192	20	96	608	62	304	608	62	304

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

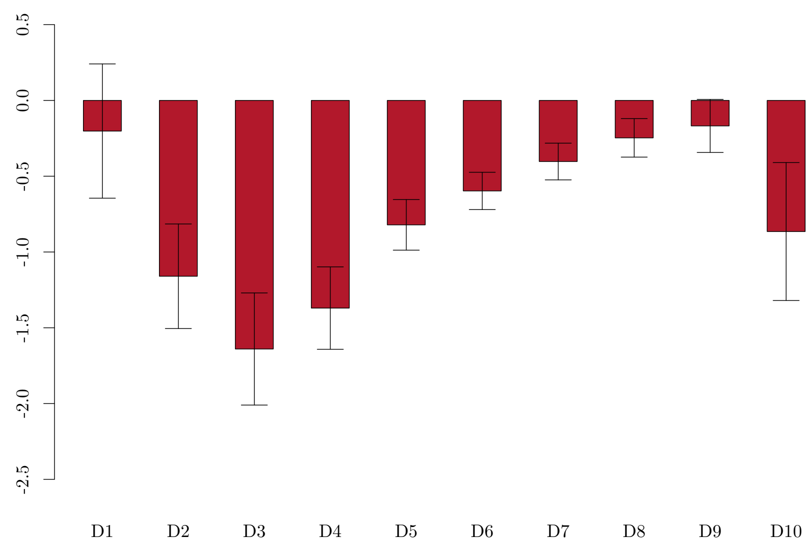
Note: The unit of observation is the ZE (*Zone d'emploi*, definition of 2010) or the *département*. The dependent variable is the evolution (in percentage points) of the related average yearly income per tax unit within the ZE or dép. as reported in the IRCOM database. Data about all incomes are computed over restriction zero (i.e. all cities, excluding those with fewer than 11 inhabitants). Exp. var. ΔIPW and corr. instr. have been described herinabove. All specifications include the full vector of controls (at the exception of offshorability and machine penetration indexes due to data limitation). For restrictions, we rank *départements* and ZEs according to their mean wage at the beginning of the period of estimation (2000) and distinguish between the top decile versus the five lowest deciles. Observations are weighted by the start-of-the-decade total population of the ZE. Standard errors are clustered at the level of the INSEE superzones.

Figure 51: Exposure to import competition and evolution of wages (*within*-inequality) − Predicted impact of a +$1000 rise in imports per head exposure within the *département*, on the average yearly wage of each quantile (decline in pp)

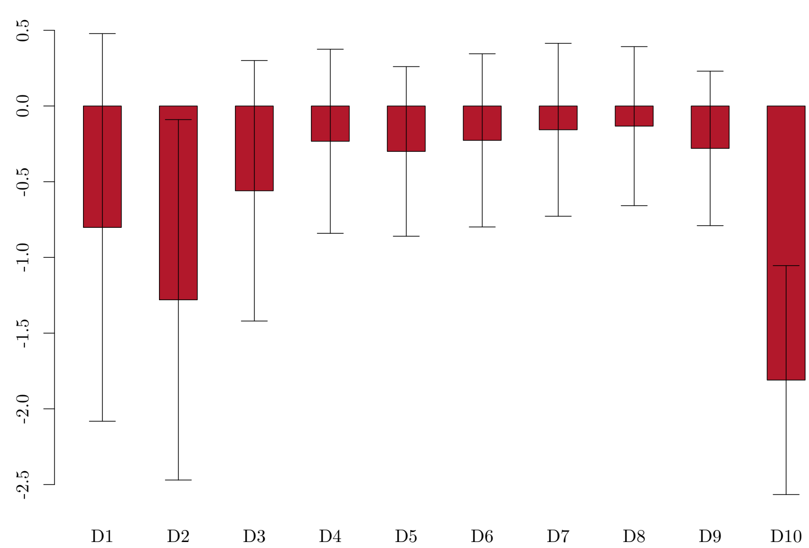
(a) All sectors



(b) Non-manufacturing



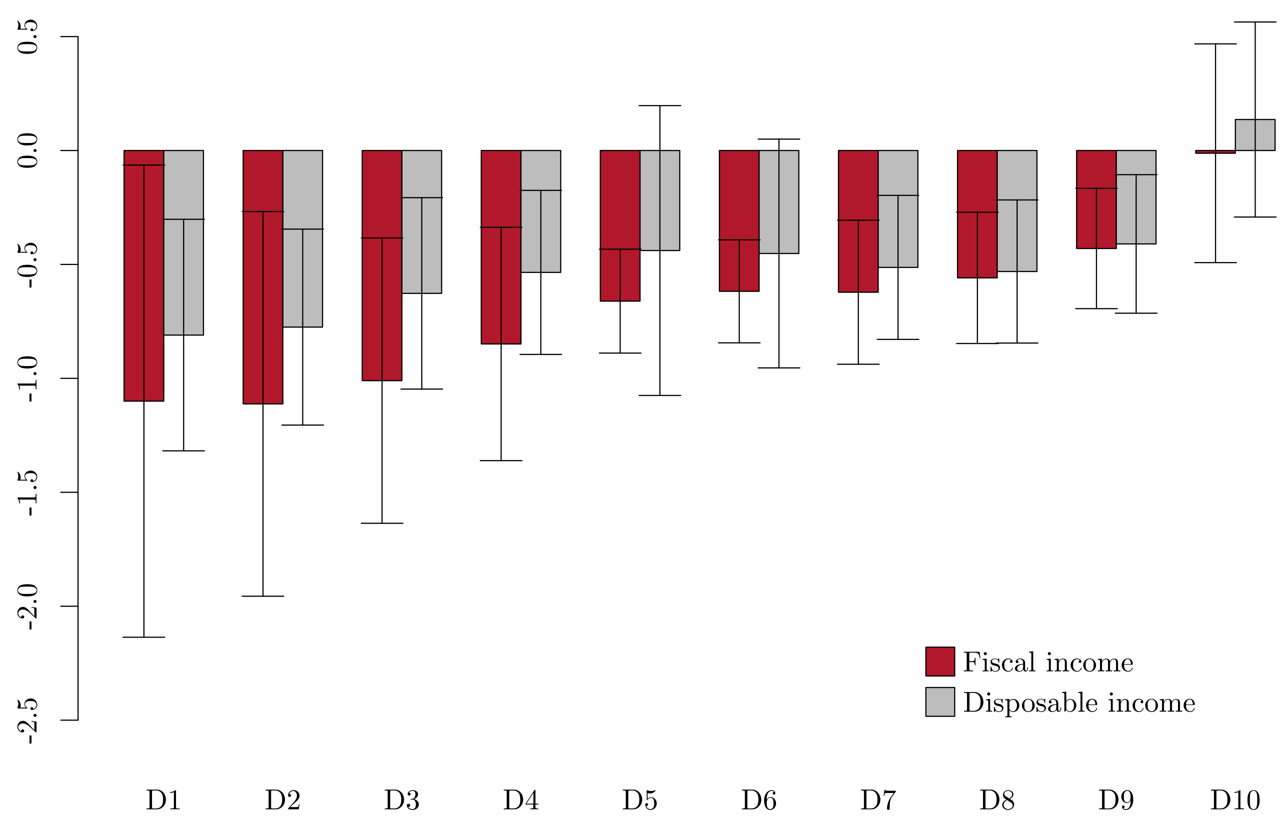
(c) Manufacturing



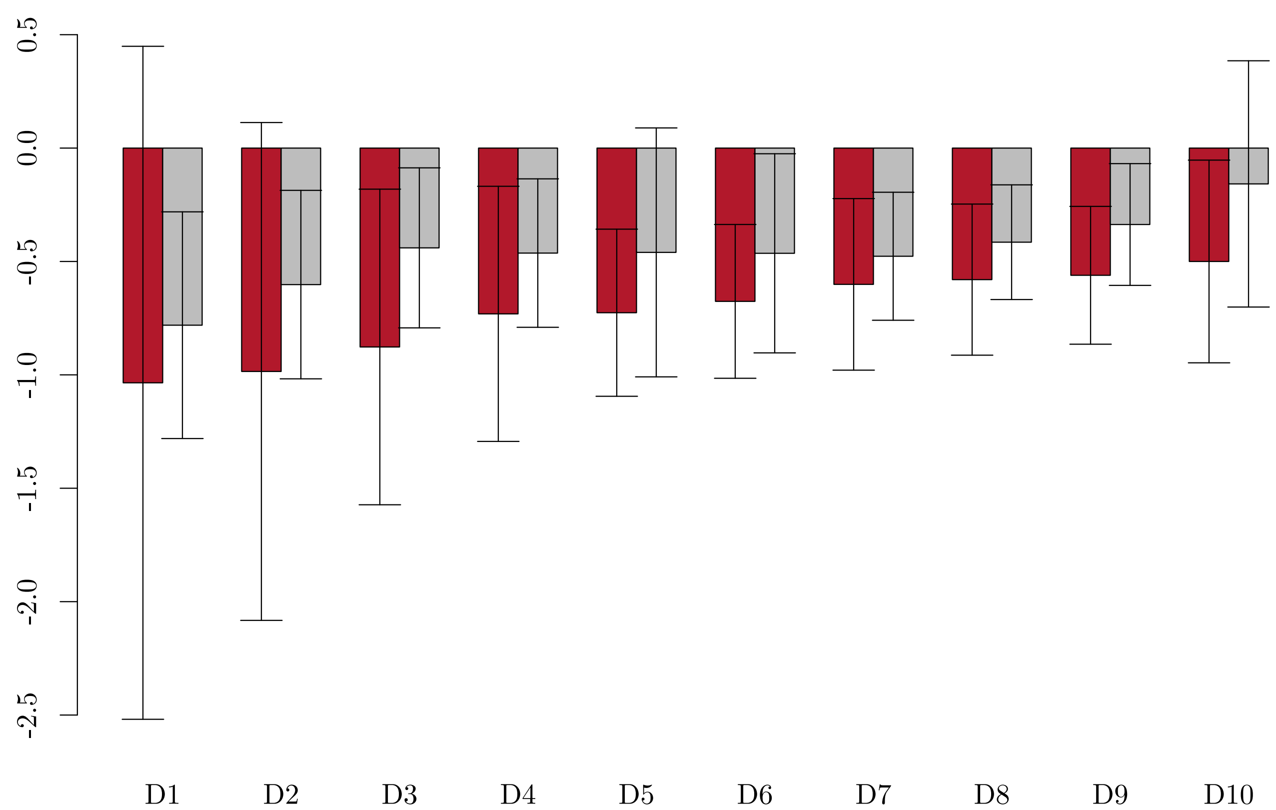
Note: The unit of interest is the *département*. The main source is the 1/12 microsample of the DADS. The dependent variable is the evolution (in pp) of the average yearly wage reported in the DADS, computed over each decile of the distribution of wages of each *département*. Exp. var. ΔIPW and corr. instr. have been described herinabove. All specifications include the full vector of controls (at the exception of offshorability and machine penetration indexes due to data limitation). Observations are weighted by the start-of-the-decade total population of the *département*. Standard errors are clustered at the level of the INSEE superzones. Bars denote the main coefficient, with the corresponding 95% conf. interval.

Figure 52: Exposure to import competition and evolution of fiscal income and disposable income (*within*-inequality) − Predicted impact of a +$1000 rise in imports per head exposure within the spatial unit, on the average yearly incomes of each quantile (in pp)

(a) Estimated at the level of the *département* (2012-2019)



(b) Estimated at the level of the ZE (2012-2017)



Note: The unit of interest is the *département* and the ZE. The main source is the Filosofi database. The dependent variable is the evolution (in pp) of the average yearly fiscal (red) and disposable (grey) incomes reported in the base, computed over each decile of the distribution of incomes of each *département* or ZE. The main explanatory variable is the index ΔIPW, described herein above, which provides an estimation of the mean rise in Chinese imports (in value) per worker of the industrial sector within each *département*. The instrument is the same ΔIPW, in which French trade data has been replaced by a control group of four countries (Japan, Germany, Spain, Switzerland) and all labour force variables are taken with a decadal lag. The period of estimation is 2012-2019. All specifications include the full vector of controls (at the exception of offshorability and machine penetration indexes due to data limitation). Observations are weighted by the start-of-the-decade total population of the *département*. Standard errors are clustered at the level of the INSEE superzones. Bars denote the main coefficient, with the corresponding 95% conf. interval.

Table 41: Exposure to import competition and evolution of fiscal income (*within*-inequality)

	Dep. : Decadal change in fiscal income (in pp)				
	Evolution of fiscal income			*Ratio T10/B50*	
	All incomes	Incomes of decile 1 to 5	Incomes of decile 10	Start-of-the-period value	Evolution
	(1)	(2)	(3)	(4)	(5)
Panel A − Estimated at the level of the département *(2012-2019)*					
Rise in imports from China per worker:					
+Full vector of controls:	-0.541^{***}	-0.885^{***}	-0.0121	-0.21^{***}	0.026^{*}
	(0.146)	(0.25)	(0.24)	(0.05)	(0.015)
R^2	0.86	0.79	0.87	0.96	0.87
F-stat	14.5^{***}	8.9^{***}	14.6^{***}	67.2^{***}	15.2^{***}
Obs.	304	304	304	304	304
Panel B − Estimated at the level of the ZE *(2012-2017)*					
Rise in imports from China per worker:					
+Full vector of controls:	-0.658^{***}	-0.818^{**}	-0.501^{**}	-0.162^{*}	0.003
	(0.181)	(0.31)	(0.23)	(0.84)	(0.01)
R^2	0.49	0.79	0.87	0.79	0.68
F-stat	8.3^{***}	8.9^{***}	9.1^{***}	33.6^{***}	18.7^{***}
Obs.	304	304	304	304	304

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: The unit of observation is the ZE (*Zone d'emploi*, definition of 2010) or the *département.* The dependent variable is the average evolution (in pp) of the share of each type of transfers within the final disposable (after-redistribution) income of a tax unit within the ZE of interest, as reported in the Filosofi database of the INSEE. The main explanatory variable is the index ΔIPW, described herein above, which provides an estimation of the mean rise in Chinese imports (in value) per worker of the industrial sector within each ZE. The instrument is the same ΔIPW, in which French trade data has been replaced by a control group of four countries (Japan, Germany, Spain, Switzerland) and all labour force variables are taken with a decadal lag. All specifications include the full vector of controls (at the exception of offshorability and machine penetration indexes due to data limitation). Observations are weighted by the start-of-the-decade total population of the ZE. Standard errors are clustered at the level of the INSEE superzones.

Robustness checks

Table 42: Exposure to imports from some countries and change in manufacturing employment at the ZE level - Alternative import partner

	Dep. : Decadal change in total manuf. employ. per working age pop. 1990-2008 (in pp)				
	China		Turkey	Poland	Germany
	(1)	(2)	(3)	(4)	(5)
Panel A. OLS estimates					
Rise in imports from specified exporter	-0.15^{***}	-0.15^{***}	0.09	0.43	0.02
per worker (in 2022 kUSD)	(0.05)	(0.05)	(0.32)	(0.46)	(0.02)
Panel B. 2SLS estimates					
Rise in imports from specified exporter	-0.19^{***}	-0.14^{***}	0.98	1.04	0.08
per worker (in 2022 kUSD)	(0.06)	(0.05)	(1.1)	(0.63)	(0.05)
R^2	0.49	0.48	0.46	0.46	0.46
F-stat	16.7^{***}	16.6^{***}	16.1^{***}	16.4^{***}	16.1^{***}
First-stage:					
Original instrument	0.83^{***}				
	(0.08)				
Extra-EU instrument		0.68^{***}	10.1^{***}	2.3^{***}	2.28^{***}
		(0.06)	(1.06)	(0.31)	(0.32)
R^2	0.91	0.98	0.98	0.97	0.91
F-stat	266^{***}	251^{***}	81.1^{***}	579^{***}	188^{***}
Obs.	608	608	608	608	608

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: The unit of observation is the ZE (*Zone d'emploi*, definition of 2010). The dependent variable is the change (in pp) of total manufacturing employment within the ZE, as a ratio of the total working age population of the zone. The main explanatory variable is the index ΔIPW, described herein above, which provides an estimation of the mean rise in Chinese imports (in value) per worker of the industrial sector (in 2022 kUSD) within each ZE. The original instrument is the same ΔIPW, in which French trade data has been replaced by a control group of four countries (Japan, Germany, Spain, Switzerland) and all labour force variables are taken with a decadal lag; the extra-EU instrument is similarly built with a control group made out of Japan, New Zealand, Australia & Canada). Observations are weighted by the start-of-the-decade total population of the ZE. All specifications contain the full vector controls of the main model. Standard errors are clustered at the level of the 10 INSEE superzones.

Table 43: Testing for reverse causality (Present exposure to China trade versus Past industrial employment decline) - 2SLS estimates

	Dep. : Decadal change in total manuf. employ. per working age pop. 1975-1990 (in pp)		
	(1)	(2)	(3)
Rise in imports from China per worker:			
1990-1999	3.4		
	(3.1)		
1999-2008		−.32	
		(.21)	
2008-2018			.21
			(.66)
R^2	0.07	0.002	0.001
F-stat	3.2*	5.2**	0.16

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: The unit of observation is the ZE (*Zone d'emploi*, definition of 2010). The dependent variable is the change (in pp) of total manufacturing employment within the ZE, divided by the total working age population of the zone. The main explanatory variable is the index ΔIPW, described herein above, which provides an estimation of the mean rise in Chinese imports per worker of the industrial sector (in 2022 kUSD) within each ZE. The instrument is the same ΔIPW, in which French trade data has been replaced by a control group of four countries (Japan, Germany, Spain, Switzerland) and all labour force variables are taken with a decadal lag. No other control is applied. Observations are weighted by the start-of-the-decade total population of the ZE. Standard errors are clustered at the level of the 10 INSEE superzones.

Table 44: Testing for reverse causality (Exposure to China trade and industrial decline with a decadal lag) - 2SLS estimates

	Dep. : Decadal change in total manuf. employ. per working age pop. 1990-1999 (in pp)			
	Exposed ZEs		All ZEs	
	(1)	(2)	(3)	(4)
Current period exposure (1990-1999)				
Rise in imports from China per worker	−1.1	0.18	−1.59	−0.49
	(1.4)	(1.1)	(1.7)	(1.4)
Start-of-the-period manuf. empl. share		−0.19***		−0.22***
		(0.03)		(0.04)
Future period exposure (1999-2008)				
Rise in imports from China per worker	−0.22	−0.22*	−0.04	−0.05
	(0.13)	(0.12)	(0.11)	(0.11)
Start-of-the-period manuf. empl. share		−0.08		−0.01
		(0.08)		(0.05)

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: The unit of observation is the ZE (*Zone d'emploi*, definition of 2010). The dependent variable is the change (in pp) of total manufacturing employment within the ZE, divided by the total working age population of the zone. The main explanatory variable is the index ΔIPW, described herein above, which provides an estimation of the mean rise in Chinese imports per worker of the industrial sector (in 2022 kUSD) within each ZE. The instrument is the same ΔIPW, in which French trade data has been replaced by a control group of four countries (Japan, Germany, Spain, Switzerland) and all labour force variables are taken with a decadal lag. Exposed ZEs are the top quartile of ZEs ranked according to their $\Delta IPW^{fr}_{1999-2008} - \Delta IPW^{fr}_{1990-1999}$. Observations are weighted by the start-of-the-decade total population of the ZE. Standard errors are clustered at the level of the 10 INSEE superzones.

Variance breakdown exercise

Autor, Dorn & Hanson's intuition to discriminate the supply-driven from the demand-driven dimension of the China shock relies on a comparison between the OLS and 2SLS estimates of the marginal impact of exposure. If we rewrite the main model and drop covariates, we get:

$$\Delta L_{it} = \beta \Delta IPW_{it} + u_{it} \tag{26}$$

Coefficients are then straightforwardly estimated with:

$$\hat{\beta}_{OLS} = \frac{Cov(\Delta IPW, \Delta L)}{Var(\Delta IPW)}, \qquad \hat{\beta}_{2SLS} = \frac{Cov(\Delta IPW_{IV}, \Delta L)}{Var(\Delta IPW_{IV})}$$

Our instrumentation is meant to isolate the exogenous from the endogenous dimension of the change in China trade exposure:

$$\Delta IPW = \Delta IPW_{IV} + \Delta IPW_{Endo}$$

In 26, an OLS estimation will therefore yield:

$$\begin{aligned}\hat{\beta}_{OLS} =& \frac{Cov(\Delta IPW_{IV} + \Delta IPW_{Endo}, \Delta L)}{Var(\Delta IPW_{IV} + \Delta IPW_{Endo})} \\ =& \frac{Cov(\Delta IPW_{IV}, \Delta L) + Cov(\Delta IPW_{Endo}, \Delta L)}{Var(\Delta IPW_{IV}) + Var(\Delta IPW_{Endo}) + 2 \times Cov(\Delta IPW_{IV}, \Delta IPW_{Endo})} \\ =& \frac{Cov(\Delta IPW_{IV}, \Delta L) + Cov(\Delta IPW_{Endo}, \Delta L)}{Var(\Delta IPW_{IV}) + Var(\Delta IPW_{Endo})}\end{aligned}$$

The simplification of the last lign being operated through the orthogonality of construction between the endogenous and exogenous part of the breakdown of ΔIPW. If we substitute with our estimated coefficients, and define a similar endogenous β, we get:

$$\hat{\beta}_{OLS} = \hat{\beta}_{IV} \times \frac{Var(\Delta IPW_{IV})}{Var(\Delta IPW_{IV}) + Var(\Delta IPW_{Endo})} + \hat{\beta}_{Endo} \times \frac{Var(\Delta IPW_{Endo})}{Var(\Delta IPW_{IV}) + Var(\Delta IPW_{Endo})}$$

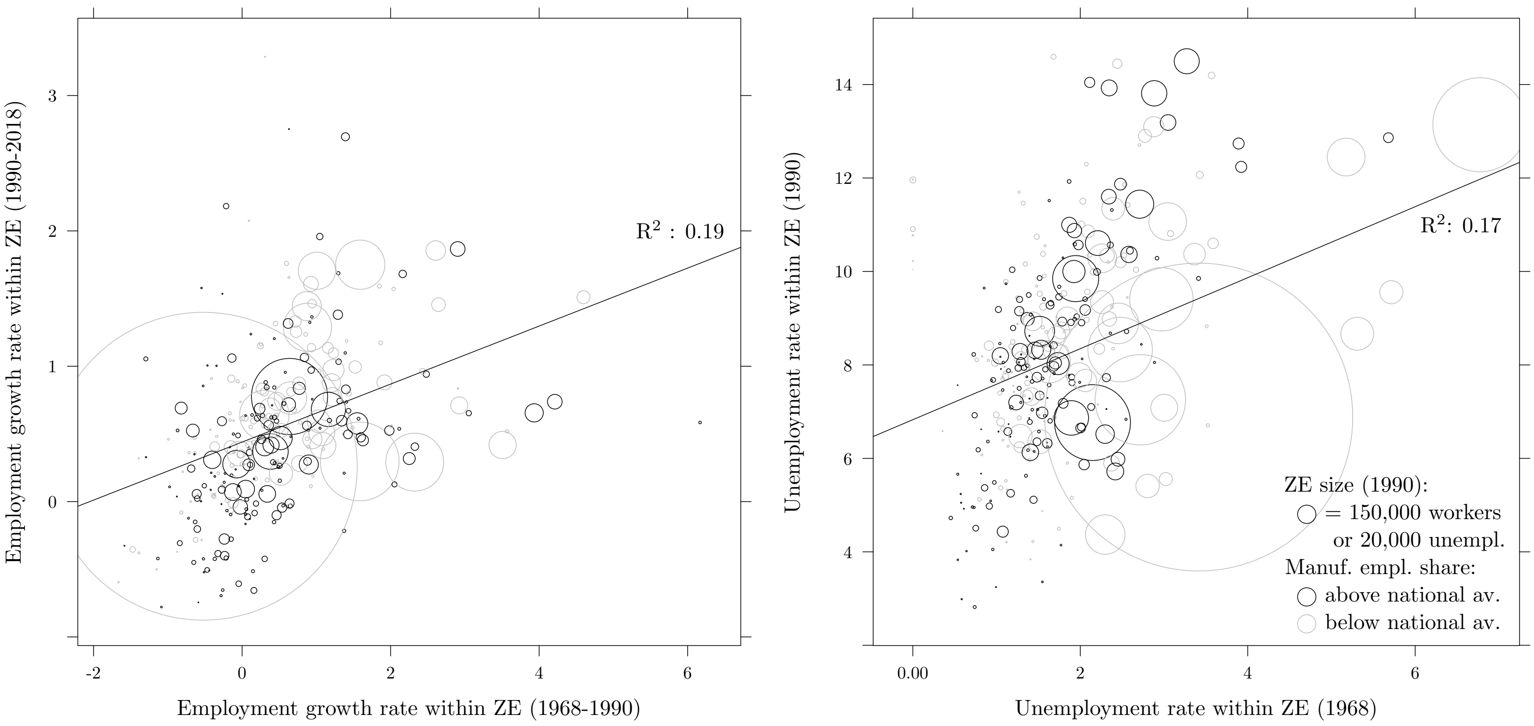
The corresponding coefficients are then taken from the data. For instance, over the decades 1999-2018, in the simplest specification with no controls, we obtained $\hat{\beta}_{OLS} = -5.66$ and $\hat{\beta}_{2SLS} = -6.27$ and we can compute $\hat{\beta}_{Endo} = 2.07$. From this, it follows that $\frac{Var(\Delta IPW_{IV})}{Var(\Delta IPW_{IV})+Var(\Delta IPW_{Endo})} \approx 0.93$.

That figure is superior to the one of [Autor, Dorn, and Hanson 2013], but very similar to the ones found in other European replications.

E Complements to the augmented Blanchard-Katz model

Supplementary figures

Figure 53: The persistence of unemployment and of employment growth within ZEs (pre-exposure)



Note: The unit of interest is the *Zone d'emploi* (the INSEE's commuting zone, 2010 definition). Data are from the INSEE's Census. Reported statistics include the average growth rate of total employment within the ZE (LHS panel) and the unemployment rate (ILO-INSEE definition, RHS panel). Circle sizes provide the related population of the ZE (total number of employed workers, or of unemployed people). A black circle indicates that the share of manufacturing within total employment of the ZE (in 1990) is above the national average (and vice versa for a grey circle). We plot the regression line of the variable on the y-axis on the variable of the x-axis, weighting by the related population of each ZE.

Supplementary tables

Table 45: Autoregressive models for some major labour and income variables at the level of the *ZE*

	Log employment change	*Unemployment rate*	*Logged Participation*	*Logged av. inc. (top 10%)*	*Logged av. inc. (bottom 50%)*
	(1)	(2)	(3)	(4)	(5)
Coefficient on lagged dependent variable					
One lag	.921	.759	.754	.321	.352
	(.011)	(.012)	(.011)	(.018)	(.017)
Two lags	−.108	−.117	− 0.137	.085	.019
	(.012)	(.011)	(.011)	(.019)	(.018)
Three lags				.109	.063
				(.018)	(.017)
Implied impulse response					
Year 1	1	0.759	1	1	1
Year 2	1.92	0.459	0.754	0.321	0.352
Year 3	2.66	0.259	0.432	0.188	0.143
Year 4	3.24	0.143	0.222	0.197	0.119
Year 5	3.69	0.08	0.05	0.114	0.067

Note: *p<0.1; **p<0.05; ***p<0.01

Note: Period of estimation is 2006-2018 (yearly) for the employment variables from the INSEE's Census, 2000-2019 for the income variables from the IRCOM set. Incomes are average yearly fiscal income computed over restr. 2 with the *gpinter* algorithm. Standard errors are in parentheses. We allowed for a specific intercept for each zone. Observations with a stationarity issue are excluded.

Table 46: Correlation between local employment growth rates and population growth rates

	Period of estimation: 2008-2018	
	Natural increase	*Migrations*
At the city level	0.19	0.42
restr. to Paris metro.	0.16	0.69
restr. to N.-W. regions	0.25	0.54
At the employment zone (ZE) level	0.28	0.36

Table 47: Av. local unemployment rate vs. Av. employment growth rates (at the ZE level)

	Average local unemployl. growth rate 1968-1999	2006-2018
	(1)	(2)
Av. employ. growth rate		
1968-1999	$-.29^{**}$	
	(.12)	
2006-2018		$-.25^{*}$
		(.15)
N	298	298
R^2	.02	.009
F-stat	5.81^{***}	2.64^{*}

Note: *p<0.1; **p<0.05; ***p<0.01

Note: Period of estimation is 2006-2018 (yearly) and 1968-1999 (5 y. interval); all data are from the INSEE's Census, retrieved at the city level and aggregated at the level of the employment zone (ZE) in the 2010 definition of a ZE. Standard errors are in parentheses.

Table 48: Regressing local employment growth rates on their national counterpart

	Empl. growth	Main model (2007-2012)			Main model (2013-2018)			Indicative variables		
	(1968-1990)	α_c	β_c	R^2	α_c	β_c	R^2	u_{08}	Vac_{08}	$\Delta w_{07,18}$
All ZEs	*+8.9%*	*−.0021*	*.59*	*.24*	*.0043*	*.12*	*.37*	*8.87%*	*6.2%*	*+5.9%*
synchronised & counter-cyclical ZEs	*+20.9%*	*.004*	*.49*	*.39*	*.005*	*.53*	*.241*	*8.3%*	*5.8%*	*+6.5%*
Lyon	+2.49%	.0081	−.013	.0016	.0046	$.989^{**}$	.581	6.78%	5.4%	+7.4%
Bordeaux	+8.73%	.0082	$.621^{**}$	.51	.011	.791	.22	9.2%	6.5%	+6.2%
Toulouse	+28.9%	.012	$.966^{***}$	.69	.011	.433	.11	8.5%	6.3%	+10.6%
Montpellier	+63.6%	.014	.09	.007	.009	1.17^{**}	.61	10.8%	4.8%	+7.4%
synchronised & Pro-cyclical ZEs	*+9.2%*	*.0006*	*1.56*	*.75*	*−.001*	*1.51*	*.17*	*9.5%*	*5.9%*	*+6.1%*
Sète	+8.6%	.006	1.19	.71	−.013	3.42^{***}	.89	13.9%	6.3%	+8.7%
La Roche-sur-Yon	+10.4%	.005	1.13^{***}	.75	−.0011	3.86	.33	8.9%	5.1%	+8.5%
Saint-Nazaire	+2.1%	.0005	1.89^{***}	.64	.006	.67	.21	9.2%	4.6%	+10.8%
Dunkerque	+41.4%	.002	3.17	.08	−.009	$.93^{*}$	.41	9.9%	4.3%	+4.1%
Hysteresic & Counter-cyclical ZEs	*−13.7%*	*−.005*	*.61*	*.37*	*−.005*	*.43*	*.17*	*9.1%*	*7.5%*	*+4.7%*
Saint-Etienne	−14.2%	−.003	.24	.08	−.003	.16	.01	7.1%	12.5%	+6.2%
Sarreguemines	−1.8%	−.004	$.91^{***}$	.86	−.004	$.65^{**}$	.52	9.6%	5.6%	+6.1%
Bar-le-Duc	−9.3%	−.015	1.56^{***}	.93	−.006	-1.81^{***}	.71	8.5%	6.8%	+3.7%
Hysteresic & Pro-cyclical ZEs	*−11.9%*	*−.007*	*1.49*	*.69*	*−.007*	*2.18*	*.31*	*9.1%*	*6.6%*	*+4.5%*
Les Sables d'Olonne	−5.5%	−.006	1.25	.15	.064	9.76	.01	10.1%	3.7%	+9.2%
Saint-Malo	−6.2%	−.005	.99	.33	−.001	1.89^{**}	.62	8.5%	5.8%	+7.9%
Belfort-Montbéliard	−4.8%	−.007	1.04^{**}	.64	−0.11	1.64	.34	9.1%	6.1%	+12.4%

*p<0.1; **p<0.05; ***p<0.01

Note: The unit of interest is the employment zone (ZE) in the 2010 definition of the term. Columns 2 to 7 report the estimation of model (20), i.e. the regression of the time series of local employment growth rates on the nationwide series; two periods are tested (2007-2012 and 2013-2018); we report the intercept, the main coefficient and its significance, and the R-squared. UR tests have been applied prior to the estimation; we had to remove 24 out of 321 ZEs for which there was an issue of non-stationarity. Column 1 reports the global growth rate of employment between 1968 and 1990. Indicative variables include the unemployment rate in 2008 u, the share of vacant accommodation Vac, and the growth rate of the average labour income per taxable unit within the zone (as reported in the IRCOM set), denoted Δw for simplicity. We provide details for 14 out of 297 estimated ZEs. We have divided the ZEs according to two criteria: it is said to be pro-cyclical (counter-cyclical) if the sum of the normalised elasticities is above (below) 0. It is said to be an hysteresis district if the growth rate of employment over 1968-1990 and the intercept α_c of the 2007-2012 estimated model are below their national average. Four types of ZEs are thus defined, for which we reported the mean values: averages are computed using as weights the number of jobs within the ZEs at the end of the period of estimation (2018). Data are from the INSEE's Census and the IRCOM database.

Robustness checks for the Brechling-Thirlwall decompozition exercice

- *Using a different variable of interest* − Data from the INSEE is a bit more detailed for unemployment (there exists series about the quarterly unemployment rate at the ZE level, yet this is not genuine data, but an estimation provided by the INSEE based on the Census, calibrated using the local quarterly unemployment figures of *Pôle emploi*; the spatial subdivision used in this series are the ZEs of 2020 and not of 2010); an alternative strategy would be to estimate:

$$\Delta u_{c,t} = \alpha_c + \beta_c \Delta u_t + \eta_{c,t} \tag{27}$$

  We find a considerably superior average R^2 (0.89) and coefficients β_c much more concentrated around 1. The global outcome of the test is less interesting than the results retrieved with model 20, but they are consistent with the predictions of the model, especially when it comes to the population dimension (the short-term adjustment on unemployment is higher in more populated and denser regions). The general picture is one of unemployment shocks being short-lived and quickly erased, as the model predicted.

- *Using different spatial units of interest* − We estimated model (20) at the city, ZE and region level, model (27) at the ZE and region level. The average R^2 and β_c remains the same whatever the spatial level. At the city level, the weighted average R^2 is 0.11. High α are found among touristic and residential areas (Les Sables d'Olonne is the highest α we estimated at the city level: $+0.32$). Profiles with low R^2 - low β are mainly found within the Parisian metropolis (in popular and bourgeois suburbs alike). High R^2 on the contrary characterise the two extremes of the distribution: fast-growing and declining cities alike. The city level provides, especially for some declining industrial cities, stunning results. Saint-Etienne is the city with the highest R^2 (0.81) and unsurprisingly, some of the highest negative β_c : -3.03 (1% significant).

F Complements to the augmented Schelling tipping point model

Estimation strategy

Choosing data, finding relevant variables

Datasets - Card, Mas & Rothstein were relying on data from the U.S. Census. A natural French equivalent is the census of the INSEE (we used the issues of 1962, 1968, 1975, 1982, 1990, 1999 and the New annual Census from 2006 onward).

Geographical unit of interest - Applying our specification over actual data requires of choice of geographical level: one must pick a *supra*-unit, the city c, and an *infra*-unit, the district i. In the authors' original setting, the *supra*-unit is the MSA (Metropolitan Statistical Area), the *infra*-unit, the census tract. Datasets from the INSEE leave us with a wide range of options: the *infra*-unit might be the *commune* or the infracommunal unit (formally, the IRIS) ; for the *supra*-unit, we might take any definition of the urban agglomeration used by official statistics, from the narrowest, known as the UU (*unité urbaine*) to the widest, the AAV (aire d'attraction des villes), or a middle definition, the ZE (Zone d'emploi) which has the comparative advantage to cover the entirety of the territory:

- The main drawback of using communes lies in the fact that, contrary to the census tract, it is a highly uneven division (in our subsample, the tiniest communes do not surpass 1000 inhabitants, while the largest unit, Toulouse, is far above 300.000). However, communes are among the oldest administrative divisions; while the IRIS are barely known outside of a limited circle of experts, communes are the pivotal level at which citizens perceive the impact of public policies and social changes (symptomatically, municipal elections are among the few elections which have resisted the decline in the turnout rate over the last decades).
- Conversely, IRIS divisions are more homogeneous. Every commune with more than 10.000 inhabitants, and most communes over 5000 is divided into IRISs, and communes below 5000 act as an IRIS in the INSEE system. They are extremely useful when studying major metropolises like Paris or Lyons, but for the remainder of the French territory, they leave us with the same issue about the size of communes mentioned herein above.

All in all, IRISes in ZEs seems to be the most natural choice, at least the one which is the most similar to the original setting. To provide one comparative example, the city of Chicago has 2.7 million inhabitants and 866 census tracts, while the ZE of Lyons has 1.2 million inhabitants, 248 communes (in our subsample) and 538 IRISes. However, since there is wide controversy among spatial econometricians about how the choice of the geographical unit might bias the estimates, we will systematically report results at the AAV, UU and *commune* level in this annex.

Controls - Another issue is the choice of the control variables in $X_{ic,t-10}$. The authors have recourse to the following ones, always computed at the tract level: unemployment rate, log(mean family income), housing vacancy rate, renter share, fraction of homes in single-unit buildings, and fraction of workers who commute using public transit (used as a proxy for working-class population). At the IRIS level, we use data from the Census of the INSEE with the following controls: unemployment, share of persons with no diploma, and share of vacant accommodation; share of migrants (people not born in France) and share of blue-collar workers are included when they are not colinear with the explanatory.

Summary statistics

We'll use two major sources from the INSEE: the Census and the RFL-Filosofi database.

The Census provides highly reliable estimates with almost no selection issues. To provide but one example, the 1999 Census contains 50066 IRISes with a non-zero population (or 57.9 million inhabitants). Once we have ensured that we can match the IRISes over time, removed those units the INSEE recommends not to use in comparative analysis, and applied our selection rules[79], we are left with 47878 units, or 57.1 million inhabitants (87% of the national population).

The RFL-Filosofi sets, since they are dependent on statistical secrecy rules, provide far less information. Bracket breakdown of income is recorded for roughly one third of IRISes, mostly urban ones. This urban subsample has a more unequal income structure than the whole nation (in 2011, 11.9% of the subsample population lies in decile 1, 11.1% in decile 10, the population share of any other decile being below 10%; i.e. in our urban subsample there's a overrepresentation of the richest and poorest households). For the 2001-2011 restriction rules, once applied, leave us with 15615 units (or 51.4 million inhabitants, 20.6 million households in 2001). A detailed overview is provided in Table 49.

In the original article, there were 110 zones in the sample, and a mean 300 *infra*-units (census tracts) within each area. In our replication, we get 305 *supra*-units, and an average 160 *infra*-units within each one.

For the RC-IRCOM dataset, we used a special restriction which concatenates civil years 2000, 2010 and 2020 (which we called R4), which originally contains 4216 cities. Once restriction rules are applied[80], we are left with 2958 cities or 18.43 million tax units.

Structural break searching algorithm

In order to estimate the ZE-specific tipping point, we rely on two different methods.

[79] The IRIS must be within a *supra*-unit, and within each area, we require a minimum of 12 IRISes.
[80] I.e. being in a ZE, and having a minimum of 6 cities within each ZE

Table 49: Summary statistics

	Origin-based strategy		*Income-based strategy*	
	1999-2010	2010-2017	2001-2011	2011-2019
	(1)	(2)	(3)	(4)
Nb. of IRISes in original dataset	50066	50885	48404	48386
Nb. of IRISes preserved over time	49764	49281	48015	41978
Nb. of IRISes in final sample	48813	48848	15615	13326
Nb. of ZEs in final sample	305	305	199	199
Mean lower group share at beginning of decade	10.1	8.6	11.1	11.9
Growth in total population over the decade	7.1	3.8	13.9	6.57
Growth in upper-group population	7.8	2.5	10.9	6.54

Growths of ind. pop. for the Census, of the tax unit pop. for the RFL sets.

Method 1 (Structural break) - In this setting, identification of the thresholds relies on the methodology of [Hansen 2000]. For each supra-region or ZE c, we define a search interval $[m_1, m_2]$ where:

$$m_1 = \min(l_{ic,t-10})$$
$$m_2 = \max(l_{ic,t-10}) \times \mathbb{I}[l_{ic,t-10} < 0.5]$$

Candidate tipping points are all the values of the sequence defined by $v_{n+1} = v_n + 1/1000, v_0 = 0$ that lie within $[m_1, m_2]$.

For each candidate point l^* within each area, the dependent variable $\Delta u_{ic,t}$ is regressed over a dummy equal to one if $l_{ic,t-10} > l^*_{ic,t-10}$, formally if $\delta_{ic,t-10} > 0$ (i.e. if the initial lower-group share is above the tipping threshold):

$$\Delta u_{ic,t} = a_c + \alpha \mathbb{I}[\delta_{ic,t-10} > 0] + \varepsilon_{ic,t}$$

That simple operation is reproduced over the whole range of hypothetical tipping points $l^*_{c,t-10}$ within the extreme values of $l_{c,t-10}$. The chosen area-specific tipping point is the value of $l^*_{c,t-10}$ which maximises the R-squared of the model (provided that coefficient α is negative)[81].

Method 2 (Fixed point) - Card, Mas & Rothstein propose a complementary, non-standard method to the classical approach of Hansen, which, according to them, performs better for lower-size cities. The intuition is that if we take the average of $\Delta u_{ic,t}$ over the zone, and center our dependent variable around this mean, the evolution of the upper-group population shall be positive in districts below the tipping point, and negative for districts which are beyond the tipping point. I.e., if we plotted our centred dependent variable vs. the lower group share, we should see a piece-wise function that is equal to 0 at l^*. More formally, Card and his coauthors rely on a smooth approximation of the differential variable:

$$\Delta u_{ic,t} - E[\Delta u_{ic,t}|c]$$

We fit that variable to a quartic polynomial in $l_{ic,t-10}$; to obtain a function $R(l_{t-10})$. The algorithm of method 2 then finds the roots of that polynomial, provided it is below 0.5. If multiple roots are identified, the algorithm picks the one point at which the first-order derivative of $R()$ is the lowest (most negative).

Method 1 algorithm returns a missing value when it cannot find any structural break associated with a negative drop. As to method 2 algorithm, it creates a *NA* when $R()$ has no root.

Exploratory results - Over the INSEE's census

One important limitation of empirical strategies tackling the issue of spatial segregation is the element of arbitrariness in the choice of the dominant and of the discriminated groups. There exists auto-aggregative algorithms which infer the relevant groups from the patterns of the spatial data themselves; we applied the one of [Louf and Barthelemy 2016] at the IRIS level for a fiscal-income-based strategy, yet, as in their own application, it leads us to use three main aggregates (the bottom 50% of the fiscal income distribution, the top 40%, and a little middle 10% in-between) which do not perform well in a tipping setting.

We therefore empirical tested several alternative definitions of the upper and lower group, comparing two issues of the INSEE's Census (1999 and 2010) and two issues of the RFL-Filosofi set (2001 and 2011).

Over the Census, group labels include the place of birth (for simplicity of language, we will improperly call *natives* those who are born on French territory, and *migrants* those born in a foreign country), the 8-classes SES scale of the INSEE (especially, the shares of white-collar and blue-collar workers, of employees and intermediate occupations ; we use the label *working-class* for the sum of blue-collar and employee population). Due to data limitations, the dependent variable must focus on each separate item : SES or place of birth (the interaction of the two being absent from the 2010 series). The explanatory on the contrary can rely on such an interaction (we can use as *explanans*

[81] Actually, in some parts of the replication STATA code provided by Jesse Rothstein, they are using the maximisation of the t-stat as the main criterion to choose the tipping point. We tested this criterion on some sets, but this alteration tends to make the threshold estimation less congruent across methods. Since their published article mentions the R-squared as the only criterion, consistent with the method described by Hansen 2000, we keep the original setting.

the share of working-class migrants in 1999 for instance).

We always report the result of the full pooled model, with Δu as the dependent, and as explanatory variables, a quartic polynominal on δ, a dummy equal to one when $\delta > 0$, *supra*-units-specific fixed effects, and the full vector of controls. We require a minimum of 12 infra-units within the geographical supra-units (6 when the infra-unit is the commune). Controls in these specifications include: share of vacant accommodation, unemployment, share of people with no diploma, and when it is not redundant with the explanatory, share of migrants and share of working-class people. Over the RFL-Filosofi set, we add an extra control, the mean fiscal income of the household within the IRIS. Reported statistics are weighted by supra-unit (ZE) total population. Standard errors are clustered at the supra-unit level.

Table 50: Full model with alternative upper/lower groups and alternative spatial units (I)

Upper gr.	Lower gr.	M1/2 thr.	Method 1			Method 2			Obs.
			F-stat	d	t-stat	F-stat	d	t-stat	
			—INSEE's Census - ZE / IRIS—						
Natives	Migrants	10.4/11.6	4.6	-2.29	2.44	4.78	-3.84	4.36	16271
Natives	Work.-clas. migr.	4.1/5.2	8.2	-2.2	3.7	8.3	-3.1	5.7	id.
White-col.	Work.-clas.	18.8/23.9	5.4	$+0.1$	0.67	7.8	-0.5	1.63	id.
White-col.	Work.-clas. migr.	3.1/4.2	4.9	-0.03	0.2	7.6	-0.09	0.2	id.
Intermediate	Work.-clas.	16.3/19.9	2.7	$+0.55$	1.37	2.6	-0.33	0.84	id.
Intermediate	Work.-clas. migr.	3.1/5.6	2.7	-0.67	3.9	2.8	-0.68	4.1	id.
			—INSEE's Census - AAV / IRIS—						
Natives	Migrants	8.2/10.4	4.9	-3.05	4.8	5	-3.69	5.9	12903
Natives	Work.-clas. migr.	2/3.4	3.2	-5.1	6.5	3.1	-3.6	4.9	id.
White-col.	Work.-clas.	27.1/31.2	2.3	-2.1	3.4	2.2	$+0.67$	1.4	id.
White-col.	Work.-clas. migr.	2.3/6.1	3.7	-0.17	1.1	3.3	-0.26	1.7	id.
Intermediate	Work.-clas.	16.3/19.9	2.7	$+0.55$	1.37	2.6	-0.33	0.84	id.
Intermediate	Work.-clas. migr.	2.9/3.4	3.9	-1.1	5.13	3.3	-0.71	3.41	id.
			—INSEE's Census - UU / IRIS—						
Natives	Migrants	7.3/11.3	4.1	-6.37	2.81	3.9	-6.51	3.61	5084
Natives	Work.-clas. migr.	2.5/5.2	2.1	-8.1	5.12	1.9	-1.68	1.32	id.
White-col.	Work. clas.	18.1/21.7	1.4	$+0.04$	0.13	1.4	$+0.07$	2.4	id.
White-col.	Work.-clas. migr.	2.2/24.1	4.2	-2.36	1.85	4.1	-1.3	1.45	id.
Intermediate	Work. clas.	19.1/26.3	1.8	-4.7	0.67	1.7	-26.2	4.7	id.
Intermediate	Work. clas. migr.	2.3/5.4	1.9	-1.19	1.1	1.9	-1.6	2.24	id.
			—INSEE's Census - ZE / Commune—						
Natives	Migrants	7.52/9.66	7.4	-0.23	0.41	7.6	-3.39	5.43	11587

Over the RFL-Filosofi set, available group labels include the ten deciles of the national structure of fiscal income.

Table 51: Full model with alternative upper/lower groups and alternative spatial units (II)

Upper gr.	Lower gr.	M1/2 thr.	Method 1			Method 2			Obs.
			F-stat	d	t-stat	F-stat	d	t-stat	
			—RFL-Filosofi - ZE / IRIS—						
Decile 2	Decile 1	14.4/15.81	8.7	-0.29	2.7	8.9	-0.56	4.4	5205
D3	D1	17.6/13.7	10.2	$+0.1$	0.93	10.7	0.04	0.38	id.
D4	D1	15.37/12.6	11.1	-0.46	3.7	10.8	-0.34	1.53	id.
D5	D1	12.3/14.8	8.7	-0.4	4.1	8.5	-0.59	5.2	id.
D6	D1	11.6/13.8	8.9	-0.68	6.2	8.8	-0.45	2.9	id.
D7	D1	10.6/14.1	8.5	-0.53	4.5	11.9	-0.69	9.2	id.
D8	D1	9.4/10.9	7.1	-0.64	4.1	13.7	-0.71	8.42	id.
D9	D1	11.31/11.7	6.9	-0.61	6.8	12.4	-0.62	7.46	id.
D10	D1	6.6/20.1	19.6	-0.41	1.41	20.9	-0.84	4.14	id.
D6 to 9	D2 to 5	40.1/37.5	17.12	-0.43	1.3	17.1	-0.63	1.81	id.
			—RFL-Filosofi - AAV / IRIS—						
D6 to 9	D1	9.4/10.3	12.8	-1.26	3.1	12.9	-1.74	4.2	4746
D2 to 5	D1	13.8/19.4	4.2	-1.5	3.8	4.6	-0.6	1.15	id.
D10	D1	6.7/18.2	7.7	-0.65	4.3	7.6	-0.5	3.5	id.
D6 to 9	D2 to 5	44.1/46.2	8.1	-0.2	0.56	7.6	-0.5	3.4	id.
			—RFL-Filosofi - UU / IRIS—						
D6 to 9	D1	9.33/9.4	12.8	-2.37	4.67	12.9	-0.94	2.28	3348
D2 to 5	D1	14.3/38.7	11.9	-0.23	0.35	11.5	-1.79	2.8	id.
D10	D1	14.4/22.5	6.4	-0.29	1.1	6.3	-1.41	2.3	id.
D6 to 9	D2 to 5	28.9/36.8	14.5	-1.53	-4.67	13.9	-0.03	0.04	id.
			—IRC - AAV / Commune—						
D6 to 9	D1	7.6/8.1	4.8	-0.85	0.14	5.6	-1.56	2.51	1930
D2 to 5	D1	7.7/11.4	5.7	-0.33	1.1	5.8	-0.07	0.18	id.
D10	D1	7.3/7.6	6.2	-0.55	2.25	5.9	-0.59	2.9	id.
D6 to 9	D2 to 5	37.6/32.5	6.6	$+0.15$	0.34	6.7	-0.36	0.76	id.

Towards an alternative strategy based on incomes

More interesting, the RFL-Filosofi allows us to search for tipping behaviour based on income. For these *income-based*-thresholds, we'll define the upper group as the *Middle 40%* (these households, the disposable income of which is comprised between decile 5 and 9 of the national distribution of income) and the lower group as the *Bottom 10%* (these households, the fiscal income of which is below decile 1 of the national distribution of income)[82]. We did not define those class arbitrarily: as we show in our annex, these two classes are the only one on which consistent tipping is identified[83]. The rationale behind this reaction seems relatively straightforward: To sum it up :

- The bottom 50% are trapped into an unwilling sedentarity. When the social conditions of their district starts to deteriorate, they are unable to move away. We detect some departure trends in districts with very high levels of underprivileged population, but no tipping or discontinuity. Constrained on the credit market, working-class families do not have the financial leverage to back an individualistic-optimizing residential strategy;
- The top 10% barely react to the rise of local migrant or underprivileged population. They are living in relatively preserved districts. Besides, these groups have close control over local politics, and are able to influence housing policies to maintain population homogeneity[84].
- The middle 40% are the only ones which react with tipping. They do not have enough control on local policies to prevent the social evolution of the neighbourhood, but sufficient leverage to leave when conditions deteriorate.

As we see in table 52, thresholds based on income are of very similar magnitude than thresholds based on origin, around the national average of each lower group. However, standard deviations are far lower for *income*-based estimates, and cross-correlation less robust; tipping based on income seems more systematic, less concentrated in certain areas (estimated *origin*-base-thresholds are often abnormally low in areas with high conservative or far-right vote shares):

Table 52: *Income*-based-tipping - Estimated tipping points

	2001-2011		2011-2019	
	Structural break (1)	Fixed point (2)	Structural break (3)	Fixed point (4)
Mean	11.52%	13.44%	11.85	13.71
SE	5.19	7.17	6.12	8.65
Without identified threshold	0	0	0	0
Correlations				
2001-2011 Structural break	1.00			
2001-2011 Fixed point	0.37	1.00		
2011-2019 Structural break	0.11	0.12	1.00	
2011-2019 Fixed point	0.06	0.11	0.26	1.00

Points are expressed in share of bottom 10% pop. in district. Summary stats are unweighted.

Table 53 provides estimates for the tipping behaviour around *income*-based estimated thresholds. For each method, we provide the average drop in the middle-class population (those belonging to the national middle 40%) when the share of underclass tax units (those belonging to the national bottom 10%) is beyond a area-specific tipping threshold computed by our two methods.

The zone of Lyon provides a good example of income-based tipping dynamics. Among popular suburbs at the East of the city, the most deprived QPV still maintain a relatively stable middle-class, which seems unwilling to leave the district. At the heart of Bron-Parilly, the middle 40% account for 20.2% of the population in 2001, and there are only 8 departures coming from that group over the next decade. The sharpest declines in Δu are to be found further South, in individual houses districts which are adjacent to a well known QPV. One very symbolic example is found in Vénissieux - Les Minguettes, a priority district known to be the birthplace of one of the main antiracist activist movement of the late 1980s. In the IRIS lying at the heart of that suburb, the bottom 10% make out 29.4% of the population, and 33.1% of the population is not born in France ; but the sharpest tipping behaviour from the middle 40% are found in nearby single-home districts: (Charreard and Chassagnon with resp. -6.9 and -6.7pp over 2000-2010) which have remained relatively mixed neighborhoods (with l_{2001} at respectively 8.9 and 9.2%).

[82]The RFL-Filosofi dataset provides the deciles of fiscal income for the whole country, and for almost all IRISes which lie within a urban unit. It is then possible to use the interpolation method pioneered by [Blanchet, Fournier, and Piketty 2017] to determine the share of each national group within each district. To give but one example, over the whole zone of Lyon, the bottom 10% make out exactly 10.04% of the population, the middle 40%, 40.86. But in the most deprived priority districts, the figures are almost reversed: in Bron-Parilly, the bottom 10 provide 30% of the local population, the middle-class, 11%.

[83]Here again, we tested a wide range of specifications, the most important ones reported in our annex. The main outcome of that exercise is that tipping is exclusively observed coming from the middle 40% as they react to the bottom 10% population. The middle-class does not react with tipping to the dynamics of any other group (be it defined by income, SES, place of birth, or interaction between these last two); conversely, the bottom 10% population triggers no tipping from another group, whatever its definition.

[84]The list of communes which do not respect the ceilings of local public housing mandated by the 2000 SRU law, recently updated by [Ministère chargé du logement 2020], is almost comprised of communes with the highest of local population earning more than decile 10, if not centile 100 of the national structure of income

Table 53: *Income*-based-tipping - Regression discont. model for middle 40% pop. change around the tipping pt

	Dependent var.: Change in middle-class population in the district from $t-10$ to t					
	Method 1 - Structural break			Method 2 - Fixed point		
	Base	F.E.	Full	Base	F.E.	Full
	(1)	(2)	(3)	(4)	(5)	(6)
2001-2011 decade						
Beyond tipping point (coef. d)	-3.38pp***	-2.22pp***	-1.93pp***	-3.42pp***	-2.06pp***	-1.75pp***
SE	(0.63)	(0.54)	(0.64)	(0.49)	(0.54)	(0.64)
Observations	5205	5205	5205	5205	5205	5205
R^2	6.4%	25.7%	29.1%	7.6%	25.7%	29.1%
F-stat	70.6	8.6	9.8	8.97	8.57	9.81
2011-2019 decade						
Beyond tipping point (coef. d)	-2.85pp***	-3.46pp***	-2.46pp***	-2.81pp***	-3.93pp***	-2.39pp***
SE	(0.36)	(0.35)	(0.34)	(0.34)	(0.35)	(0.33)
Observations	4442	4442	4442	4442	4442	4442
R^2	8.7%	21.9%	33.1%	5.2%	22.3%	33%
F-stat	126.2	9.8	16.7	71.7	10.1	16.6
ZE fixed effects		X	X		X	X
Controls			X			X

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: The un. of obs. is the IRIS. The dep. var. is the growth in middle 40% pop. (def. as these tax units, the fisc. inc. of which lies betwe. dec. 5 and 9 of the nat. struct.) within the IRIS over $t, t-10$, as perc. of the base pop. at $t-10$. The main exp. var. is a dummy equal to one if the share of bottom 10% pop. (simil. def.) is beyond the ZE-specif. estimated tipping pt. All spec. include a quartic polyn. in the devia. of lower-group shares from the tipping pt, plus ZE fixed eff. Controls include unempl. rate, blue-collar sh., sh. of migrants, no diploma sh., vacant accom. sh. and log mean fisc. inc. SE are clustered at the ZE level.

Robustness checks

Higher-order polynomial in control variables

Table 54: Sensitivity of the tipping coefficient to flexible controls (*origin*-based tipping)

	(1)	(2)	(3)	(4)	(5)	(6)
1999-2010	-2.29^{**}	-2.19^{**}	-2.11^{**}	-2.18^{**}	-2.02^{***}	-1.83^{**}
	(0.94)	(0.88)	(0.86)	(0.93)	(0.92)	(0.91)
4-th order polynominal in:						
Unemployment		X				X
Share with no diploma			X			X
working-class share				X		X
Vacant accommodations rate					X	X

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: The specification is still 24, plus quartic polynomials in the indicated variable. We report exclusively the results from the first method of estimation of the tipping points, results for method 2 being similar. Standard errors are in parentheses.

Table 55: Sensitivity of the tipping coefficient to flexible controls (*income*-based tipping)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
2001-2011	-1.93^{***}	-1.31^{***}	-1.39^{***}	-1.29^{***}	-1.61^{***}	-1.45^{***}	-1.47^{***}	-0.82^{**}
	(0.64)	(0.41)	(0.38)	(0.38)	(0.41)	(0.39)	(0.39)	(0.39)
4-th order polynominal in:								
Unemployment		X						X
Logged mean income			X					X
Migrant share				X				X
Share with no diploma					X			X
working-class share						X		X
Vacant accommodations rate							X	X

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: See fig. 54

Distance to a priority district

Table 56: Tipping reaction by distance from nearest priority district (*income*-base tipping)

	Dist. to ZUS		Dist. to QPV		Dist. to new QPV		By nearby spillovers	
	2001-2011	2011-2019	2001-2011	2011-2019	2001-2011	2011-2019	2001-2011	2011-2019
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Main effect: beyond tipping point	-2.73^{**}	-3.68^{***}	-3.21^{***}	-3.32^{***}	-2.27^{***}	-2.35^{***}	-1.63^{***}	-2.83^{***}
	(1.01)	(0.44)	(0.87)	(0.41)	(0.77)	(0.44)	(0.54)	(0.34)
Interacted: Beyond TP × Nearest priority district is 1-3 km away	$+0.92^{*}$	$+1.43^{***}$	$+1.28^{**}$	$+0.76^{**}$	$+0.69$	-0.28		
	(0.48)	(0.34)	(0.55)	(0.33)	(0.45)	(0.44)		
Total tipping effect	-1.81^{***}	-2.25^{***}	-1.93^{***}	-2.56^{***}	-1.57^{***}	-2.64^{***}		
	(0.47)	(0.37)	(0.48)	(0.38)	(0.46)	(0.38)		
Interacted: Beyond TP × Nearest priority district is >3 km away	$+1.79$	$+0.81^{*}$	$+3.61^{***}$	$+0.91^{**}$	$+0.93^{*}$	-0.5		
	(1.15)	(0.41)	(1.05)	(0.44)	(0.55)	(0.39)		
Total tipping effect	-0.94^{*}	-2.87^{***}	$+0.41$	-2.41^{***}	-1.34^{**}	-2.85^{***}		
	(0.49)	(0.39)	(0.54)	(0.44)	(0.46)	(0.37)		
Interacted: Beyond TP × None of neighbours with $l > l^*$							$+1.29^{*}$	$+1.59^{**}$
							(0.73)	(0.66)
Total tipping effect							-0.33	-1.24^{*}
							(1.18)	(0.69)

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: The specification is still 24, now fully interacted with the indicated tract characteristic. In column 7 and 8, *neighbours* are the four closest IRISes (computed in terms of distance from centroid to centroid).

Replication on the mobility series of the INSEE

One critical assumption of our identification strategy revolves around the evolution of the upper group share $\Delta u_{i,t-10,t}$:

- We assume that this evolution is driven exclusively by departures and arrivals in district i;
- Yet it might also be explained through the inner dynamics of the upper group. Maybe native population recedes because of its demographic dynamics; maybe local middle-class population recedes because it has been impoverished over the decade after a negative chock specific to the region.

Actually, the INSEE's Census provides one way to determine which effects predominates. The *Mobilité* series allow us to know, for a very large sample of the national population, in which *commune* one person was living at the beginning of a reference period, and in which *commune* that person lives at the end of that period. That reference interval has changed over time: the 1999 Census recorded mobility over 10 years (i.e. between 1989 and 1999) ; the 2006,2007,2008 Censuses, over 5 years, and the most recent series, since 2013, over one year (there was a break in the series between 2008 and 2013). These sets offer no information on individual incomes, but data about the origin of the person is provided since 2006 onward.

We therefore picked the 2006 issue and replicated the very same *origin*-based tipping identification strategy, with two special differences :

- 1. We are know working on a representative subsample of 19.8 million people (roughly one third of the national population);
- 2. The *supra*-unit is still the ZE, but the *infra*-unit is the *commune*[85]

Except for these two points, there are no other differences with the original strategy. $l_{ic,t}$ is still defined as the local share of the commune's population who is of foreign origin (not born on French territory). We estimate the tipping points with the very same algorithms, and the explanatory is still a dummy equal to one if $l_{ic,t}$ is above the ZE-specific tipping threshold. Our dependent is still the evolution of the native population (i.e. the number of arrivals, minus the number of departures of native individuals, as a share of the base population at the beginning of the reference period). Our reference interval is 2001-2006. Controls are not computed on the subsample of the mobility set; we matched each communal observation with the corresponding value of the control variable in the 1999 Census[86].

The outcome of that robustness exercise is displayed in table 57.

[85] Restriction rules are not relevant there since, even with our 12 infra-units minimum per supra-unit, we can still preserve all communes.

[86] It is a unsatisfactory strategy, but a second best in this context. The Mobility series record individual characteristics, not at the beginning, but at the end of the reference period, meaning if we took the control variables from that set, we would have the values of 2006, not of 2001. Besides, some control variables, like the share of vacant accommodation, cannot be extracted from the set.

Table 57: *Origin*-based-tipping - Regression discontinuity model for change of native share around the tipping point

	Dependent var.: Change in native population in the district					
	Method 1 - Structural break			Method 2 - Fixed point		
	Base	F.E.	Full	Base	F.E.	Full
	(1)	(2)	(3)	(4)	(5)	(6)
2001-2006 semi-decade						
Beyond tipping point (coef. d)	-2.94^{***}	-3.04pp^{***}	-2.95pp^{***}	-3.62pp^{***}	-4.59pp^{***}	-4.48pp^{***}
SE	(0.44)	(0.48)	(0.48)	(0.43)	(0.49)	(0.49)
Observations	17378	17378	17378	17378	17378	17378
R^2	0.5%	4.3%	4.4%	0.6%	4.6%	4.7%
F-stat	18.1	2.51	2.56	21.1	2.71	2.76
ZE fixed effects		X	X		X	X
Controls			X			X
Average estimated T.P.		7.95%			11.03%	

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: The unit observation is the commune. The dependent variable is the growth of the native population (defined as these persons who are born on French territory) within the commune from t to $t-10$, as percentage of the base population at date $t-10$. The main explanatory variable is a dummy equal to one if the share of migrant population (those persons who are not born on French territory) is beyond the ZE specific estimated tipping threshold. All specifications also include a quartic polynomial in the deviation from the district's migrant share from the local tipping point, plus zone fixed effects. The vector of controls, drawn from the INSEE's Census, includes unemployment rate, share of working-class people, share of persons with no diploma, and share of vacant accommodation. Reported statistics are weighted by total population. Standard errors are clustered at the ZE level.

G Complements to the political economy framework

Setting

Table 58: Definition of political aggregates

Pres. elect.	1995	2002	2007	2012	2017
Far-right	J.-M. Le Pen	J.-M. Le Pen	J.-M. Le Pen	M. Le Pen	M. Le Pen
	P. de Villiers	B. Megret	P. de Villiers		N. Dupont-Aignan
Right	J. Chirac	J. Chirac	N. Sarkozy	N. Sarkozy	F. Fillon
		A. Madelin			
Centre-right	E. Balladur	F. Bayrou	F. Bayrou	F. Bayrou	E. Macron
Centre-left	L. Jospin	L. Jospin	S. Royal	F. Hollande	B. Hamon
	D. Voynet	N. Mamere	D. Voynet	E. Joly	
		C. Taubira			
Far-left	R. Hue	A. Laguiller	O. Besancenot	J.-L. Mélenchon	J.-L. Mélenchon
	A. Laguiller	J.-P. Chevènement	M.-G. Buffet	P. Poutou	P. Poutou
		O. Besancenot	A. Laguiller	N. Arthaud	N. Arthaud
		R. Hue	J. Bové		
		D. Gluckstein			

Note: Little candidates whose political identity is ambiguous have been dropped.

Supplementary results

Table 59: Simple model for conservative vote lead in 2nd rounds (2007 & 2012 pres. elections)

	Dep. var.: Right-wing vote lead in 2nd rounds of pres. elec.			
	OLS		Pooled OLS	FD panel
	2007	2012	2007,2012	2007,2012
	(1)	(2)	(3)	(4)
Income regressors				
Ratio T10/B50	1.39***	2.28***	2.78***	−1.34***
	(0.31)	(.42)	(.29)	(.19)
Av. income within city	.00094***	.0011***	.00061***	.00096***
(2021 euros)	(.00018)	(.00018)	(.00013)	(.00008)
Spatial regressors				
Density (p. per km^2)	−.0014***	−.0013***	−.0013***	−.005***
	(.0005)	(.0005)	(.0005)	(.001)
Distance to a metropolis (km)	−.021	−.033	−.029	
	(.024)	(.025)	(.024)	
Share of commuters	−.026	−.0075	−.025	.0274
	(.097)	(.008)	(.02)	(.024)
Economic regressors				
Unemployment rate	−1.02***	−1.18***	−1.75***	−.092
	(.19)	(.23)	(.16)	(.06)
Share of insecure jobs	.065	.15	.27	−.0008
	(.17)	(.17)	(.15)	(.021)
Blue-collar share	.091	.38	.13	.09**
	(.24)	(.28)	(.25)	(.04)
Cultural regressors				
Share of highly educated	−.37***	−.48***	−.49***	−.13
	(.104)	(.12)	(.104)	(.02)
Share of retired people	.23*	.302**	.14	−.038
	(.23)	(.14)	(.13)	(.03)
Share of immigrants	.39**	.135	.28	−.41***
	(.187)	(.18)	(.19)	(.12)
Within catholic realm	−.96	.0199	−.84	
	(2.15)	(2.24)	(2.22)	
R^2	0.26	0.26	0.29	0.83
F-stat	133.4***	133.4***	313.1***	2033.3***

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: The unit of inter. is the *commune*. Elect. data are from the CDSP sets, income data from the IRCOM. Obs. are weighted by tot. Census pop., and SE clustered at the level of the *département*.

Table 60: Simple model for the FN-RN vote shares in 1st rounds (2007 to 2017 pres. elections)

	Dep. var.: Vote share of the far-right in corresponding presidential election (round 1)							
	Pooled OLS				FD panel			
	2002-2007		2012-2017		2002-2007		2012-2017	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Income regressors								
Ratio T10/B50	−.223**	−.048	−.220	−.233**	−.260*	−.127	−.710***	−.485***
	(0.109)	(0.091)	(0.154)	(0.118)	(0.136)	(0.086)	(0.152)	(0.143)
Av. income within city	−.0002***	−.00000	−.0005***	−.00001	0.0001	0.0001	0.001***	0.0004***
	(0.00003)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Ratio T10/B50	0.041	0.054	−.489**	−.419***	0.164*	0.236***	0.058	0.049
× Beyond tipping point	(0.106)	(0.085)	(0.220)	(0.136)	(0.089)	(0.081)	(0.086)	(0.083)
Av. income within city	0.00000	0.00000	0.00001**	0.00001**	−.00000	−.00000*	−.00001**	−.00000*
× Beyond tipping point	(0.00000)	(0.00000)	(0.00001)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Spatial regressors								
Density		−.0004***		−.0001		0.0004		−.004***
		(0.0001)		(0.0002)		(0.0005)		(0.001)
Distance to nearest metropolis		−.013*		−.018**				
		(0.007)		(0.007)				
Share of commuters		0.058***		0.235***		0.048***		0.144***
		(0.008)		(0.032)		(0.011)		(0.025)
Economic regressors								
Unemployment		0.313***		0.339***		−.066		−.050
		(0.059)		(0.081)		(0.080)		(0.043)
Share of blue-collar pop.		0.147**		−.004		0.249***		−.109***
		(0.059)		(0.080)		(0.076)		(0.030)
Cultural regressors								
Share of highly educated		−.135***		−.319***		0.110***		−.150***
		(0.031)		(0.036)		(0.036)		(0.020)
Share of retired people		−.091***		−.268***		0.197***		−.100***
		(0.032)		(0.041)		(0.040)		(0.020)
Late dechristianization		−1.115*		−1.862**				
		(0.665)		(0.800)				
Observations	8,302	8,302	8,302	8,302	8,302	8,302	8,302	8,302
R^2	.348	.470	.261	.535	.738	.771	.621	.668
F Statistic	887.2***	566.3***	584.9***	734.1***	2337.3***	1268.8***	1358.8***	756.7***

*p<0.1; **p<0.05; ***p<0.01

Note: The unit of interest is the *commune*. Electoral data are from the CDSP-Sciences Po sets, income data from the IRCOM. Observations are weighted by total start-of-the-period Census population. Standard errors in parentheses are clustered at the level of the *département*.

Robustness checks

Table 61: An example of interpolation: predicting the structure of the vote with three different combinations of sources − Difference between the first and last deciles of the distribution of electoral districts along the corresponding variable as to the av. conservative vote lead (1988 pres. elec. rd 2) depending on interpolating methods

Interpolation method	Av. fiscal income		Ratio T10/B50	
	Raw	Controlled	Raw	Controlled
Recensement-1990 / ERF-1990	−21.3	−14.7	−28.4	−3.8
Recensement-1990 / ERF-1984	−19.6	−25.2	−32.1	−6.2
Recensement-1982 / ERF-1984	−21.7	−10.2	−22.2	−3.9

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: Electoral data are from the CDSP-Sciences Po sets, income data from the IRCOM base. Controlled series have been factored out of the marginal effect of the Boulard index, the share of working-class people, the share of rural population.

Table 62: Controlling for the choice of spatial unit − Difference between the first and last deciles of some distributions in av. conservative vote lead (2007 pres. elec. rd 2)

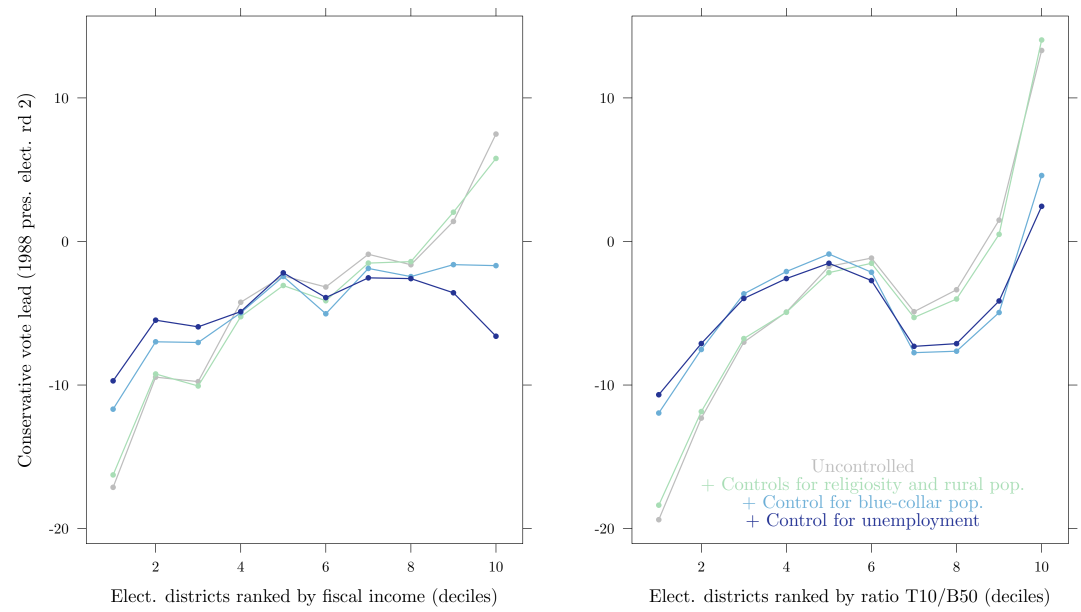
Type of distributions	Av. fiscal income		Ratio T10/B50	
	Raw	Controlled	Raw	Controlled
Distribution of départements	22.2	15.1	10.7	5.9
Distribution of circ. électorales	11.2	5.3	13.1	16.3
Distribution of communes	29.7	18.1	19.1	6.5

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

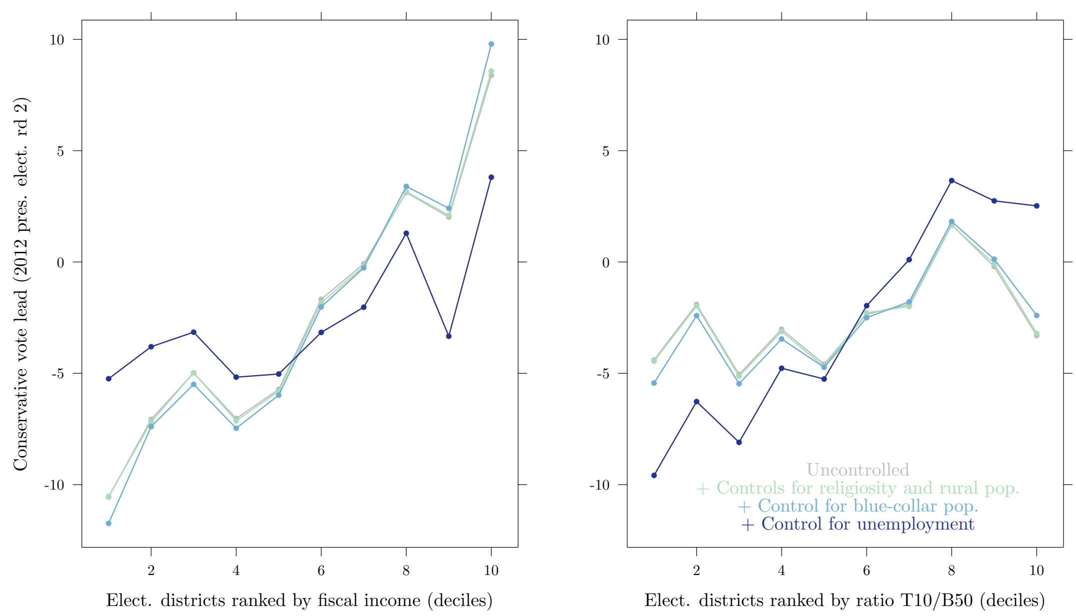
Note: Electoral data are from the CDSP-Sciences Po sets, income data from the IRCOM base. Controlled series have been factored out of the marginal effect of the Boulard index, the share of retired persons, the share of working-class people, and of the unemployment rate.

Figure 54: Income, income inequality, and the left-right cleavage at the regional level − Sensibility to controls

(a) 1988 presidential election rd 2



(b) 2012 presidential election rd 2



Note: The unit of int. is the *circonscription législative*. We use the pol. geo. at the time of the elect.; we follow the CDSP sets for vote shares and con. tab. *communes-circ*. Income data are from IRCOM, or, for elec. prior to 1995, interpol. from the Census (see our annex A). On the y-axis: vote lead of the conserv. RPR-UMP cand. in the 2nd rd of cor. pres. elec. On the x-axis, circ. ranked along: 1. Av. fiscal inc. of their inhab.; 2. The w.-av. of the ratio T10/B50 of fisc. inc. of all *communes* within the circ. Raw data is plotted in light grey. We then successively factor out the effect of some major controls (unempl. rate, rural pop. sh., blue-collar pop. sh., and share of the pop. living in a parish with high churchgoing rates as defined by [Boulard 1982]). x-axis quantiles are weig. by nb of vote cast.

Table 63: Sensibility to controls of coefficients plotted in figure 30

	Dep. var.: Vote lead of the conservative candidate (presidential election, 2nd round) in the electoral district											
	1974 presid. elec.						*1981 presid. elec.*					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Av. fisc. income	−0.51	−0.002	0.45	0.85**	2.39**	2.48**	−0.07	0.19	0.68**	0.72**	1.19***	0.13
	(0.28)	(0.05)	(0.47)	(0.43)	(0.97)	(0.99)	(0.23)	(0.24)	(0.29)	(0.31)	(0.32)	(0.49)
Ratio T10/B50		1.81	−1.55	−0.44	5.38	4.86		2.41**	−0.01	0.85	7.16***	4.39**
		(1.74)	(2.21)	(2.13)	(3.62)	(3.67)		(1.05)	(1.46)	(1.39)	(1.68)	(1.31)
Sh. rural pop.			1.13***	0.91**	1.03**	0.89*			1.15**	0.81*	0.29	−0.28
			(0.37)	(0.42)	(0.39)	(0.66)			(0.46)	(0.48)	(0.51)	(0.62)
Religiosity				0.13***	0.13***	0.11***				0.09***	0.09***	0.08***
				(0.03)	(0.03)	(0.02)				(0.03)	(0.03)	(0.02)
Sh. blue-collar pop.					0.92**	0.77					1.38***	0.86***
					(0.44)	(0.51)					(0.31)	(0.27)
Unemployment						−2.36						−2.54***
						(1.52)						(0.82)
R^2	0.015	0.024	0.069	0.172	0.197	0.214	0.001	0.05	0.099	0.185	0.238	0.273

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: Dep. var. is the vote lead of the conserv. candidate (V. Giscard d'Estaing) in the 2nd round of the cor. presidential election. Regres. are weigh. by the tot. Census household pop, and SEs are clustered at the *département*-level. Other parameters similar to fig. 54.

Testing some hypotheses

Local provision of public goods and political outcomes

In raw descriptive statistics, we find little evidence that *départements* with high FN-RN vote shares or high Yellow Vests activity in December 2018 are less equipped or less administered. In first-differences, a rise of the administrative rate (i.e. the total number of civil servants per 1000 inhabitants, State, local and medical personnel included, military personnel excluded) is associated with a rise of the FN-RN vote, even once controlled for basic socio-demographic variables.

We also tried to replicate the strategy of [Dippel et al. 2017] using the provision of public goods as a mediating variable between the import exposure shock and the political outcomes. Actually, if we use the evolution of the administrative rate over 2012-2018 as a dependent in model 3, whether with $\Delta IPW_{1999,2008}$ or $IPW_{2008,2018}$, we get a second-stage with a *p*-value of the Fisher test largely above 0.1. We get the same problem if we use a simple OLS specification regressive the rate on the home ΔIPW, with or without controls. The main coefficient retrieved is positive equal to +0.84 when we use the lagged decade, negative equal to −0.011 when we use the simultaneous decade.

Table 64: Raw correlation between local provision of public goods and some political outcomes (*département*-level)

	Dep. var.: Yellow vest activity (nb of protests per 100k inhab.) in December 2018 or vote shares of M. Le Pen in presidential elections			
	YW act.	Le Pen 2017-r1	Le Pen 2017-r2	Δ Le Pen 2012/17-r1
	(1)	(2)	(3)	(4)
State spending	0.033	0.38	0.25	
(euros per inh., 2016)	(0.03)	(0.28)	(0.21)	
R^2	0.02	0.03	0.03	
Δ Nb of civil serv. per 1k inhab.	0.146^{**}			0.84^{**}
(2012-2018)	(0.07)			(0.35)
R^2	0.07			0.08
Nb of schools per 100k inhab.	0.034^{***}	0.187	0.12	
(2016)	(0.004)	(0.16)	(0.11)	
R^2	0.41	0.14	0.13	
Nb of gen. practitioners per 100k inhab.	0.007	−0.04	−0.04	
(2016)	(0.008)	(0.11)	(0.08)	
R^2	0.002	0.005	0.004	
Nb of proximity shops per 100k inhab.	0.019^{***}	0.136^{*}	0.08	
(2016)	(0.06)	(0.08)	(0.06)	
R^2	0.23	0.13	0.11	
Nb of cultural equip. per 1k inhab.	0.09	-1.59^{**}	-1.18^{**}	
(2016)	(0.11)	(0.78)	(0.52)	
R^2	0.07	0.08	0.06	
Membership in associations	0.126^{***}	0.52	0.28	
(2016)	(0.04)	(0.33)	(0.28)	
R^2	0.11	0.02	0.01	
Time (in min.) to reach a pool	0.033	0.38	0.25	
of local services *(2016)*	(0.02)	(0.28)	(0.21)	
R^2	0.02	0.03	0.03	

Sign. thr. : *p<0.1; **p<0.05; ***p<0.01

Note: The unit of observation is the *département*. Electoral data are from the CDSP-Sciences Po databases, the Yellows vests data has been described in figure 36c. All explanatory variables are from the INSEE's Census or BPE bases. The membership in associations index is proxied by the number of persons within sport associations as a ratio of the total population. Observations are weighted by total Census population. Reported specifications include no control. Reported S.E. are not clustered.

H Spatial econometrics robustness exercises

With almost all our empirical models, when we apply the Moran test to the residuals, we systematically reject the null, meaning there is strong spatial correlation, which is not surprising for the type of phenomena we are studying. Our correction strategy, congruent with what the literature advises in such cases (see especially [Loonis and Bellefon 2018]), consists in modelising our main specifications as a peer effects model *à la* [Manski 1993], distinguishing zones by their level of proximity to one another.

Manski model

Building a the weight matrix

In order to do so, we need a weight matrix. There is a wide debate in the econometric literature about the influence of the choice of the weight matrix on the estimation strategy, though recent research suggests that the sensibility of estimates to this choice is generally overestimated [Lesage and R. Pace 2014]. As a precaution, we shall use four different ones, built with the following options:

- A Delaunay triangulation on the coordinates of the centroids of the zones;
- A matrix based on contiguities between zones[87];
- A matrix based on the 3 nearest neighbours[88];
- A matrix based on distance (when the unit of interest in the ZE, for instance, the weights decrease by the square of the distance, and are equal to zero beyond a radius of 150km);

A spatial overview of the three first methods is provided in figure 55. For all types of matrices, weights are normalised row-wise so that the sum of each line is equal to one. The matrix of weights is always denoted W.

Main model

The Manski framework has been recently transcribed in spatial econometrics, notably by [Elhorst 2010]. In this setting, the model is specified in matrix form as:

$$\begin{aligned} Y &= \rho \cdot WY + X \cdot \beta + WX \cdot \theta + u \\ u &= \lambda \cdot Wu + \varepsilon \end{aligned} \tag{28}$$

Where Y is the matrix of the dependent variable (Δu in our case), X the matrix of explanatory and control variables, and W the weight matrix defined herein above. The interpretation of the coefficients is derived from Manski:

- ρ captures the *endogenous effects* (the retroactive impact of the choice variable; in our main specification 3, it would be a decline of industrial employment in one ZE caused by the industrial decline of neighbouring ZEs);
- θ captures the *exogenous effects* of each explanatory variable (the impact of neighbouring district's characteristics on the choice variable ; here, it would an industrial decline in one ZE driven by the change in some control variable of another ZE);
- λ captures the *correlated effects* (the impact of the wider context);

Choice of the relevant spatial model

It is well know that Manski models cannot be estimated and interpreted directly, and require extra hypotheses about the nullity of at last one of those coefficients.

From now on, we'll be relying on a simple OLS version of model (3), with the evolution of import exposure over the second decade $\Delta IPW_{1999,2008}$ as the main explanatory variable, plus the full set of start-of-the-period values of controls mentioned in table 3; the default dependent for now is the evolution of manufacturing employment within the ZE over 1999-2008 (in pp).

The usual strategy, drawing from [Elhorst 2010] is to start with two robust Lagrange multiplier tests, taking $\rho = 0$ and $\lambda = 0$ as null hypotheses. With our main specification, at a 5% risk, we reject the second but not the first null. In this case, it is advised to estimate separately the framework as a SDM - Spatial Durbin Model (in which we assume that $\lambda = 0$, but allow ρ and θ to be non-null) and as a SEM - spatial error model (based on the hypothesis that ρ and θ are equal to zero). In the ensuing likelihood ratio test, we reject the hypothesis of a common factor between these two models (the null $\theta = -\rho\beta$). We therefore opt for a SDM model.

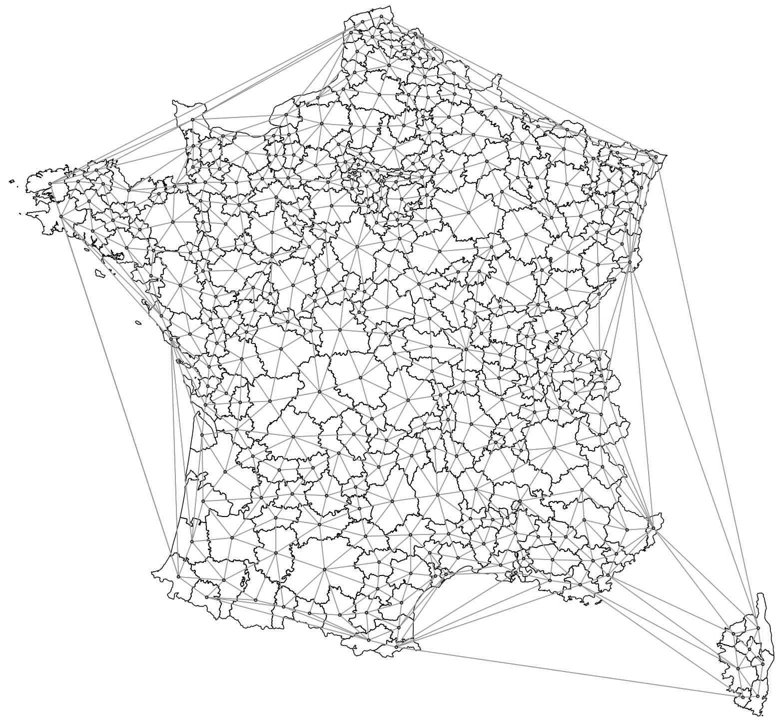
In such a specification, it is impossible to interpret directly coefficients. In a model where $\rho \neq 0$, there is a contamination effect of the industrial decline in one ZE in the decline in the neighbouring ones. Similarly, when

[87] In our computations, two districts are considered to be contiguous if they have at least one point in common; this rule is called the *queen* contiguity in spatial econometrics, a metaphorical comparison with the moves of the queen in a chess game.

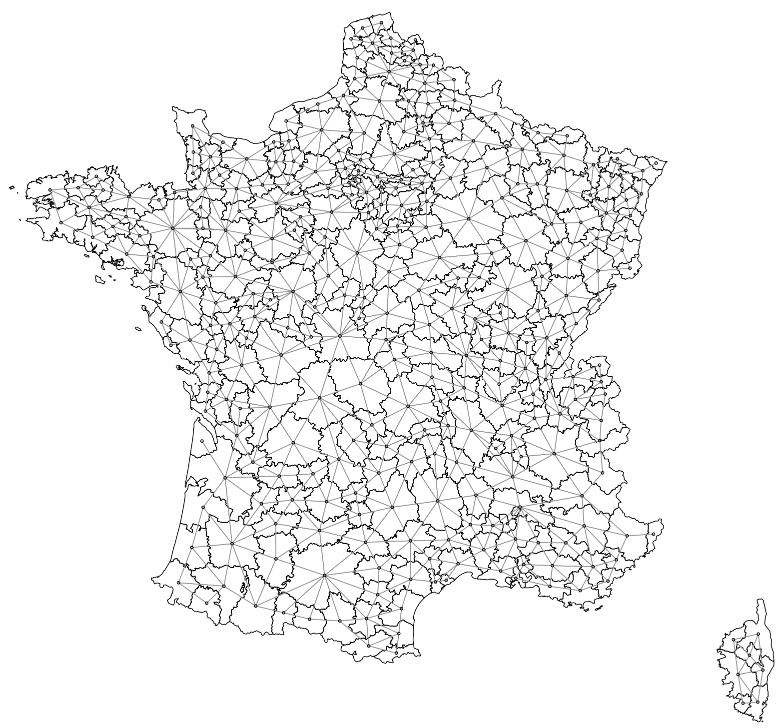
[88] For these three first types of matrix, if one district has n neighbours, each neighbouring district gets a weight of $1/n$, the remaining ones, a zero weight.

Figure 55: Three methods to build a weight matrix for the proximity of ZEs

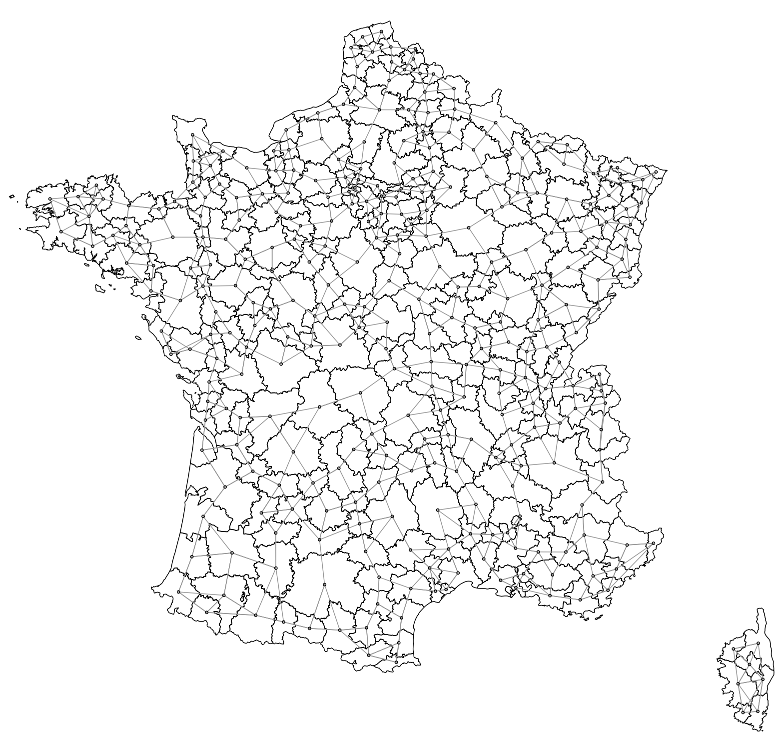
(a) Delaunay triangulation of the centroids



(b) Contiguity - Common border



(c) Closest three neighbours



$\theta \neq 0$, the industrial decline of a zone can be driven by the control variables of neighbouring zones. Formally, if we use subscript r to index our explanatory variables, that specification writes:

$$
\begin{aligned}
& Y = \rho \cdot WY + X \cdot \beta + X \cdot \theta + \varepsilon \\
\Longleftrightarrow & (1-\rho W)Y = X \cdot \beta + WX \cdot \theta + \varepsilon \\
\Longleftrightarrow & Y = (1-\rho W)^{-1} X \cdot \beta + (1-\rho W)^{-1} WX \cdot \theta + (1-\rho W)^{-1} \varepsilon \\
& Y = \sum_{r=1}^{R} (1-\rho W)^{-1} \beta_r X_r + \sum_{r=1}^{R} (1-\rho W)^{-1} W \theta_r X_r + (1-\rho W)^{-1} \varepsilon \\
& Y = \sum_{r=1}^{R} (1-\rho W)^{-1} (I_n \beta_r + W \theta_r) X_r + (1-\rho W)^{-1} \varepsilon
\end{aligned}
$$

The matrix $(1-\rho W)^{-1}(I_n\beta_r + W\theta_r)$ is specific to each explanatory variable r; we can rename it $S_r(W)$. It is of size $N \times N$, providing, for that variable, for each district, its direct impact, and indirect impact on any other district. Formally:

$$
Y = \begin{pmatrix} \Delta u_{1,t-10,t} \\ \Delta u_{2,t-10,t} \\ \vdots \\ \Delta u_{N,t-10,t} \end{pmatrix}, X_r = \begin{pmatrix} x_{1,r} \\ x_{2,r} \\ \vdots \\ x_{N,r} \end{pmatrix}, S_r(W) = \begin{pmatrix} S_r(W)_{1,1} & S_r(W)_{1,2} & \dots & S_r(W)_{1,N} \\ S_r(W)_{2,1} & S_r(W)_{2,2} & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ S_r(W)_{N,1} & S_r(W)_{N,2} & \dots & S_r(W)_{N,N} \end{pmatrix}
$$

Following [Lesage and R. K. Pace 2009], the usual interpretative solution is to rely on two measures: 1. The average of the diagonal coefficient (or equivalently, the trace of the matrix divided by N) which provides the *average direct impact* of one district on itself for one explanatory variable; 2. The average of all coefficients, which provides the average total impact; differenced with the average direct impact, it gives the *average indirect impact*; intuitively, it provides, for one average district, the average impact of the N marginal changes of r in all zones, minus the average direct impact.

The usual estimation strategy relies on Markov chain Monte Carlo methods to approximate the distribution of these two effects; with 1000 repetitions, we can build reliable confidence intervals, which are displayed in table 65.

Results of the estimation of the spatial Durbin model

As we see, with almost all choices of matrices of weights, we find a sizeable negative direct impact of the non-instrumented ΔIPW on the local industrial decline, an impact which is significantly different from zero with a 2.5% risk. As to indirect/retroactive impacts, their significance is critically dependent on the choice of the weight matrix; they become significant when we use a more restrictive definition of neighbourhood.

Table 65: Controlling for spatial autocorrelation

	Dep. var.: Change in total manuf. employment (pp)			
Matrix of weighs	Delaunay	Contiguity Q	Closest 3	Distance
	(1)	(2)	(3)	(4)
Exposure to import competition per worker				
OLS estimates	-1.42	-1.42	-1.42	-1.42
	(0.96)	(0.96)	(0.96)	(0.96)
- *Moran test stat.*	7.42***	8.67***	7.51***	11.13***
Spatial Durbin model				
- *Av. direct impact*	**-3.06**	**-3.03**	**-3.67**	**-3.48**
	[-5.16,-0.87]	[-5.03,-0.89]	[-5.57, -1.75]	[-5.64,-1.33]
- *Av. indirect-retroactive impact*	-2.91	0.53	**-5.99**	**-16.08**
	[-11.46, 5.36]	[-7.85, 8.47]	[-10.93,-0.95]	[-32.39, -0.65]

Note: The main specification in an OLS version of model (3), with the evolution of import exposure over the second decade $\Delta IPW_{1999,2008}$ as the main explanatory variable, plus the full set of start-of-the-period values of controls mentioned in table 3; the dependent is the evolution of manufacturing employment within the ZE over 1999-2008 (in pp)., and rewritten as a Spatial Durbin model (where we allow the value of the dependent variable of one district to be influenced by the values of the dependent and the explanatory variables of neighbouring districts), with four different matrices of weights indicated in each column. We use 1000 iterations of Markov chain Monte Carlo methods to produce an estimation of the average direct and indirect (or retroactive) impact of our main explanatory on the local evolution of the upper-group population. We reported in brackets the confidence interval for the estimated impact, for a 2.5% risk. We put in bold coefficients, the confidence interval of which does not comprise zero.

List of Figures

List of Tables