Time-Varying Agglomeration Economies and Aggregate Wage Growth

Clémence Berson, Pierre-Philippe Combes, Laurent Gobillon, Aurélie Sotura

Discussion Paper

TIME-VARYING AGGLOMERATION ECONOMIES AND AGGREGATE WAGE GROWTH

Clémence Berson, Pierre-Philippe Combes, Laurent Gobillon, and Aurélie, Sotura

SCIENCES PO ECONOMICS DISCUSSION PAPER

No. 2023-07

Time-varying agglomeration economies and aggregate wage growth[a]

Clémence Berson[b]
Banque de France and European Central Bank

Pierre-Philippe Combes[c]
Sciences Po-CNRS

Laurent Gobillon[d]
Paris School of Economics-CNRS and Banque de France

Aurélie Sotura[e]
Banque de France

December 14, 2023

Abstract

We quantify the effects of city agglomeration economies on labour earnings in France over a forty-year period using individual wage panel data. We first delineate cities at every date to consider changes in their footprint over time. We then estimate a daily wage specification that includes time-varying city effects while controlling for observed and unobserved individual heterogeneity. We regress these city effects on agglomeration variables every year, and assess how changes in values and returns to agglomeration variables affect the evolution of daily wages. We find a negligible role for changes in values, but an important role for changes in returns. There is also significant heterogeneity across cities, even among large cities of similar sizes. We propose a theoretical model in which agglomeration economies affect both population and city area. A calibration exercise shows that changes in returns to agglomeration economies are not enough to generate variations in population and city area influencing significantly aggregate labour earnings. This result is consistent with the negligible role of changes in values found in our empirical investigation.

JEL Codes: R23, J31, J61

Keywords: Agglomeration economies, growth, wages

[a]The scientific output reflects the opinions of the authors and does not necessarily express the view of the Banque de France or the European Cental Bank. We are grateful to Clément Gorin for his help with the data and Alban Roger for outstanding research assistance. We also thank Clément Bosquet, Gilles Duranton, Frédéric Robert-Nicoud and Maximilian Von Ehrlich for their useful comments. Access to some confidential data, on which is based this work, has been made possible within a secure environment offered by CASD – Centre d'accès sécurisé aux données (Ref. 10.34724/CASD). Laurent Gobillon acknowledges the support of the EUR grant ANR-17-EURE-0001 and the ORA grant ANR-20-ORAR-0007.

[b]Banque de France and European Central Bank, 31, rue Croix des Petits Champs 75001 Paris, France, e-mail: Clemence.BERSON@banque-france.fr, webpage: https://sites.google.com/site/clemenceberson/home.

[c]Sciences Po, Economics Department, 28, Rue des Saints-Pères, 75007 Paris, France, e-mail: ppcombes@gmail.com, website: https://sites.google.com/view/pierrephilippecombes/home/. Also affiliated with the Centre for Economic Policy Research.

[d]PSE-CNRS, 48 Boulevard Jourdan, 75014 Paris, France, e-mail: laurent.gobillon@psemail.eu, website: http://laurent.gobillon.free.fr/. Also affiliated with the Centre for Economic Policy Research and the Institute for the Study of labour (IZA).

[e]Banque de France, 31, rue Croix des Petits Champs 75001 Paris, France, e-mail: Aureli.e.SOTURA@banque-france.fr.

1 Introduction

Over the last fifty years, there has been a steady growth of cities over the world although cities do not grow at the same rate (Duranton and Puga, 2014). Many countries have seen the rise of urban giants concentrating population and employment, which can generate agglomeration economies. Returns to agglomeration economies may also have evolved with the rise of services industries and the improvement of transports. The aggregate productivity and labour earnings for a given distribution of city sizes may thus vary over time.

In this paper, we study the contribution of city agglomeration economies to labour earnings in France over the 1976-2015 period using wage panel data. Our contribution is fourfold. We first delineate cities every year to properly consider the evolution of city sizes. We then estimate how returns to agglomeration economies on wages evolve over time. We assess to what extent the evolution of aggregate labour earnings is affected by changes in returns to city agglomeration variables and changes in their values. Finally, we model a system of cities with agglomeration economies to study the equilibrium effects of changes in agglomeration economies on labour earnings. We derive and evaluate formulas for both the direct effect stemming from changes in returns to agglomeration variables, and the indirect effect coming from the influence of these changes on values of agglomeration variables.

The evolution of wage disparities across places has been assessed descriptively and studied in a macroeconomic perspective. Bauluz *et al.* (2023) quantify the evolution of wage disparities across local labour markets within and between countries in North America and Western Europe over the last fourty years. Butts *et al.* (2022) focus on the changes in the urban wage premium since 1940 in the US, and find that it decreased until the eighties before stabilizing. Giannone (2022) and Eckert *et al.* (2022*a*) reconcile this empirical evidence with the spatial diffusion of technology and skill-biased technological change that has mostly affected large cities. We depart from this literature by considering a system of cities with varying footprint and employment density, consistently with urban models. We are interested in the impact of changes in city characterictics and their returns on wages.

Our work complements a large literature studying the effects of agglomeration economies on productivity and wages (see Combes and Gobillon, 2015; Ahlfeldt and Pietrostefani, 2019; Duranton and Puga, 2020). In particular, Combes *et al.* (2008) estimate static agglomeration economies by regressing individual wages on city employment density and area taking into account the spatial sorting of individuals on observables and unobservables.[1] Their work has been extended to consider learning effects in cities that may be transferrable to some extent when moving to another city (de la Roca and Puga, 2017; Koster and Ozgen, 2021; Eckert *et al.*, 2022*b*; Card *et al.*, 2023). Contributions in a historical perspective are scarce because data on firms and wages are often missing. An exception is Combes *et al.* (2011) who estimate the effects of agglomeration economies on past value added. Whereas returns to agglomeration economies are usually considered to be constant over time, we evaluate how they evolve over a forty-year period.

[1]See Moretti (2013) and Diamond (2016) for frameworks involving spatial sorting depending on education, and Diamond and Gaubert (2022) for a survey and discussion of spatial sorting and its evolution over time.

Our approach considers changes in the city size distribution which has been extensively studied. According to Zipf (1949), the proportion of cities greater than a given size threshold would be inversely proportional to that threshold, which yields that the larger the size, the smaller the number of cities. Whether this law holds empirically is subject to debate (Gabaix and Ioannides, 2004). Latest evidence provided by Dittmar (2019) shows that it would be verified for Western European countries from the beginning of the 19^{th} century. Several explanations have been put forward to explain Zipf law, including random growth (Gabaix, 1999), and specific distributions of first-nature geography endowments (Krugman, 1996) or entrepreneurial talents (Behrens *et al.*, 2014). We do not assess the shape of city size distribution, but rather consider the impact on wages of its evolution over time as given in our data.

City sizes are curbed by land use constraints and building codes that limit the construction of new buildings (Gyourko and Molloy, 2015). Constraints affect the distribution of city sizes and thus aggregate productivity. Hsieh and Moretti (2019) show that there is an important misallocation of labour across US cities due to restrictions to the new housing supply. Duranton and Puga (2023) show that building restrictions limit growth. Finally, Duranton *et al.* (2015) emphasize the role that local policies in property markets can have on the misallocation of building inputs participating to production. We do not study building contraints *per se*, but their consequences are implicitely considered when studying agglomeration economies through our use of city footprint and density that may be affected.

Changes in returns to agglomeration economies can generate people and firm migration flows between cities that may reinforce spatial wage disparities in the same way urban policies may influence the development of cities (Glaeser and Gottlieb, 2008; Kline and Moretti, 2014; Gaubert, 2018). Indeed, urban policies can foster agglomeration economies in some cities due to increases in population and employment. Such increases can come from the immigration of people and firms from other cities that may experience losses of agglomeration economies. The impact of urban policies then depend on gains and losses incurred by cities. We consider such equilibrium effects when studying the evolution of aggregate labour market earnings since we sum the effects of changes in agglomeration economies across cities.

For our study, we need time-varying delineations of cities to accomodate variations in their footprint due to changes in agglomeration economies, and local productivity and amenity shocks. We delineate cities every year from 1976 onwards using the dartboard approach proposed by de Bellefon *et al.* (2021). For that purpose, we consider a grid of metropolitan France in 200m $\times$ 200m pixels for which we compute the building volume density any given year from information on building footprint, height and construction year from CEREMA and BDTOPO data. We consider that pixels are "urban" if their building volume density is significantly larger than that obtained when reallocating randomly building volume densities over the territory. We define urban areas as sets of contiguous urban pixels, and determine whether they have "cores" defined as pixels significantly denser than random among the set of urban pixels. We define cities as urban areas with at least one core. Overall, our approach allows us to consider expansions, absorptions and fusions of cities over time. Interestingly, we show that large cities may have different fate. In particular, Lille that concentrated mining, textile and heavy industies has not evolved much, whereas Marseille on the French Riviera has expanded a lot and absorbed many urban areas with and without cores during our period of interest. This is consistent with an increasing interest for consumption amenities, especially nice weather and coasts (Rappaport and Sachs, 2003; Rappaport, 2007, 2009).

Equipped with our yearly delineations of cities, we conduct estimations of time-varying agglomeration economies using administrative panel data (*Déclarations Anuelles des Données Sociales*-DADS) over the 1976-2015 period. In a first step, we regress individual daily wages on city-year fixed effects while controlling for individual observed and unobserved heterogenity, and industry. In a second step, we regress estimated city-year fixed effects on city variables (i.e. density, area and market potential). Importantly, we consider that coefficients of these variables may vary over time. We find that the elasticity of wages with respect to city density (resp. area) increases over time from 0.011 to 0.042 (resp. 0.007 to 0.022). These estimates are barely affected when instrumenting agglomeration variables with city historical characteristics and geological features. Using a log-wage growth decomposition, we also show that changes in returns have a large effect on wage growth, whereas changes of values for agglomeration variables do not. Indeed, even if cities evolved, changes are slow compared to cross-section disparities across cities.

Finally, we propose a theoretical model to evaluate the indirect effects of changes in returns to city density and area at equilibrium. Indeed, changes in these returns affect agglomeration economies and thus the respective attractiveness of cities depending on their characteristics. As a result, they trigger migrations across cities that impact local levels of density and area, and thus agglomeration economies. More precisely, the modelling consists in incorporating the effects of agglomeration economies on wages in a system of monocentric cities. Agglomeration economies are specified as a Cobb-Douglas function in city density and area, in line with our empirical specification. The model cannot be solved analytically, but we are able to conduct comparative statics for wages with respect to returns to agglomeration variables and to quantify not only the direct effect of changes in these returns, but also the indirect effect stemming from the influence of these changes on the values of agglomeration variables. With a calibration exercise, we show that the indirect effect is rather small, which is consistent with the negligible role of changes in values of agglomeration variables found in our empirical investigation.

The rest of the paper is as follows. Section 2 explains how we delineate cities every year. Section 3 presents our data and stylized facts on the distribution of wages and agglomeration variables. Section 4 details our empirical strategy and Section 5 discusses our results. Section 6 proposes robustness checks when changing the definition of cities. Section 7 presents our model of a system of cities involving agglomeration economies as well as its quantification. Finally, Section 8 concludes.

2 Delineating cities over time

Time variations in returns to agglomeration econonomies, as well as city productivity and amenity shocks, affect the relative attractiveness of cities over the territory. Consequently, city housing demands are affected, and this puts upward or downward pressure on city footprint and dwelling sizes, which both affect city density (see model in Section 7). We thus need delineations of cities that vary over time to capture changes in city sizes but have inter-temporal consistency in the way they are determined. Their construction also needs to be transparent enough to allow meaningful interpretations of changes in city physical size and population. In France, an official delineation of cities was introduced in 1990 only, and changed in 1999 and 2010 using rules that are quite opaque. These delineations do not cover our whole period of study and inter-temporal consistency is far from granted. Consequently, we rather rely on our own delineations of cities based on building continuity that vary over time. For that purpose, we apply the statistical approach proposed by de Bellefon *et al.* (2021) to a dataset obtained from the match between the 2020 land files (*Fichiers Fonciers*) and 2020 3D modelling of territory and infrastructures (*BDTOPO*).[2]

2.1 Data sources

Land files include detailed information on land register, land use and property rights. In particular, it gives for each parcel, its cadastral identifier and information on the building that sits on it. For each building, information includes the construction year, the footprint and height, but the last two variables are plagued with measurement errors. Hopefully, BDTOPO provides both footprint and height for buildings with much more accuracy (but not the construction year). We thus match buildings in the land files and BDTOPO data. The matching procedure is the following. From the parcel identifier in land files, we get the parcel limits from a parcel shapefile. We then consider that a building in BDTOPO data belongs to a parcel in the land files if its centroid is located within its limits. When it is the case, we can match it with the building located in the parcel in the land files. For a building matched this way, we can then use the construction year from land files, and the footprint and height from BDTOPO data. In fact, we will use the volume of buildings, defined as footprint times height.

Using the construction year, we can determine which buildings are on the territory every year over the 1975-2015 period. There is a bias due to some buildings being destroyed over time and not appearing in the land files. At the same time, we do not expect this bias to be large over the last 50 years. Moreover, as we will see, our statistical procedure to determine cities smooths measurement errors.

For each year, it is possible to divide the French territory into 200 x 200 meter squares such that, for each square, we have the total footprint and volume for buildings present that year. We complement these data with

[2]The 2020 land files are provided by the Center for Studies and Expertise on Risks, Environment, Mobility and Planning (*Centre d'études et d'expertise sur les risques, l'environnement, la mobilité et l'aménagement*, ie CEREMA). The 3D modelling of territory and infrastructures is provided by the French Institute of Geography (*Institut National de l'Information Géographique et Forestière*, i.e. IGN-F).

census municipal populations in 1975 and 2015. In these years, the population in a municipality is allocated to the 200m x 200m squares (labelled as "pixel" below) that covers it proportionally to the total volume of their buildings. Gridded population data will be used to characterise delineated cities at the beginning and the end of our time period.

2.2 Algorithm to delineate cities

We then apply the statistical approach proposed by de Bellefon *et al.* (2021) to delineate cities every year. This delineation is conducted from building volume density such that it captures both building continuity and intensity. Of course, variants are possible using building footprint or population density.

The approach can be decomposed into five steps. The first one consists in constructing a raster of 200m x 200m pixels with values of building volume density for mainland France. In the second step, we smooth the building volume density of each pixel. In particular, this step smooths measurement errors in the land files due to the demolition of buildings. In the third step, we compute and smooth counterfactual building volume densities for each pixel by randomly redistributing observed pixel building densities (with replacement) across buildable pixels. In the fourth step, we consider that a pixel is urban if its actual smoothed density is above the 95^{th} percentile of smoothed counterfactual densities computed for that pixel. In the fifth step, we determine the urban areas as sets of contiguous urban pixels. They are named after the most populated municipality with which they overlap. As we have location information at the municipality level in our wage data, we need delineation of cities that aggregate municipalities rather than squares. In practice, we consider that a municipality is in an urban area if more than 50% of its population is located in urban area pixels. Our urban areas are thus aggregates of urban municipalities, and the physical area of an urban area is the sum of municipal areas.

This approach leads to many urban areas, some of which are small and not particularly dense in terms of buildings or population. We will therefore consider only urban areas which a dense core that we label "cities". The procedure to determine whether urban areas exhibit at least a core is the following. Within all the urban areas previously delineated, we generate counterfactual distributions of building volume densities by randomly redistributing buildings across pixels. We then smooth as previously both the observed and counterfactual distributions, and we consider that a pixel belongs to an urban area core if its smoothed observed density is above the 95^{th} percentile of smoothed counterfactual densities. It is then possible to consider only urban areas which have at least one pixel that belongs to a core.

2.3 Description of delineated cities

Cities are delineated every year over the 1976-2015 period. To use our delineation algorithm, we first need to determine non-buildable pixels. For that purpose, we determine the 99^{th} percentile for the proportion of pixels covered by water, elevation and the average slope. We then consider that all pixels with a proportion of water, an elevation or a slope above their 99^{th} percentile are non-buildable (see de Bellefon *et al.*, 2021, for more details). For the smoothing procedure, we use a bi-square kernel with a 2.1-kilometer bandwidth.[3] More details on bandwidth choice are provided in Section 6.

In Table A.1, we report descriptive statistics on delineated urban areas at every census year. Interestingly, the number of delineated urban areas has decreased over time. There are two possible reasons for that. The first reason is that the delineation of cities in a given year is based on a relative criterion, i.e. it depends on building densities over the whole territory in that year. If the territory is not dense, it is likely that any place with some concentration of buildings will emerge as an urban area. Reciprocally, if the territory is quite dense, concentration needs to be quite large for an urban area to be detected. Average density over the territory has increased over time due to a population increase and new housing construction. As a consequence, the number of urban areas might have decreased. The second reason is that population has concentrated over time in some urban areas and more buildings have been built at the peripheries of these urban areas. In some cases, these peripheries have ended up touching other urban areas which have been absorbed.

When using our year-specific delineations, the number of urban areas is very large, around $2,700-3,100$ every census year. Whatever the year, the 95^{th} percentile of population size looks quite small and most cities do not seem relevant due to their small size. For that reason, we will focus on cities, i.e. urban areas with at least one urban core. Table 1 shows that there are much fewer of them since their number is around $290-310$ every census year, with a non-linear evolution over time. Their 95^{th} percentile of population size is way larger, at around 250,000. Also importantly, the median population size is still sizable at around 27,000 every census year.

[3]The bandwidth is such that, for each pixel, smoothing takes into account 10 pixels on each of its side.

Table 1: Descriptive statistics on delineated cities (ie. urban areas with cores)

Var	Min	p25	p50	mean	p75	p95	Max
Panel A : 1975 (286)							
Population	172	14605	29392	108177	62066	257064	9083917
Area	4.68	42.83	65.40	113.63	101.52	272.92	2911.36
Density	10.04	243.65	519.73	652.52	908.69	1671.53	3120.16
Panel B : 1982 (311)							
Population	117	12385	26834	103116	57399	257262	9372229
Area	3.48	42	63	116.83	104.08	318.60	3377.16
Density	8.02	223.74	477.28	571.06	811.57	1352.46	2775.18
Panel C : 1990 (309)							
Population	153	13077	27987	108650	60157	266602	9862985
Area	8.32	43.92	70.64	126.52	113.56	327.55	3519.44
Density	7.49	234.32	466.01	553.16	784.48	1315.58	2802.43
Panel D : 1999 (303)							
Population	198	13511	27140	114972	62875	277733	10116852
Area	10.12	45.56	74.28	134.97	120.02	339.76	3647.48
Density	8.14	239.61	458.91	545.06	733.31	1297.33	2773.66
Panel E : 2006 (297)							
Population	81	14776	28474	122495	66407	295295	10666306
Area	5.60	46.12	73.48	140.88	120.60	425.29	3696.40
Density	6.91	247.74	457.40	547.12	750.77	1245.85	2885.59
Panel F : 2015 (308)							
Population	187	13068	27327	120231	66156	283045	11078022
Area	10.80	47.02	75.86	146.20	130.87	418.69	3759.88
Density	5.63	244.07	408.28	485.42	663.86	1091.40	2946.38

Note: We consider year-specific delineations when reporting descriptive statistics on cities in a given year. Note that some delineated cities are small in terms of population and corresponding places are not considered as cities according to main adminstrative definitions. At the same time, this does not matter for us since the smallest cities do not include any observation of our DADS dataset and our city regressions are weighted by the number of workers (which means that small cities play a negligible role).

In Figures A.1 and A.2, we represent graphically the four largest cities in 1975 and 2015 to get a visual of their evolution over the period. Cities are represented in blue stripes (with cores in deep blue), and urban areas without core are represented in red. We also draw boundaries of municipalities "included" in cities (i.e. such that at least 50% of their population is included in a city). Figures A.1.a and A.1.b show that Paris has grown significantly, and has absorbed several urban areas without core. This is also the case for Lyon (Figures A.1.c and A.1.d) that has absorbed not only urban areas without core, but also cities. Marseille is an interesting case because it has evolved a lot between 1976 and 2015 (Figures A.2.a and A.2.b). Its area has increased a lot with the absorption of cities, urban areas without cores, and rural areas. In particular, it is possible to check that Marseille absorbed Aix-en-Provence in 1989. Finally the area of Lille has not evolved much (Figures A.2.c and A.2.d). This is not suprising since its growth has been slowed down by the decline of mines, and heavy and textile industries.

3 Data

3.1 Wage dataset

The main data used in the estimations is the Annual Social Data Declarations database (*Déclarations Annuelles des Données Sociales*, DADS) for the 1976-2015 period. These administrative panel data are collected from employers and self-employed in France for pensions, benefits and tax proposes (see Combes *et al.*, 2008, for more details). Information is available at the job level, i.e. at the cell defined as an individual in an establishment in a given year.

Data include an individual identifier, the identifier of the municipality where the workplace is located, the socio-professional category of the worker, the part-time/full-time status, the number of working days, the net wage (deflated by the consumer price index such that it is in constant euros) and the industry at the 4-digit level. Our main measure of earnings is the daily wage computed as the ratio between the net wage and the number of working days. We use an aggregate industry classification in 3-digit industries (i.e. NAF114). There are significant changes in the industry classification in 1993 and 2009 such that it is not possible to obtain a codification that is stable over time. Consequently, we use distinct industry classifications before 1993, between 1993 and 2003, and after 2009.

Information is available for jobs in manufacturing and services in the private and semi-public sectors for all employees born in October of even years over the 1976-2001 period, and for all employees born in October of even years or the 2^{nd}, 3^{rd}, 4^{th}, 5^{th} of January, or the 1^{st}, 2^{nd}, 3^{rd}, 4^{th} of April, July or October over the 2002-2015 period For our analysis, we restrict our attention to individuals aged 18-65 born in October of even years to avoid overweighting the recent period. We only keep their main job every year which is defined as the job with the highest net wage. We retain full-time jobs in the private sector such that duration and net wage are strictly positive. Resulting jobs in the agriculture and fishery industry and in the banking industy are excluded.[4] Our final

[4]Agriculture and fishery industry is normally not covered by the data. We exclude the remaining workers in that industry. An issue for the banking industry is that data are declared at the regional level rather than at the establishment one at the beginning of the

sample includes 18,619,578 observations.

Our sample is used to compute employment densities and market potentials for every city and every year. For a given city, employment density is the ratio between the employment computed from our sample and footprint. The market potential is the sum of employment densities in other cities weighted by the inverse of distance (Harris, 1954), the city itself being excluded from this sum. The distance between two cities is that between their centres, where the center of a city is defined as the barycentre of city hall coordinates for municipalities within the city, weighting by municipal employment. Note that this distance depends on time because the delineation of cities depends on time.

3.2 Stylized facts

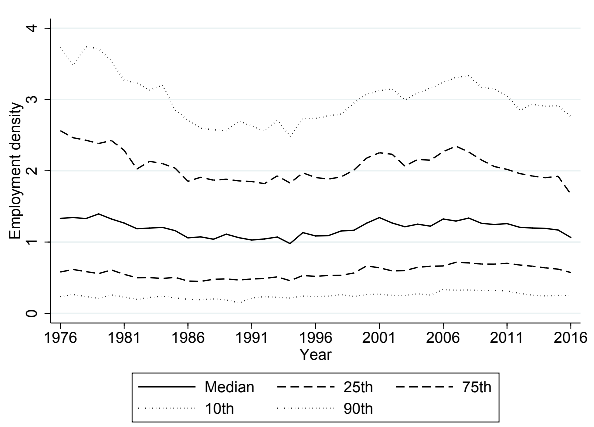
We now provide descriptive statistics on cities over the 1976-2015 period. We represent moments of their distributions for employment density, area, market potential and wages over time (see Figure 1). As shown by Figure 1.a, density exhibits cycles and does not have any increasing trend. This is not surprising since, for growing cities, even if centers may become denser, new peripheries resulting from the absorption of rural and urban areas are likely to be less dense. Hence, the evolution of employment density can go both ways for a given city. This is well illustrated by Figure 2.a that represents the evolution of density for the four largest cities. In particular, density for Marseille increases before decreasing a lot as the city grows and absorbs other urban areas.

The growth of area for large cities is shown by Figure 1.b as the 75^{th} and 90^{th} centiles of area increase in a sizable way between 1976 and 2015. Still, there is heterogeneity among large cities. Whereas area grows a lot for Paris, Lyon and Marseille as shown by Figure 2.b, Lille area does not evolve much during our period of interest. Market potential exhibits cycles consistent with those of density for centiles of the distribution (Figure 1.c) and for the four largest cities (Figure 2.c). Finally, wages (in constant euros) increase over time for centiles of the distribution (Figure 1.d) and for the four largest cities (Figure 2.d) without any specific trend in terms of wage disparities.

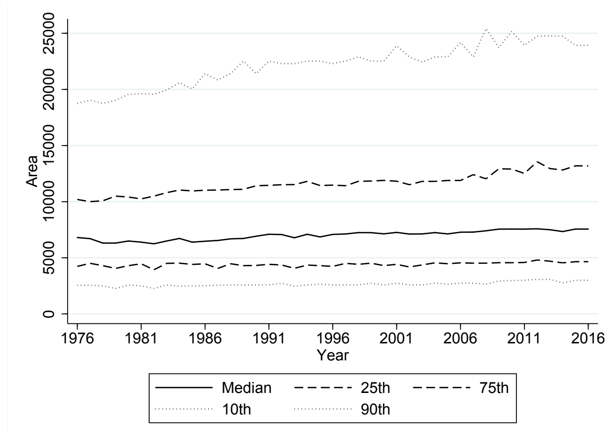
panel. This is why we drop that industry.

Figure 1: Moments of city distributions for our delineated cities

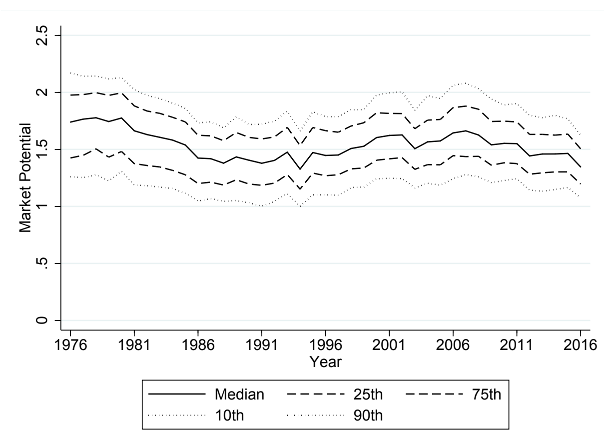
1.a: Employment density



1.b: Area



1.c: Market potential

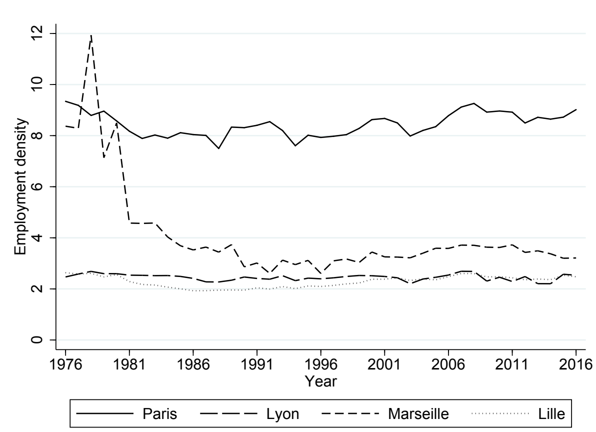


1.d: Average wage

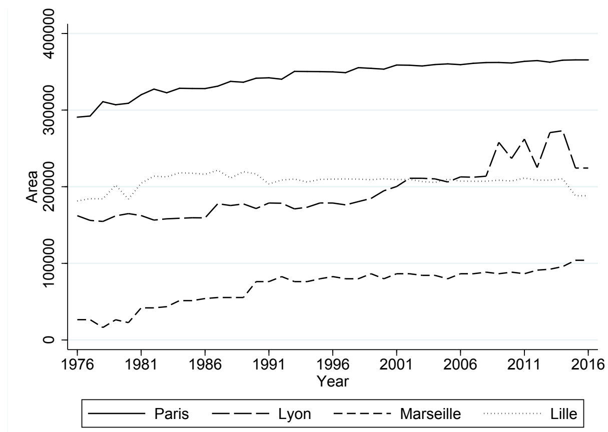
Note: Employment density is computed as the ratio between DADS employment and area (in km^2). Market potential is the sum of employment densities divided by distances, excluding the city itself. Wage is in constant euros.

Figure 2: Evolution of city variables for Paris, Lyon, Lille and Marseille

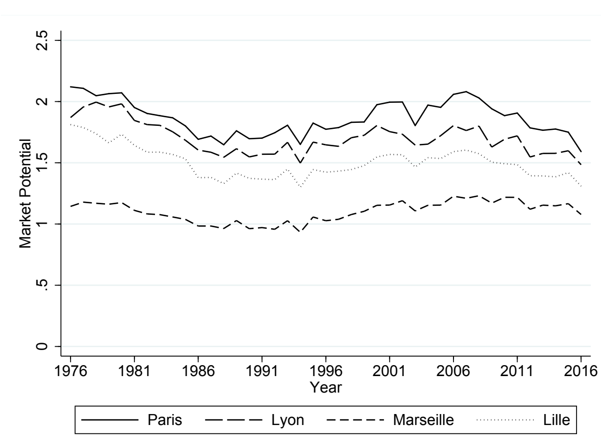
1.a: Employment density



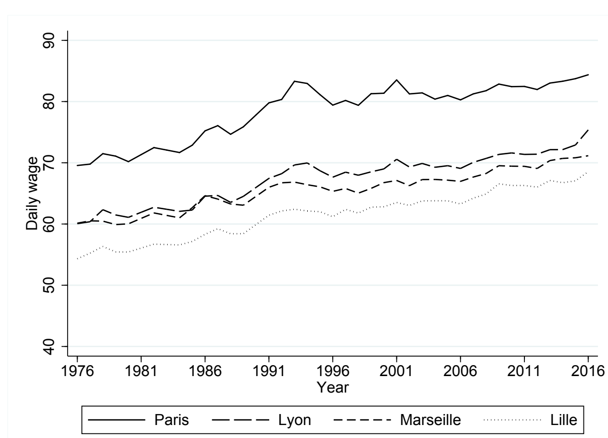
1.b: Area



1.c: Market potential



1.d: Average wage



Note: Employment density is computed as the ratio between DADS employment and area (in km^2). Market potential is the sum of employment densities divided by distances, excluding the city itself. Wage is in constant euros.

4 Empirical strategy

Equipped with our time-varying delineations of cities, we quantify how changes of agglomeration economies affected labour earnings over the past fourty years in France using the following empirical strategy. We first estimate a log-wage specification that involves city-year effects, while taking into account the sorting of individuals with respect to observed and unobserved characteristics across cities. City-year effets are then regressed on our agglomeration variables. We finally turn to a decomposition of the evolution of average daily wage into changes of composition effects, agglomeration variables and returns to these variables.

4.1 Specification

Denote $\mathcal{C}$ the set of cities, $\mathcal{U}$ the set of urban areas without core and $\mathcal{R}$ the rural area. In our empirical analysis, we distinguish among the types of locations (city, urban area without core or rural area) where individuals work.[5] For those in cities, we consider the specific effect of the city where the workplace is located. Our empirical specification is the following:

$$\ln w_{i,t} = X_{i,t}\beta + 1_{\{(i,t)\in\mathcal{C}\}}\left[\sum_{c=1}^{C} 1_{\{c(i,t)=c\}}\gamma_{c,t}\right] + 1_{\{(i,t)\in\mathcal{U}\}}\gamma_t^{\mathcal{U}} + 1_{\{(i,t)\in\mathcal{R}\}}\gamma_t^{\mathcal{R}} + \mu_{s(i,t),t} + u_i + \varepsilon_{i,t} \tag{1}$$

where $w_{i,t}$ is the daily wage of an individual i in year t, C is the number of cities, $c(i,t)$ (resp. $s(i,t)$) is the city (resp. industry) where individual i works in year t, $X_{i,t}$ are time-varying individual variables (in practice, squared age), $\mu_{s,t}$ is an industry-year fixed effect, γ_t^k is a year-specific location-k fixed effect, u_i is an individual fixed effect and $\varepsilon_{i,t}$ is the residual.[6]

We then investigate further the city effects with the specification:

$$\gamma_{c,t} = Z_{c,t}\theta_t + \delta_t + \eta_{c,t} \tag{2}$$

where $Z_{c,t}$ are time-varying variables at the city level which effects vary over time, δ_t is a year fixed effect and $\eta_{c,t}$ is a city error term capturing city unobserved effects such as the influence of amenities. We use three variables to capture agglomerations economies for a given city: employment density, area and Harris market potential constructed from densities (but excluding the city itself). As city area is included in the specification, the effect of density is measured while holding area constant. The market potential variable captures access to the market constituted of all other cities.

[5] Note that the set of cities ($\mathcal{C}$), the set of urban areas without core ($\mathcal{U}$) and the rural area ($\mathcal{R}$) vary empirically over time during our period of interest. Since variations are very minor, we do not index $\mathcal{C}$, $\mathcal{U}$ and $\mathcal{R}$ by t for simplicity.

[6] As the definition of industries changes over time due to classification changes in 1993 and 2009, we use three sets of fixed effects $\{1,...S\}$, $\{S+1,...,S+S'\}$ and $\{S+S'+1,...,S+S'+S''\}$ such that $s(i,t)$ is included in the first set before 1993, in the second set from 1993 to 2008, and in the third set after 2008.

The model is estimated in two stages. First, we estimate equation (1) using data at the individual level and recover in particular some estimates of city-year fixed effects. We then evaluate equation (2) by regressing these estimates on agglomeration variables and time fixed effects. Since all coefficients depend on time, estimation can be conducted year by year. We weight the second-stage regression with the city-year number of workers. This is done for two reasons. First, we have adopted the perspective of individuals re-aggregated at the city and national level. We want the estimated effects of agglomeration variables to be consistent with this individual perspective. Second, city-year effects used as dependent variable in the second stage are estimates with a sampling error. Our weights give more importance to city-year effects of large cities that are estimated more accurately and this makes the second-stage estimates more accurate. This also means that we recover effects of agglomeration variables using variations for larger cities rather than for the whole set of cities. In an econometric sense, it amounts to estimating the effects of agglomeration variables locally at large city sizes.

4.2 Instrumentation

As density, market potential and area are potentially endogenous, we instrument them in the second stage with historical variables constructed from past censuses in line with Combes *et al.* (2008). We resort to EHESS historical population data that give population counts for all municipalities in France for every censuses over the 1793-2006 period. We reaggregate these data at the city level to construct our instruments. Since we consider time-varying delineations of cities, the reaggregation of municipalities is specific to the year that is considered. More precisely, instruments are the logarithms of population densities in 1793, 1800, 1836 and 1856, market potentials for the same years and a peripherality index defined as the market potential obtained when fixing density to one for every city (to measure peripherality). We expect the explanatory power of these instruments to be large because there is inertia of local housing stocks that induces inertia of local population. At the same time, we expect these instruments to verify the exclusion restrictions because the production processes have changed a lot over 150 years with the rise of services industries and there have been several disruptions due to wars.

An alternative set of instruments consists in the aggregates of small-scale soil information at the city level (Rosenthal and Strange, 2008; Combes *et al.*, 2010). We aggregate data from the European Soil Database (ESDB) compiled by the European Soil Data Centre that originally comes as a raster of 1 km x 1 km cells. These instruments are expected to have some explanatory power because geology is likely to influence the productivity in the agricultural sector and should thus explain the location of first settlements which might have turned into cities. At the same time, the exclusion restrictions are expected to be satisfied because there is no clear link between geology and the production processes of the modern era. In practice, we use as instruments the proportions of the city area by levels of depth to rock, soil erodability, hydrogeological class, subsoil mineralogy, and topsoil organic carbon content.

4.3 Assessing the role of agglomeration economies in wage evolutions

Our main goal is to assess how changes in average daily wages between two periods, say $t-1$ and t, are related to the location type (city, urban area without core, rural) and agglomeration economies within cities. For that purpose, we first decompose national log-wage growth into two components:

$$\overline{\log w}_t - \overline{\log w}_{t-1} = \sum_{k \in \{\mathcal{C},\mathcal{U},\mathcal{R}\}} (p_{k,t} - p_{k,t-1}) \overline{\log w}_{k,t-1} + \sum_{k \in \{\mathcal{C},\mathcal{U},\mathcal{R}\}} p_{k,t} \left(\overline{\log w}_{k,t} - \overline{\log w}_{k,t-1}\right) \tag{3}$$

where $\overline{\log w}_t$, $\overline{\log w}_{k,t}$ and $p_{k,t}$ denote respectively the average log-wage in year t, the average log-wage in type-k locations in year t, and the related proportion of workers. On the right-hand side, the first sum captures the time variations in the allocation of workers between the three location types, and the second sum captures the evolution of log-wages in every location type.

We are particularly interested in cities and we thus focus on the evolution of average log-wage in cities, $\overline{\log w}_{\mathcal{C},t} - \overline{\log w}_{\mathcal{C},t-1}$. Denote by $p_{\mathcal{C},c,t}$ the proportion of workers located in city c at date t. The evolution of average log-wage in cities can be decomposed in the following way:

$$\overline{\log w}_{\mathcal{C},t} - \overline{\log w}_{\mathcal{C},t-1} = \sum_c (p_{\mathcal{C},c,t} - p_{\mathcal{C},c,t-1}) \overline{\log w}_{\mathcal{C},c,t-1} + \sum_c p_{\mathcal{C},c,t} \left(\overline{\log w}_{\mathcal{C},c,t} - \overline{\log w}_{\mathcal{C},c,t-1}\right) \tag{4}$$

where $\overline{\ln w}_{\mathcal{C},c,t}$ is city-year average of log-wage. The first right-hand side sum captures the change in the distribution of workers across cities, holding constant the average log-wage in cities. The second sum captures the changes in log-wage in every city, holding constant the proportions of workers in cities.

We now detail the causes of log-wage evolution for any given city c. We first insert expression (2) into equation (1) and average the resulting expression at the city level for a given year:

$$\overline{\ln w}_{\mathcal{C},c,t} = \bar{X}_{c,t}\beta + Z_{c,t}\theta_t + \bar{\mu}_{c,t} + \bar{u}_{c,t} + \delta_t + \eta_{c,t} \tag{5}$$

where $\bar{X}_{c,t}$, $\bar{u}_{c,t}$ and $\bar{\mu}_{c,t}$ denote respectively city-year averages of individual variables, individual fixed effects and industry effects. Note that, since city-year fixed effects are introduced in equation (1), the city-year average of first-stage residuals is zero by construction and thus does not intervene in this equation. From expression (5), we get the following decomposition of city log-wage growth into four components:

$$\overline{\log w}_{\mathcal{C},c,t} - \overline{\log w}_{\mathcal{C},c,t-1} = (M_{c,t} - M_{c,t-1}) + Z^*_{c,1}(\theta_t - \theta_{t-1}) + \left(Z^*_{c,t} - Z^*_{c,t-1}\right)\theta_{t-1} + (\eta_{c,t} - \eta_{c,t-1}) \tag{6}$$

where $Z^*_{c,t} = Z_{c,t} - Z^*$ with Z^* the value of city variables for a reference city, and $M_{c,t} = \bar{X}_{c,t} + \bar{u}_{c,t} + \bar{\mu}_{c,t} + Z^*\theta_t + \delta_t$ the composition effect for city a in year t. The first right-hand side term captures changes in composition effects (related to age, individual unobservables and industry structure) and time effects on city wage growth. It cannot be decomposed further because individual fixed effects, time fixed effects and linear age effects cannot be disentangled due to identification issues. The second one corresponds to the effect of changes in returns to agglomeration variables. The third one captures the impact of changes in the values of agglomeration variables. Note that terms two and three constitute a Oaxaca-Blinder decomposition of spatial effects. Finally, the fourth term corresponds to the evolution of city unobserved effects.

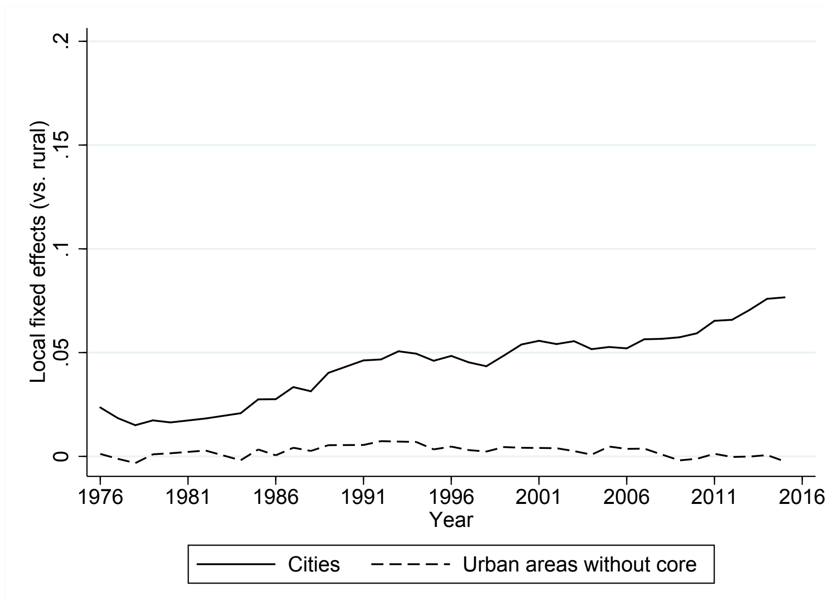
We introduced a reference city to make a meaningful assessment for the effect of changes in returns to agglomeration variables as these variables are in logarithmic form, and changing the measurement unit for area changes their values. Considering the difference between cities and the reference makes this issue disappears since logarithm differences are immune to changes of the measurement unit. For the reference, we consider a fictitious city which values for all the agglomeration variables are the minima. This way, differences of agglomeration variables with the reference city are all positive, and the effect of changes in returns to agglomeration variables captures effects for cities having agglomeration economies that are larger than for the reference. For instance, if returns to density are increasing over time, it captures the average effect of an increase in returns for being denser than the mininum.

Our decomposition can be quantified for each city and then aggregated weighting by the city proportions of workers to obtain a more detailed decomposition of the second component of $\overline{\log w}_{\mathcal{C},t} - \overline{\log w}_{\mathcal{C},t-1}$ (see equation 4).

5 Results

We first estimate the log-wage specification that involves location-year fixed effects and controls for individual observed and unobserved characteristics, as well as industry (equation (1)). For now, we focus on the estimates for the effects of being in a rural area, $\gamma_t^{\mathcal{R}}$, in an urban area without core, $\gamma_t^{\mathcal{U}}$, and in a city, $\gamma_t^{\mathcal{C}} \equiv \sum_{c=1}^{C} p_{a,t}\gamma_{a,t}$. Figure 3 graphs $\gamma_t^{\mathcal{U}} - \gamma_t^{\mathcal{R}}$ and $\gamma_t^{\mathcal{C}} - \gamma_t^{\mathcal{R}}$ as a function of time. Urban areas without core are characterized by yearly effects that remain close to those of rural areas. This is not really surprising since these urban areas are usually quite small and agglomerations economies are not really expected. By contrast, yearly effects for cities are above those for rural areas and the difference is increasing over time, suggesting an increase in agglomeration economies. These results can be contrasted with those obtained when individual unobserved heterogeneity is ignored. Figure A.3 shows that differences in local effects between cities and rural areas are larger in that case, and there is no increasing trend. The contrast between the two figures suggests a strong sorting of workers with higher unobserved skills in cities but a decrease of this sorting over time.

Figure 3: Yearly effects of working in an urban area without core or a city relatively to the rural area



Note: We represent differences between yearly effects of being in an urban area without core or in a city and yearly effects of being in the rural area. Yearly effects of being in a city are yearly averages of city-year fixed effects, weighting by the yearly number of individuals working in cities. There are no yearly effects for 1981, 1983 and 1990 since data are missing for those years, and we therefore consider instead interpolated values from neighbouring years.

We then turn to the estimation results when regressing estimated city-year fixed effects on agglomeration variables (equation (2)), with the constraint that the effects of city variables are constant over time. This gives us a benchmark that is comparable to regressions conducted in previous studies on agglomeration economies. We first report results when omitting individual fixed effects in the first-stage regression. In column (1), we regress estimated city-year fixed effects on density and find a positive elasticity of wages that decreases when introducing area as shown by column (2). The elasticity of wages with respect to density and area are respectively 0.051 and 0.024, suggesting that there are agglomeration economies related to both density and area. When introducing individual fixed effects in the first-stage regression (column 4), these two elasticities decrease to 0.033 and 0.014 due to a positive sorting of individuals with larger unobserved skills into denser and larger cities. These estimates are in line with past studies (Combes *et al.*, 2008; de la Roca and Puga, 2017).

When considering both density and market potential in the specification while controlling for area, the estimated elasticity of wage with respect to density is smaller because of a positive correlation between density and market potential, whether individuals fixed effects are excluded (column 3) or included (column 6). In the most complete specification that involves individual fixed effects, the wage elasticity for density, area and market potential are respectively 0.025, 0.014 and 0.065.

When we rather regress estimated city-year fixed effects on market potential, density and area, we find a wage elasticity for market potential larger than for density in presence (resp. absence) of individual unobserved heterogeneity, as it reaches 0.065 (resp. 0.063).

Table 2: Second stage OLS regression results

	(1)	(2)	(3)	(4)	(5)	(6)
Density	0.084***	0.051***	0.044***	0.052***	0.033***	0.025***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
Area		0.024***	0.024***		0.014***	0.014***
		(0.001)	(0.001)		(0.001)	(0.001)
Market potential			0.063***			0.065***
			(0.003)			(0.003)
Individual FE	No	No	No	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.772	0.858	0.872	0.915	0.934	0.945
N	10974	10974	10970	10969	10969	10965

Note: Standard errors are in parentheses. Explanatory variables are all in logarithm. Regressions are weighted by the number of workers. *: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$.

We also assess to what extent our estimates are affected by endogeneity issues instrumenting density, area and market potential with historical and soil variables. Table 3 reports IV results when including individual fixed effects in the specifications. We start with a specification that includes only density and area. Columns 1-3 show that, when we instrument with historical variables, soil ones or both sets, the estimated coefficient for density is a bit larger but that for area remains more or less the same. As reported in columns 4-6, estimates for density coefficient are a bit lower when adding and instrumenting market potential. In the most complete specification where our three city variables are instrumented with both historical and soil instruments, the wage elasticity for density, area and market potential are respectively 0.031, 0.014 and 0.050.

Table 3: Second stage IV regressions results

	(1)	(2)	(3)	(4)	(5)	(6)
Density	0.043***	0.049***	0.043***	0.027***	0.046***	0.031***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)
Area	0.012***	0.013***	0.012***	0.016***	0.013***	0.014***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Market potential				0.050***	0.015***	0.050***
				(0.003)	(0.005)	(0.003)
Historical IV	Yes	No	Yes	Yes	No	Yes
Soil IV	No	Yes	Yes	No	Yes	Yes
KP F-stat.	95.71	96.07	88.81	59.55	62.15	68.50
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	10679	10921	10679	10679	10917	10679

Note: Standard errors are in parentheses. Explanatory variables are all in logarithm. Regressions are weighted by the number of workers. *: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$. Historical instruments: logarithms of population densities in 1793, 1800, 1836 and 1856, and market potentials for the same years. Soil instruments: shares of the city area by levels of depth to rock, soil erodability, hydrogeological class, subsoil mineralogy, and topsoil organic carbon content.

We then turn to our estimation results based on equation (2) when coefficients of density, area and market potential depend on time, and individual fixed effects are included. Figures 4.a and 4.c show that the estimated effects of density (resp. area) increase over time from 0.011 to 0.042 (resp. 0.007 to 0.022) whereas, according to Figure 4.e, there is no clear pattern for the estimated effects of market potential. Interestingly, when omitting individual fixed effects in the first stage, the patterns are quite different (Figures A.4.a, A.4.c and A.4.e). In particular, there is no upward trend for the estimated effects of density and area after 1995. This suggests some variations in the spatial sorting of individuals across time.

We investigate this sorting further by representing the yearly correlation between city quantities and individual fixed effects computed at the worker-year level. Interestingly, Figure A.5.a shows that the correlation between city-year fixed effects and individual fixed effects increases until 1990, then remains stable and finally decreases after 2000. Turning to agglomeration variables, the correlation between density (resp. area) and individual fixed effects is positive, decreases slightly after 1992, before decreasing more abruptly after 2002 (Figures A.5.b and A.5.c). Since the correlation between density (resp. area) and individual fixed effects is positive, taking into account individual unobserved heterogeneity lowers density and area estimates. This lowering gets smaller and smaller after year 1992, because the correlation between density (resp. area) and individual fixed effects decreases over time after 1992. The density and area estimates have a humped-shape profile with a small plateau ending up in 2000 in absence of individual fixed effects. Introducing individual fixed effects "corrects" the trend from the middle of the plateau (around 1992) and makes it increasing instead of decreasing. Because of that, the overall profile of density and area estimates ends up being increasing. The correlation between market potential and individual fixed effects is rather stable until 2002 and then decreases (Figure A.5.c). Overall, correlation patterns are consistent with the changes in the estimated coefficients of agglomeration variables obtained when introducing individual fixed effects.

Still, for the interpretation of correlations, we have the issue that individual effects capture both unobserved skill effects and age effects.[7] In particular, average age for years present in the panel decreases as one enters the labour market closer to the end of the panel, and this blurs the interpretation related to individual effects. We try to isolate unobserved skill effects considering the individual residuals obtained when regressing the sum of individual fixed effect and squared age on age, squared age and year fixed effects. We label them "net individual effects". Yearly correlations between city quantities and net individual effects have shapes similar to those obtained when considering individual fixed effects, but their values are larger. In particular, the correlations involving density and area are still decreasing after 1992. This suggests a decrease over time in the spatial sorting according to unobserved individual skills.

One may also wonder whether important changes in the profiles of estimated coefficients when introducing individual fixed effects could be an artefact due to periods involved in their identification. Indeed, in 1976 (resp. 2015), only observations from 1976 onward (resp. 2015 downward) for a given individual participate to the identification of the density, area and market potential coefficients for year 1976 (resp. 2015). For a given year y between 1976 and 2015, coefficients are identified thanks to observations at year y, as well as years before and after y. To investigate the existence of a possible bias due to edge effects, we re-estimate our specification when considering only the first four observations for individuals appearing at least four times in the panel. Indeed, considering short time spans for individuals should lessen edge effects although restricting the estimations to individuals appearing at least four times in the panel may lead to sample selection. Making such restriction changes profiles of estimated coefficients that are now rather decreasing over time when individual fixed effects are not introduced, especially for density and area (Figure A.6.b, A.6.d and A.6.f), suggesting sample selection in our robustness check. But profiles are "corrected" in the same direction as with the whole sample when introducing individual fixed effects, as estimated coefficients are then increasing over time for density and market potential, and rather stable for area (Figure A.6.a, A.6.c and A.6.e).

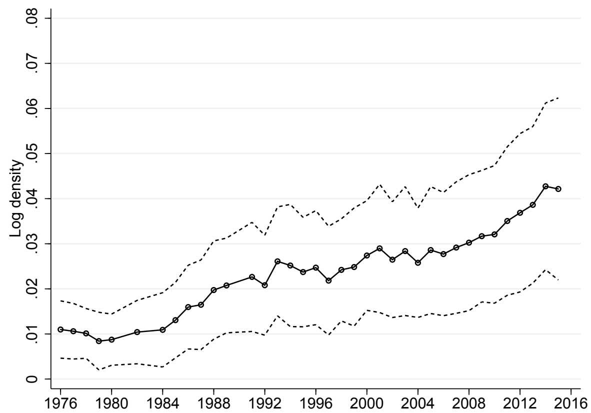
Finally, we assess the importance of endogeneity issues by instrumenting agglomeration variables in the second-stage equation with historical and geological variables. Results represented in Figures 4.b, 4.d and 4.e are very

[7]Indeed, the linear effect of age cannot be identified seperately from individual fixed effects and time effects captured by city-year and industry-year fixed effects.

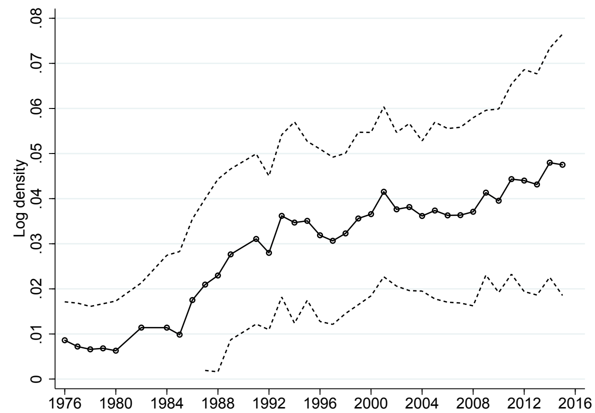
close to those obtained without instrumentation (Figures 4.a, 4.c and 4.e).

Figure 4: Estimated yearly coefficients of city variables with individual effects in the first-stage specification, with and without instrumentation with historical and geological variables

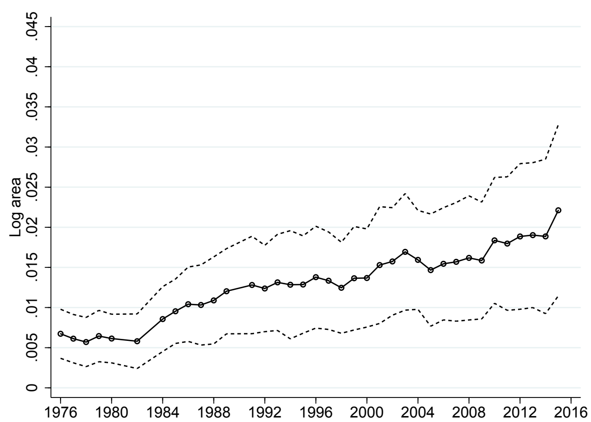
4.a: Density, no instrumentation



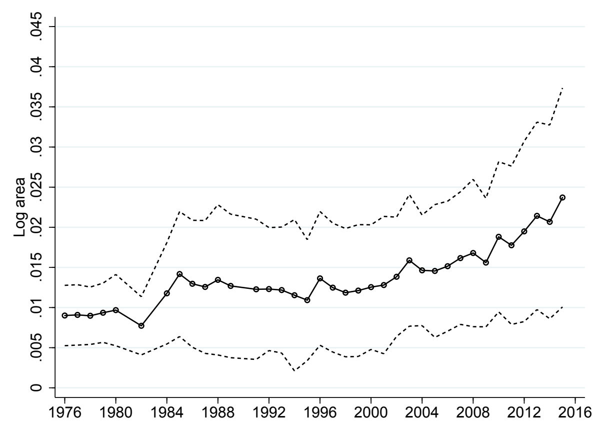
4.b: Density, instrumentation



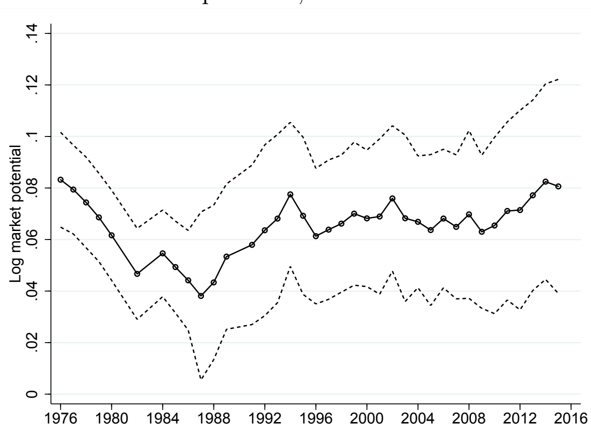
4.c: Area, no instrumentation



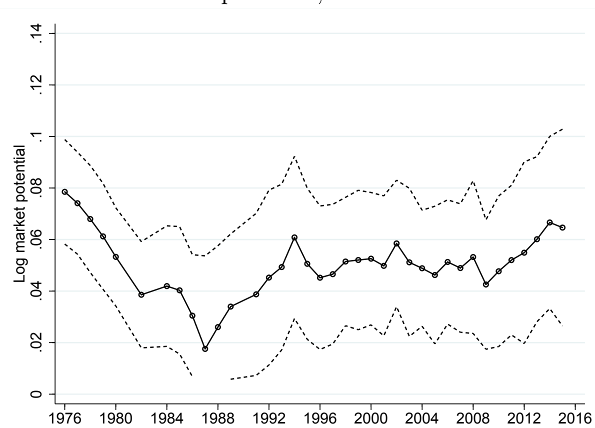
4.d: Area, instrumentation



4.e: Market potential, no instrumentation



4.f: Market potential, instrumentation



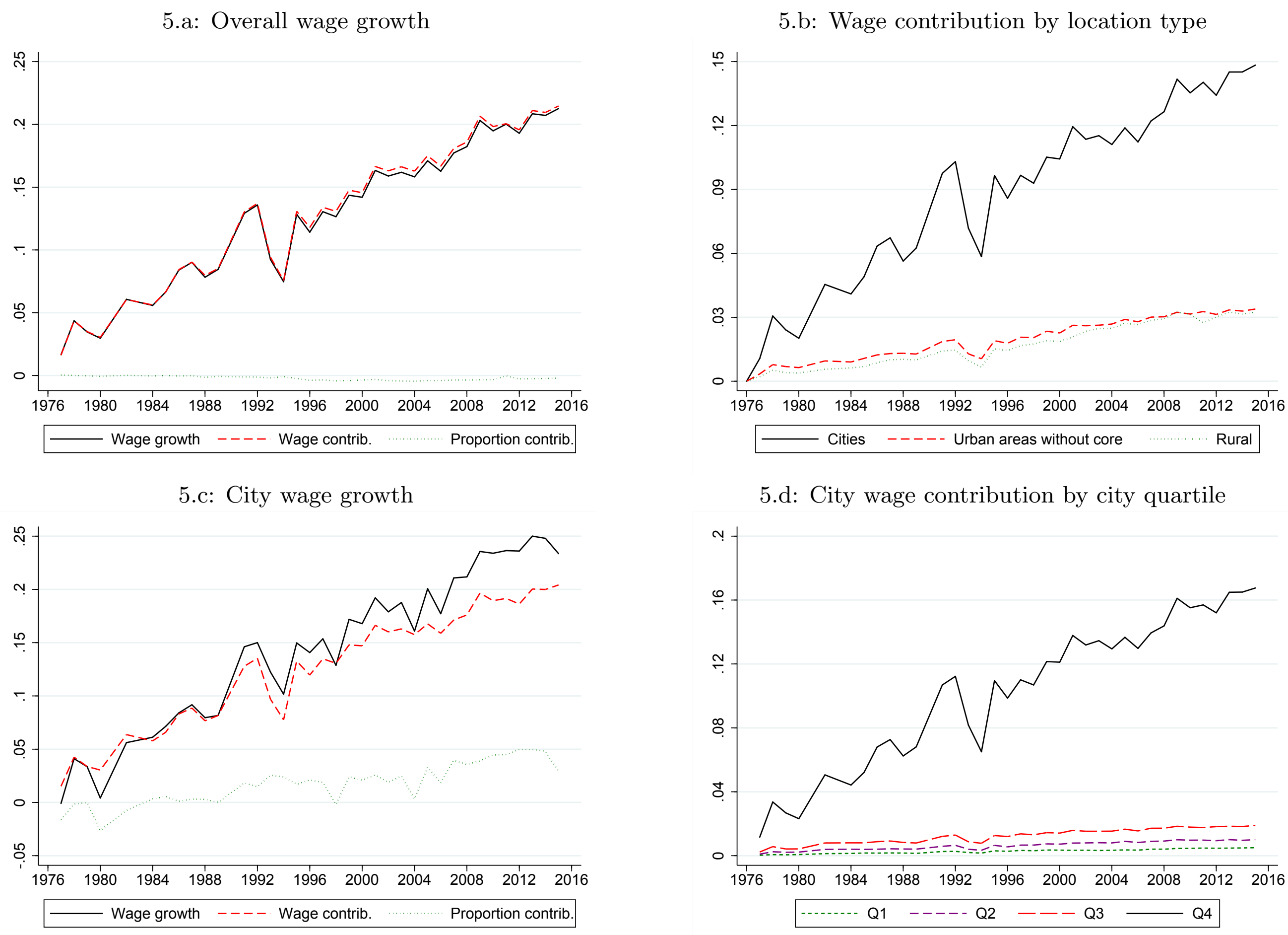
Note: Estimated coefficients are represented by bullets and linked with a plain line, and bounds of confidence intervals are represented in dots. Historical instruments: logarithms of population densities in 1793, 1800, 1836 and 1856, and market potentials for the same years. Soil instruments: shares of the city area by levels of depth to rock, soil erodability, hydrogeological class, subsoil mineralogy, and topsoil organic carbon content.

5.1 Decompositions of wage growth

We now consider wage growth defined as the difference in average log-wage between any given year and 1976. We quantify the contributions to this growth of wage evolutions for the three location types and the reallocation of workers across location types (equation 3). Figure 5.a shows that wage growth comes only from wage evolutions in the three location types whereas reallocation effects are negligible whatever the time horizon. Figure 5.b gives contributions related to wage evolutions for the different location types and shows that cities contribute the most.

We then consider city wage growth and quantify the contributions of wage evolutions in the different cities and the reallocation of workers across cities (equation 4). Figure 5.c shows that city wage growth is mostly driven by wage evolutions in the different cities. The reallocation of workers has an influence that is negligible for short-run wage growth but it grows larger and becomes significant, while remaining small, for wage growth between 1976 and any year after 2006. Figure 5.d decomposes the contribution related to wage evolutions by city quartile of DADS employment in 1976. Most of this contribution comes from the fourth quartile that is composed of larger cities.

Figure 5: Decomposition of wage growth between 1976 and any given year



Note: In panel a: "Wage growth": average log-wage growth (relatively to 1976); "Wage contrib.": contribution of wage evolutions in the three location types (rural, urban areas without core and cities); "Proportion contrib.": contribution of changes in the proportions of workers in the three location types.

We then evaluate the contributions to city wage growth of changes in values and returns to agglomeration variables. These contributions are computed as the sums of those obtained for cities (see equation (6)), weighting by the city proportion of workers (among workers living in cities) at the initial date. Figure 6.a shows that the contribution of changes in values is negative and small whatever the time horizon. There are two main reasons for this result. First, time variations in agglomeration variables are rather small compared to cross-section differences across cities. Second, even if some cities expand to a large extent over the period, this usually occurs together with a decrease in their density. Hence, an increase in area agglomeration economies is usually compensated by a decrease in density agglomerations economies which is sometimes larger.

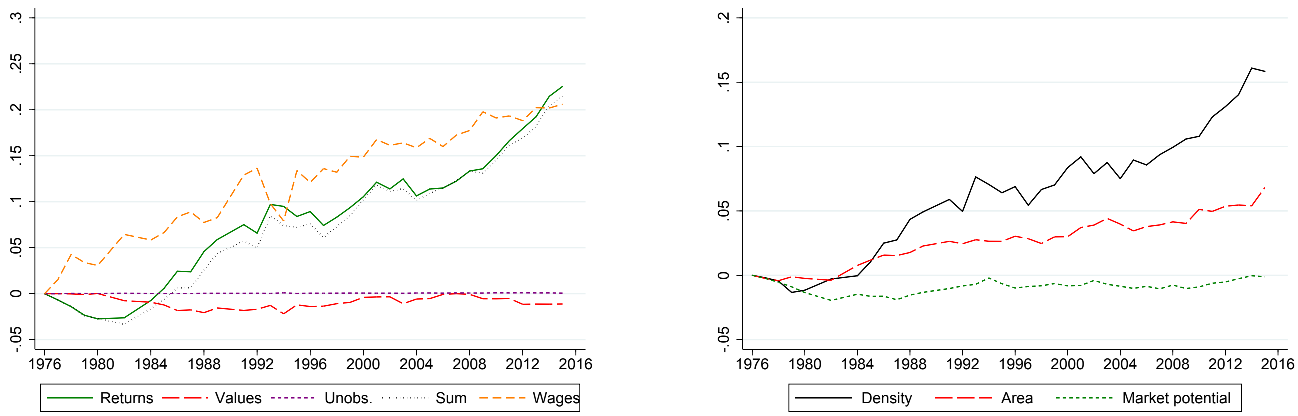
By contrast, returns to agglomeration variables have an important positive impact on wage growth between 1976 and any year from 1985 onwards.[8] This impact is increasing over time consistently with the wage increase observed in cities. Before 1985, the contribution of changes in returns is negative. We can assess which agglomeration variables play the most important role in the contribution of changes in returns. Figure 6.b shows that the change in returns to density has the largest impact, followed by that for area. The large impacts on medium-run city wage growth for density and area come from both the significant increase in their returns and the large differences in values that often occur between cities and the reference city (which has the minimum density and area). By contrast, the change in returns to market potential has a small negative impact on city wage growth between 1976 and any given year. This occurs both because the return to market potential is large in 1976 compared to that in other years, and there are only small differences in the values of market potential between cities and the reference city. Actually, it is the negative contribution of market potential before 1985 that explains the negative contribution of changes in return before that date.

[8]Note that we do not comment on the contribution of changes in city unobservables to city wage growth. Indeed, it is close to zero by construction because the yearly weighted average of city unobservables is zero due to the introduction of time fixed effects in the second-stage specification of the model (equation 2).

Figure 6: Decomposition of city wage growth between 1976 and any given year

6.a: Contributions to city wage growth

6.b: Contributions of changes in returns

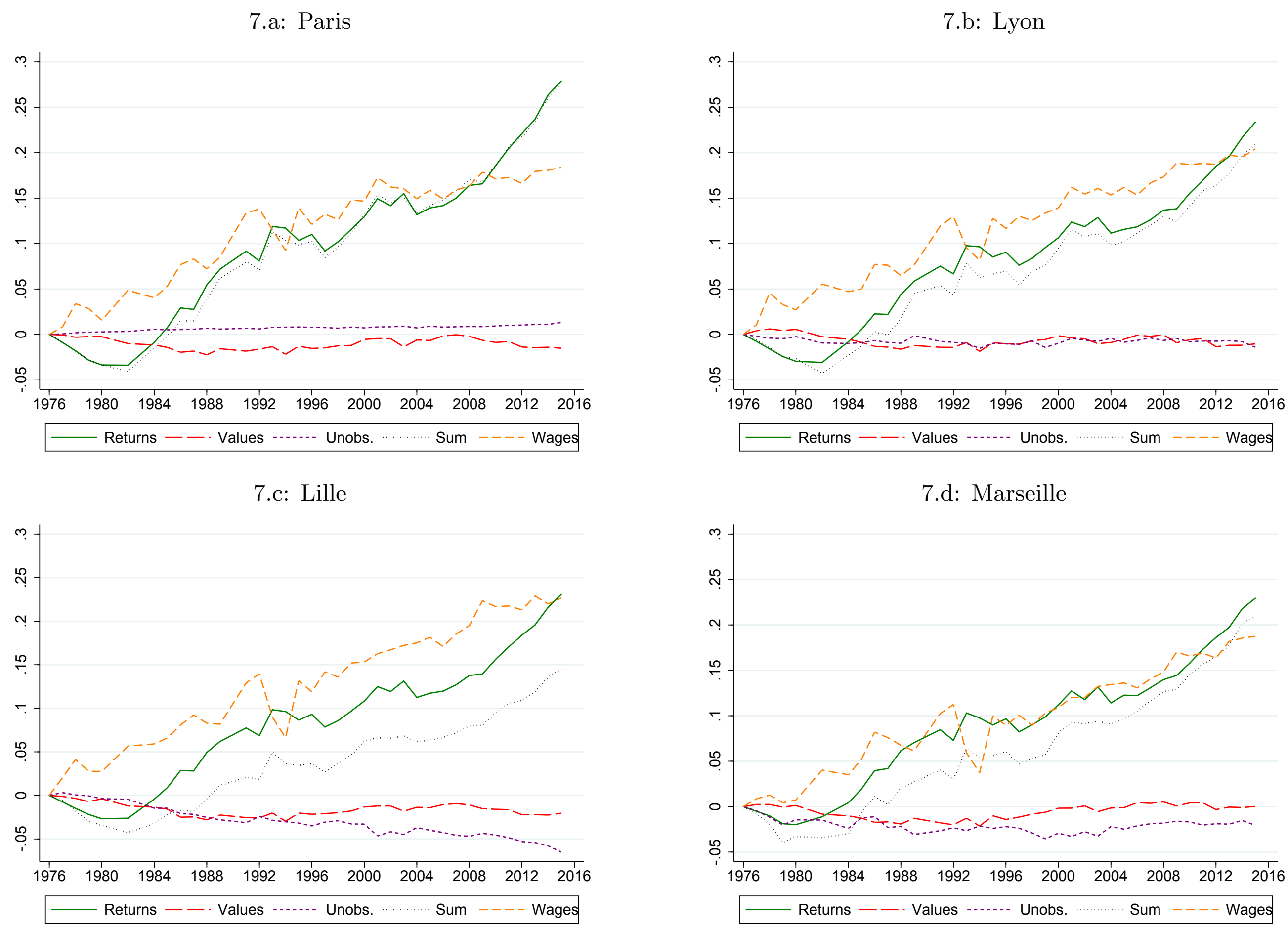


Note: In panel a: “Returns”: Contributions of changes in returns of city variables to city wage growth; “Values”: Contribution of changes in their values; “Unobs.”: Contribution of changes in city unobservables; “Sum”: Sum of these three contributions; “Wages”: city average log-wage growth (relative to 1976). In panel b: “Density” (resp. “Area” and “’Market Potential”: Contribution of changes in returns to density (resp. area and market potential). Values are missing for $t = 1981, 1983, 1990$.

So far, we have presented results for the decomposition when aggregating all the cities, but it is also possible to consider the decomposition by city size group. For that purpose, we replicate the decomposition exercise separately for each quartile of DADS city employment in 1976, obtained when weighting cities by the number of employed workers. Not suprisingly, Figure A.7 shows that the contribution of the change in returns to agglomeration variables increases with the quartile. Indeed, for the computation of this contribution, changes in returns are multiplied by values of density and area which are more important for cities in higher quartiles. One can also note some positive non negligible contributions of city unobservables in lower quartiles for periods of more than 10 years. This suggests a slightly higher increase of the productivity in smaller cities than in larger ones.

We now consider separately contributions for the four largest cities in quartile 4, i.e. Paris, Lille, Lyon and Marseille (Figures 7.a-7.d). Not surprisingly, the contribution of the changes in returns to agglomeration variables is the largest for Paris because the values for density and area, with which changes in returns are multiplied, are the highest for that city (Figures 2.a and 2.b). Interestingly, even if Marseille has grown a lot over the last 40 years, its contribution of the changes in values of agglomeration variables is not much different from that of other large cities. In particular, this is due to the decrease of its density that has gone together with the large increase in its area. Finally, note that Lille has experienced a decrease in city unobservables, suggesting a decrease in productivity due to the decline of textile, mining and steel industries.

Figure 7: Decomposition of wage growth between 1976 and any given year, for the four largest cities



Note: "Returns": Contributions of changes in returns of city variables to log-wage growth; "Values": Contribution of changes in their values; "Unobs.": Contribution of changes in city unobservables; "Sum": Sum of these three contributions; "Wages": average city log-wage growth.

6 Robustness checks

6.1 Alternative definitions of cities

We have made several choices to delineate cities. In this section, we consider robustness checks when changing the way cities are defined. In particular, delineation of cities varies over time and there are some fusions/absorption of cities over the period that can be quite important. This happens for instance when Marseille absorbs Aix-en-Provence during our period of study. Marseille experiences a discrete jump in its area and density at the time of absorption. One can wonder to what extent time variations of delineations affect our results. We conducted a first robustness check considering that the delineation of cities is constant over time and fixed to the one obtained in 2015. Figure A.9 shows that the time-varying profiles of estimates for coefficients for agglomeration variables are barely affected. Somehow, this is not suprising because these coefficients are estimated in second-stage using cross-section variations across cities within each year, and these variations are much larger than time variations that may be caused by changes in the delineations of cities.

Moreover, the literature has proposed alternative ways of constructing cities. Here, cities are defined based on the continuity of built areas, but one may rather consider commuting patterns and aggregate municipalities iteratively when they send a proportion of their commuters above a given threshold to the rest of the city (see for instance Duranton, 2015; Bosker *et al.*, 2021). Actually, one may even consider a mix of the two approaches such that the delineation of urban areas is based on both continuity of built areas and commuting patterns. This is the case for the urban area definition of the French Institute of Statistics (INSEE). Indeed, according to that definition, a urban unit is defined based on the continuity of built areas, and municipalities are aggregated to that urban unit according to commuting flows to form an urban area. Proposing time-varying delineations of cities based on different notions is beyond the scope of the paper. Nevertheless, we can vary the scope of cities by changing the bandwidth of the kernel used for smoothing distributions of building density in our delineation algorithm. Indeed, the larger the bandwidth, the more building densities will be smoothed across squares. Squares mildly built close to dense squares will have a larger smoothed building density and will be more likely to be considered as urban. Consequently, peripheral municipalities from which workers commute to the city center will end up being included in the city. In pratice, we chose a rather large bandwidth of 2.1km to include such municipalities.

We now consider a smaller bandwidth of 1km to have a definition of cities more closely related to building continuity. Table A.2 shows that this change of bandwidth yields more cities in both 1975 and 2015, but on average smaller cities. This is actually not surprising since, when the bandwidth decreases, there is less smoothing in urban areas when assessing the existence of cores, and it is thus easier to detect cores and consider the corresponding urban areas to be cities. Moreover, some cities may end up being disaggregated because the smoothing of building volume density in low-density areas between cities involves fewer dense squares belonging to cities. Finally, peripheral areas of cities may have smaller smoothed building volume densities and may end up being considered as rural, which explains why cities may end up being smaller. Smoothing with a 1km bandwidth yields increasing profiles for the estimated coefficients of density and area, whether or not agglomeration variables are instrumented, in line with results obtained with a 2.1km bandwidth (Figure A.8). Still, estimated density coefficients are smaller, suggesting that a 2.1km bandwidth is more appropriate to delineate cities relevant to measure agglomeration economies.

We also experiment using the 2010 definitions of urban areas proposed by INSEE.[9] When considering these urban areas as cities, we find that the profiles of estimated coefficients are very close to those obtained in our benchmark when running ordinary least squares. When instrumenting, increasing slopes for estimated coefficients of density and area differ to some extent, with the slope being less steep for density but steeper for area. This can be explained by different correlations between density and area depending on the definition of cities, and different relationships between agglomeration variables and instruments.

6.2 Dynamic agglomeration effects

We also conduct a robustness check when considering dynamic agglomeration effects, such that workers can benefit from experience in large cities which is at least partly transferrable when moving elsewhere (de la Roca and Puga, 2017). Indeed, there might be an omitted-variable bias in our benchmark estimations due to our agglomeration variables being correlated with time spent in cities. We introduce past experience in rural area, past experience in cities without core, as well as past experiences in every quartile of city size interacted with dummies for being in rural area, any city without core and every quartiles of city size.[10] Interestingly, the positive time trends for the estimated coefficients of density and area remain, whether or not we instrument the agglomeration variables, but the sizes of the effects are smaller (Figures A.11).

7 Theoretical insights

In this section, we propose a city model to quantify the extent to which city land size and density are influenced by variations in their returns that affect agglomeration economies. We also want to quantify how these changes affect wages. Our framework starts with the standard monocentric city model with the specificity that wages are endogenous and depend on agglomeration economies. We characterize the equilibrium for the closed city and its open version, and conduct comparative statistics to establish quantitative relationships between changes of city land size and population density and changes of their returns. We finally explain how these relationships can be brought to the data. All developments and proofs are relegated in Appendix B.

[9] There also exist past definitions for those geographic units but cities are not delineated in a consistent way across time and this is why we stick to the delineation of cities in a single year and consider that it is constant over time.

[10] Here, a quartile in a given year is the unweighted quartile of city size for that year.

7.1 The model

Consider a linear city c such that all the jobs are exogenously located in a central business district (CBD) in $x = 0$. Workers are located on a segment $[0, \overline{x}_c]$ where $\overline{x}_c$ is the city fringe, and a worker located at distance x from the CBD commutes for a linear monetary cost $\tau_c x$. Every worker earns an endogenous wage w_c that depends on agglomeration economies such that:

$$w_c = A_c \left(N_c/L_c\right)^{\beta} L_c^{\alpha} \tag{7}$$

where A_c is the city total factor productivity, and L_c and N_c are respectively the city land size and population, such that $N_c > L_c > 1$.[11] There are density agglomeration economies with elasticity parameter β such that $0 < \beta < 1$, and land size agglomeration economies with elasticity parameter α such that $0 < \alpha < 1$. An individual consumes a numeraire z and land ℓ at price $R_c(x)$. Utility is Cobb-Douglas and given by $U(\ell, z) = B_c \ell^a z^{1-a}$ where B_c captures city-specific consumption amenities, and the budget constraint is $w_c - \tau_c x = z + R_c(x)\,\ell$. Workers maximize their utility under budget constraint choosing their location, and how much numeraire and land they want to consume.

The fringe is determined by the equality $R_c(\overline{x}_c) = \underline{R}$ where $\underline{R}$ is the agricultural land price. Since the city is linear, its fringe is such that $L_c = \overline{x}_c$, and the city population verifies the equilibrium equation:

$$N_c = \int_0^{\overline{x}} n_c(x)\, dx \tag{8}$$

where $n_c(x)$ is the population density at distance x.

7.2 Equilibrium and comparative statics

7.2.1 Closed cities

When city c is closed, its population N_c is fixed and only its land size L_c can vary. The model has a unique solution and it is possible to conduct comparative statics when varying the elasticities of wages with respect to land size and population density, α and β. We obtain:

$$\frac{\partial \log L_c}{\partial \log \alpha} = -\frac{\partial \log (N_c/L_c)}{\partial \log \alpha} = \frac{\alpha}{1+\beta-\alpha} \log L_c \tag{9}$$

$$\frac{\partial \log L_c}{\partial \log \beta} = -\frac{\partial \log (N_c/L_c)}{\partial \log \beta} = \frac{\beta}{1+\beta-\alpha} \log \left(\frac{N_c}{L_c}\right) \tag{10}$$

$$\frac{\partial \log N_c}{\partial \log \alpha} = \frac{\partial \log N_c}{\partial \log \beta} = 0 \tag{11}$$

The elasticity of land size with respect to α (equation 9) is positive because of an income effect. Indeed, incomes increase due to additional land agglomeration economies, which makes the aggregate land demand increase. Con-

[11] This implicitely imposes conditions on parameters at equilibrium.

versely, the elasticity of population density with respect to α is negative since population remains fixed whereas land consumption increases. The magnitude of elasticities depends on land size L_c, α and β. Indeed, land size L_c determines how much income is affected by a change in the intensity of land agglomeration economies. Moreover, a change in land size generates additional gains from an increase in land agglomeration economies (captured by α), but also losses from a decrease in density agglomeration economies (captured by β). It is possible to comment on variations of land size and population density when β varies (equation 10) in the same way. Interestingly, increasing the intensity of density agglomeration economies makes population density decrease. This is again due to the income effect that makes land size increase whereas population is held fixed.

We can then turn to variations of wages due to changes in the returns to land size and population density. Indeed, deriving the logarithm of the wage expression (7), we get:

$$\frac{\partial \log w_c}{\partial \alpha} = \log L_c + \frac{\beta}{\alpha}\frac{\partial \log (N_c/L_c)}{\partial \log \alpha} + \frac{\partial \log L_c}{\partial \log \alpha} \tag{12}$$

$$\frac{\partial \log w_c}{\partial \beta} = \log\left(\frac{N_c}{L_c}\right) + \frac{\alpha}{\beta}\frac{\partial \log L_c}{\partial \log \beta} + \frac{\partial \log (N_c/L_c)}{\partial \log \beta} \tag{13}$$

There are direct effects due to changes in land size and density agglomeration economies (first right-hand side terms) and indirect effects due to changes in land size and population density (respectively second and third right-hand side terms). Inserting equations (9), (10) and (11) into these expressions, we get:

$$\frac{\partial \log w_c}{\partial \alpha} = \log L_c + \frac{\alpha-\beta}{1+\beta-\alpha}\log L_c \tag{14}$$

$$\frac{\partial \log w_c}{\partial \beta} = \log\left(\frac{N_c}{L_c}\right) + \frac{\alpha-\beta}{1+\beta-\alpha}\log\left(\frac{N_c}{L_c}\right) \tag{15}$$

Our estimations give small values for α and β (below 0.05). We can see from these two expressions that indirect effects due to changes in land size and population (second right-hand side terms) are then negligible compared to direct effects due to changes in agglomeration economies (first right-hand side terms).

7.2.2 Open cities

We can then turn to the open city case. Population is free to move across cities and, at the equilibrium, utility is the same in every city. It is again possible to conduct comparative statics. We obtain when varying α:

$$\frac{\partial \log L_c}{\partial \log \alpha} = \frac{\alpha}{1+\beta-\alpha}\left[\log L_c + \frac{N_c M_c}{L_c}\frac{\partial \log N_c}{\partial \log \alpha}\right] \tag{16}$$

$$\frac{\partial \log (N_c/L_c)}{\partial \log \alpha} = \frac{\alpha}{1+\beta-\alpha}\left[-\log L_c + \left(\frac{1+\beta-\alpha}{\alpha} - \frac{N_c M_c}{L_c}\right)\frac{\partial \log N_c}{\partial \log \alpha}\right] \tag{17}$$

where:

$$\frac{N_c M_c}{L_c} = \frac{a}{\mu_c}\left(\frac{w_c}{\tau_c L_c} - 1\right) + \beta \text{ with } \mu_j = \frac{1}{1 - \underline{R}/R_c(0)} \tag{18}$$

In expressions (16) and (17), the first right-hand side terms are the same as in the closed-city case. Still, there are now additional terms capturing migration effects. They involve the elasticity of population with respect to α which expression needs to be determined. Consider for now a city c such that there is in-migration ($\partial \log N_c / \partial \log \alpha > 0$). The elasticity of land size is positively influenced by this in-migration since it creates additional land demand. The city population may evolve either way as there is both additional population and additional land demand.

The elasticity of city population for a given city, say $c = 1$, verifies:

$$\frac{\partial \log N_1}{\partial \log \alpha} = \alpha \left(\frac{N_1 Q_1}{w_1 - \tau_1 L_1}\right)^{-1} \left(\sum_{c=1}^{C} \frac{w_c - \tau_c L_c}{Q_c}\right)^{-1} \sum_{c=1}^{C} \frac{w_c - \tau_c L_c}{Q_c} \left(\log L_c - \log L_1\right) \tag{19}$$

where:

$$\frac{N_c Q_c}{w_c - \tau_c L_c} = \beta - \frac{a}{\mu_c}\left[1 + (\beta - \alpha)\frac{w_c}{\tau_c L_c}\right] \tag{20}$$

Equation (19) shows that the elasticity of city-1 population depends on a weighted average of land-size differences between city 1 and other cities. Indeed, in- or out-migration results from changes in land agglomeration economies in every city. Importantly, this equation involves $Q_c / (w_c - \tau_c L_c)$ for every city, which depends only on quantities that can be computed as shown by expression (20). These quantities are the city population N_c, the land budget share a, the ratio between land prices at the fringe and at the center $\underline{R}/R_c(0)$, the share of commuting costs at the fringe $\tau_c L_c / w_c$, and agglomeration economies parameters α and β that are recovered from the estimations.

We can then turn to variations of wages. Inserting equations (16) into the expression for variations of wages with respect to α given by equation (12), we get:

$$\frac{\partial \log w_c}{\partial \alpha} = \log L_c + \frac{\alpha - \beta}{1 + \beta - \alpha} \log L_c + \left(\beta + \frac{\alpha - \beta}{1 + \beta - \alpha} \frac{N_c M_c}{L_c}\right) \frac{1}{\alpha} \frac{\partial \log N_c}{\partial \log \alpha} \tag{21}$$

where the expression of $\frac{\partial \log N_c}{\partial \log \alpha}$ is given by equation (19). There is now a third term compared to the closed city case, that comes from migrations between cities affecting both city land size and density.

Expressions for elasticities of city land size and population density as well as variations of wages with respect to β are very similar and are detailed in appendix. Interestingly, none of our expressions for elasticities depends on production and consumption amenities A_c and B_c that would be hard to quantify. Production amenities disappear because they are introduced multiplicatively in the wage function and the computation of elasticities makes intervene the derivative of log-wage (that involves the derivative of log-production amenity effect which is zero). Consumption amenities disappear because they enter multiplicatively the utility function, and the derivative of the between-city equilibrium then involves ratios of consumption amenity effects between city pairs that can be replaced by ratios of wages net of commuting costs (using the between-city equilibrium equation).

7.3 Bringing the model to the data

Empirically, we estimate a log-wage equation at the individual level and then conduct a decomposition of the evolution of average log-wage at the city level between two dates. For city c, this evolution is denoted: $\log w_{c,t} - \log w_{c,t-1}$. Indexing model parameters by t for year t, we can isolate the evolution due to a change in agglomeration economies related to land size and population density, and the rest:

$$\log w_{c,t} - \log w_{c,t-1} = \log w_{c,t}\left(\alpha_t, \beta_t, A_t, B_t\right) - \log w_{c,t-1}\left(\alpha_{t-1}, \beta_{t-1}, A_{t-1}, B_{t-1}\right) \tag{22}$$

$$+ \left[\log w_{c,t} - \log w_{c,t}\left(\alpha_t, \beta_t, A_t, B_t\right)\right] \tag{23}$$

$$- \left[\log w_{c,t-1} - \log w_{c,t-1}\left(\alpha_{t-1}, \beta_{t-1}, A_{t-1}, B_{t-1}\right)\right] \tag{24}$$

where:

$$\log w_{c,t}\left(\alpha_t, \beta_t, A_t, B_t\right) = \log A_{c,t} + \beta_t \log\left(N_{c,t}/L_{c,t}\right) + \alpha_t \log L_{c,t} \tag{25}$$

is the log-wage equation specified in the model where $A_t = (A_{1,t}, ..., A_{C,t})'$ and $B_t = (B_{1,t}, ..., B_{C,t})'$. Note that wage is impacted by all city-specific total factor productivities and consumption amenities since they affect city land size and population at the equilibrium.

Considering that dates $t-1$ and t are close, we can write that:

$$\log w_{c,t}\left(\alpha_t, \beta_t, A_t, B_t\right) - \log w_{c,t-1}\left(\alpha_{t-1}, \beta_{t-1}, A_{t-1}, B_{t-1}\right) \approx d\log w_{c,t}\left(\alpha_t, \beta_t, A_t, B_t\right) \tag{26}$$

In particular, we are interested in the evolutions of log-wages when the values of agglomeration parameters α and β vary (and we want to leave aside their evolutions when city-specific total factor productivities and consumption amenities vary). Using equations (24) and (26), and a Taylor first-order approximation, it is possible to show that:

$$\log w_{c,t} - \log w_{c,t-1} = \frac{\partial \log w_{c,t}}{\partial \alpha}\left(\alpha_t - \alpha_{t-1}\right) + \frac{\partial \log w_{c,t}}{\partial \beta}\left(\beta_t - \beta_{t-1}\right) + r_{c,t} \tag{27}$$

where $r_{c,t}$ is a residual that captures discrepancies between observed and theoretical wages, as well as theoretical wage variations due to changes in total factor productivities and consumption amenities. The model allows the recovery of partial derivatives on the right-hand side of the decomposition (27), and they are given by equations (14) and (15) in the closed city case, and equations (21) and (B.123) in the open city case. Note that expressions make intervene $\log L_c$ that is affected by the choice of measurement units. In line with our empirical application, we consider city land size and density relatively to our reference city (i.e. rather than the theoretical objects $\log L_c$ and $\log(N_c/L_c)$, we consider $\log\left(L_c/L^*\right)$ and $\log\left(N_c/L_c/(N/L)^*\right)$ where L_c and N_c/L_c are values observed in the data, and L^* and $(N/L)^*$ are empirical minimum values for land size and density across cities). Implicitely, it means that there are no agglomeration economies in the reference city, and that agglomeration economies start with values higher than those for the reference city.

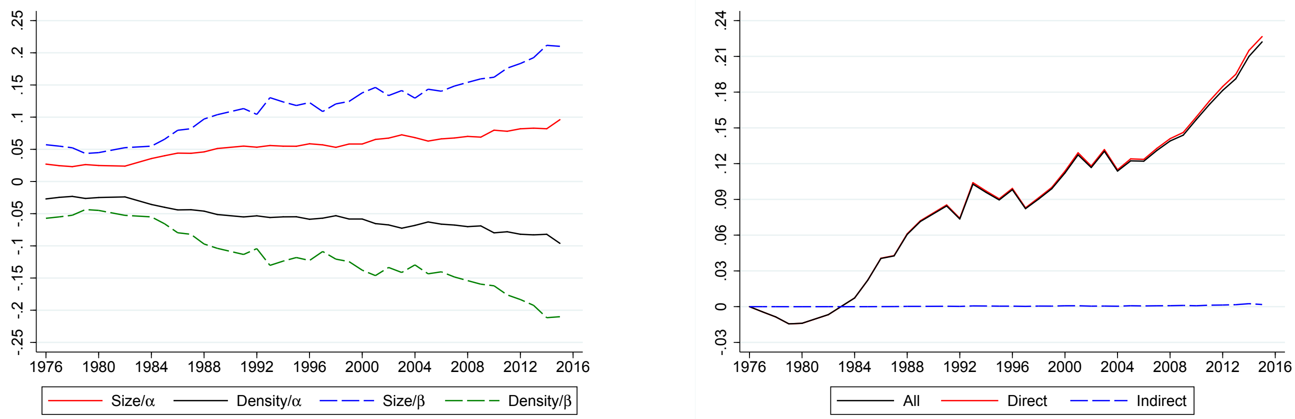
For parameters related to agglomeration economies, α_t and β_t, we consider the values obtained from our estimations. For quantities $\frac{\tau_{c,t} L_{c,t}}{w_{c,t}}$, $\mu_{c,t}$ and a_t, a first natural step is to consider that their values are the same for all cities. Interestingly, in that case, it is possible to check that the population-weighted averages of indirect migration effects on log-area, log-density and log-wages involved in equations (16), (17) and (21) are zero. In particular, this means that productivity increases in some cities due to positive changes of agglomeration variables are compensated by productivity decreases in other cities due to negative changes of those variables.

Figure 8.a represents the area and density elasticities weighted by the number of employed workers for the closed and open monocentric city models. They are identical in the two cases and quite low. This is consistent with agglomeration variables not varying much when changing their returns. It suggests that indirect effects of returns to agglomeration variables on wage growth should be small, which is confirmed by Figure 8.b where the indirect effects on the log-wage difference between 1976 and any other year is close to zero. These calibration results are in line with our empirical findings.

Figure 8: Elasticities of agglomeration variables and wage growth predicted by the model

6.a: Elasticities

6.b: Wage growth



Note: In panel a: Weighted averages of area and density elasticties with respect to returns to agglomeration variables α and β, where the weight is the number of employed workers. Values of these elasticities are equal in the cases of the open and closed monocentric city models. In panel b: difference in log-wage between 1976 and any given year.

8 Conclusion

In this paper, we assessed the effect of urbanisation on the evolution of wages, focusing on the role of agglomeration economies. We considered separately the effects of changes in the values of agglomeration variables and changes in their returns. We showed that, even if some cities grew significantly, their growth was not enough to affect aggregate labour earnings through agglomeration economies in a sizable way. We also documented an increase in the returns to density and area which greatly affected wages. Finally, we modelled a system of cities and showed that changes in returns to agglemeration economies do not affect enough values of density and area to significantly influence aggregate labour earnings, which is consistent with our empirical findings.

Overall, disparities in city sizes and the emergence of urban giants spans over centuries if not millenia. There is path dependence in development with buildings lasting for a long time period and such that population adjusts slowly across the territory. Our results show that forty years are, by far, not enough to generate population evolution in agglomerations that may impact wages significantly. Population changes are small compared to cross-section disparities in population across cities. Morover, most of the French population lives in cities, and the growth of some cities may be counterbalanced with the decline of some others. This means that gains in agglomeration economies for growing cities may go along with losses in declining ones. National population growth may play a role since city growth may be achieved without population leaving some cities, but natural and migratory balances are quite limited.

The country has also experienced important structural changes in the last three centuries, in particular with transitions from agriculture to manufacturing and services, and the improvement of tranport means. These changes have influenced the nature and intensity of agglomeration economies, and we showed that they impacted wages in a sizable way over only a few decades. Future research could investigate how different types of agglomeration economies evolved. This could inform public authorities on what margins may be influenced to impact productivity and labour earnings.

References

Ahlfeldt, Gabriel and Elisabetta Pietrostefani. 2019. The economic effects of density: A synthesis. *Journal of Urban Economics,* 111:93–107.

Bauluz, Luis, Paweł Bukowski, Mark Fransham, Annie Lee, Margarita Lopez Forero, Filip Novokmet, Sebastien Breau, Neil Lee, Clément Malgouyres, Moritz Schularick, and Gregory Verdugo. 2023. Spatial Wage Inequality in North America and Western Europe: Changes Between and Within Local Labour Markets 1975-2019. CEPR Working Paper 18381.

Behrens, Kristian, Gilles Duranton, and Frédéric Robert-Nicoud. 2014. Productive cities: Sorting, selection, and agglomeration". *Journal of Political Economy,* 122(3):507–553.

Bosker, Marteen, Jane Park, and Mark Roberts. 2021. Definition matters. metropolitan areas and agglomeration economies in a large-developing country. *Journal of Urban Economics,* 125. 103275.

Butts, Kyle, Taylor Jaworski, and Carl Kitchens. 2022. The urban wage premium in a historical perspective. NBER Working Paper 31387.

Card, David, Jesse Rothstein, and Moises Yi. 2023. Location, Location, Location. NBER Working Paper 31587.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon. 2008. Spatial wage disparities: Sorting matters! *Journal of Urban Economics,* 63(2):723–742.

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, and Sébastien Roux. 2010. Estimating agglomeration economies with history, geology, and worker effects. In Edward Glaeser (ed.) *Agglomeration economics.* The University of Chicago Press, 15–66.

Combes, Pierre-Philippe and Laurent Gobillon. 2015. The empirics of agglomeration economies. In Gilles Duranton, Vernon Henderson, and William Strange (eds.) *Handbook of regional and urban economics, Volume 5A*. North Holland: Elsevier, 247–348.

Combes, Pierre-Philippe, Miren Lafourcade, Jacques Thisse, and Jean-Claude Toutain. 2011. The rise and fall of spatial inequalities in France: A long-run perspective. *Explorations in Economic History,* 48(2):243–271.

de Bellefon, Marie-Pierre, Pierre-Philippe Combes, Gilles Duranton, Laurent Gobillon, and Clément Gorin. 2021. Delineating urban areas using building density. *Journal of Urban Economics,* 125. 103226.

de la Roca, Jorge and Diego Puga. 2017. Learning by working in big cities. *The Review of Economic Studies,* 84(1):106–142.

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

Diamond, Rebecca and Cécile Gaubert. 2022. Spatial sorting and inequality. *Annual Review of Economics,* 14:795–819.

Dittmar, Jeremiah. 2019. Historical Change in European City Populations: The Emergence of Zipf's Law. Working Paper.

Duranton, Gilles. 2015. A proposal to delineate metropolitan areas in Colombia. *Economia & Desarrollo,* 75(0):169–210.

Duranton, Gilles, Ejaz Ghani, Arti Grover Goswami, and William Kerr. 2015. The Misallocation of Land and Other Factors of Production in India. World Bank, Policy Research Working Paper 7221.

Duranton, Gilles and Diego Puga. 2014. The growth of cities. In Philippe Aghion and Steven Durlauf (eds.) *Handbook of economic growth, Volume 2.* North Holland: Elsevier, 781–853.

Duranton, Gilles and Diego Puga. 2020. The Economics of Urban Density. *Journal of Economic Perspective,* 34(3):3–26.

Duranton, Gilles and Diego Puga. 2023. Urban Growth and its Aggregate Implications. *Econometrica,* Forthcoming.

Eckert, Fabian, Sharat Ganapati, and Conor Walsh. 2022*a*. Urban-Biased Growth: A Macroeconomic Analysis. NBER Working Paper 30515.

Eckert, Fabien, Mads Hejlesen, and Conor Walsh. 2022*b*. The return to big-city experience: Evidence from refugees in Denmark. *Journal of Urban Economics,* 130(103454).

Gabaix, Xavier. 1999. Zipf's law for cities: an explanation. *The Quarterly journal of economics* 114(3):739–767.

Gabaix, Xavier and Yannis Ioannides. 2004. The evolution of city size distributions. In Vernon Henderson and Jacques Thisse (eds.) *Handbook of economic growth, Volume 4.* North Holland: Elsevier, 2341–2378.

Gaubert, Cécile. 2018. Firm sorting and agglomeration. *American Economic Review,* 108(11):3117–3153.

Giannone, Elisa. 2022. Skill-Biased Technical Change and Regional Convergence. Working Paper.

Glaeser, Edward and Joshua Gottlieb. 2008. The Economics of Place-Making Policies. *Brookings Papers on Economic Activity, Economic Studies Program, The Brookings Institution* 39(1):155–253.

Gyourko, Joseph and Raven Molloy. 2015. Regulation and housing supply. In Gilles Duranton, Vernon Henderson, and William Strange (eds.) *Handbook of regional and urban economics, Volume 5*. North Holland: Elsevier, 1289–1337.

Harris, C. D. 1954. The market as a factor in the localization of industry in the united states. *Annals of the Association of American Geographers* 44:315–348.

Hsieh, Chang-Tai and Enrico Moretti. 2019. Housing constraints and spatial misallocation. *American Economic Journal: Macroeconomics* 11(2):1–39.

Kline, Patrick and Enrico Moretti. 2014. People, places, and public policy: Some simple welfare economics of local economic development programs. *Annual Review of Economics* 6(1):629–662.

Koster, Hans and Ceren Ozgen. 2021. Cities and Tasks. *Journal of Urban Economics,* 126(103386).

Krugman, Paul. 1996. Confronting the mystery of urban hierarchy. *Journal of the Japanese and International economies* 10(4):399–418.

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

Rappaport, Jordan. 2007. Moving to nice weath. *Regional Science and Urban Economics,* 3:375–398.

Rappaport, Jordan. 2009. The increasing importance of quality of life. *Journal of economic geography,* 9(6):779–804.

Rappaport, Jordan and Jeffrey Sachs. 2003. The united states as a coastal nation. *Journal of Economic Growth,* 8:5–46.

Rosenthal, Stuart and William Strange. 2008. The attenuation of human capital spillovers. *Journal of Urban Economics,* 64(2):373–389.

Zipf, Georges. 1949. *Human Behaviour and the Principle of Least Effort.* Addison-Wesley, Reading, MA.

A Additional descriptive statistics on urban areas and cities

Table A.1: Descriptive statistics on all delineated urban areas

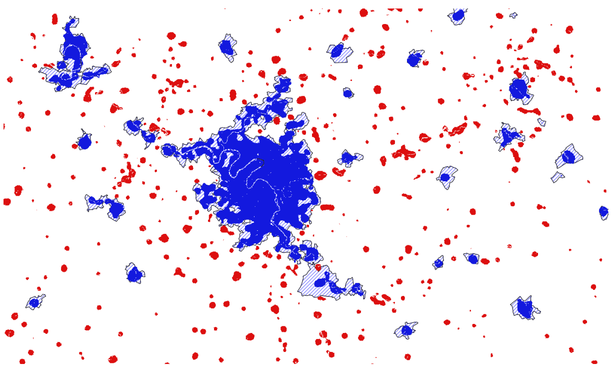
Var	Min	p25	p50	mean	p75	p95	Max
Panel A : 1975 (2930)							
Population	9	1096	2252	13409	4849	28625	9083917
Area	0.44	13.84	22.92	34.93	38	87.38	2911.36
Density	0.91	55.18	101.92	198.40	219.95	699.96	3120.16
Panel B : 1982 (2940)							
Population	29	1172	2373	13804	5135	28578	9372229
Area	0.44	14.24	23.44	36.48	39.07	92.58	3377.16
Density	1.67	58.43	107.64	195.50	223.36	666.42	2775.18
Panel C : 1990 (2922)							
Population	12	1230	2491	14493	5366	30913	9862985
Area	0.44	14.28	23.56	37.74	39.61	97.76	3519.44
Density	1.21	62.70	114.78	195.38	226.45	642.54	2802.43
Panel D : 1999 (2835)							
Population	9	1306	2636	15431	5674	32345	10116852
Area	0.44	14.42	23.84	38.86	40.28	97.14	3647.48
Density	0.62	65.77	124	199.83	237.72	626.82	2773.66
Panel E : 2006 (2872)							
Population	12	1436	2835	16388	5956	33380	10666306
Area	0.44	14.44	23.86	39.75	40.88	99.45	3696.40
Density	0.81	71.17	133.14	206.51	246.12	646.77	2885.59
Panel F : 2015 (2729)							
Population	14	1439	2931	16995	6087	35176	11078022
Area	0.44	14.60	24.24	41.45	41.84	104.46	3759.88
Density	0.88	72.14	135.93	199.44	247.89	592.87	2946.38

Table A.2: Descriptive statistics on delineated cities in 1975 and 2015

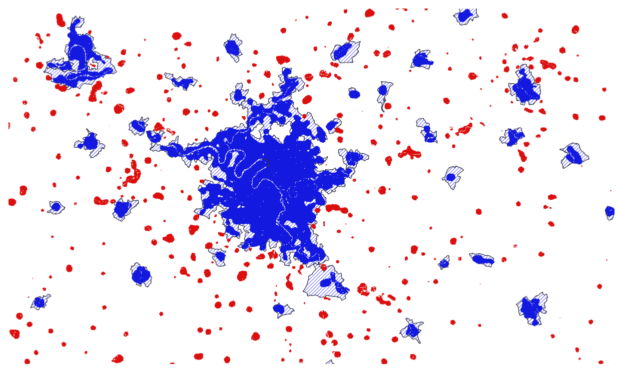
Panel A: 1975, bandwidth 2.1km, 286 cities							
	Min	p25	p50	mean	p75	p95	Max
Population	172	14,605	29,392	10,817	62,066	257,064	9,083,917
Area	4	42	65	113	101	272	2,911
Density	10.04	243.65	519.73	652.52	908.69	1671.53	3120.16
Panel B: 1975, bandwidth 1km, 565 cities							
	Min	p25	Median	Mean	p75	p95	Max
Population	462	5,564	11,825	55,037	32,269	176,757	8,593,972
Area	3	25	39	59	61	159	2,150
Density	10.04	186.20	361.09	564.32	760.54	1,670.70	3,995.78
Panel C: 2015, bandwidth 2.1km, 308 cities							
	Min	p25	p50	mean	p75	p95	Max
Population	187	13,068	27,327	120,231	66,156	283,045	11,078,022
Area	10	47	75	146	130	418	3,759
Density	5.63	244.07	408.28	485.42	663.86	1091.40	2946.38
Panel D: 2015, bandwidth 1km, 573 cities							
	Min	p25	Median	Mean	p75	p95	Max
Population	225	5,564	12,291	63,603	33,089	176,131	10,636,760
Area	3	27	44	77	75	199	3,039
Density	6.17	160.56	317.51	432.61	604.07	1156.02	3500.09

Figure A.1: City delineations for Paris and Lyon, 1975 and 2015

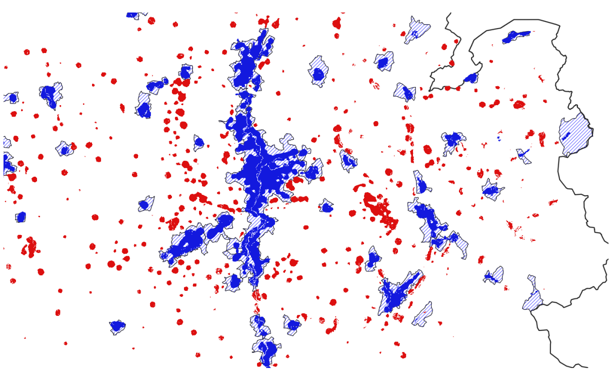
A.1.a: Paris, 1975



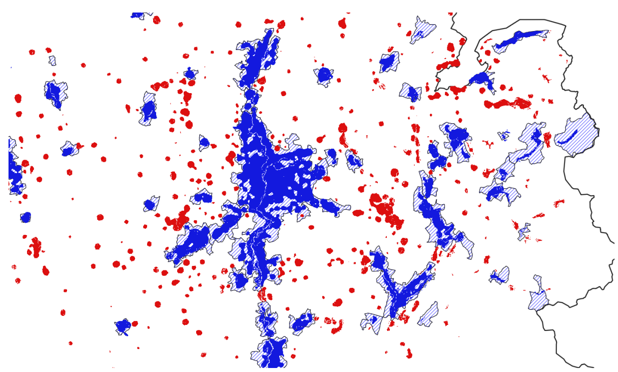
A.1.b: Paris, 2015



A.1.c: Lyon, 1975



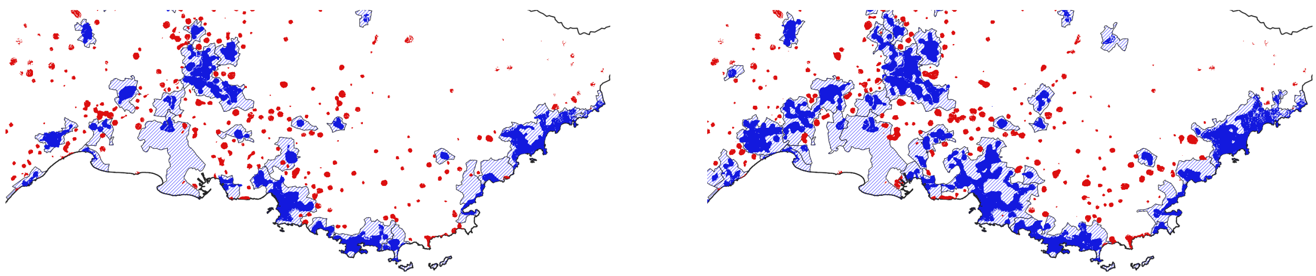
A.1.d: Lyon, 2015



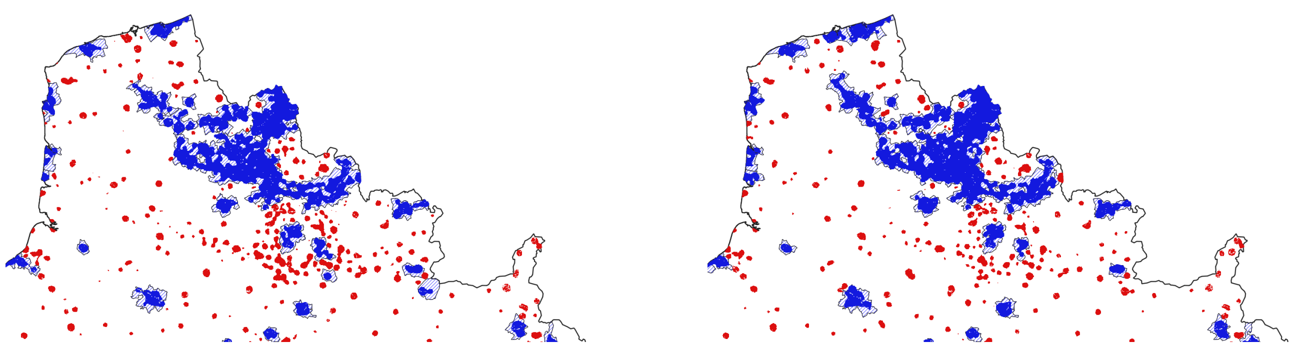
Note: Urban areas are obtained with our delineation algorithm run separately in 1975 and 2015 using a 2.1km bandwidth. Urban areas with cores (cities) are in blue and urban areas without core are in red. Borders of municipalities including part of a city are in black, and the area of the municipalities not covered by city are in blue stripes.

Figure A.2: City delineations for Lille and Marseille, 1975 and 2015

A.1.a: Marseille, 1975 | A.1.b: Marseille, 2015

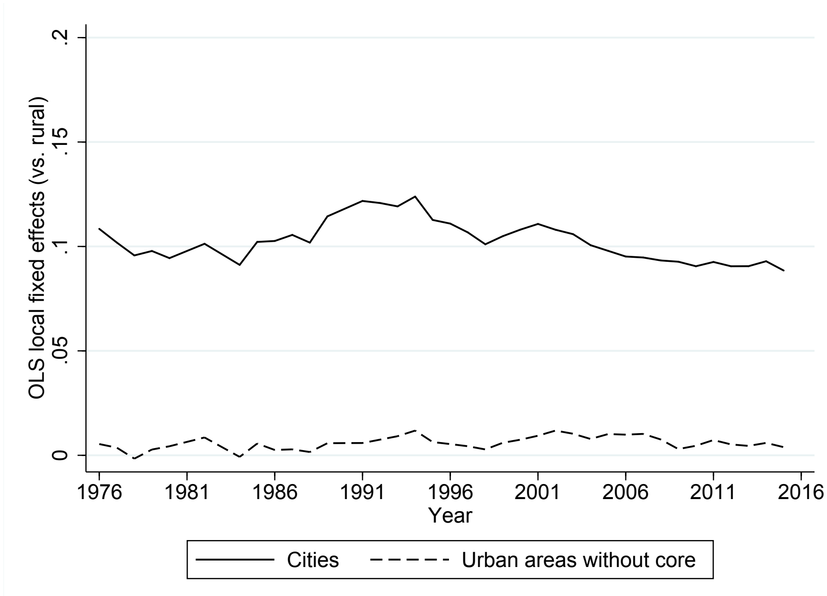


A.2.c: Lille, 1975 | A.2.d: Lille, 2015



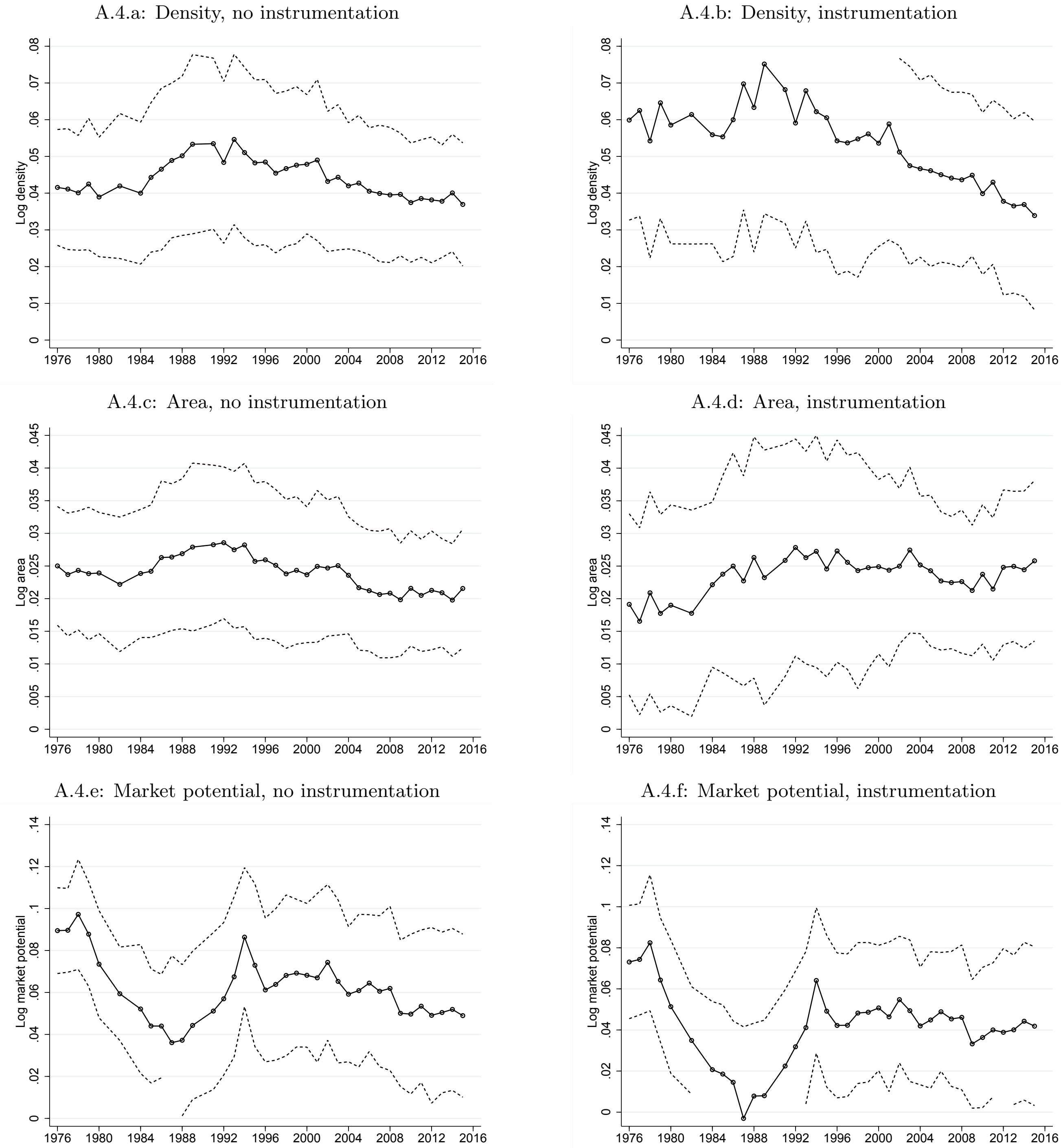
Note: Urban areas are obtained with our delineation algorithm run separately in 1975 and 2015 using a 2.1km bandwidth. Urban areas with cores (cities) are in blue and urban areas without core are in red. Borders of municipalities including part of a city are in black, and the area of the² municipalities not covered by city are in blue stripes.

Figure A.3: OLS yearly effects of working in an urban area without core or a city relatively to rural areas, individual unobserved heterogeneity is ignored



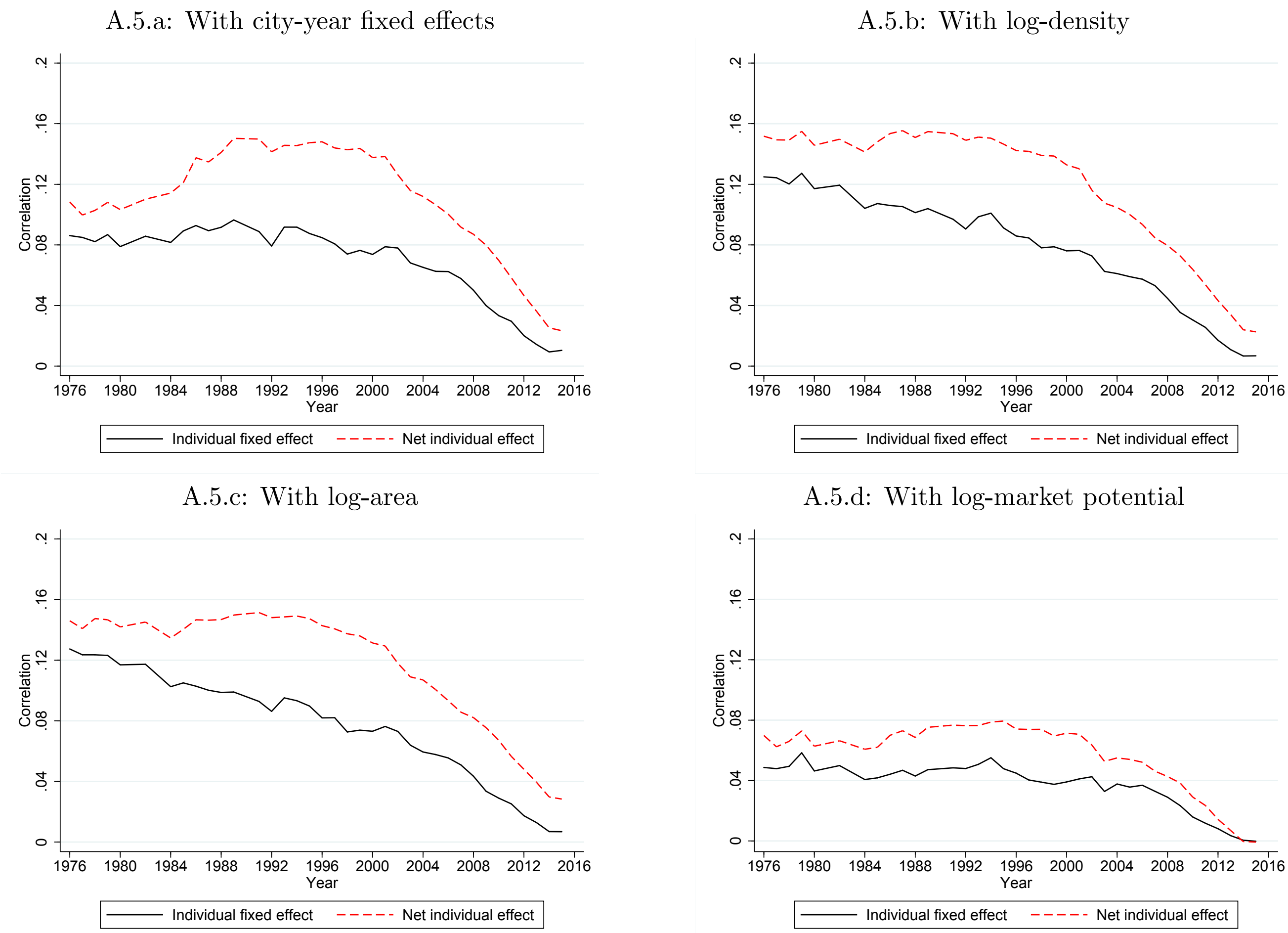
Note: We represent differences between yearly effects of being in an urban area without core or in a city and yearly effects of being in a rural area. Yearly effects of being in a city are yearly averages of city-year fixed effects, weighting by the yearly number of individuals working in cities.

Figure A.4: Estimated yearly coefficients of city variables without individual effects in the first-stage specification, with and without instrumentation with historical and geological variables



Note: Estimated coefficients are represented by bullets and linked with a plain line, and bounds of confidence intervals are represented in dots (with interruptions if bound values are outside the y-axis grid). Historical instruments: logarithms of population densities in 1793, 1800, 1836 and 1856, and market potentials for the same years. Soil instruments: shares of the city area by levels of depth to rock, soil erodability, hydrogeological class, subsoil mineralogy, and topsoil organic carbon content.

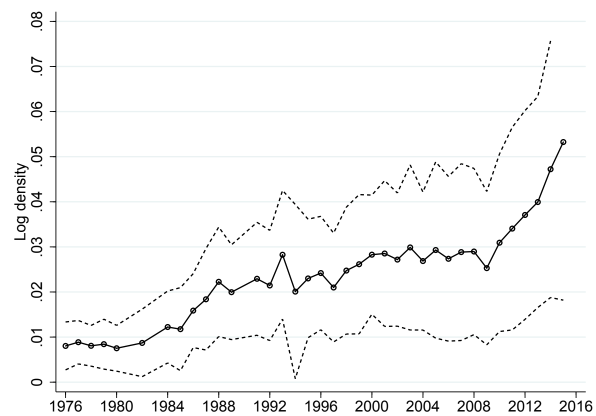
Figure A.5: Correlation between individual fixed effects and city fixed effects or city variables



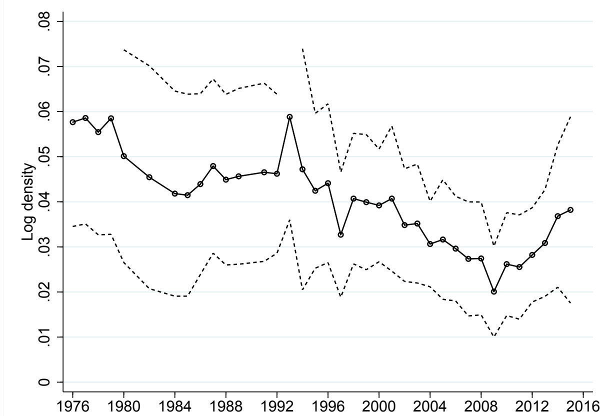
Note: "Individual fixed effect": Correlation between the city quantity given in panel title and individual fixed effects. "Net individual effect": Correlation between the city quantity given in panel title and the net individual effects defined as individual averages of residuals obtained when regressing the sum of individual fixed effect and squared age on age, squared age and year fixed effects. Observations are at the worker-year level.

Figure A.6: Estimated yearly coefficients of city variables with or without individual fixed effects in the first-stage specification, when restricting the sample to the first four observations of individuals appearing at least four times in the panel

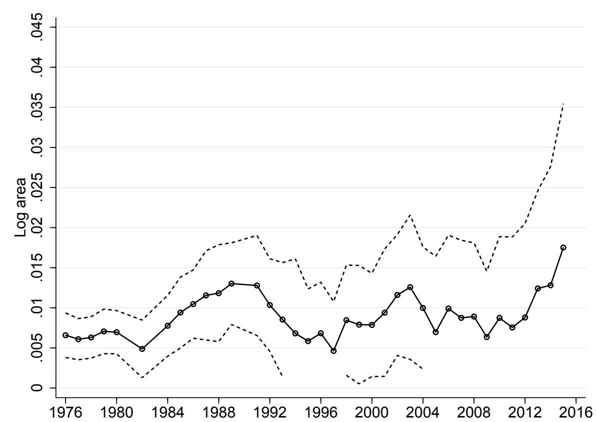
A.6.a: Density, with individual fixed effects



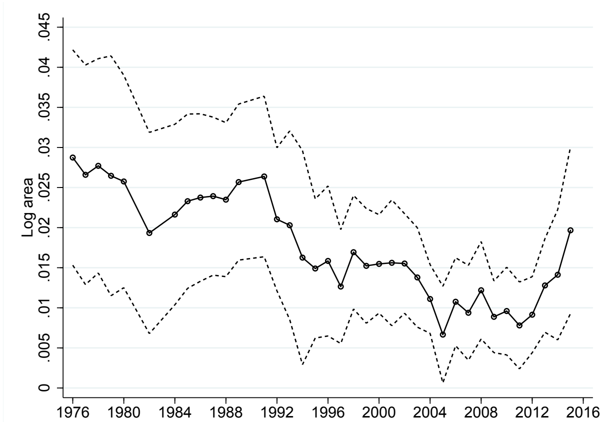
A.6.b: Density, without individual fixed effects



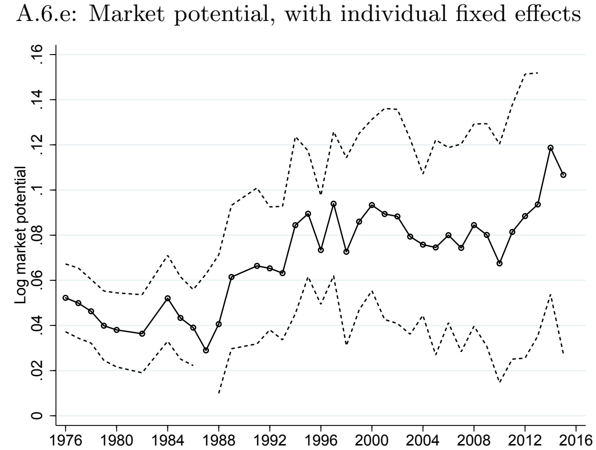
A.6.c: Area, with individual fixed effects



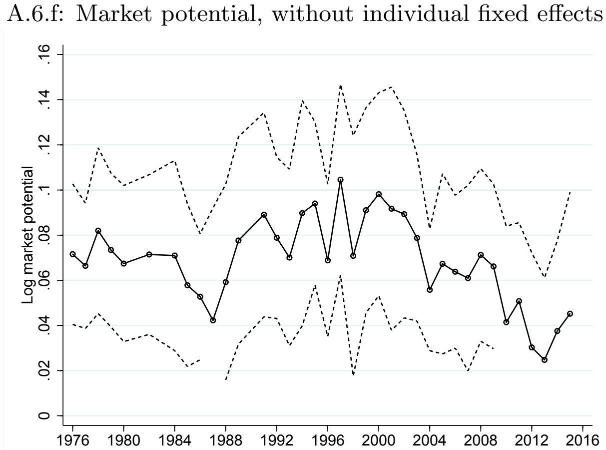
A.6.d: Area, without individual fixed effects



A.6.e: Market potential, with individual fixed effects

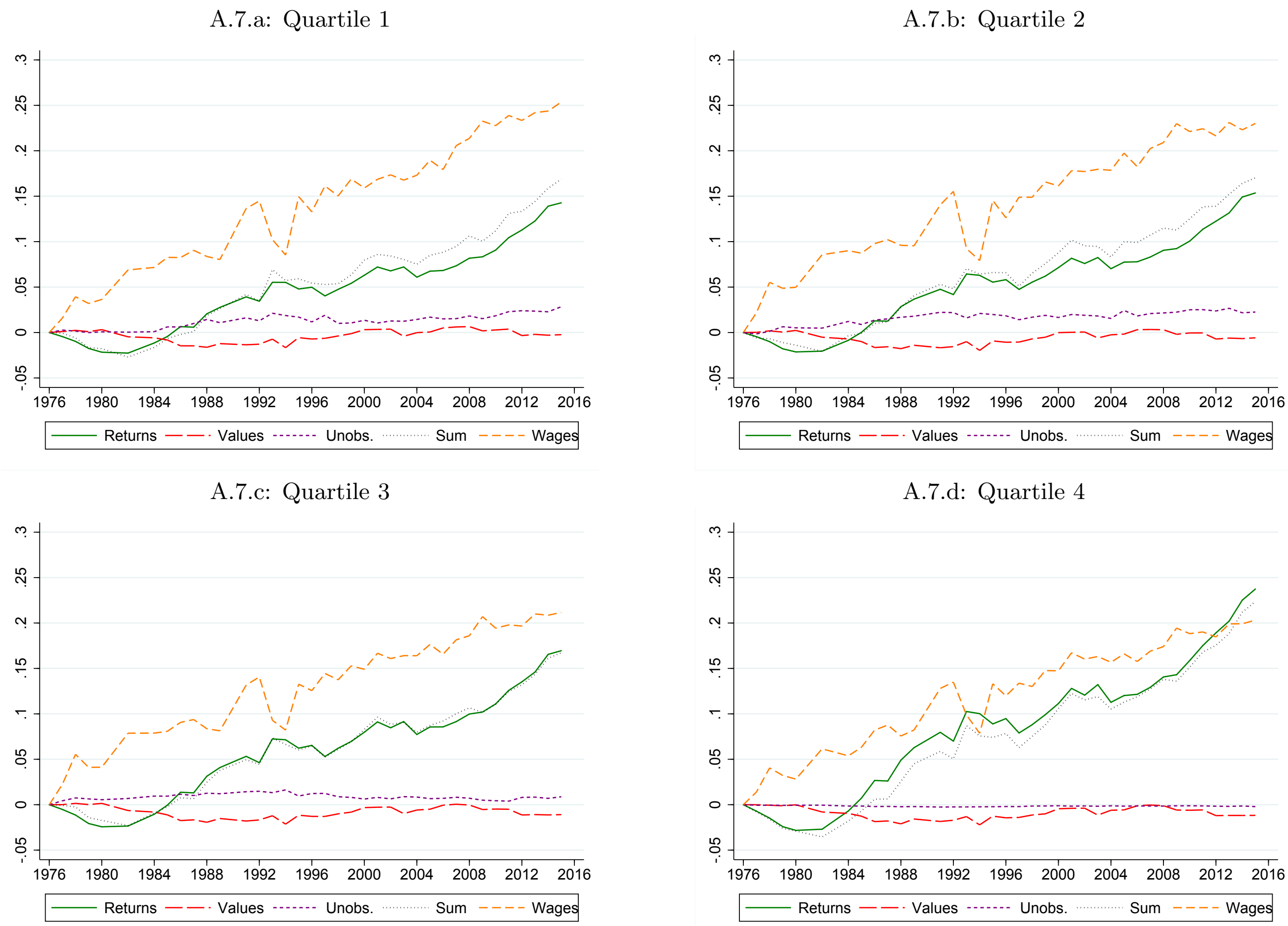


A.6.f: Market potential, without individual fixed effects



Note: Estimated coefficients are represented by bullets and linked with a plain line, and bounds of confidence intervals are represented in dots (with interruptions if bound values are outside the y-axis grid). Historical instruments: logarithms of population densities in 1793, 1800, 1836 and 1856, and market potentials for the same years. Soil instruments: shares of the city area by levels of depth to rock, soil erodability, hydrogeological class, subsoil mineralogy, and topsoil organic carbon content.

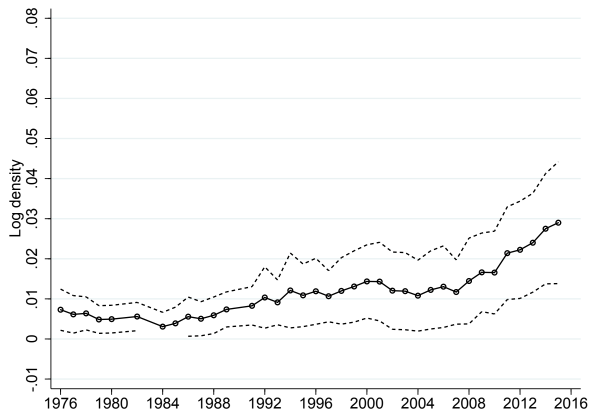
Figure A.7: Decomposition of city wage growth between 1976 and year t, with $1976 \leqslant t \leqslant 2015$, by city employment quartile



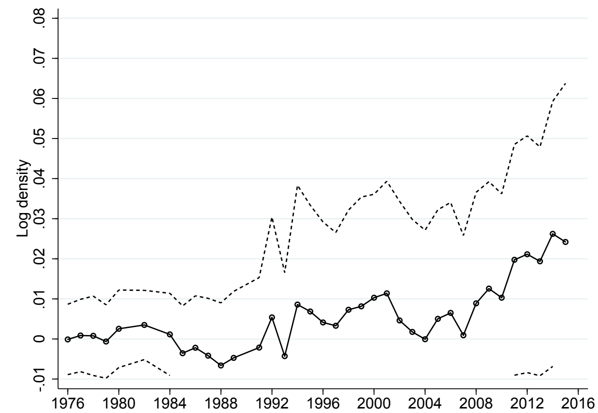
Note: "Returns": Contributions of changes in returns of city variables to log-wage growth; "Values": Contribution of changes in their values; "Unobs.": Contribution of changes in city unobservables; "Sum": Sum of these three contributions; "Wages": log-wage growth.

Figure A.8: Estimated yearly coefficients of city variables when there are individual fixed effects in the first-stage specification, building volume densities in pixels smoothed using a 1km bandwidth

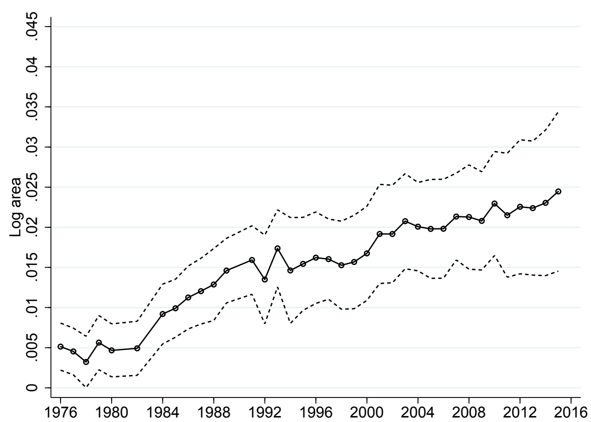
4.a: Density, no instrumentation



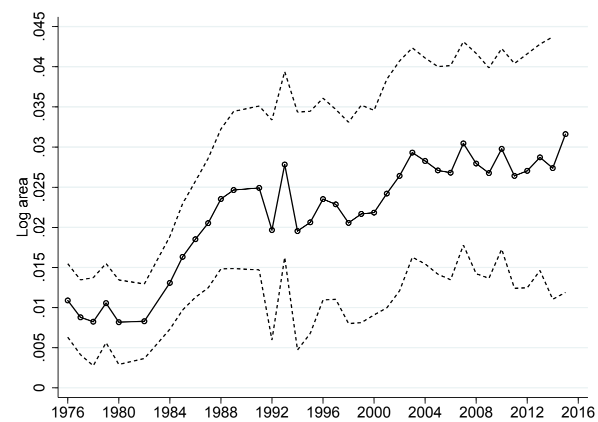
4.b: Density, instrumentation



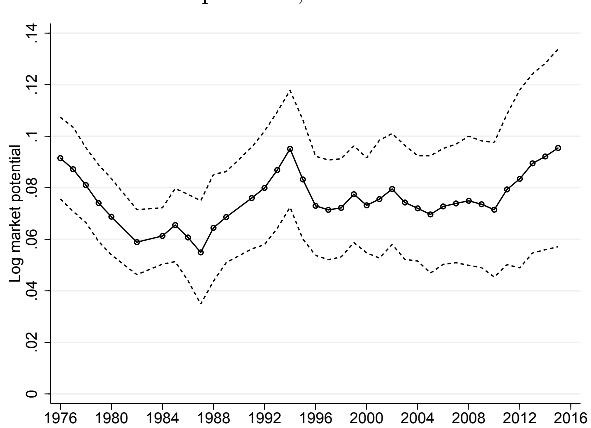
4.c: Area, no instrumentation



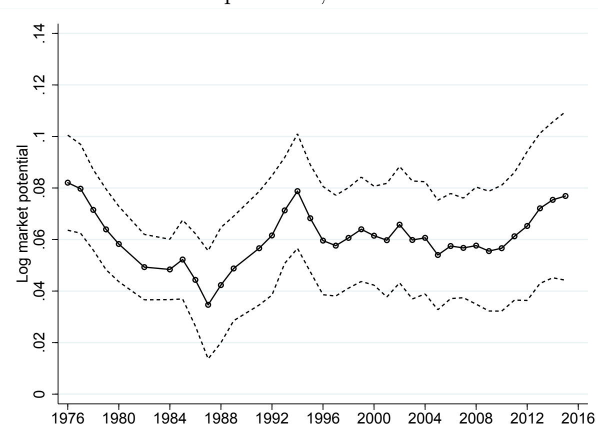
4.d: Area, instrumentation



4.e: Market potential, no instrumentation



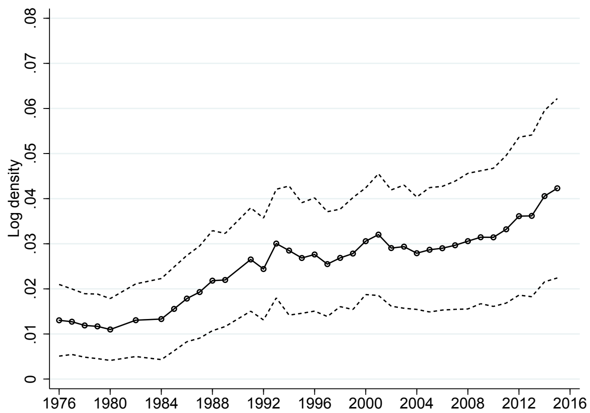
4.f: Market potential, instrumentation



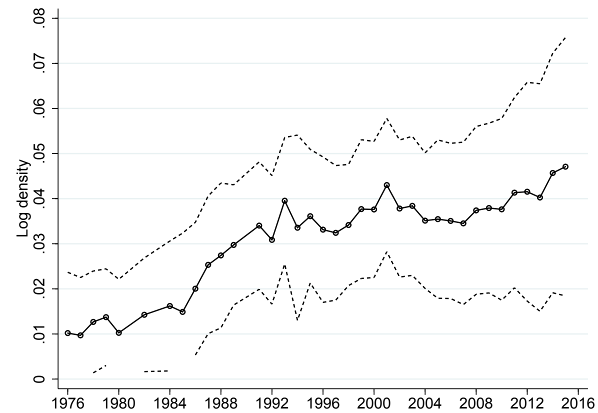
Note: Estimated coefficients are represented by bullets and linked with a plain line, and bounds of confidence intervals are represented in dots (with interruptions if bound values are outside the y-axis grid). Historical instruments: logarithms of population densities in 1793, 1800, 1836 and 1856, and market potentials for the same years. Soil instruments: shares of the city area by levels of depth to rock, soil erodability, hydrogeological class, subsoil mineralogy, and topsoil organic carbon content.

Figure A.9: Estimated yearly coefficients of city variables when there are individual fixed effects in the first-stage specification, 2015 delineation of cities for all years

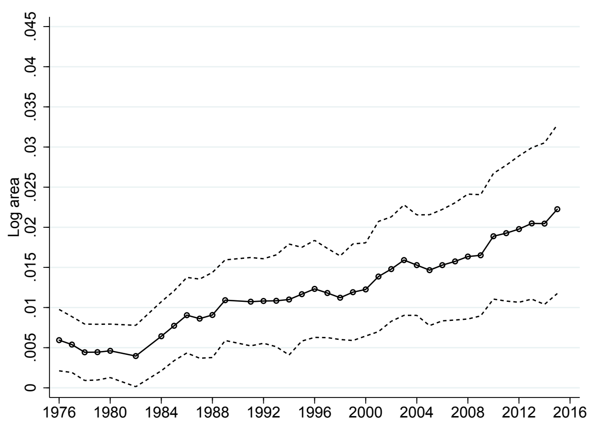
4.a: Density, no instrumentation



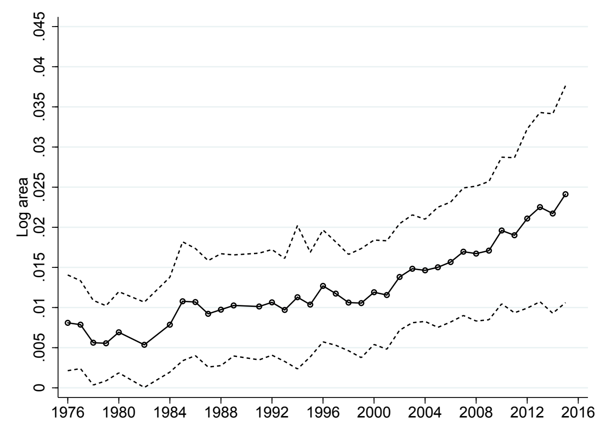
4.b: Density, instrumentation



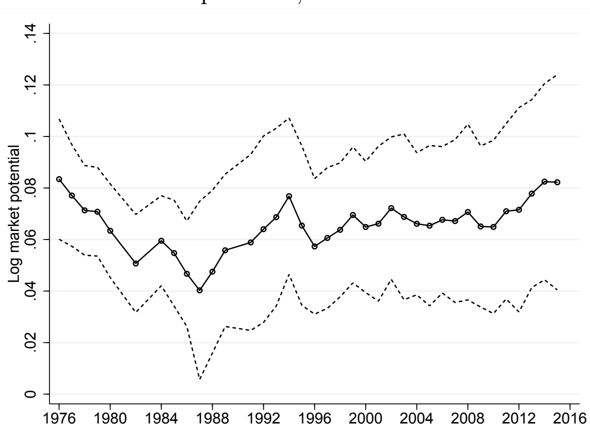
4.c: Area, no instrumentation



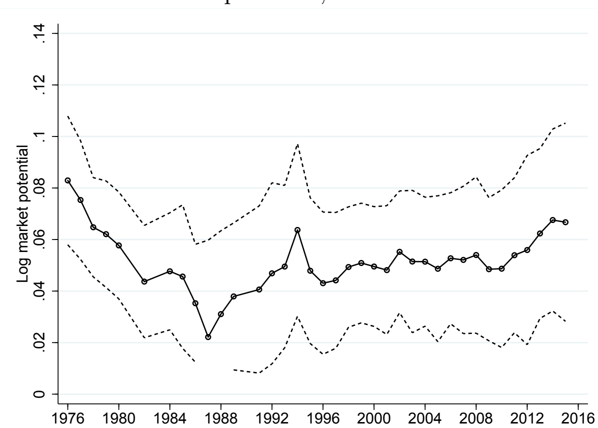
4.d: Area, instrumentation



4.e: Market potential, no instrumentation



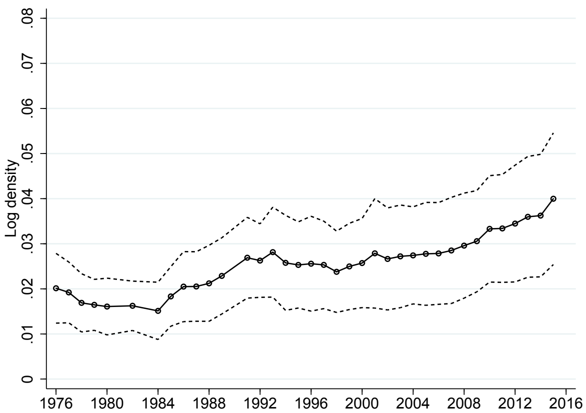
4.f: Market potential, instrumentation



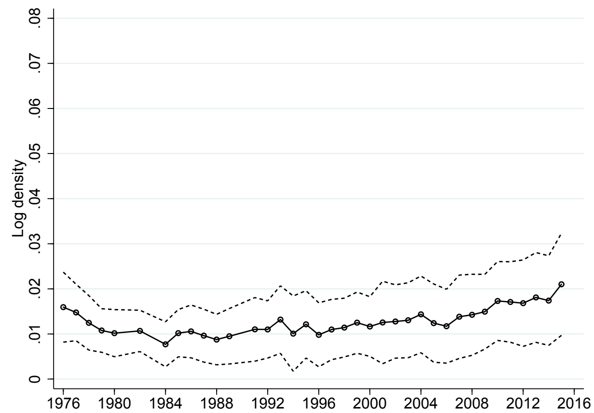
Note: Estimated coefficients are represented by bullets and linked with a plain line, and bounds of confidence intervals are represented in dots (with interruptions if bound values are outside the y-axis grid). Historical instruments: logarithms of population densities in 1793, 1800, 1836 and 1856, and market potentials for the same years. Soil instruments: shares of the city area by levels of depth to rock, soil erodability, hydrogeological class, subsoil mineralogy, and topsoil organic carbon content.

Figure A.10: Estimated yearly coefficients of city variables when there are individual fixed effects in the first-stage specification, INSEE 2010 urban area definition of cities for all years

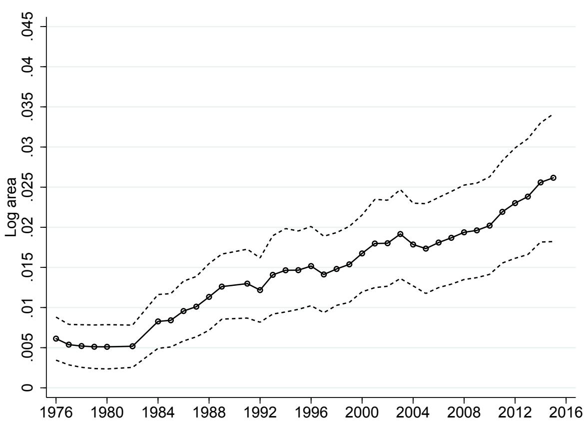
A.10.a: Density, no instrumentation



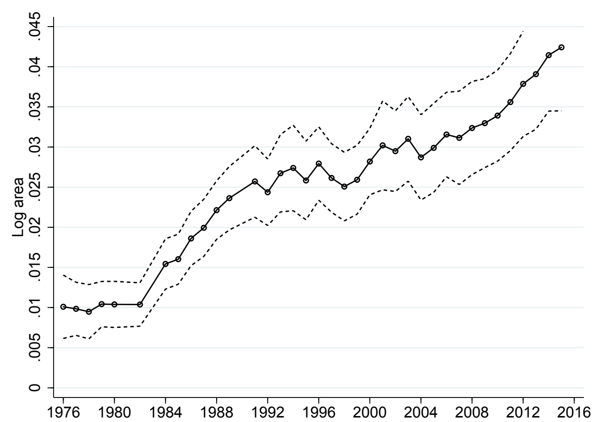
A.10.b: Density, instrumentation



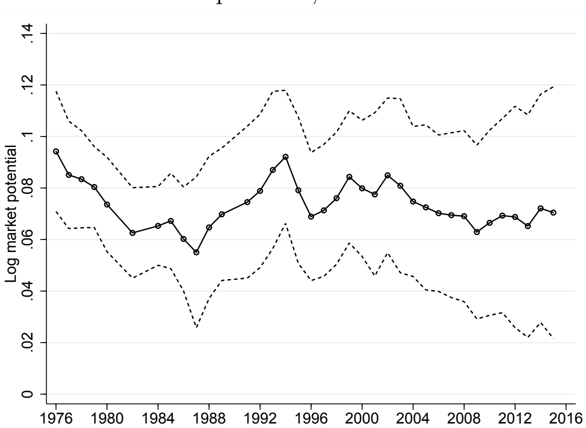
A.10.c: Area, no instrumentation



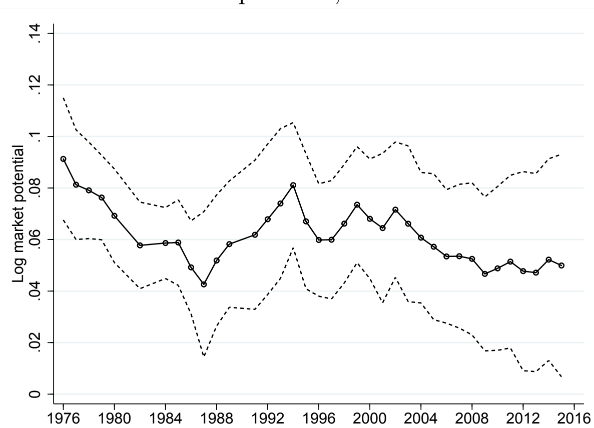
A.10.d: Area, instrumentation



A.10.e: Market potential, no instrumentation



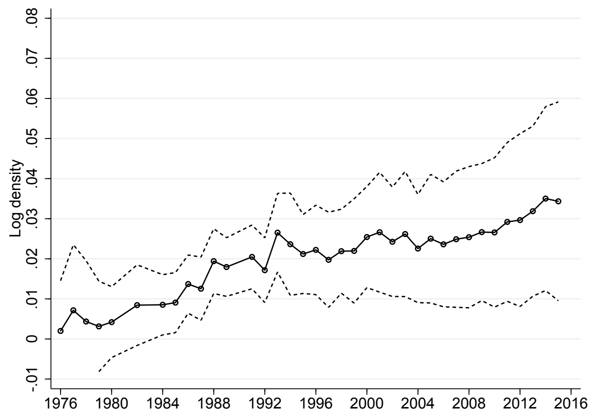
A.10.f: Market potential, instrumentation



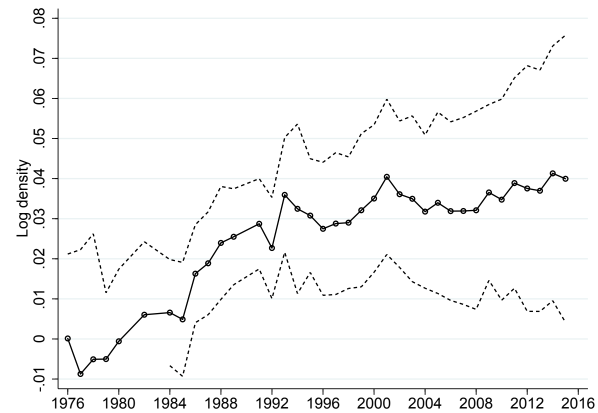
Note: Estimated coefficients are represented by bullets and linked with a plain line, and bounds of confidence intervals are represented in dots. Historical instruments: logarithms of population densities in 1793, 1800, 1836 and 1856, and market potentials for the same years. Soil instruments: shares of the city area by levels of depth to rock, soil erodability, hydrogeological class, subsoil mineralogy, and topsoil organic carbon content.

Figure A.11: Estimated yearly coefficients of city variables when there are individual fixed effects and learning effects in the first-stage specification

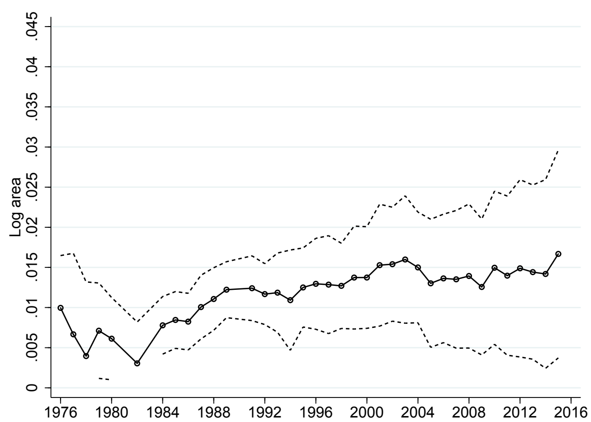
A.11.a: Density, no instrumentation



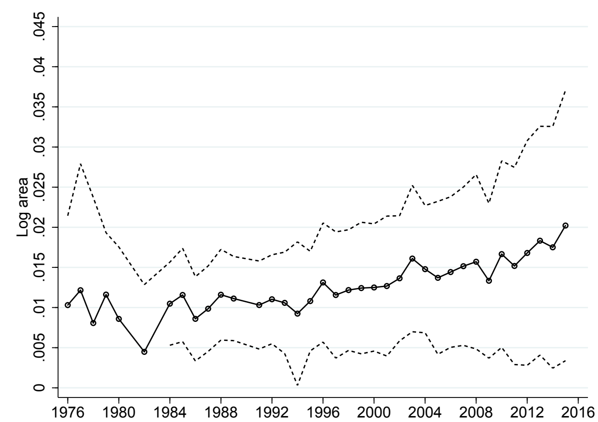
A.11.b: Density, instrumentation



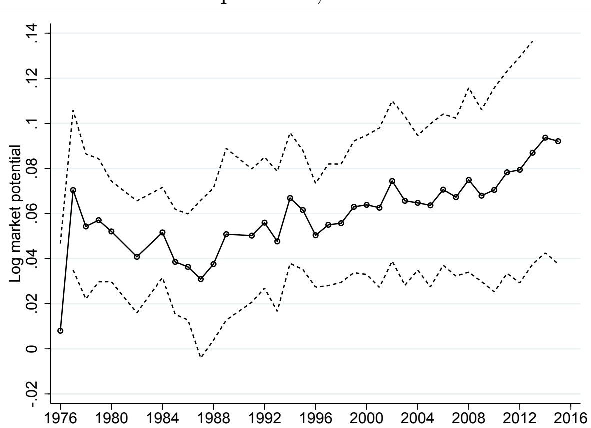
A.11.c: Area, no instrumentation



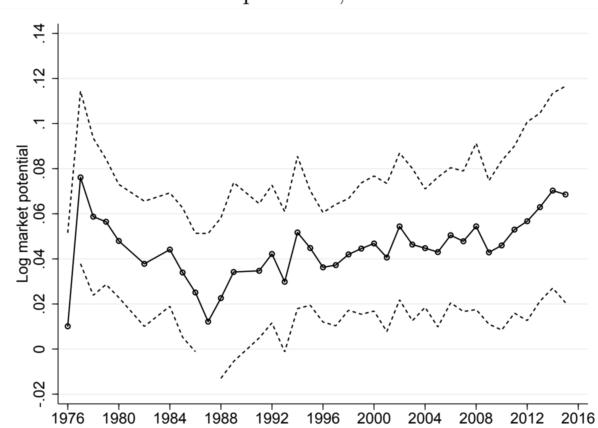
A.11.d: Area, instrumentation



A.11.e: Market potential, no instrumentation



A.11.f: Market potential, instrumentation



Note: Estimated coefficients are represented by bullets and linked with a plain line, and bounds of confidence intervals are represented in dots(with interruptions if bound values are outside the y-axis grid). Historical instruments: logarithms of population densities in 1793, 1800, 1836 and 1856, and market potentials for the same years. Soil instruments: shares of the city area by levels of depth to rock, soil erodability, hydrogeological class, subsoil mineralogy, and topsoil organic carbon content.

B Appendix: model

In this Appendix, we provide detailed developments for our monocentric model and establish formulas given in the main text. Workers maximize their utility under budget constraint. The first order condition is given by:

$$R_c(x) = \frac{\partial U}{\partial \ell} / \frac{\partial U}{\partial z} \tag{B.28}$$

From this equation and the budget constraint, we get the optimal consumption quantities:

$$\begin{aligned} \ell_c(x) &= a(w_c - \tau_c x)/R_c(x) && \text{(B.29)} \\ z_c(x) &= (1-a)(w_c - \tau_c x) && \text{(B.30)} \end{aligned}$$

At spatial equilibrium, all the individuals within the city get the same utility $\overline{u}$:

$$\begin{aligned} U(\ell_c(x), w_c - \tau_c x - R_c(x)\ell_c(x)) &= \overline{u} && \text{(B.31)} \\ B_c[a(w_c - \tau_c x)/R_c(x)]^a [(1-a)(w_c - \tau_c x)]^{1-a} &= \overline{u} && \text{(B.32)} \\ B_c(w_c - \tau_c x)/R_c(x)^a &= \overline{u} && \text{(B.33)} \end{aligned}$$

Deriving equation (B.31) with respect to x, we get:

$$\frac{\partial U}{\partial \ell}\frac{\partial \ell_c(x)}{\partial x} - \frac{\partial U}{\partial z}R_c(x)\frac{\partial \ell_c(x)}{\partial x} + \frac{\partial U}{\partial z}\left[-\tau_c - \frac{\partial R_c(x)}{\partial x}\ell_c(x)\right] = 0 \tag{B.34}$$

Using the first-order condition (B.28), the first two terms cancel out and we are left with the Alonso-Muth condition:

$$\frac{\partial R_c(x)}{\partial x} = -\frac{\tau_c}{\ell_c(x)} \tag{B.35}$$

The fringe is determined by the equality $R_c(\overline{x}_c) = \underline{R}$ where $\underline{R}$ is the agricultural land price. Land occupied by individuals is the segment $[0, \overline{x}_c]$ where $\overline{x}_c$ is the city fringe, such that $L_c = \overline{x}_c$, and city population verifies the equilibrium equation:

$$N_c = \int_0^{\overline{x}} n_c(x)\,dx = \int_0^{\overline{x}} \frac{1}{\ell_c(x)}dx = -\frac{1}{\tau_c}\left[R_c(\overline{x}) - R_c(0)\right] \tag{B.36}$$

where $n_c(x)$ is the population density at distance x (equation to the ratio between land supply 1 divided by land demand per individual $\ell_c(x)$).

B.1 Closed-city case

When city c is closed, its population N_c is fixed and only its land size $L_c = \overline{x}_c$ can vary. We conduct comparative statics with respect to a change in agglomeration economies parameter, α or β. Deriving equation (7), we get:

$$\frac{\partial \log w_c}{\partial \alpha} = \log L_c + (\alpha - \beta) \frac{1}{L_c} \frac{\partial L_c}{\partial \alpha} \tag{B.37}$$

$$\frac{\partial \log w_c}{\partial \beta} = \log \left(\frac{N_c}{L_c} \right) + (\alpha - \beta) \frac{1}{L_c} \frac{\partial L_c}{\partial \beta} \tag{B.38}$$

Importantly, agglomeration economies vary not only because of the change in parameter, but also because of the resulting change in city land size that affects productivity through agglomeration economies (which is an equilibrium effect).

B.1.1 Elasticities with respect to parameter α

We first consider variations in α, and then turn to variations in β. We derive the expression for $R_c(x)\,\ell_c(x)$ given by equation (B.29), which yields:

$$\ell_c(x) \frac{\partial R_c}{\partial \alpha} + R_c(x) \frac{\partial \ell_c}{\partial \alpha} = a \frac{\partial w_c}{\partial \alpha} \tag{B.39}$$

$$\frac{\partial \ell_c}{\partial \alpha} = \frac{1}{R_c(x)} \left[a \frac{\partial w_c}{\partial \alpha} - \ell_c(x) \frac{\partial R_c}{\partial \alpha} \right] \tag{B.40}$$

$$= \frac{1}{R_c(x)} \left[a \frac{\partial w_c}{\partial \alpha} - a (w_c - \tau_c x) \frac{1}{R_c(x)} \frac{\partial R_c}{\partial \alpha} \right] \tag{B.41}$$

Changes in land consumption are the sum of two terms: an income effect and a substitution effect (i.e. individuals substitute the composite good for land if land becomes too costly). This is explained at length by Duranton and Handbury (2022) in the case of a change in commuting costs due to working from home.

At the equilibrium, utility is equal in any given location x and the center. From equation (B.33), this equality can be rewritten in the following way:

$$(w_c - \tau_c x) / R_c(x)^a = w_c / R_c(0)^a \tag{B.42}$$

We derive the land market clearing condition given by equation (B.36), holding N fixed, since we are in the closed city case. We get:

$$\frac{\partial R_c(0)}{\partial \alpha} = \frac{\partial R_c(\overline{x}_c)}{\partial \alpha} \tag{B.43}$$

Finally, deriving the fringe condition $R_c(\overline{x}_c) = \underline{R}_c$, we obtain $\partial R_c(\overline{x}_c) / \partial \alpha = 0$, and thus:

$$\frac{\partial R_c(0)}{\partial \alpha} = 0 \tag{B.44}$$

Deriving the logarithm of expression (B.42) with respect to α and using the equality (B.44) gives:

$$\frac{1}{w_c - \tau_c x}\frac{\partial w_c}{\partial \alpha} - a\frac{1}{R_c(x)}\frac{\partial R_c}{\partial \alpha} = \frac{1}{w_c}\frac{\partial w_c}{\partial \alpha} \tag{B.45}$$

or equivalently:

$$a\frac{1}{R_c(x)}\frac{\partial R_c}{\partial \alpha} = \left(\frac{1}{w_c - \tau_c x} - \frac{1}{w_c}\right)\frac{\partial w_c}{\partial \alpha} \tag{B.46}$$

$$= \frac{\tau_c x}{w_c(w_c - \tau_c x)}\frac{\partial w_c}{\partial \alpha} \tag{B.47}$$

Changes in wages due to changes in agglomeration economies are capitalised into land prices. The increase in land prices is larger in percentage as one gets further away from the CBD. Inserting expression (B.47) into equation (B.41) gives:

$$\frac{\partial \ell_c}{\partial \alpha} = \frac{1}{R_c(x)}\left(a\frac{\partial w_c}{\partial \alpha} - \frac{\tau_c x}{w_c}\frac{\partial w_c}{\partial \alpha}\right) = \frac{1}{R_c(x)}\left(a - \frac{\tau_c x}{w_c}\right)\frac{\partial w_c}{\partial \alpha} \tag{B.48}$$

Importantly, $\frac{\partial \ell_c}{\partial \alpha}$ is of the same sign as $\frac{\partial w_c}{\partial \alpha}$ as long as the share of commuting costs in wages $\tau_c x / w_c$ is lower than the share of land in spendings. Usually, one considers that $a \approx .3$ and commuting costs must thus be very large for the substitution effect to dominate.

We have $L_c = \overline{x}_c$, and thus: $\frac{\partial L_c}{\partial \alpha} = \frac{\partial \overline{x}_c}{\partial \alpha}$. Deriving the equality $R_c(\overline{x}_c) = \underline{R}_c$ and using the Alonso-Muth condition (B.35), we get:

$$\frac{\partial R_c(\overline{x}_c)}{\partial \alpha} + \frac{\partial \overline{x}_c}{\partial \alpha}\frac{\partial R(\overline{x}_c)}{\partial x} = 0 \tag{B.49}$$

$$\frac{\partial R_c(\overline{x}_c)}{\partial \alpha} - \frac{\tau_c}{\ell_c(\overline{x}_c)}\frac{\partial \overline{x}_c}{\partial \alpha} = 0 \tag{B.50}$$

Then, using that fact that $L_c = \overline{x}_c$, as well as equations (B.47) and (B.29), we get:

$$\frac{\partial L_c}{\partial \alpha} = \frac{\ell_c(\overline{x}_c)}{\tau_c}\frac{\partial R_c(\overline{x}_c)}{\partial \alpha} \tag{B.51}$$

$$\frac{\partial L_c}{\partial \alpha} = \frac{R_c(\overline{x}_c)\,\ell_c(\overline{x}_c)\,\overline{x}_c}{w_c(w_c - \tau_c \overline{x}_c)}\frac{1}{a}\frac{\partial w_c}{\partial \alpha} \tag{B.52}$$

$$\frac{\partial L_c}{\partial \alpha} = \frac{L_c}{w_c}\frac{\partial w_c}{\partial \alpha} \tag{B.53}$$

Importantly, whereas land consumption in some locations within the city may vary in the opposite way of wages as a increases (for specific values of the parameters), aggregate land consumption always varies in the same way. Put differently, at the aggregate level, the substitution effect is dominated by the income effect. Interestingly, equation (B.53) can also be obtained by deriving the logarithm of the equality between utilities in the center and at the fringe given by equation (B.42). Indeed, derivation of its logarithm gives:

$$\frac{1}{w_c - \tau_c L_c}\left(\frac{\partial w_c}{\partial \alpha} - \tau_c \frac{\partial L_c}{\partial \alpha}\right) = \frac{1}{w_c}\frac{\partial w_c}{\partial \alpha} \tag{B.54}$$

$$\frac{\partial w_c}{\partial \alpha} - \tau_c \frac{\partial L_c}{\partial \alpha} = \left(1 - \frac{\tau_c L_c}{w_c}\right)\frac{\partial w_c}{\partial \alpha} \tag{B.55}$$

$$-\tau_c \frac{\partial L_c}{\partial \alpha} = -\frac{\tau_c L_c}{w_c}\frac{\partial w_c}{\partial \alpha} \tag{B.56}$$

$$\frac{\partial L_c}{\partial \alpha} = \frac{L_c}{w_c}\frac{\partial w_c}{\partial \alpha} \tag{B.57}$$

Then, inserting expression (B.53) into equation (B.37), we obtain:

$$\frac{1}{w_c}\frac{\partial w_c}{\partial \alpha} = \log L_c + (\alpha - \beta)\frac{1}{L_c}\frac{L_c}{w_c}\frac{\partial w_c}{\partial \alpha} \tag{B.58}$$

$$[1 + (\beta - \alpha)]\frac{1}{w_c}\frac{\partial w_c}{\partial \alpha} = \log L_c \tag{B.59}$$

Since $\alpha < 1$, $\beta > 0$ and $L_c > 1$, we have $\frac{\partial w_c}{\partial \alpha} > 0$, i.e. wages increase as agglomeration economies with respect to city land size increases (while holding population constant). From equations (B.53) and (B.59), we get variations in city land size:

$$\frac{\partial L_c}{\partial \alpha} = \frac{1}{1 + \beta - \alpha} L_c \log L_c \tag{B.60}$$

$$\frac{\partial \log L_c}{\partial \log \alpha} = \frac{\alpha}{1 + \beta - \alpha}\log L_c \tag{B.61}$$

City land size varies to a larger extent when the intensity of density agglomeration economies β is smaller and the intensity of land agglomeration economies α is larger. In particular, as land size increases, population density decreases (since population is constant) and this lowers density agglomeration economies. The larger β, the larger the loss.

It is then easy to establish a relationship for the variations of density with α since population is fixed. We have, using (B.61):

$$\frac{\partial \log (N_c/L_c)}{\partial \log \alpha} = -\frac{\partial \log L_c}{\partial \log \alpha} = -\frac{\alpha}{1 + \beta - \alpha}\log L_c \tag{B.62}$$

We can then turn to variations of wages. Inserting expressions (B.62) into equation (B.37), we get:

$$\begin{aligned} \frac{\partial \log w_c}{\partial \alpha} &= \log L_c + \frac{1}{\alpha}(\alpha - \beta)\frac{\partial \log L_c}{\partial \log \alpha} \\ &= \log L_c + \frac{\alpha - \beta}{1 + \beta - \alpha}\log L_c \end{aligned}$$

B.1.2 Elasticities with respect to parameter β

We now consider a change in β. The only difference when establishing a formula for variations in city land size comes from a difference in the formula of wage variations. By analogy, we have similarly to equation (B.53):

$$\frac{\partial L_c}{\partial \beta} = \frac{L_c}{w_c}\frac{\partial w_c}{\partial \beta} \tag{B.63}$$

Inserting expression (B.63) into equation (B.38) yields:

$$\frac{1}{w_c}\frac{\partial w_c}{\partial \beta} = \log\left(\frac{N_c}{L_c}\right) + (\alpha - \beta)\frac{1}{w_c}\frac{\partial w_c}{\partial \beta} \tag{B.64}$$

or equivalently:

$$[1 + (\beta - \alpha)]\frac{1}{w_c}\frac{\partial w_c}{\partial \beta} = \log\left(\frac{N_c}{L_c}\right) \tag{B.65}$$

From equations (B.63) and (B.65), we get variations in city land size:

$$\frac{\partial L_c}{\partial \beta} = \frac{1}{1+\beta-\alpha} L_c \log\left(\frac{N_c}{L_c}\right) \tag{B.66}$$

$$\frac{\partial \log L_c}{\partial \log \beta} = \frac{\beta}{1+\beta-\alpha}\log\left(\frac{N_c}{L_c}\right) \tag{B.67}$$

and finally, we get:

$$\frac{\partial \log\left(N_c/L_c\right)}{\partial \log \beta} = -\frac{\partial \log L_c}{\partial \log \beta} = -\frac{\beta}{1+\beta-\alpha}\log\left(\frac{N_c}{L_c}\right) \tag{B.68}$$

We turn agian to variations of wages. Inserting expressions (B.68) into equation (B.38), we get:

$$\begin{aligned}\frac{\partial \log w_c}{\partial \beta} &= \log\left(\frac{N_c}{L_c}\right) + \frac{1}{\beta}(\alpha-\beta)\frac{\partial \log L_c}{\partial \log \beta} \\ &= \log\left(\frac{N_c}{L_c}\right) + \frac{\alpha-\beta}{1+\beta-\alpha}\log\left(\frac{N_c}{L_c}\right)\end{aligned}$$

B.2 Open-city case

We now consider an economy with C cities denoted by $c \in \{1, 2, ..., C\}$. We allow city population to vary due to migrations between cities. We conduct comparative statics with respect to a change in income parameter, α or β. Deriving equation (7), we get:

$$\frac{\partial \log w_c}{\partial \alpha} = \beta\frac{\partial \log\left(N_c/L_c\right)}{\partial \alpha} + \log L_c + \alpha\frac{\partial \log L_c}{\partial \alpha} \tag{B.69}$$

$$\frac{\partial \log w_c}{\partial \beta} = \beta\frac{\partial \log\left(N_c/L_c\right)}{\partial \beta} + \log\left(\frac{N_c}{L_c}\right) + \alpha\frac{\partial \log L_c}{\partial \beta} \tag{B.70}$$

Compared to the closed city case, there are additional effects resulting from the adjustment of city population to a variation in the value of an agglomeration economies parameter. The change of city population affects productivity through a change in density agglomeration economies. As before, we conduct the rest of the exercise for variations in α and then turn to variations in β.

Utility should be the same for any two cities at equilibrium, in particular at the fringe. Using the expression of indirect utility given by equation (B.33) evaluated at the fringe $\overline{x}_c = L_c$ and the equality $R_c(L_c) = \underline{R}$ (agricultural land price being assumed to be the same at the fringe of every city), we get:

$$B_c(w_c - \tau L_c) = B_1(w_1 - \tau_1 L_1) \tag{B.71}$$

Using the equality (B.42) for the fringe and the city center, as well as expression (B.36), we also get:

$$\begin{aligned} (w_c - \tau_c L_c)/R_c(L_c)^a &= w_c/R_c(0)^a && \text{(B.72)} \\ (w_c - \tau_c L_c)/\underline{R}^a &= w_c/(\underline{R} + \tau_c N_c)^a && \text{(B.73)} \end{aligned}$$

We then end up with a system of $3C$ equations for $3C$ unknowns (w_c, L_c, N_c) for $c \in \{1, ..., C\}$:

$$\begin{aligned} B_1(w_1 - \tau_1 L_1) &= B_c(w_c - \tau_c L_c) \text{ for } c \in \{2, ..., C\} && \text{(B.74)} \\ \frac{w_c - \tau_c L_c}{\underline{R}^a} &= \frac{w_c}{(\underline{R} + \tau_c N_c)^a} \text{ for } c \in \{1, ..., C\} && \text{(B.75)} \\ \Sigma_{c=1}^C N_c &= N && \text{(B.76)} \\ w_c &= A_c(N_c/L_c)^\beta L_c^\alpha \text{ for } c \in \{1, ..., C\} && \text{(B.77)} \end{aligned}$$

(with parameters considered to be such that we have: $N_c > L_c > 1$).

B.2.1 Elasticities with respect to parameter α

Deriving expressions (B.74) and (B.76), as well as the logarithm of expression (B.75) with respect to α gives:

$$\begin{aligned} B_1\left(\frac{\partial w_1}{\partial \alpha} - \tau_c \frac{\partial L_1}{\partial \alpha}\right) &= B_c\left(\frac{\partial w_c}{\partial \alpha} - \tau_c \frac{\partial L_c}{\partial \alpha}\right) && \text{(B.78)} \\ \frac{1}{w_c - \tau L_c}\left[\frac{\partial w_c}{\partial \alpha} - \tau_c \frac{\partial L_c}{\partial \alpha}\right] &= \frac{1}{w_c}\frac{\partial w_c}{\partial \alpha} - \frac{\tau_c a}{\underline{R} + \tau_c N_c}\frac{\partial N_c}{\partial \alpha} && \text{(B.79)} \\ \sum_{c=1}^C \frac{\partial N_c}{\partial \alpha} &= 0 && \text{(B.80)} \end{aligned}$$

We are first going to insert the expression for the change in agglomeration economies (B.69) into the with-city equilibrium derivative (B.79):

$$\begin{aligned}
\frac{w_c}{w_c-\tau_c L_c}\frac{1}{w_c}\frac{\partial w_c}{\partial \alpha}-\frac{\tau_c}{w_c-\tau_c L_c}\frac{\partial L_c}{\partial \alpha} &= \frac{1}{w_c}\frac{\partial w_c}{\partial \alpha}-\frac{\tau_c a}{\underline{R}+\tau_c N_c}\frac{\partial N_c}{\partial \alpha} && \text{(B.81)}\\
\frac{\tau_c L_c}{w_c-\tau_c L_c}\frac{1}{w_c}\frac{\partial w_c}{\partial \alpha}-\frac{\tau_c}{w_c-\tau_c L_c}\frac{\partial L_j}{\partial \alpha} &= -\frac{\tau_c a}{\underline{R}+\tau_c N_c}\frac{\partial N_j}{\partial \alpha} && \text{(B.82)}\\
\frac{\tau_c L_c}{w_c-\tau_c L_c}\left[\beta\frac{1}{N_c}\frac{\partial N_c}{\partial \alpha}+\log L_c+(\alpha-\beta)\frac{1}{L_c}\frac{\partial L_c}{\partial \alpha}\right]-\frac{\tau_c}{w_c-\tau_c L_c}\frac{\partial L_c}{\partial \alpha} &= -\frac{\tau_c a}{\underline{R}+\tau_c N_c}\frac{\partial N_c}{\partial \alpha} && \text{(B.83)}\\
\beta\frac{\tau_c L_c}{w_c-\tau_c L_c}\frac{1}{N_c}\frac{\partial N_c}{\partial \alpha}+\frac{\tau_c L_c}{w_c-\tau_c L_c}\log L_c+(\alpha-\beta-1)\frac{\tau_c}{w_c-\tau_c L_c}\frac{\partial L_c}{\partial \alpha} &= -\frac{\tau_c a}{\underline{R}+\tau_c N_c}\frac{\partial N_c}{\partial \alpha} && \text{(B.84)}\\
(1+\beta-\alpha)\frac{\partial L_c}{\partial \alpha} &= L_c\log L_c+M_c\frac{\partial N_c}{\partial \alpha} && \text{(B.85)}
\end{aligned}$$

where

$$M_c = a\frac{w_c-\tau_c L_c}{\underline{R}+\tau_c N_c}+\beta\frac{L_c}{N_c} \tag{B.86}$$

and finally, rearranging the terms, we get:

$$\frac{\partial \log L_c}{\partial \log \alpha} = \frac{\alpha}{1+\beta-\alpha}\log L_c+\frac{1}{1+\beta-\alpha}\frac{N_c M_c}{L_c}\frac{\partial \log N_c}{\partial \log \alpha} \tag{B.87}$$

This expression for the variation of city land size is similar to the one obtained in the closed city case that is given by equation (B.61), except that there is now the additional term $\frac{\alpha}{1+\beta-\alpha}\frac{N_c M_c}{L_c}\frac{\partial N_c}{\partial \alpha}$ with $M_c > 0$ due to migrations between cities. In particular, M_c captures the effect of increasing land prices that makes land less attractive if city population increases (since M_c is smaller when $R_c(0) = \underline{R}+\tau_c N_c$ is larger), and the additional effect of density agglomeration economies that makes people want to consume more land as their income is higher. In fact, the migration term $\frac{\partial N_c}{\partial \alpha}$ can be positive or negative, depending on whether cities become more or less attractive with respect to each other when land agglomeration economies change. We can then insert the expression for the change in agglomeration economies (B.69) into the between-city equilibrium derivative (B.78):

$$\begin{aligned}
& B_1\left[\beta\frac{w_1}{N_1}\frac{\partial N_1}{\partial \alpha}+w_1\log L_1+(\alpha-\beta)\frac{w_1}{L_1}\frac{\partial L_1}{\partial \alpha}-\tau_1\frac{\partial L_1}{\partial \alpha}\right]\\
= \;& B_c\left[\beta\frac{w_c}{N_c}\frac{\partial N_c}{\partial \alpha}+w_c\log L_c+(\alpha-\beta)\frac{w_c}{L_c}\frac{\partial L_c}{\partial \alpha}-\tau_c\frac{\partial L_c}{\partial \alpha}\right] && \text{(B.88)}\\
& B_1\left[\beta\frac{w_1}{N_1}\frac{\partial N_1}{\partial \alpha}+w_1\log L_1+\left[(\alpha-\beta)w_1-\tau_1 L_1\right]\frac{1}{L_1}\frac{\partial L_1}{\partial \alpha}\right]\\
= \;& B_c\left[\beta\frac{w_c}{N_c}\frac{\partial N_c}{\partial \alpha}+w_c\log L_c+\left[(\alpha-\beta)w_c-\tau_c L_c\right]\frac{1}{L_c}\frac{\partial L_c}{\partial \alpha}\right] && \text{(B.89)}
\end{aligned}$$

Inserting the expression of city land size (B.87) into this equation, we get:

$$\begin{aligned}
& B_1\left[\beta\frac{w_1}{N_1}\frac{\partial N_1}{\partial\alpha}+w_1\log L_1+\frac{(\alpha-\beta)\,w_1-\tau_1L_1}{1+\beta-\alpha}\frac{1}{L_1}\left(L_1\log L_1+M_1\frac{\partial N_1}{\partial\alpha}\right)\right] \\
= \; & B_c\left[\beta\frac{w_c}{N_c}\frac{\partial N_c}{\partial\alpha}+w_c\log L_c+\frac{(\alpha-\beta)\,w_c-\tau_cL_c}{1+\beta-\alpha}\frac{1}{L_c}\left(L_c\log L_c+M_c\frac{\partial N_c}{\partial\alpha}\right)\right] && \text{(B.90)} \\
& B_1\left[\left(\beta\frac{w_1}{N_1}+\frac{(\alpha-\beta)\,w_1-\tau_1L_1}{1+\beta-\alpha}\frac{M_1}{L_1}\right)\frac{\partial N_1}{\partial\alpha}+\frac{w_1-\tau_1L_1}{1+\beta-\alpha}\log L_1\right] \\
= \; & B_c\left[\left(\beta\frac{w_c}{N_c}+\frac{(\alpha-\beta)\,w_c-\tau_cL_c}{1+\beta-\alpha}\frac{M_c}{L_c}\right)\frac{\partial N_c}{\partial\alpha}+\frac{w_c-\tau_cL_c}{1+\beta-\alpha}\log L_c\right] && \text{(B.91)} \\
& B_1\left[\left(\beta\,(1+\beta-\alpha)\,\frac{w_1}{N_1}-\left[\tau_1L_1+(\beta-\alpha)\,w_1\right]\frac{M_1}{L_1}\right)\frac{\partial N_1}{\partial\alpha}+(w_1-\tau_1L_1)\log L_1\right] \\
= \; & B_c\left[\left(\beta\,(1+\beta-\alpha)\,\frac{w_c}{N_c}-\left[\tau_cL_c+(\beta-\alpha)\,w_c\right]\frac{M_c}{L_c}\right)\frac{\partial N_c}{\partial\alpha}+(w_c-\tau_cL_c)\log L_c\right] && \text{(B.92)}
\end{aligned}$$

We define Q_c with the following equation:

$$Q_c=\beta\,(1+\beta-\alpha)\,\frac{w_c}{N_c}-\left[\tau_cL_c+(\beta-\alpha)\,w_c\right]\frac{M_c}{L_c} \tag{B.93}$$

The term Q_c captures the influence of in/out migration on individual utility. It is the sum of two terms. The first one is positive and captures the effect of an increase in density agglomeration economies. The second one can be either positive or negative, and comes from the change in city land size due to the in/out migration of workers. A negative effect comes from the increase in commuting costs. An additional effect, which can be positive or negative comes from the difference in the changes of density and area agglomeration economies.

We consider that parameters are such that, for every c, we have $Q_c\neq 0$. Substituting expression (B.93) into equation (B.92), and using equality (B.74) to make ratios of consumption amenity effects disappear, gives:

$$\begin{aligned}
B_1\left[Q_1\frac{\partial N_1}{\partial\alpha}+(w_1-\tau_1L_1)\log L_1\right] &= B_c\left[Q_c\frac{\partial N_c}{\partial\alpha}+(w_c-\tau_cL_c)\log L_c\right] && \text{(B.94)} \\
\frac{B_1Q_1}{B_cQ_c}\frac{\partial N_1}{\partial\alpha} &= \frac{\partial N_c}{\partial\alpha}+\frac{w_c-\tau_cL_c}{Q_c}\log L_c-\frac{B_1}{B_c}\frac{w_1-\tau_1L_1}{Q_c}\log L_1 && \text{(B.95)} \\
\frac{w_c-\tau_cL_c}{w_1-\tau_1L_1}\frac{Q_1}{Q_c}\frac{\partial N_1}{\partial\alpha} &= \frac{\partial N_c}{\partial\alpha}+\frac{w_c-\tau_cL_c}{Q_c}\left(\log L_c-\log L_1\right) && \text{(B.96)}
\end{aligned}$$

Summing over all c and using the population derivative given by equation (B.80), we obtain:

$$\begin{aligned}
\frac{Q_1}{w_1-\tau_1L_1}\left(\sum_{c=1}^{C}\frac{w_c-\tau_cL_c}{Q_c}\right)\frac{\partial N_1}{\partial\alpha} &= \sum_{c=1}^{C}\frac{w_c-\tau_cL_c}{Q_c}\left(\log L_c-\log L_1\right) && \text{(B.97)} \\
\frac{\partial\log N_1}{\partial\log\alpha} &= \alpha\frac{w_1-\tau_1L_1}{N_1Q_1}\left(\sum_{c=1}^{C}\frac{w_c-\tau_cL_c}{Q_c}\right)^{-1}\sum_{c=1}^{C}\frac{w_c-\tau_cL_c}{Q_c}\left(\log L_c-\log L_1\right) && \text{(B.98)}
\end{aligned}$$

Hence, the change in city-1 population is a weighted average of differences in initial land size between city 1 and every city. From expression (B.87), we can also deduce variations for population density. We have:

$$\frac{\partial \log(N_c/L_c)}{\partial \log \alpha} = \frac{\partial \log N_c}{\partial \log \alpha} - \frac{\partial \log L_c}{\partial \log \alpha} \tag{B.99}$$

$$= \frac{\partial \log N_c}{\partial \log \alpha} - \frac{\alpha}{1+\beta-\alpha} \log L_c - \frac{\alpha}{1+\beta-\alpha} \frac{N_c M_c}{L_c} \frac{\partial \log N_c}{\partial \log \alpha} \tag{B.100}$$

$$= -\frac{\alpha}{1+\beta-\alpha} \log L_c + \left(1 - \frac{\alpha}{1+\beta-\alpha} \frac{N_c M_c}{L_c}\right) \frac{\partial \log N_c}{\partial \log \alpha} \tag{B.101}$$

and we can use equation (B.98) to develop this expression and get an expression for this elasticity that can be computed in our empirical analysis.

We now turn to variations of wages. Inserting expressions (B.87) into equation (B.69), we get:

$$\frac{\partial \log w_c}{\partial \alpha} = \log L_c + \frac{\beta}{\alpha} \frac{\partial \log N_c}{\partial \log \alpha} + \frac{\alpha-\beta}{\alpha} \frac{\partial \log L_c}{\partial \log \alpha} \tag{B.102}$$

$$= \log L_c + \frac{\beta}{\alpha} \frac{\partial \log N_c}{\partial \log \alpha} + \frac{\alpha-\beta}{\alpha} \left(\frac{\alpha}{1+\beta-\alpha} \log L_c + \frac{1}{1+\beta-\alpha} \frac{N_c M_c}{L_c} \frac{\partial \log N_c}{\partial \log \alpha}\right) \tag{B.103}$$

$$= \log L_c + \frac{\alpha-\beta}{1+\beta-\alpha} \log L_c + \left(\beta + \frac{\alpha-\beta}{1+\beta-\alpha} \frac{N_c M_c}{L_c}\right) \frac{1}{\alpha} \frac{\partial \log N_c}{\partial \log \alpha} \tag{B.104}$$

where the expression of $\frac{\partial \log N_c}{\partial \log \alpha}$ is given by equation (B.98).

B.2.2 Elasticities with respect to parameter β

We now consider a change in β. We are first going to insert the expression for the change in agglomeration economies (B.70) into the counterpart of the expression of the with-city equilibrium derivative (B.79) considered when deriving with respect to β rather than α. We obtain:

$$\frac{w_c}{w_c - \tau_c L_c} \frac{1}{w_c} \frac{\partial w_c}{\partial \beta} - \frac{\tau_c}{w_c - \tau_c L_c} \frac{\partial L_c}{\partial \beta} = \frac{1}{w_c} \frac{\partial w_c}{\partial \beta} - \frac{\tau_c a}{\underline{R} + \tau_c N_c} \frac{\partial N_c}{\partial \beta} \tag{B.105}$$

$$\frac{\tau_c L_c}{w_c - \tau_c L_c} \frac{1}{w_c} \frac{\partial w_c}{\partial \beta} - \frac{\tau_c}{w_c - \tau_c L_c} \frac{\partial L_j}{\partial \beta} = -\frac{\tau_c a}{\underline{R} + \tau_c N_c} \frac{\partial N_j}{\partial \beta} \tag{B.106}$$

$$\frac{\tau_c L_c}{w_c - \tau_c L_c} \left[\beta \frac{\partial \log(N_c/L_c)}{\partial \beta} + \log\left(\frac{N_c}{L_c}\right) + \alpha \frac{\partial \log L_c}{\partial \beta}\right] - \frac{\tau_c}{w_c - \tau_c L_c} \frac{\partial L_c}{\partial \beta} = -\frac{\tau_c a}{\underline{R} + \tau_c N_c} \frac{\partial N_c}{\partial \beta} \tag{B.107}$$

$$\beta \frac{\tau_c L_c}{w_c - \tau_c L_c} \frac{1}{N_c} \frac{\partial N_c}{\partial \beta} + \frac{\tau_c L_c}{w_c - \tau_c L_c} \log\left(\frac{N_c}{L_c}\right) + (\alpha - \beta - 1) \frac{\tau_c}{w_c - \tau_c L_c} \frac{\partial L_c}{\partial \beta} = -\frac{\tau_c a}{\underline{R} + \tau_c N_c} \frac{\partial N_c}{\partial \beta} \tag{B.108}$$

$$(1 + \beta - \alpha) \frac{\partial L_c}{\partial \beta} = L_c \log\left(\frac{N_c}{L_c}\right) + M_c \frac{\partial N_c}{\partial \beta} \tag{B.109}$$

and finally, rearranging the terms, we get:

$$\frac{\partial \log L_c}{\partial \log \beta} = \frac{\beta}{1+\beta-\alpha} \log\left(\frac{N_c}{L_c}\right) + \frac{\beta}{1+\beta-\alpha} \frac{N_c M_c}{L_c} \frac{\partial \log N_c}{\partial \beta} \tag{B.110}$$

This equation is similar to equation (B.87) except for the presence of $\log(N_c/L_c)$ rather than $\log(L_c)$ in the first right-hand term, and the presence of parameter β at the numerator of the two right-hand side terms rather than α.

We now compute the expression corresponding to equation (B.98) when considering variations in β rather than α. The counterpart of equation (B.78) is:

$$B_1 \left(\frac{\partial w_1}{\partial \beta} - \tau_1 \frac{\partial L_1}{\partial \beta}\right) = B_c \left(\frac{\partial w_c}{\partial \beta} - \tau_c \frac{\partial L_c}{\partial \beta}\right) \tag{B.111}$$

Inserting equation (B.70) into this expression, we obtain:

$$\begin{aligned}
& B_1 \left[\beta \frac{w_1}{N_1} \frac{\partial N_1}{\partial \beta} + w_1 \log\left(\frac{N_1}{L_1}\right) + (\alpha-\beta) \frac{w_1}{L_1} \frac{\partial L_1}{\partial \beta} - \tau_1 \frac{\partial L_1}{\partial \beta}\right] \\
= \quad & B_c \left[\beta \frac{w_c}{N_c} \frac{\partial N_c}{\partial \beta} + w_c \log\left(\frac{N_c}{L_c}\right) + (\alpha-\beta) \frac{w_c}{L_c} \frac{\partial L_c}{\partial \beta} - \tau_1 \frac{\partial L_c}{\partial \beta}\right] && \text{(B.112)} \\
& B_1 \left[\beta \frac{w_1}{N_1} \frac{\partial N_1}{\partial \beta} + w_1 \log\left(\frac{N_1}{L_1}\right) + \left[(\alpha-\beta) \frac{w_1}{L_1} - \tau_1\right] \frac{\partial L_1}{\partial \beta}\right] \\
= \quad & B_c \left[\beta \frac{w_c}{N_c} \frac{\partial N_c}{\partial \beta} + w_c \log\left(\frac{N_c}{L_c}\right) + \left[(\alpha-\beta) \frac{w_c}{L_c} - \tau_1\right] \frac{\partial L_c}{\partial \beta}\right] && \text{(B.113)}
\end{aligned}$$

Inserting the expression of city land size (B.110) into this equation, we get:

$$\begin{aligned}
& B_1 \left[\beta \frac{w_1}{N_1} \frac{\partial N_1}{\partial \beta} + w_1 \log\left(\frac{N_1}{L_1}\right) + \left[(\alpha-\beta) \frac{w_1}{L_1} - \tau_c\right] \left[\frac{L_1}{1+\beta-\alpha} \log\left(\frac{N_1}{L_1}\right) + \frac{1}{1+\beta-\alpha} M_1 \frac{\partial N_1}{\partial \beta}\right]\right] && \text{(B.114)} \\
= \quad & B_c \left[\beta \frac{w_c}{N_c} \frac{\partial N_c}{\partial \beta} + w_c \log\left(\frac{N_c}{L_c}\right) + \left[(\alpha-\beta) \frac{w_c}{L_c} - \tau_c\right] \left[\frac{L_c}{1+\beta-\alpha} \log\left(\frac{N_c}{L_c}\right) + \frac{1}{1+\beta-\alpha} M_c \frac{\partial N_c}{\partial \beta}\right]\right] && \text{(B.115)} \\
& B_1 \left[\left(\beta \frac{w_1}{N_1} + \frac{(\alpha-\beta) w_1 - \tau_1 L_1}{1+\beta-\alpha} \frac{M_1}{L_1}\right) \frac{\partial N_1}{\partial \beta} + \frac{w_1 - \tau_1 L_1}{1+\beta-\alpha} \log\left(\frac{N_1}{L_1}\right)\right] && \text{(B.116)} \\
= \quad & B_c \left[\left(\beta \frac{w_c}{N_c} + \frac{(\alpha-\beta) w_c - \tau_c L_c}{1+\beta-\alpha} \frac{M_c}{L_c}\right) \frac{\partial N_c}{\partial \beta} + \frac{w_c - \tau_c L_c}{1+\beta-\alpha} \log\left(\frac{N_c}{L_c}\right)\right] && \text{(B.117)} \\
& B_1 \left[\left(\beta(1+\beta-\alpha) \frac{w_1}{N_1} - [\tau_1 L_1 + (\beta-\alpha) w_1] \frac{M_1}{L_1}\right) \frac{\partial N_1}{\partial \beta} + (w_1 - \tau_1 L_1) \log\left(\frac{N_1}{L_1}\right)\right] && \text{(B.118)} \\
= \quad & B_c \left[\left(\beta(1+\beta-\alpha) \frac{w_c}{N_c} - [\tau_c L_c + (\beta-\alpha) w_c] \frac{M_c}{L_c}\right) \frac{\partial N_c}{\partial \beta} + (w_c - \tau_c L_c) \log\left(\frac{N_c}{L_c}\right)\right] && \text{(B.119)}
\end{aligned}$$

The other developments are straightforward and follow those when there are variations in α. We then end up with the expression:

$$\frac{\partial \log N_1}{\partial \log \beta} = \beta \frac{w_1 - \tau_1 L_1}{N_1 Q_1} \left(\sum_{c=1}^{C} \frac{w_c - \tau_c L_c}{Q_c} \right)^{-1} \sum_{c=1}^{C} \frac{w_c - \tau_c L_c}{Q_c} \left[\log \left(\frac{N_c}{L_c} \right) - \log \left(\frac{N_1}{L_1} \right) \right] \tag{B.120}$$

This expression is similar to the one obtained when there are variations in α that is given by equation B.98), except for terms in brackets in the right-hand side that are of the form $\log(N_c/L_c)$ rather than $\log(L_c)$, and the whole expression on the right-hand side is multiplied by β rather than α.

We finally turn to variations of wages. Inserting expressions (B.110) into equation (B.70), we get:

$$\begin{aligned} \frac{\partial \log w_c}{\partial \beta} &= \frac{\partial \log N_c}{\partial \log \beta} + \log \left(\frac{N_c}{L_c} \right) + \frac{\alpha - \beta}{\beta} \frac{\partial \log L_c}{\partial \log \beta} && \text{(B.121)} \\ &= \log \left(\frac{N_c}{L_c} \right) + \frac{\partial \log N_c}{\partial \log \beta} + \frac{\alpha - \beta}{\beta} \left[\frac{\beta}{1 + \beta - \alpha} \log \left(\frac{N_c}{L_c} \right) + \frac{1}{1 + \beta - \alpha} \frac{N_c M_c}{L_c} \frac{\partial \log N_c}{\partial \log \beta} \right] && \text{(B.122)} \\ &= \log \left(\frac{N_c}{L_c} \right) + \frac{\alpha - \beta}{1 + \beta - \alpha} \log \left(\frac{N_c}{L_c} \right) + \left(\beta + \frac{\alpha - \beta}{1 + \beta - \alpha} \frac{N_c M_c}{L_c} \right) \frac{1}{\beta} \frac{\partial \log N_c}{\partial \log \beta} && \text{(B.123)} \end{aligned}$$

where the expression of $\frac{\partial \log N_c}{\partial \log \beta}$ is given by equation (B.120). This expression is very similar to the one obtained when deriving with respect to α that is given by equation (B.104). Once the derivative of city populations have been replaced by their expressions, the difference is that terms $\log L_c$ are replaced by terms $\log \left(\frac{N_c}{L_c} \right)$.

B.3 Computation of expressions

B.3.1 Decomposition

We first explain give some details on how we get the decomposition of interest provided in the main text. In this subsection, model parameters are indexed by t for year t as we are interested in wage evolution over time and parameters can change. Considering that dates $t-1$ and t are close, we can write that:

$$\begin{aligned} & \log w_{c,t} (\alpha_t, \beta_t, A_t, B_t) - \log w_{c,t-1} (\alpha_{t-1}, \beta_{t-1}, A_{t-1}, B_{t-1}) && \text{(B.124)} \\ \approx\ & d \log w_{c,t} (\alpha_t, \beta_t, A_t, B_t) && \text{(B.125)} \\ =\ & \frac{\partial \log w_{c,t}}{\partial \alpha} d\alpha_t + \frac{\partial \log w_{c,t}}{\partial \beta} d\beta_t + \frac{\partial \log w_{c,t}}{\partial A_t} dA_t + \frac{\partial \log w_{c,t}}{\partial B_t} dB_t && \text{(B.126)} \\ \approx\ & \frac{\partial \log w_{c,t}}{\partial \alpha} (\alpha_t - \alpha_{t-1}) + \frac{\partial \log w_{c,t}}{\partial \beta} (\beta_t - \beta_{t-1}) + \frac{\partial \log w_{c,t}}{\partial A} (A_t - A_{t-1}) + \frac{\partial \log w_{c,t}}{\partial B} (B_t - B_{t-1}) && \text{(B.127)} \end{aligned}$$

where $\frac{\partial \log w_{c,t}}{\partial A} = \left(\frac{\partial \log w_{c,t}}{\partial A_1}, ..., \frac{\partial \log w_{c,t}}{\partial A_C} \right)$ and $\frac{\partial \log w_{c,t}}{\partial B} = \left(\frac{\partial \log w_{c,t}}{\partial B_1}, ..., \frac{\partial \log w_{c,t}}{\partial B_C} \right)$.

In particular, we are interested in the evolution of log-wage when the values of agglomeration parameters α and β vary. Hence, inserting expression (B.127) into equation (24), we obtain our decomposition of interest given by eqution (27) where $\frac{\partial \log w_{c,t}}{\partial A} (A_t - A_{t-1}) + \frac{\partial \log w_{c,t}}{\partial B} (B_t - B_{t-1})$ enters the residual $r_{c,t}$.

B.3.2 Expressions

We need to bring expressions (B.87), (B.110), (B.98) and (B.120) to the data. For that purpose, we are first going to rewrite them as functions of quantities for which we may be able to find an empirical counterpart. We have values in the data for N_j and L_j, and we will have estimates for α and β. We need values for the quantities N_cM_c/L_c and $(w_c - \tau_c L_c)/Q_c$. Using expression (B.86), we get:

$$\frac{N_cM_c}{L_c} = \frac{N_c}{L_c} a \frac{w_c - \tau_c L_c}{\underline{R} + \tau_c N_c} + \beta \tag{B.128}$$

This expression makes intervene $R_c(0) = \underline{R} + \tau_c N_c$ that can be rewritten such that:

$$R_c(0) = \frac{R_c(0)}{\tau_c N_c} \tau_c N_c = \mu_c \tau_c N_c \tag{B.129}$$

where:

$$\mu_c = \frac{R_c(0)}{R_c(0) - \underline{R}} = \frac{1}{1 - \underline{R}/R_c(0)} \tag{B.130}$$

and provided that the ratio $\underline{R}/R_c(0)$ can be computed from the data, we can compute μ_c. Inserting expression (B.129) into equation (B.128), we get:

$$\begin{aligned} \frac{N_cM_c}{L_c} &= \frac{N_c}{L_c} a \frac{w_c - \tau_c L_c}{\mu_c \tau_c N_c} + \beta \\ &= \frac{a}{\mu_c}\left(\frac{w_c}{\tau_c L_c} - 1\right) + \beta \end{aligned} \tag{B.131}$$

Provided that the ratio $\tau_c L_c/w_c$ can be computed from the data, we can compute N_cM_c/L_c.

Inserting expressions (B.128) and (B.129) into equation (B.93), we obtain:

$$N_cQ_c = \beta(1+\beta-\alpha) w_c - [\tau_c L_c + (\beta-\alpha) w_c] \frac{N_c}{L_c}\left(a\frac{w_c - \tau_c L_c}{\mu_c \tau_c N_c} + \beta \frac{L_c}{N_c}\right) \tag{B.132}$$

$$= \beta(1+\beta-\alpha) w_c - [\tau_c L_c + (\beta-\alpha) w_c]\left(a\frac{w_c - \tau_c L_c}{\mu_c \tau_c L_c} + \beta\right) \tag{B.133}$$

$$= \beta(w_c - \tau_c L_c) - [\tau_c L_c + (\beta-\alpha) w_c] \frac{a}{\mu_c \tau_c L_c}(w_c - \tau_c L_c) \tag{B.134}$$

Hence:

$$\frac{Q_c}{w_c - \tau_c L_c} = \frac{1}{N_c}\left[\beta - \frac{a}{\mu_c}\left[1 + (\beta-\alpha)\frac{w_c}{\tau_c L_c}\right]\right] \tag{B.135}$$

and this expression can be computed for given values for parameters, μ_c and $\tau_c L_c/w_c$, but also N_c.

To sum up, we are able to compute the elasticities of land size and population density with respect to α and β from estimated parameters for α and β, the housing budget share a, the city share of costliest transport cost in wages $\tau_c L_c / w_c$, the city ratio between land prices at the fringe and at the center $\underline{R}/R_c\,(0)$, and city land size and population L_c and N_c.