Italian Economic Journal (2023) 9:789–827
https://doi.org/10.1007/s40797-022-00202-6

RESEARCH PAPER - EUROPE AND ITALY

Urban Non-urban Agglomeration Divide: Is There a Gap in Productivity and Wages?

Eleonora Bartoloni[1,2] · **Andrea Marino**[3,4] · **Maurizio Baussola**[5] · **Davide Romaniello**[5]

Received: 29 May 2021 / Accepted: 17 June 2022 / Published online: 23 July 2022

Abstract
This paper investigates the productivity–wage relation using a novel and integrated employer-employee database covering the entire population of non-financial firms' plants in one of the most developed regions in Europe, i.e., the Italian Lombardy region. We suggest that although a growing literature shows that locating in urban areas yields substantial productivity gains due to agglomeration economies, the interaction between productivity and wages is ultimately the key to ascertaining the true advantage of the high densely populated areas. By adopting an empirical specification that allows us to explore interaction effects between localization and the sector of activity at the establishment level, we find that agglomeration economies play a significant but conditional role in affecting productivity and wage differentials while also controlling for firm-specific factors (in particular, job-related characteristics) and selection effects. The estimated impacts are heterogeneous across sectors, depending on their technological features. The effect of locating in *High-density* urban areas on the productive-wage gap is significantly positive only in highly knowledge-intensive services sectors; for firms supplying less technologically sophisticated services and for manufacturing plants, the impact is either not significant or negative. Locating in *Rural* areas generally exerts a downward (or not significant) impact on the productivity-wage gap.

Keywords Labour productivity · Productivity–wage gap · Agglomeration divide · Knowledge intensive sectors

✉ Maurizio Baussola
maurizio.baussola@unicatt.it

1 Italian National Institute of Statistics (ISTAT), Rome, Italy

2 Eurostat, Luxembourg, Luxembourg

3 Italian National Institute of Statistics (ISTAT), Genova, Italy

4 DISPI, University of Genoa, Genova, Italy

5 Dipartimento di Scienze Economiche e Sociali, Università Cattolica del Sacro Cuore, Piacenza, Italy

JEL Classification R12 · R10 · R30 · J24 · J31

1 Introduction

The dynamics of labour productivity have always been crucial to understanding the developmental pace of a country or region, or the growth path of industries. Technological change is viewed as the driver of productivity enhancement, together with the ability of firms to exploit the innovative opportunities that are then made available. The close relationship between productivity and technological change became contradictory in the early 1980s, in that productivity growth declined despite the spread of information and communication technologies. Later periods also saw an inconsistent path, with first a new increase and then a drop in productivity, even with the extensive use of technologies related to the internet (OECD 2019). With this framework, discrepancies between countries, regions, and industries significantly increase, thus encouraging studies aimed at explaining such differentials. In particular, regional and intra-regional investigations have focused on the role of agglomeration economies and the local endowment of factors that positively affect productivity.

A complementary issue to productivity patterns is the dynamics of wages. The simultaneous analysis of these two variables is relevant because their misalignment raises issues related to competitiveness, inflation, and income distribution.

The macroeconomic picture for OECD countries indicates a significant and constant decoupling of labour productivity growth and wage growth from the mid-1990s. This has resulted in a significantly lower growth rate for the latter. However, Italy was characterized by lower growth rates for both productivity and wages during the 2000s, and oddly, the productivity growth rate was sometimes lower than that of wages, thus adding additional challenges to the entire economy (Sharpe and Uguccioni 2017; OECD 2018).

It is therefore crucial to provide in-depth analyses of such a divergence that take into consideration both territorial and industry patterns. For this reason, in this paper we analyse the productivity–wage relationship in the manufacturing and services industries in the Italian region of Lombardy. Using matched employer–employee database built at the establishment level, we investigate the determinants of both productivity and wages, and the possible misalignment in the productivity–wage gap across urban and nonurban areas.

This region represents an ideal setting to test some crucial hypotheses concerning the productivity–wage relationship. One should recall that it is one of the so-called four motors of Europe, together with Baden-Württemberg, Rôhne-Alpes, and Catalonia. These regions represent core European areas characterized by a high level of per capita income, industrialization, and innovativeness. In addition, Lombardy has a diversified economy in which both manufacturing and services play a significant role, together with a diversified geographical pattern in terms of urbanization. For these reasons, we think that the results obtained for this macro-area may provide useful information for international comparisons and for a generalization to the Italian economy more broadly.

It is worth highlighting that this disaggregated analysis is relevant because one of the most significant and striking pieces of evidence regarding growth patterns in OECD countries is the patchy territorial characteristics of growth, even within regions characterized by higher average living standards (Cheshire 2019). In particular, an urban–rural divide is clearly observed, requiring a more in-depth analysis based specifically on firm behaviour in different industries and areas.

We analyse the characteristics of urbanization within Lombardy through the lens of the DegUrba classification developed by Eurostat at the European level (European Union et al. 2021). The DegUrba methodology combines population size and population density attributes, starting from a very granular geographic disaggregation of the territory. Thus, it represents an interesting tool for investigating possible territorial gaps, even within an urbanized region like Lombardy, in which almost 45% of municipalities belong to rural (nonurban) areas, representing nearly 60% of the region's area.

This paper may, therefore, represent the first step for further comparative analyses aiming to thoroughly investigate patterns of productivity and wages in the most developed European areas, thus providing useful information for policy intervention.

Local differences in productivity and wages may have different explanations. So far, the empirical evidence shows that agglomeration economies may enhance the economic performance of firms operating in urban areas in terms of productivity and wage growth (Beaudry and Schiffauerova 2009). It has been argued that firms in the highest-density areas may find it easier to access diversified job experiences, thus increasing 'workers" abilities (Duranton and Puga 2004; Rosenthal and Strange 2004). However, together with the potential advantages of urbanization, negative impacts (Henderson 1974) may originate from competition among firms for the skilled workforce, which can lead to higher production costs and further upward pressure on wages (Zheng 2001; Combes and Duranton 2006).

The implicit assumption of most empirical investigations on agglomeration economies is the competitive wage determination mechanism, according to which wage differences across workers only reflect productivity differences due to individual characteristics or working conditions. Thus, potential misalignment between productivity and wages is not explicitly considered, with both measures used alternatively (Combes et al. 2008a, b; Ahrend et al. 2014).

However, wage variation not justified by productivity gains has increasingly received empirical attention (Hellerstein et al. 1999; Van Biesembroeck 2014). Individual heterogeneity, on both the worker and the firm side, may play a role in the correlation between productivity and wages at the firm level, together with the bargaining power of workers and the prevalent bargaining mechanism (McGuinness and O'Connell 2010; Rusinek and Tojerow 2014). In continental European countries, wage bargaining is more centralized and, thus, the relevance of individual bargaining is lower (Caju et al. 2008).

Given these considerations, the aim of this contribution is twofold. We intend to assess how urbanization, together with industry firm-specific factors and worker characteristics, affects productivity and wages in manufacturing and service establishments. We also aim to verify the misalignment of the productivity–wage gap across local areas with different levels of urbanization. Indeed, the typical result that productivity is positively related to urbanization is complemented and challenged by considering the productivity–wage gap.

Our analysis represents an original contribution to the current debate in various ways. Firstly, it considers a census of manufacturing and service establishments with payroll-registered employees. Secondly, it investigates the local distribution of these establishments using a spatial grid harmonized at the European level, and this is—to our knowledge—the first attempt to apply this kind of analysis to Italian industries. Thirdly, the availability of information on worker heterogeneity and firm-specific characteristics allows us to disentangle their direct impacts on productivity, wages, and the relevant gaps from the observed role played by the specific location context.

The results indicate that the degree of urbanization plays a significant but conditional role in affecting productivity and wage differentials at the local level. Indeed, this effect depends on the technological and knowledge-based resources characterising the industrial mix within urban and nonurban agglomerations. It is also highlighted that in some cases, there exists a negative gap between productivity and wages that, therefore, should be considered in depicting the real and comprehensive advantage of urbanization.

The paper proceeds as follows. The next section discusses the literature on geographical differences in productivity and wages. Section 3 presents the data and the level of geographical disaggregation used. Section 4 provides the empirical analysis, while Sect. 5 provides further robustness evidence. Section 6 offers some concluding remarks.

2 Space, Productivity, and Wages

A large body of literature has focused on different patterns of firm performance according to location. Hence, micro-grounded investigations are particularly suited to verifying the role of geographical factors that may affect productivity and wages, together with industry and worker characteristics. We aim to further investigate these factors, taking into consideration the main results prevailing within the reference literature.

2.1 Localization and Urbanization

The spatial economics literature has investigated the existence of productivity (or wage) premia due to agglomeration economies in different contexts and using various methodological tools to identify the forces behind the advantages of agglomeration (Combes and Gabillon 2015). As concerns the impact of agglomeration economies on productivity, if agglomeration advantages outweigh agglomeration costs, the impact of an urban location should be positive. However, besides positive agglomeration externalities, congestion costs negatively affect firm performance (Henderson 1974), thus representing a centrifugal force. These negative externalities may be due to congestion (Sweet 2014), pollution, crime, high housing and land rents, or increased labour costs because of competition among firms for the skilled workforce (Combes and Duranton 2006).

Among the vast literature, the analyses by Henderson (2003), Martin et al. (2011), Baldwin et al. (2008), and Anderson and Loof (2011) provide estimates of the impact of urbanization—and agglomeration in general—on firm productivity in different economic contexts and industries. Although these studies show a clear productivity gain as long as urban density increases, it also arises that firm costs may be affected by agglomeration diseconomies, thus reducing their profitability (Jennen and Verwijmeren 2010; Stavropoulos and Skures 2016; Bartoloni and Baussola 2021). Bartoloni and Baussola (2021) find an urban-nonurban productivity divide in a sample of subregional Italian areas, but this premium vanishes when considering profitability, thus calling for a negative role played by diseconomies (costs) of agglomerations. This result is robust to the local aggregations used, presenting different spatial density characteristics. A similar finding was also confirmed for the Tokyo metropolitan area in a study by Zheng (2001), in that diseconomies of agglomeration are significant and mainly reflect high housing and land prices, long commuting times, and low environmental quality.

However, as wages are a significant component of a firm's production costs, it is therefore worth analysing whether agglomeration matters for describing wage determination at the firm/plant level. Ahrend et al. (2014) underline the crucial role of urban areas—more specifically, high-density metropolitan areas across five OECD countries—in affecting wages and, according to their reasoning, productivity. An explicit investigation of a productivity–wage gap is not considered, thereby assuming a straightforward relationship between these two variables.

Our investigation aims to highlight differences in firm productivity and wages according to location in terms of population density. We adopt a modelling approach aligned with the analysis proposed by Kampelmann et al. (2018) for Belgian regions. Using the institutional framework of wage-setting rules in the background, i.e. the relevance of national contract agreements (NCAs) in the wage-setting mechanism in continental Europe (in their case, Belgium), they derive the impact on productivity and wages conditional on workers' human capital, gender, and age, and taking into account industry characteristics. The institutional framework that contributes to determining wages, i.e. unionization and collective bargaining, therefore lies in the background of the present analysis.

When investigating the possible misalignment of the productivity–wage gap across urban and nonurban areas, one should also consider some individual sorting across space that may arise both because workers tend to locate in the areas where their specific skills are most required and because firms can decide to localize in the areas better served by their particular managerial abilities.

Sorting effects may play a significant role, as described in Gaubert (2018), as firm decisions regarding optimal locations may be crucially affected by city size. This may cause endogeneity issues, which we attempt to tackle as described in Sect. 5. In addition, potential measurement error induced by the specific localization areas used for capturing agglomeration effects may cause additional bias concerns in the econometric estimation of the agglomeration effect. The empirical literature has addressed this issue using various methodologies that, nevertheless, fail to provide conclusive indications about the induced potential bias. (Ciccone 2002; Rice et al. 2006; Combes et al. 2008a, b). However, it is worth stressing that although our database's cross-sectional nature does not enable us to describe the sorting mechanism fully, we undertake instrumental variables estimations that consider firm location decisions. Sectoral specificities are also relevant given that the characteristics and composition of the workforce crucially depend on the technological context in which a firm operates. Nevertheless, technological spillovers within a firm's sector of activity can directly affect business performance. They also represent one of the mechanisms behind agglomeration economies (Duranton and Puga 2004). Unfortunately, measuring the specific contribution of technological spillovers to agglomeration economies and their indirect impact on firm performance is not easy. The empirical literature provides some attempts, the results of which cannot be generalized because of the very specific contexts in which they are undertaken (Rosenthal and Strange 2001; Ellison et al. 2010; Carlino and Kerr 2015).

Our analysis assumes that the technological and knowledge-based resources characterizing the industry to which a firm belongs may directly affect firm productivity and wages. In addition, we assume a moderator effect on localization as we interact the set of dummy variables capturing sector-specific effects with the localization variables. In line with previous evidence on Italy's urban local labour market areas (Di Giacinto et al. 2020), we expect that the productivity and wage premia of establishments localized in urban areas are higher in technology-intensive and knowledge-based sectors compared to other manufacturing and services activities.

In general, we expect more productive firms to pay a higher average wage premium (Card and al. 2018). However, such a premium varies significantly and may crucially depend on the prevailing wage-setting rules. In our context, where the typical NCAs establish wage floors and individual firm bargaining play a less relevant role in comparison with other institutional frameworks (e.g. the Anglo-Saxon industrial relation context), the elasticity between productivity and wages may be milder.

2.2 Worker and Firm Characteristics

The empirical literature has extensively investigated the role of worker characteristics in determining productivity and wage premia. The traditional human capital approach to wage determination has been challenged since the contribution by Oi (1962) in which labour is described as a quasi-fixed factor. Therefore, one cannot expect a one-to-one correspondence between productivity and wages; as such, fixed costs may vary significantly across occupations and are crucially determined by the degree of specific human capital required to accomplish job tasks.

In addition, efficiency wage theory (Akerlof and Yiellen 1986) and the other information-based approaches to wage determination (Shapiro and Stiglitz 1984) help explain discrepancies between the effective and implicit market equilibrium wage. Of course, it is not our intention to review the literature on this issue as this is beyond the scope of this analysis. However, it is worth recalling these approaches in order to better interpret the results of the estimates we present in the following sections. More recent investigations have indeed focused on firm (plant) estimations of occupational pay and productivity.

Indeed, the availability of matched employer–employee datasets has allowed the joint estimation of production and wage equations to determine the relationship between productivity and wage premia. Following the study by Hellerstein et al. (1999), a vast strand of empirical literature has investigated the impact of worker characteristics such as education, age, types of labour contracts, and gender. The evidence has pointed out the presence of wage premia relative to productivity levels for older workers in different European countries (Haegeland and Klette 1999; Crépon et al. 2003; Cataldi et al. 2011; Van Ours and Stoeldraijer 2011). Furthermore, scholars have investigated gender as a cause for wage discrimination, but with more controversial results. Hellerstein and Neumark (2009) find that the gender wage gap systematically exceeds the productivity gap in the US, while other studies do not find evidence of gender discrimination (Haegeland and Klette 1999; Hellerstein and Neumark 1999; Crepon et al. 2003).

Worker skills as proxied by accumulated training have been found to be significant in explaining productivity and wage differentials. Dearden et al. (2006) show that the productivity effect of training substantially exceeds the wage effect in U.K. manufacturing, while Konings and Vanormelingen (2015) find a significant positive gap in productivity–wage premia for the services sector in Belgium, but not for manufacturing.

When considering education, evidence of a productivity–wage gap is less strong. Haegeland and Klette (1999) find that the wage premium for education is in line with the productivity premium, whereas Hellerstein and Neumark (2009) find support for the positive effect of education on productivity in the U.S. manufacturing industry for workers with a college education.

The empirical analysis of different types of work contracts (temporary vs permanent; part-time vs full-time) is often associated with the debate on labour market flexibility, and in general, employment protection legislation (EPL) (Bertola et al. 2000), which has significantly affected labour market adjustment in OECD countries.

We do not intend to enter into this very controversial debate. However, it is worth stressing that despite significant evidence on the gap between permanent and temporary workers not related to job or individual characteristics (Brown and Sessions 2003; De la Rica 2004; Bosio 2009; Comi and Grasseni 2012), much less evidence is available on the productivity–wage relationship. One example is the work of Garnero et al. (2014) based on matched employer–employee panel data on Belgian private-sector firms, which provides evidence that both fixed-term and part-time contracts exert stronger positive effects on productivity than on wages.

Together with worker heterogeneity, firm-specific heterogeneity also plays a crucial role in explaining productivity differentials, as reported in the findings by Syverson (2011) and Foster et al. (2008). Firm size, international openness, and innovation, although analysed through different frameworks, appear as crucial determinants of business productivity.

Hence, it is worth recalling that the core aim of our study is to test and quantify the urban-nonurban productivity divide and its main determinants related to firm and industry characteristics. Simultaneously, we explicitly consider spatial wage discrepancies and the productivity–wage gap. It is therefore possible to verify that urban density matters for productivity gains, but conditional on industry characteristics, i.e.

technological features. In other words, the positive relationship between urbanization and productivity should be reconsidered in light of the gap with wages. For this reason, we explicitly consider this issue and derive a comprehensive picture of the effective gain from urbanization.

3 Data Description

3.1 General

We use ISTAT's Frame territoriale SBS (Structural Business Statistics), i.e. the Italian business register, which integrates establishment data using (i) the SBS register, the main data source on structural and economic characteristics for the total population of Italian enterprises, and (ii) the Statistical Archive of Active Firms (ASIA-UL), the statistical register of business establishments. Information on job quality stems from the ASIA Employment Archive, a matched employer–employee dataset from which we derive additional variables on the demographic and job-related characteristics of employment at the firm level.[1]

The original data source covers more than 4.7 million establishments operating in all industry and service activities (excluding the financial sector as well as some personal and household services) in 2016 and generating almost 716 billion euros of value added. More than 850 thousand establishments are localized in Lombardy. They generate 186 billion euros of value added, corresponding to about one-fourth of the national total. Our analysis is performed on the manufacturing and services industries. Thus, we exclude establishments operating in the extraction and utilities sector and in the construction sector, whose characteristics are not directly comparable to those operating in manufacturing or services. The share of the excluded sectors is 12% in terms of local units and 9% in terms of value added. From this regional database, we extract a subsample of establishments with payroll-registered employees. The dataset used in the present analysis is composed of almost 260 thousand establishments and represents 35% of the total number of establishments in the region. The share is lower for manufacturing (11%) compared to services (30%).

We adopt sector aggregation according to the level of technology and knowledge intensity (Eurostat).[2] Although based on the NACE Rev. 2, this aggregation of activities is better suited to capturing the differences related to firms' technological capabilities and the availability of skills within the productive units.

3.2 The DegUrba Classification

The localization of productive units is defined at the municipality level (Local Administrative Units, LAU2), the degree of urbanization of which is detected using the DegUrba methodology. The DegUrba classification was set up by Eurostat using a

[1] See the Appendix for details.

[2] See the reference metadata in Euro SDMX Metadata Structure (ESMS) Annex 3 (high-tech aggregation by NACE Rev. 2); https://ec.europa.eu/eurostat/cache/metadata/en/htec_esms.htm.

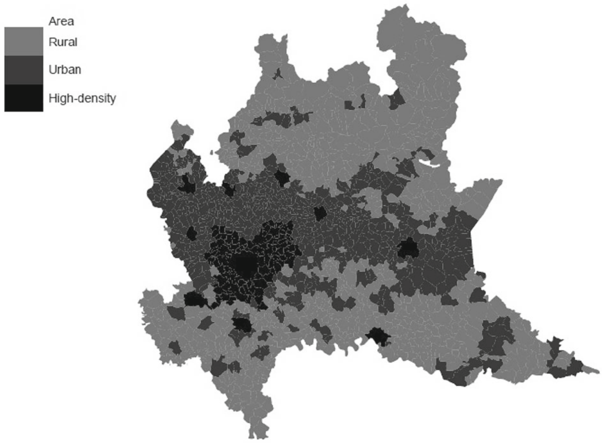

Fig. 1 DegUrba clusters in the Lombardy region

combination of population size and population density thresholds, and it has been adopted by various surveys conducted at the European level. To avoid the so-called modifiable areal unit problem (Arbia 2012), i.e. the distortion caused by using local administrative units varying in size and/or shape, the DegUrba classification is based on a map of the regional territory with a square grid cell of 1 km^2. The advantage of using this classification is that all territorial cells have the same shape and size, thus producing a classification that is comparable across space and more stable along time (European Union et al., 2021).

By grouping the grid cells using a combination of population density, population size, and contiguity (neighbouring cells) criteria, the methodology creates a classification of municipalities into three groups: high-density population areas (at least 50% of the population lives in high-density urban centres), intermediate urban areas (at least 50% of the population lives in urban clusters), and rural areas (at least 50% of the population lives in rural grid cells).

An additional advantage of the DegUrba classification is that it is updated periodically to consider changes in spatial unit boundaries over time. Thus, the derived territorial map is not static and could also represent an interesting tool—harmonized internationally—to analyse the impact of urbanization changes on business performance. To our knowledge, our work is the first attempt to use this kind of spatial aggregations in the analysis of the productivity–wage relation at the local level.

Figure 1 provides a map of the areas of interest, and Table 1 reports some descriptive statistics.

3.3 Descriptive Analysis

In this section, we present a descriptive analysis of the main aggregates at the establishment level: labour productivity, wages, and the relative gap. Labour productivity

Table 1 Basic characteristics of DegUrba categories (totals and % by aggregations)

	Municipalities	Population (thousands)[a]	Surface (Km^2)	Firms' local units (thousands)	Value added (billions)	Employees (thousands)	Payroll-registered (thousands)
Lombardy Region	1,527	10,019	23,864	259,337	147,4	2,561	2,308
Urban high density	8.2	40.4	7.8	49.5	55.3	52.5	53.6
Urban Intermediate	47.1	47.8	34.2	42.2	38.1	40.4	39.7
Rural	44.7	11.8	58	8.3	6.6	7.1	6.7

[a]Residents as of the 1st of January, 2017

Manufacturing and services activities within the subsample of establishments with payroll-registered employees is computed as the ratio of value added to the total number of employees. The individual wage is given by the ratio of gross wages to the number of payroll-registered employees. We assume that independent employees, representing 26% of the region's employment, have the same gross wage as salaried employees. External employment, including agency workers, are not included in the computation of labour cost. The cost of external employment is part of intermediate consumption; thus, it does not contribute to value added (Arnaldi et al. 2016, p. 65).

The aggregate evidence (Table 2), based on median averages, shows that both productivity and wages are higher in manufacturing (42.9 thousand and 24.0 thousand EUR, respectively) than in the services industry (31.3 thousand and 19.0 thousand EUR). One should note, however, that financial services are excluded from our investigation, which will affect this comparison. Differences are higher in productivity than in wages, and we observe a higher productivity–wage gap in manufacturing than in services (18.3 thousand and 12.9 thousand, respectively).

These patterns are clearly correlated to location. We use an index representation to better disclose differences (Fig. 2; Lombardy average = 1): establishments in high-density areas are better off than those located in other areas. However, this advantage is not equally shared between services and manufacturing activities. The productivity gain due to urban localization is mainly concentrated in services, whereas it is far less pronounced in manufacturing. The services activities localized in high-density areas show a positive productivity–wage gap that we do not observe for manufacturing as a whole. This gap is mainly determined by their higher productivity; whereas for manufacturing in high-density areas, the lower gap seems crucially linked to lower productivity.[3] An additional sectoral breakdown[4] indicates that in the services, the positive gap is mainly concentrated in high-technology and knowledge-intensive activities, whereas in manufacturing, where the average gap is slightly negative, the medium–high and high-technology activities still gain a premium (Fig. 3).

Interestingly, we also observe differences in the spread of the data. We use kernel density estimations to investigate firm-specific heterogeneity within each area (Fig. 4). Firstly, manufacturing establishments show not only a higher median gap compared to services, but also a higher dispersion of the data. This is well described by the estimated distributions, which are taller and thinner for services compared to manufacturing establishments in all areas. Secondly, we note that the manufacturing data are equally spread across different locations, with median values being quite aligned, while in the services the spread seems to correlate with the location of units, thus indicating a higher level of heterogeneity in densely populated locations and, also, less heterogeneity in rural areas compared to urban ones. As for service establishments in high-density locations, the kernel density estimations show a lower share around the central values and a slightly higher share towards larger values. All in all, this additional evidence

[3] It is worth stressing that this descriptive evidence at the establishment level is consistent with that obtained at the aggregated level, i.e. by computing productivity and wage averages in the three types of territorial area as ratios of the reference totals. In this case, we used the entire sample of establishments localized in the region and summed up plant-level values for full-time equivalent employees, value added, and gross wages (results are available on request). Both views—at the micro- and macro-aggregated level—coherently support the evidence derived.

[4] For the sectoral breakdown according to the 'high-tech' classification, see the Appendix.

Table 2 Labour productivity, wages, and productivity–wage gap: averages values (euros)

	Manufacturing			Services			Total		
	Productivity	Wages	Gap	Productivity	Wages	Gap	Productivity	Wages	Gap
Urban high density	43,722	24,778	18,250	33,817	19,947	14,746	35,240	20,659	13,670
Urban intermediate	43,198	24,029	18,588	29,180	18,229	12,601	33,097	20,051	12,663
Rural	39,533	22,486	17,140	26,415	17,488	12,425	30,195	19,286	10,745
Total	42,952	24,042	18,309	31,264	19,065	13,649	33,786	20,256	12,937

Productivity: value added per employee (full-time equivalent). Wages: gross wages per employee (full-time equivalent). Gap: difference between productivity and wages; median values

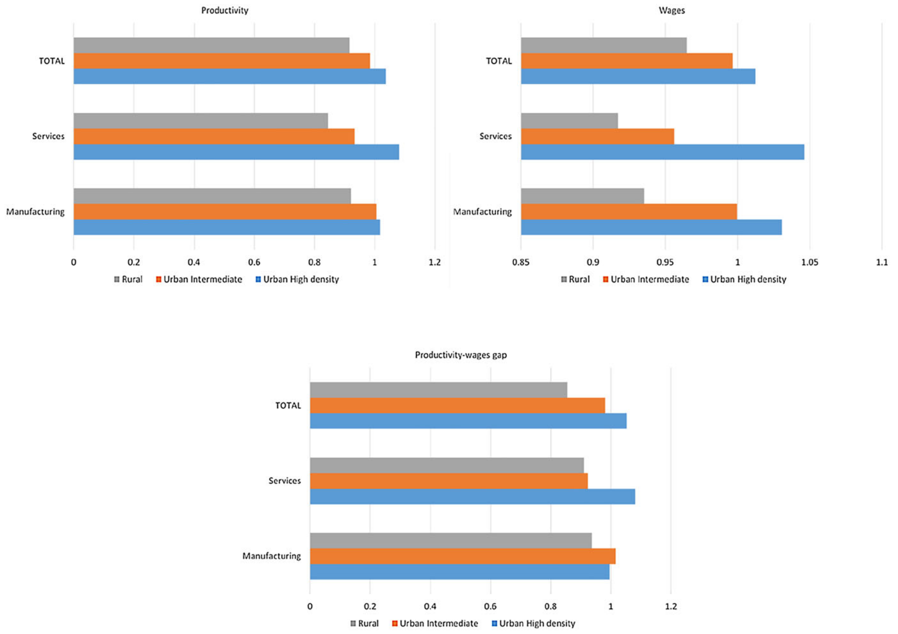

Fig. 2 Establishment-level patterns of productivity and wages (medians): between-area variation. Lombardy = 1

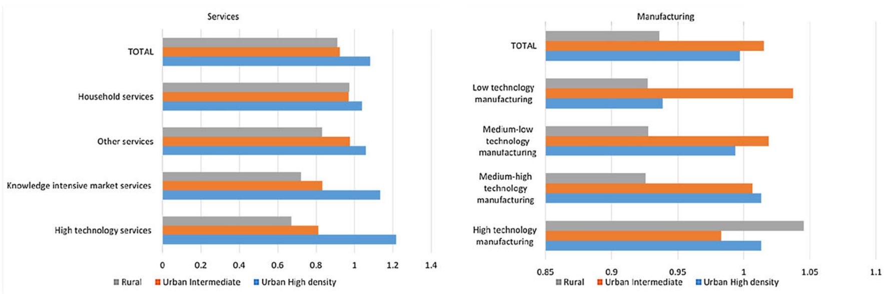

Fig. 3 Productivity–wage gap by sector: between-area variation. Lombardy = 1

reveals a higher degree of firm heterogeneity linked to localization factors within the services industry.

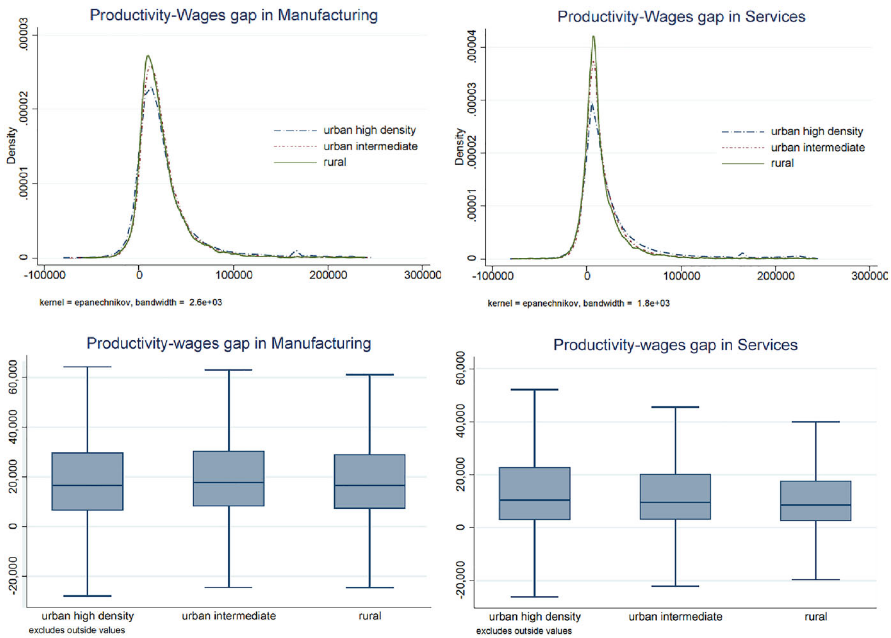

Fig. 4 Productivity–wage gap: kernel density estimation and box plot

4 The Empirical Model

4.1 Methodology

We separately estimate three equations for labour productivity (log Y), wages (log W), and the productivity–wage gap (log(Y/W)) at the plant level. We are aware that our investigation reflects a static view of the spatial effect on productivity and wages and that sorting effects may be operational (De la Roca and Puga 2017; Gaubert 2018). We discuss and tackle this issue in Sect. 5, although the static nature of the data is a limitation. However, the availability of matched employer–employee characteristics enable us to separately consider five main determinants:

(i) urban density;
(ii) industry characteristics related to different technological levels;
(iii) plant characteristics which capture effects related to increasing returns: productive scale (six size classes) and the international openness of the firm (a dummy variable indicating whether the establishment is part of an exporting firm);
(iv) employee characteristics: gender and age structure, level of education, and type of labour contract (temporary, open-ended);
(v) a proxy of the slack in the local labour market (the employment rate).

Thus, the empirical specification can be represented as follows:

$$y_i = \beta_0 + \beta_1 U_a + \beta_2 U_s + \beta_3 (U_a U_s) + \beta_4 X_i + \beta_5 I_a + \varepsilon_i, \qquad (1)$$

where y_i generically indicates the dependent variable of interest for establishment *i,* and U_a and U_s are the vectors of dummy variables corresponding to area and sector characteristics to which the establishment belongs. X_i is a vector of plant-specific variables, while I_a indicates a vector of additional area-level characteristics.$\beta_1, \ldots \beta_5$ are the corresponding vectors of coefficients, and ε_i is the error term that accounts for potential dependence based on spatial proximity.[5] With this specification, we can explore interaction effects between localization (the agglomeration characteristics of the area) and the sector of activity at the establishment level, together with spatial autocorrelation.

In addition, the inclusion of individual characteristics that may directly affect performance at the establishment level allows us to control for differences in terms of both worker and firm characteristics. Finally, we assume that vector I captures other demand or supply factors affecting the local labour market, in our case proxied by the employment rate. Descriptive statistics and further explanation are reported, respectively, in Table 3 and in the Appendix.

We first present OLS estimates of Eq. (1), which, nevertheless, report standard errors modified according to Conley (1999).[6] In Sect. 5, we discuss a complementary estimation that explicitly considers simultaneity issues and, therefore, shows instrumental variable (IV) estimates. With low inflation rates characterizing both consumption and production prices, we use nominal values for both productivity and wages.[7]

In Table 4, we present the results for the overall sample of establishments operating in the manufacturing and service industries. In this set of estimates, the localization dummies are interacted with two dummies related to the services and manufacturing industries, with the latter as the reference. In Tables 5 and 6, we present results for manufacturing and services separately. In both sets of estimates, we further explore the industry impact by interacting the localization dummy with the establishments classified according to their technology- and knowledge intensity (low-technology manufacturing and household services are the reference categories for manufacturing and services, respectively). In all estimations, the intermediate urban area is used as the reference category.

4.2 Results: All Establishments

The estimates based on the overall sample show that, in general, being localized in the highest-density areas significantly affects productivity when the interaction with

[5] We use a spatial error representation following Conley (1999). We thank an anonymous referee for this suggestion. Conley (1999) method accounts for spatial correlation using coordinates. See the following note for further details.

[6] Following the recent technique proposed by Colella et al. (2019), we imposed a threshold of 50 km. This means that the errors of each municipality are assumed to be correlated with the ones of all other municipalities that are located within a radius of 50 km from it. We implemented the same specifications also with a threshold of 100 km without substantial modification.

[7] We also estimated our models using real values for our focus variables. We adopted two alternatives to obtain real values: in the first, we used the value-added deflator for both productivity and wages; in the other, we used the consumer price index for wages. We obtained coefficients very close to the regressions with nominal values.

Table 3 Descriptive statistics for the variables used in the empirical model

Variable	Mean	Std. Dev
Productivity (Log)	10.390	0.898
Wage (Log)	9.869	0.558
Gap (Log)	0.521	0.722
Urban high density	0.481	0.500
Urban intermediate	0.431	0.495
Rural	0.088	0.283
HT manuf	0.008	0.092
MHT manuf	0.050	0.218
MLT manuf	0.061	0.240
LT manuf	0.067	0.250
HITS	0.042	0.200
KWNMS	0.162	0.369
Other services	0.387	0.487
Household services	0.109	0.311
Size 1–9	0.832	0.373
Size 10–19	0.094	0.292
Size 50–249	0.047	0.211
Size 20–49	0.024	0.153
Size 250–499	0.002	0.044
Size 500 +	0.001	0.028
Internationalization	0.202	0.401
Aged 30–49	0.538	0.329
Aged 50 +	0.230	0.289
High education	0.109	0.223
Males	0.529	0.389
Temporary contracts	0.140	0.250
Employment rate	50.880	1.100

See the Appendix for variable descriptions and the sectoral aggregations based on NACE Rev. 2

services activities is considered (Table 4). Service activities localized in high-density clusters show a 10% gain in productivity, as indicated by the interacted coefficient,[8] thus partially offsetting the general loss of the services industry (–12.1%). Conversely, establishments in rural areas show a 4.1% reduction in productivity, without any significant difference between services and manufacturing activities (see the interacted coefficient *rural*services,* which is not significant).

[8] In a linear model, the overall effect of, say, U_s on the dependent variable must be computed as the algebraic sum of the stand-alone coefficient β_2 and the coefficient of the interacted term β_3.

Table 4 OLS estimation: manufacturing and services establishments

Variable	Productivity	Wages	Gap
Urban high-density	– 0.0146*	0.0309***	– 0.0456***
	[0.0171]	[0.00949]	[0.0118]
Rural	– 0.0414**	– 0.0463***	0.00477
	[0.0155]	[0.00974]	[0.00915]
Services	– 0.121***	– 0.0663***	– 0.0552***
	[0.00687]	[0.00584]	[0.00506]
Urban high– density*Services	0.100***	0.0328***	0.0658***
	[0.0292]	[0.0119]	[0.0186]
Rural*Services	– 0.0151	0.0229	– 0.0381***
	[0.0192]	[0.0158]	[0.0118]
Size 10–19	0.236***	0.183***	0.0545***
	[0.00659]	[0.00500]	[0.0106]
Size 20–49	0.267***	0.223***	0.0460***
	[0.00566]	[0.00803]	[0.00984]
Size 50–249	0.289***	0.265***	0.0252*
	[0.00807]	[0.00999]	[0.0131]
Size 250–499	0.300***	0.280***	0.0202
	[0.0304]	[0.00947]	[0.0282]
Size 500 +	– 0.263***	0.186***	– 0.450***
	[0.0422]	[0.0212]	[0.0419]
Internazionalization	0.453***	0.289***	0.166***
	[0.00590]	[0.00443]	[0.00462]
Aged 30–49	0.326***	0.217***	0.109***
	[0.0310]	[0.0204]	[0.0115]
Aged 50 +	0.467***	0.291***	0.176***
	[0.0691]	[0.0402]	[0.0298]
High education	0.768***	0.542***	0.225***
	[0.0458]	[0.0534]	[0.00999]
Males	0.163***	0.294***	– 0.128***
	[0.0482]	[0.0358]	[0.0134]
Temporary contracts	– 0.212***	– 0.210***	– 0.00202
	[0.0204]	[0.0181]	[0.0283]
Employment rate	0.0135***	0.00868*	0.00487**
	[0.0062]	[0.00513]	[0.00212]
Constant	9.170***	8.980***	0.190*
	[0.315]	[0.258]	[0.107]
Observations	254,689	254,689	254,689
R-squared	0.166	0.252	0.026

Conley-robust standard errors in brackets. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$
Reference area: Intermediate density. Reference Size: less than 10 employees

Table 5 OLS estimation: manufacturing establishments

Variable	Productivity	Wages	Gap
Urban high density	− 0.0226	0.0349**	− 0.0575***
	[0.0214]	[0.0148]	[0.0145]
Rural	− 0.0408**	− 0.0366**	− 0.00422
	[0.0206]	[0.0166]	[0.0103]
HT manuf	0.192***	0.0678***	0.125***
	[0.0251]	[0.0150]	[0.0171]
Urban high density*HT manuf	0.0702*	0.0598***	0.0104
	[0.0372]	[0.0231]	[0.0178]
Rural*HT manuf	− 0.0283	0.0157	− 0.0440
	[0.0635]	[0.0305]	[0.0625]
MHT manuf	0.319***	0.190***	0.129***
	[0.0149]	[0.00932]	[0.00968]
Urban high density*MHT manuf	0.0512**	0.00754	0.0436***
	[0.0241]	[0.0129]	[0.0160]
Rural*MHT manuf	− 0.0178	− 0.0298	0.0120
	[0.0226]	[0.0196]	[0.0162]
MLT manuf	0.282***	0.155***	0.127***
	[0.0127]	[0.0100]	[0.00545]
Urban high density*MLT manuf	0.0339*	− 0.0175	0.0515***
	[0.0184]	[0.0130]	[0.0167]
Rural*MLT manuf	− 0.0211	− 0.0244	0.00333
	[0.0262]	[0.0219]	[0.0160]
Size 10–19	0.223***	0.153***	0.0695***
	[0.00746]	[0.00383]	[0.00618]
Size 20–49	0.316***	0.227***	0.0886***
	[0.0111]	[0.00709]	[0.00635]
Size 50–249	0.425***	0.311***	0.114***
	[0.00899]	[0.00918]	[0.0124]
Size 250–499	0.520***	0.402***	0.118***
	[0.0346]	[0.0218]	[0.0365]
Size 500 +	− 0.337***	0.345***	− 0.682***
	[0.111]	[0.0708]	[0.143]
Internationalization	0.350***	0.192***	0.157***
	[0.0168]	[0.00876]	[0.0105]
Aged 30–49	0.0934***	0.176***	− 0.0821***
	[0.0284]	[0.0159]	[0.0163]
Aged 50 +	0.117***	0.270***	− 0.153***

Table 5 (continued)

Variable	Productivity	Wages	Gap
	[0.0286]	[0.0154]	[0.0200]
High education	0.662***	0.527***	0.136***
	[0.0673]	[0.0525]	[0.0272]
Males	0.315***	0.389***	– 0.0739***
	[0.0150]	[0.0108]	[0.00963]
Temporary contracts	– 0.0503**	– 0.105***	0.0545**
	[0.0253]	[0.0186]	[0.0225]
Employment rate	0.0124**	0.00576	0.00661***
	[0.00514]	[0.00399]	[0.00255]
Constant	9.197***	9.027***	0.170
	[0.261]	[0.203]	[0.131]
Observations	54,491	54,491	54,491
R-squared	0.244	0.349	0.053

Conley-robust standard errors in brackets. ***$p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Reference area: Intermediate density. Reference Size: less than 10 employees

As concerns the wage variable, high-density areas show a positive contribution (+ 3.1%) to wages.[9] It is worth noting that average wages in services are lower than in manufacturing. However, the latter's sector-specific characteristics imply that the wage premium in high-density areas is amplified; on the other hand, a negative impact on compensation in rural clusters prevails (–4.6%).

Looking at the productivity–wage gap estimates, the dummy coefficient for services is negative (–5.5%), implying that the profit margin reduces compared to manufacturing. However, localization matters, as the contribution to the productivity–wage gap when the services dummy is interacted with high-density areas is positive (+ 6.6%) and tends to completely offset the negative impact previously described.

These interactions between localization and industry activity are estimated controlling for individual characteristics—included in the X vector of additional variables—that may directly impact productivity. One should recall that the set of size dummies clearly indicates that establishment productivity is positively affected by an increasing scale of production. Taking the 'fewer than 10 employees' category as the reference group, productivity increases with size, but not in the last group (> 500 employees) representing large enterprises. In this group of plants, decreasing returns are operational, implying a 26.3% reduction in productivity compared to small firms. However, one should consider the fact that larger enterprises may have multiple plants, and therefore, productivity at the firm level may indeed increase with size as internal scale economies may be in operation.

[9] Although not directly comparable, it is worthwhile noting that Ahlfeldt and Pietrostefani (2019) reported an elasticity of 0.04 between wages and density in their extensive review of the effects of density on a large set of outcomes.

The impact of plant size on labour compensation is positive. We observe an increase in wages as one moves towards larger size classes, with the highest impact registered in the medium-large sizes (250–499 employees), but with a slower pace for establishments with 500 employees or more. Conversely, this group of establishments shows a positive effect on labour compensation that is lower than that observed for the other size classes, except for the 10–20 employees category. Nevertheless, it is enough to determine a negative impact on the productivity–wage gap in the larger class size (–4.5%). As we have previously emphasized, discrepancies between productivity and wages reflect different conditions regarding the nature of labour as a quasi-fixed factor, information asymmetries, and bargaining power, with the latter typically increasing with (plant) firm size.

Regarding labour characteristics, we find a significant gender gap that is probably related to the positive correlation between part-time contracts and the proportion of females employed at a firm. A one percentage point (p.p.) increase in the proportion of male employment at a firm determines a 0.16% gain in productivity.

Age has a positive impact on productivity as well, as having employees in the middle-aged group (30–49 and 50 + years of age) improves productivity. In this case, a one p.p. increase in the proportion of employees in these age classes entails, respectively, a 0.33% and 0.47% increase in productivity.

However, temporary employment leads to a decrease in productivity of 0.21%. In contrast, human capital proxied by education level shows the highest impact among the labour input variables: a one p.p. increase in the share of highly educated workers determines an increase in productivity of 0.77%.[10]

Finally, a local labour market's slack is captured by its corresponding employment rate.[11] This variable reflects the condition of labour demand and, therefore, the potential relevance of labour market slack. In other words, a higher employment rate reflects tightness of the labour market and, thus, the extent to which firms compete to acquire better workers. The estimates suggest a significant positive impact that is, nevertheless, mild (0.01%).

4.3 Results by Sector of Activity

The medium–low-, high-, and medium–high-technology establishments represent 70% of Lombardy manufacturing in terms of total employment. Our estimates show that these establishments exhibit higher levels of productivity when compared to their low-tech counterparts, on average, with the medium–high category showing the highest impact (+ 32%) (Table 5). These plants gain an additional increase when located in the highest-density areas. In this case, the gain is greater for establishments in high-tech sectors (a further + 7% increase). Manufacturing establishments in rural areas do not show significant differences in productivity compared to intermediate urban

[10] This result is coherent with the evidence on the skill premium in urban and nonurban commuting zones of Italian local labour markets obtained by Accetturo et al. (2019).

[11] Other potential proxies could be used in order to capture the effect of local labour market conditions. Given the high disaggregation of our data at the local level, we can only count on labour force census statistics collected at the municipality level or the local labour market level. We use this latter also to take into account a referee's comment.

locations, which is the reference category. As previously stressed for the overall sample of manufacturing and service establishments, even in the manufacturing industry localization in the highest-density areas does not represent a driver of productivity per se, but it does in combination with sectoral specificities. Conversely, agglomeration economies have a distinctive role in wage setting. Our estimates show that in the highest-density areas, wages increase by 3.5% compared to intermediate locations, with a substantial further increase for high-tech establishments (+ 6%). In addition, in rural areas wages are lower, on average (–3.7%), without significant contribution induced by sectoral characteristics. The impact concerning the productivity–wage gap is negative in high-density areas because of the impact of wages. It is positive for all three technology levels, in comparison with the lowest-tech industries; however, when interacted with the high-density dummy, it is positive when the medium- and medium–low- tech levels are considered.

As for services (Table 6), being localized either in the highest-density or rural areas does not affect productivity. However, when one considers the interaction with sectoral dummies, a significant loss of productivity is observed when the establishments are localized in rural areas.

Interestingly, the knowledge-intensive market services (KWNMS) show the highest gain in productivity (+ 38%), with an additional increase for those localized in high-density areas (+ 14.7%). In addition, the high-technology services (HITS) localized in densely populated areas show a significant increase in productivity as well.

In contrast to the general (reducing) effect of agglomeration on the productivity–wage gap previously described, the gain concerning productivity for establishments operating within these knowledge-intensive services is substantial, giving rise to an increase in the productivity–wage gap, particularly in high-density locations. More specifically, we observe a further 11.5% increase, which adds to the stand-alone KWNMS dummy (+ 13%), and a milder impact of the HITS sectors, which depends on the combination of the stand-alone HITS dummy and the dummy interacted with urbanization.

Other significant differences between manufacturing and services arise when considering the direct impact of firm-specific characteristics. In particular, temporary employment and labour quality affect productivity evenly. This latter effect is greater and entails a 6.6% increase in productivity and a positive wage gap for highly educated workers.

We also control for firm internationalization by considering the impact of a dummy variable reflecting whether a firm exports goods or services. We are aware that such an inclusion may cause endogeneity issues, which should be confronted with pairwise relevant omitted variable concerns. Facing these two issues, we decided to include this variable, as it is a relevant firm-specific characteristic that crucially affects behaviour and performance. However, we have taken into consideration the endogeneity issue in the sectoral IV specification that we discuss in the next section.

Of course, manufacturing companies are more exposed to international competition, and on the whole, the degree of internationalization is also higher in this sector. Therefore, it is reasonable to expect that the impact on productivity is positive and possibly higher for the few services firms competing in international markets, compared to their counterparts operating only in a national context. One should note that

Table 6 OLS estimation: services establishments

Variable	Productivity	Wages	Gap
Urban high density	− 0.00303	0.0435***	− 0.0465***
	[0.0107]	[0.00800]	[0.0145]
Rural	0.0159	− 0.0276	0.0435**
	0.243***	0.196***	0.0461***
HITS	0.272***	0.382***	− 0.110***
	[0.0142]	[0.00995]	[0.0138]
Urban high density*HITS	0.167***	0.0491***	0.118***
	[0.0401]	[0.0118]	[0.0313]
Rural*HITS	− 0.252***	− 0.0230	− 0.229***
	[0.0344]	[0.0201]	[0.0316]
KWNMS	0.380***	0.250***	0.130***
	[0.0162]	[0.00917]	[0.0142]
Urban high density*KWNMS	0.147***	0.0322	0.115***
	[0.0477]	[0.0203]	[0.0292]
Rural*KWNMS	− 0.109***	− 0.0247	− 0.0847**
	[0.0340]	[0.0206]	[0.0340]
Other services	0.121***	0.181***	− 0.0599***
	[0.0183]	[0.0141]	[0.00959]
Urban high density*Other services	0.0313*	0.00139	0.0299*
	[0.0163]	[0.00943]	[0.0164]
Rural*Other services	− 0.0522**	0.00702	− 0.0592**
	[0.0258]	[0.0172]	[0.0238]
Size 10–19	0.243***	0.196***	0.0461***
	[0.0128]	[0.00291]	[0.0134]
Size 20–49	0.236***	0.223***	0.0130
	[0.00875]	[0.00720]	[0.0127]
Size 50–249	0.183***	0.236***	− 0.0529***
	[0.0129]	[0.0107]	[0.00674]
Size 250–499	0.147***	0.205***	− 0.0580**
	[0.0227]	[0.0104]	[0.0266]
Size 500 +	− 0.309***	0.104***	− 0.413***
	[0.0537]	[0.0166]	[0.0527]
Internationalization	0.550***	0.322***	0.228***
	[0.00719]	[0.00392]	[0.00502]
Aged 30–49	0.300***	0.192***	0.108***
	[0.0189]	[0.0183]	[0.00609]
Aged 50 +	0.474***	0.262***	0.212***

Table 6 (continued)

Variable	Productivity	Wages	Gap
	[0.0566]	[0.0425]	[0.0152]
High education	0.662***	0.478***	0.184***
	[0.0258]	[0.0451]	[0.0249]
Males	0.0950**	0.219***	− 0.124***
	[0.0422]	[0.0355]	[0.00753]
Temporary contracts	− 0.197***	− 0.209***	0.0128
	[0.0297]	[0.0160]	[0.0341]
Employment rate	0.0152**	0.0110**	0.00419
	[0.00664]	[0.00558]	[0.00267]
Constant	8.851***	8.675***	0.176
	[0.338]	[0.283]	[0.135]
Observations	200,198	200,198	200,198
R-Squared	0.172	0.243	0.043

Conley-robust standard errors in brackets. ***$p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Reference area: Intermediate density. Reference Size: less than 10 employees

the share of exporters in the services sector is far below that in manufacturing (5% vs 30%).

The estimates suggest that internationalization significantly affects productivity in both manufacturing and services, but the impact on the latter is higher (+ 55%). In both cases, the productivity–wage gap is positive, although it is larger in services (+ 22.8%) than in manufacturing (+ 12.2%).

5 Endogeneity Issues and Sector Heterogeneity

5.1 General

The evidence presented in the previous sections suggests that the influence exerted by urbanization in terms of productivity, wages, and the productivity–wage gap may substantially vary—in both sign and size—across economic sectors according to their level of technological intensity. In this section, we shed further light on the impact of heterogeneity by computing estimates at a sectoral level; this allows us to evaluate the effects of location for the entire set of activities (whereas the sectoral effects presented in the previous results are measured with respect to a chosen reference sector). Moreover, we explicitly address the issue of potential endogeneity bias in the OLS estimates of Tables 4, 5, 6 through instrumental variables (IV) estimation.

Different sources of endogeneity bias may be at work in our estimation framework, which relates the degree of urbanization to productivity and wages. First, as Combes and Gobillon (2015) explain, when estimating the effects of location on outcome variables like productivity and wages, endogeneity bias due to omitted variables may be at work both at the local economy level and at the individual level. Omitted-variable bias may occur at the local-economy level if the specification does not account for differences in the local availability of amenities (such as transport infrastructure), that are likely to influence both urbanization and productivity. Similarly, omitted-variable bias may take place at the individual level if unobserved firm characteristics (e.g. managerial abilities) not accounted for in the regression specification influence both individual outcomes and how economic agents sort across locations, thereby giving rise to selection bias. 'Selection effects' (the tendency of large cities to select more efficient economic agents through competition) and 'sorting effects' (the self-selection of more productive firms and workers in larger cities, which offer higher expected profits and wages) represent—along with traditional agglomeration externalities—the main explanations economic theory currently provides for why per capita output is higher in larger cities.[12] Another source of endogeneity might come from potential measurement error in the measure of the degree of urbanization (as in our dataset, each municipality is entirely assigned to one of the three possible categories of the DegUrba classification according to algorithms that exploit spatial information collected at a more disaggregated level). Bias deriving from reverse causality is, in theory, possible as well (think, for example, of highly productive firms, whose particular ability to attract new workers substantially affects local density); however, in practice this is considered a minor concern in the literature on the impact of urban agglomeration economies (see Melo et al. 2009; Combes et al. 2010).

While the main goal of our analysis is to consistently estimate the effects of urbanization on productivity and wages, it has to be recognized that other explanatory variables included in our regression specification may raise concern due to endogeneity problems. In particular, the assumption of exogeneity of the dummy variable for exporting firms looks problematic. Economic intuition suggests that productivity itself affects the likelihood of export, and empirical papers evaluating the impact of exporter status on firm performance take into account this source of reverse-causality bias (see, for example, Costa et al. 2017, for a recent analysis of Italian firms). As a consequence, we also treat the dummy variable for exporting firms as an endogenous regressor in the IV estimates we present, in order to address the possible inconsistency of the OLS ones.[13]

[12] In a seminal contribution, Behrens et al. (2014) build up a theoretical model allowing for all of these mechanisms, and some recent empirical papers try to evaluate their importance. Using French firm-level data, Combes et al. (2012) test for the relative size of agglomeration and selection effects, concluding that the latter are irrelevant. Again using French firm-level data, Gaubert (2018) estimates that the source of the productivity premium of large cities is almost equally distributed between agglomeration economies and sorting effects.

[13] We acknowledge that the assumption of exogeneity might be questioned for other regressors. For instance, whereas including firm size into the productivity regression seems a natural way to capture the possible presence of economies (or diseconomies) of scale, it should be not overlooked that firm size itself (as measured by the number of employees) may be affected by firm performance. A similar reasoning holds for the share of the workforce with higher education. Unfortunately, the content and the cross-sectional nature of the dataset we work with prevents us from finding suitable instrumental variables to tackle the issue of reverse causality for all regressors that may generate this bias. At the same time, economic intuition suggests that simply dropping regressors such as firm size and the share of the workforce with higher education from the specification is likely to yield an even higher bias due to the omission of relevant factors. Specifically regarding firm size endogeneity, one should also note that bias should not be severe since, as

Computing estimates at the sectoral level means that we work with a simpler specification that excludes the interaction terms of the regressions in Tables 4, 5, 6. Clearly, doing so allows a substantial reduction in both the number of parameters to estimate and the number of instrumental variables required to perform IV estimation. In practice, the regression specification can be written as follows:

$$y_i = \alpha_0 + \alpha_1 X_i + \alpha_2 I_a + \gamma_1 High_{density} + \gamma_2 Rural + \gamma_3 Export + \mu_i \quad (2)$$

where X_i denotes again the vector of plant-specific variables assumed to be exogenous and α_1 the corresponding vector of coefficients; I_a refers to local characteristics, *High density* and *Rural* are the two binary measures of urbanization derived from the DegUrba classification (with an *Intermediate_density* binary measure implicitly acting as reference category); *Export* is the dichotomous variable for exporting firms. For the sake of space, in the following we will report the OLS and IV results for the urbanization variables only, but we will briefly comment on the estimates concerning the other regressors as well.[14]

5.2 Sectoral OLS Estimates

OLS estimation of specification (2) does not obviously imply any additional technical difficulty with respect to those of Tables 4, 5, 6. The main results of interest here are presented in Table 7, which again reports—along with estimated coefficients—Conley (1999) standard errors to account for possible cross-sectional dependence.

Let us first look at the manufacturing sector, starting from the upper part of the table where labour productivity is the dependent variable. It can be seen that being located in a highly urbanized area exerts a significant (and positive) impact on productivity only for firms whose production processes are characterized by a high level of knowledge intensity (HT group), while the coefficients of the other are non-significant. In contrast, the coefficient of the *Rural* regressor is systematically negative (and significantly so in the MHT and MLT sectors), suggesting that firms in areas of low urbanization experience a productivity loss compared to their competitors located elsewhere.

The coefficients reported in the middle part of Table 7 refer to the wage regressions and are generally estimated with greater precision than the productivity equations. Indeed, most coefficients are significant at conventional levels and, not surprisingly, outline a positive monotonic association between the degree of urbanization and the wage rate. Most importantly, the upward pressure exerted by urbanization on the wage rate turns out to be higher than that on productivity. As a result, firms located in highly urbanized areas are characterized by a lower productivity–wage gap, as can be seen

Footnote 13 continued
documented by Bartoloni and Baussola (2021), Italian firms fail to climb the firm size ladder (i.e. they mainly grow within their size class). As our specification considers firm size in a discrete representation, it could be thought of as being relatively stable over time. For these reasons, we have decided to maintain a specification that also includes regressors whose exogeneity is assumed but might appear debatable. We are grateful to a reviewer for pointing this out and for stimulating a more in-depth investigation of endogeneity issues.

[14] The complete results are reported in supplemental documentation and are available upon request.

Table 7 The impact of urbanization on productivity, wages, and the productivity–wage gap: OLS sectoral estimates

Sectors	Productivity							
	Manufacturing				Services			
	HT	MHT	MLT	LOT	HITS	KWNMS	Other services	Household services
High density	.06499**	.018950	.01758	− .00873	.12093***	.16416***	.01080	− .00232
	[.02595]	[.01316]	[.01351]	[.01518]	[.01675]	[.03075]	[.00997]	[.01206]
Rural	− .07126	− .05789**	− .06695***	− .034705	− .08176***	− .09296***	− .03171	.01564
	[.07512]	[.02233]	[.01643]	[.02305]	[.02427]	[.02381]	[.02475]	[.01763]
Adj. R^2	0.2837	0.1718	0.1312	0.2136	0.1138	0.0651	0.1866	0.0852
Sectors	Wages							
	Manufacturing				Services			
	HT	MHT	MLT	LOT	HITS	KWNMS	Other services	Household services
High density	.09584***	.03598***	.01963***	.04038***	.09125***	.07144***	.03926***	.05557***
	[.01719]	[.00835]	[.00729]	[.01426]	[.01077]	[.01539]	[.00711]	[.00696]
Rural	− .02561	− .06606***	− .06383***	− .03494	− .00035	− .04474***	− .02060	− .03571**
	[.03091]	[.00944]	[.01042]	[.02197]	[.01328]	[.01596]	[.02438]	[.01682]
Adj. R^2	0.4245	0.2943	0.2370	0.2941	0.2213	0.1347	0.2356	0.1068

Table 7 (continued)

Sectors	Productivity–wage gap							
	Manufacturing				Services			
	HT	MHT	MLT	LOT	HITS	KWNMS	Other services	Household services
High density	− .03084	− .01703**	− .00205	− .04912***	.02967**	.09272***	− .02846***	− .05790***
	[.02028]	[.0071]	[.00863]	[.00866]	[.01280]	[.01614]	[.00824]	[.01418]
Rural	− .04565	.00817	− .00312	.00023	− .08141***	− .04822***	− .01111	.05135***
	[.06526]	[.01699]	[.01286]	[.00975]	[.02232]	[.01605]	[.01116]	[.01827]
Adj. R^2	0.053	0.030	0.028	0.047	0.007	0.018	0.043	0.008

OLS estimates. Each regression specification includes the same covariates as in Tables 4, 5, 6. For the sake of space, only results concerning the variables measuring the degree of urbanization are reported (based on the DegUrba classification, a binary index 'Intermediate density' acts as a reference category for 'High density' and 'Rural'). Standard errors reported below coefficients control for cross-sectional dependence (Conley 1999). The complete results for each regression are reported in a supplemental documentation and are available upon request in the bottom part of Table 7. The estimated coefficient of the *High_density* regressor is systematically negative, statistically significant in the MHT and low-tech (LOT) sectors, and close to statistical significance in the HT one.[15]

We now turn to the OLS estimates concerning services. Sectoral specificities related to technological complexity emerge more clearly in this case. Indeed, in the productivity regressions for the two sub-sectors where scientific knowledge and advanced technologies play a more relevant role (HITS and KWNMS), the estimated coefficients point to a significant, positive, and monotonic relationship between the degree of urbanization and productivity. In contrast, urbanization does not seem to affect productivity in the sectors characterized by lower technological intensity ('Other Services' and 'Household Services').[16]

Similarly to what is observed for industrial sectors, the OLS coefficients of the wage regressions indicate a positive monotonic association between the degree of urbanization and the wage rate. This is observed in all sectors, regardless of the degree of technological complexity (although the coefficients of the *Rural* regressor are again estimated with less precision).

The differences between the role played by technological intensity when evaluating the marginal impact of urbanization on productivity and wages entail interesting implications in the productivity–wage-gap regressions. Here, the interest of adopting the technology-based sector classification is apparent. As with the two sub-sectors with higher technology intensity (HITS and KWNMS), the marginal impact of being located in a highly urbanized area is much greater on productivity than on wages; the coefficient of the *High_ density* dummy in the gap regressions is positive, relatively large, and statically significant. In contrast, the point estimate is negative and significant for the two remaining sectors. An intuitive interpretation of this set of results is that whereas locating in urban agglomerations exerts an upward pressure on both productivity and wages, the positive externalities related to agglomerations overcome the increase in wages (which, in turn, reflects an increase in congestion costs) only in more technologically sophisticated sectors.

5.3 Sectoral IV Estimates

A particular difficulty in tackling endogeneity in our estimation framework is that we have to employ adequate instrumental variables for categorical variables (urbanization and exporter status). Following (and extending) suggestions in Wooldridge (2010) and Angrist and Pischke (2008) for the case of a single endogenous dummy variable, we adopt the following procedure. First, we estimate probit models, regressing the endogenous variables on a proper set Z of 'instrumental variables' as well as on the other covariates (X_i and I_a) of our base specification. As far as urbanization is concerned, we actually estimate an ordered response model (ordinal probit), which provides us with the estimated probability that a firm is located in one of the three

[15] The unreported p value of the null hypothesis is 0.128.

[16] This is not surprising, as the extent of the knowledge externalities generated by agglomeration should increase with the technological complexity of the sector at hand (i.e. knowledge externalities should be larger in sectors where knowledge matters more).

possible DegUrba categories. In the case of the export dummy, we estimate a standard probit model for binary variables, obtaining the same kind of information. Then, predicted probabilities obtained from probit estimates are used as instruments in the IV estimation of the productivity and wage regressions. Such a procedure implies that the second-stage equation is exactly identified.[17]

As for the choice of variables to be included in set Z, we take advantage of hints from the existing literature. To begin with, we draw upon Costa et al. (2017) in selecting (the log of) firm age as a plausible predictor of exporter status. The basic reasoning behind this choice is that it takes time to generate the learning effects (in terms of know-how and managerial/organizational competences) required to access foreign markets; thus, on average, the probability that a firm belongs to the exporter group should be lower for younger producers.[18] A further reason to include firm age in set Z is that it also seems to be a reasonable predictor of location. The geographical distribution of producers is not random but reflects optimization decisions that take into account the economic implications of being located in areas characterized by different degrees of urbanization. To the extent that land (space) is a scarce resource, it can be argued that younger firms face tighter constraints in their location choices and are also more likely to end up with suboptimal decisions. Furthermore, we assume that a set of locality-specific factors influence the location decisions of producers across urban and nonurban areas. In particular, drawing on Di Giacinto et al. (2014) we decide to include into set Z the following variables computed in log form at the municipality level: altitude, population density in 1921, and the share of the population holding a high school degree in 1971.[19] The inclusion of these variables reflects the current practice of choosing geographical (altitude) or historical features (old values of population density and of the schooling rate) as instruments when estimating agglomeration economies (on this, see also Ciccone and Hall 1996, and Combes et al. 2008a, b). The implicit assumption is that while these features are useful to proxy the factors driving the location decisions of firms and workers, they do not exert a direct effect on current differences in productivity. Also note that for the sake of comparison,

[17] This procedure is presented in Wooldridge's (2010, ch.21, p.935–936) for the binary case. An interesting "robustness" property is that it may yield valid instruments even in cases of misspecification of the probabilistic model.

[18] Costa et al. (2017) analyze the effects of different forms of internationalization on firm performance in a large sample of Italian firms over the 2007–2010 period. They use firm age to predict the probability that a producer shifts along the taxonomy during the considered period; then, by adopting Heckman's (1979) correction procedure to deal with selection bias, they use this prediction to evaluate the impact of internationalization on labor productivity in a second-stage regression.

[19] Di Giacinto et al. (2014) analyze a large sample of Italian manufacturing firms over the 1995–2006 period. Much of their study is focused on evaluating the location effects of two different kinds of spatially concentrated areas, 'urban areas' and 'industrial districts' (defined not at the municipality level but at a more aggregate level represented by the 'local labour market areas'; Istat 2006). They find that firms located in both types of areas enjoy a productivity premium with respect to those located elsewhere ('other areas'), the gain being significantly larger for producers located in cities. Though their mapping of the territory is different from ours, the instruments they select seem appropriate in our case as well. Also Buzzacchi et al. (2021) use population in 1921 as an instrumental variable to estimate the productivity implications of density. Their reasoning is that differences in population density mainly stem from differences in local characteristics, such as land fertility, that are not very relevant for modern industries and services and are persistent over time ("[…] *past population density was high in places high land fertility. Once cities are created, they usually display a strong persistence through time* […]").

this set Z is held fixed across both sectors and equations. Thus, for each sector the predicted probabilities we use to compute the second-stage regression are the same, regardless of the dependent variable.

Let us now turn to the IV results reported in Table 8. To begin with, the instrumental variables we build for each sector do not appear to be weak. Indeed, the conventional F-statistics reported at the bottom of the table are systematically much larger than the conventional threshold value of 10, below which—since the contribution of Stock and Yogo (2005)—weak identification is usually considered a potential issue.

Some relevant differences emerge between the OLS and IV estimates, and our comment essentially focuses on these, starting again from the manufacturing sectors. The first striking difference concerns the coefficient of *High_density* in the productivity regression of the HT group, which is no longer significant and even changes sign (from positive to negative) in the IV estimation; in the other manufacturing groups, the coefficient of the same regressor does not change sign with respect to the OLS estimate and turns out to be significant in the MHT and LOT groups. Rural locations yield either not significant or negative productivity effects. Coefficients in the wage regressions are estimated with much lower precision than in the OLS estimate; those that are significant are consistent with the idea that labour costs are greater at higher levels of urbanization. IV estimates concerning the productivity–wage gap—the main variable of interest in our paper—confirm that locating in urban agglomerations implies a lower cost of competitiveness for firms belonging to industries characterized by low technological intensity (LOT). A similar disadvantage emerges for producers who belong to the most technologically-intensive group (HT) and are located in rural areas.

As for services, IV estimates concerning the effect of locating in highly urbanized areas on productivity and wages are consistent with the OLS ones. In the most knowledge-intensive industries (HITS and KWNMS), the positive productivity effects related to urban agglomerations persist also when controlling for possible selection bias or other sources of endogeneity. In addition, the upward pressure on labour compensation emerges regardless of any distinction based on knowledge intensity. As a consequence, the IV estimates of the effects of locating in highly urbanized areas on the productivity–wage gap are very similar to the OLS ones and again display significant heterogeneity according to the degree of knowledge intensity.

More interesting differences with respect to the OLS results come from the IV estimates of the *Rural* regressor. First, when controlling for selection bias producers in knowledge-intensive sectors located in rural areas do not necessarily experience a loss of productivity (compared to the reference category of an intermediate level of urbanization). Secondly, whereas in the OLS results of the wage regressions the estimated coefficients of *Rural* are always negative, in the IV computations they systematically always take on the opposite sign. Moreover, they turn out to be statistically significant in the 'Other services' sector and not far from statistical significance in the HITS and KWNMS sectors (the unreported p values of the null hypothesis are 0.11 and 0.15, respectively). A possible interpretation of these differences involves, on the one hand, the failure of OLS to control for selection bias and, on the other hand, the lower supply of infrastructure, transport, and other amenities in rural areas. In other words, the negative OLS coefficient is consistent with the fact that less urbanized areas host less efficient producers (who pay lower wages, on average); by controlling

Table 8 The impact of urbanization on productivity, wages, and the productivity–wage gap: IV sectoral estimates

Sectors	Productivity							
	Manufacturing				Services			
	HT	MHT	MLT	LOT	HITS	KWNMS	Other services	Household services
High density	− .08706	.07900**	.00522	− .09423***	.17770***	.49572***	.03702	− .01956
	[.07015]	[.03862]	[.03565]	[.02964]	[.03746]	[.15006]	[.03596]	[.03163]
Rural	− .37291**	.00299	− .09047***	.02198	− .07998	.17384	.20748**	.03479
	[.18353]	[.06091]	[.03365]	[.04444]	[.05698]	[.18225]	[.08736]	[.06785]
Sectors	Wages							
	Manufacturing				Services			
	HT	MHT	MLT	LOT	HITS	KWNMS	Other services	Household services
High density	.00667	.09067***	− .00121	.02334	.141689***	.26236***	.09494***	.10389***
	[.06515]	[.02918]	[.01401]	[.01654]	[.02805]	[.07761]	[.02605]	[.01313]
Rural	.06742	− .01427	− .09928***	.04085	.06156	.15683	.19747***	.03942
	[.10447]	[.03475]	[.02587]	[.04016]	[.03849]	[.10935]	[.06553]	[.03897]
Sectors	Productivity–wage gap							
	Manufacturing				Services			
	HT	MHT	MLT	LOT	HITS	KWNMS	Other services	Household services
High density	− .09373	− .01166	.00642	− .11756***	.03601*	.23337***	− .05792***	− .12346***

Table 8 (continued)

Sectors	Productivity–wage gap							
	Manufacturing				Services			
	HT	MHT	MLT	LOT	HITS	KWNMS	Other services	Household services
	[.08369]	[.01608]	[.02671]	[.01912]	[.02073]	[.07310]	[.01428]	[.03328]
Rural	− .44033**	.01727	.00881	− .01887	− .14154**	.01701	.01001	− .00463
	[.20119]	[.03843]	[.02703]	[.02120]	[.05687]	[.07918]	[.03128]	[.05224]
F.st.- High density	71.97	40.25	56.51	106.23	349.21	134.5	99.63	99.39
p- value	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]
F.st.- Rural	36.51	68.2	140.52	149.8	412.95	81.84	149.32	103.46
p-value	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]
N. obs	2478	14,562	17,873	19,578	12,236	44,604	113,133	30,225

Reported IV estimates refer to the same regression specification as in Table 7 but computed assuming that the variables measuring urbanization and exporter status are endogenous. 'High rural' and 'Rural' are instrumented by predicted probabilities computed from an ordinal probit model, where the categorical variable measuring urbanization from the DegUrba classification is regressed on a firm-specific variable (age), some local characteristics measured at the municipality level (population density in 1921, altitude, and the share of the population holding a high school degree in 1971), as well as the other covariates in Table 7 assumed to be exogenous. The dummy for exporters is instrumented by predicted probabilities obtained in the same way. 'F.st.' refers to usual first-stage F-tests to evaluate instrument relevance. Standard errors reported below coefficients control for cross-sectional dependence (Conley, 1999). The complete IV estimates are reported in supplemental documentation and are available upon request this selection effect, the IV estimator is able to highlight the need for firms located in rural areas to pay higher wages, *ceteris paribus*, to attract workers to a less desirable location. In the case of the HITS sector, this 'wage premium' entails a loss in terms of cost competitiveness, as witnessed by the negative and significant coefficient of *Rural* estimated in the productivity–wage-gap regression.

Finally, we briefly discuss the other covariates' estimated (and unreported) coefficients. To begin with, the results concerning the variable assumed to be exogenous are broadly consistent with those in Tables 5 and 6. Moreover, in both the OLS and IV sectoral estimates, the coefficient of the endogenous export dummy is generally -as expected- positive and significant, leaving out the case of the manufacturing MLT sector (where it fails to achieve significance in the productivity-wage IV regression). A possible criticism underlying these estimates is that firm age may not be a suitable instrumental variable for the export dummy (as selection effects imply older producers are more productive). While we are aware of this (but also lack valid alternatives to instrument exporter status), we have to stress that the main focus of our paper is the impact exerted by agglomeration upon the productivity-wage differential. Thus, as a robustness check, we have re-run all OLS and IV estimates while excluding the export dummy from the regressors list (as well as firm age from the IV set) and have found results similar to those presented in Tables 7 and 8. Since the impact of urbanization does not depend on whether the export dummy is dropped from the specification or not, we have preferred to present the estimates obtained when it is included. Theoretical and empirical results suggest that exporter status may yield higher productivity rather than being simply due to selection (see, Fryges and Wagner (2008) on the "learning-by-exporting" hypothesis). In such a case, the choice itself of dropping the export index may entail an omitted variable bias.[20]

To sum up, the evidence of the IV sectoral estimates strengthens the idea that once both productivity and wages are taken into account, the location advantages of urban agglomerations concern only firms producing services characterized by high technological intensity. In contrast, for suppliers of services active in less technologically sophisticated sectors, the productivity benefits of locating in urban areas are likely to be overwhelmed by higher labour costs. Also, estimates imply that, on average, industrial establishments operating in high-density areas either enjoy no advantage or even face a competitiveness loss (as the upward pressure on wages is higher than that on productivity).[21]

[20] The supplementary estimates obtained when excluding the dummy for exporter status are available upon request.

[21] It is worth adding that adopting different methodologies (like those based on matching or regression adjustment) to deal with selection bias yields results similar to those based on instrumental variables estimation. In particular, in the supplemental documentation we report results based on the propensity score match (after excluding the export dummy). It should be noted, however, that such a methodology requires transforming our urbanization index into a dichotomous measure. Relying on the IV estimator does not entail such a loss of information.

6 Conclusions

We have analysed the productivity–wage relation using a novel and integrated database that considers establishment information for subregional areas in the Italian region of Lombardy. The classification of the degree of urbanization we have applied is established at the European level and is thus suitable for international comparisons. In particular, we have investigated whether an urban-nonurban agglomeration divide exists and have estimated the impact of industry- and firm-specific effects.

The crucial issue is whether density matters, as is generally acknowledged when analysing productivity or wage differentials in urban and nonurban areas. However, the answer to this question is controversial since, ultimately, the productivity–wage gap is the key indicator of an urban advantage.

Density per se does not positively affect productivity, thus negatively impacting the productivity–wage gap. However, agglomeration significantly impacts productivity and wage differentials at the local level when considering the technological and knowledge-based resources characterizing the industrial mix within urban and nonurban agglomerations.

We adopt an industry classification that enables us to identify manufacturing and service activities according to technological and knowledge-intensity features, in order to better capture the potential interactions between geographical proximity and the transmission of knowledge spillovers. Moreover, the introduction of firm-level variables allows us to control the moderating role of firm-related effects, and particularly those concerning job characteristics.

When an interaction with dummy variables reflecting an industry's technological level is considered, a positive effect is operational in manufacturing. This pattern is also confirmed in services and is mainly driven by high-technology and knowledge-intensive manufacturing plants.

As concerns wages, localization in high-density areas shows a positive and significant effect on compensation in both manufacturing and services. Nevertheless, sectoral characteristics may further widen this effect when associated with agglomeration economies. When located in high-density areas, manufacturing plants in high-technology industries show an additional increase in compensation. Plants operating in high-technology services localized in high-density areas get a further wage increase, whereas the extra premium for knowledge-intensive services is not significant. In other words, wages in the KWNMS sectors are higher overall, without any premium related to localization.

The results regarding the productivity–wage gap complement this evidence. As a result of the productivity pattern, the wage gap is negative in manufacturing and services in high-density areas. Nonetheless, the gap reverts to positive when considering the interaction with high-tech industries, although this effect is more prominent in services. This fact can be rationalized on the grounds that manufacturing plants do not show a substantial gain in productivity due to agglomeration economies. The greatest gain is obtained by establishments in high-technology industries, also controlling for the quality of a firm's labour force. However, this gain is aligned with wages, thus affecting the wage gap only marginally. It is worth noting that these results are robust to specifications that take into account sorting issues, i.e. endogenous firm localization decisions, as confirmed by the complementary IV estimates.

All in all, our study suggests that density matters only if one considers specific sectors—mainly those providing technologically advanced services. By viewing the whole picture of the possible advantages of urbanization, i.e. taking into account the productivity–wage gap, this conclusion is reinforced even more as a negative gap prevails in urban areas unless interactions with industry effects are considered. Conversely, nonurban areas show a clear overall disadvantage in productivity and wages, which threatens to widen the gap with urban areas or, more generally, affect living conditions in the former.

One should also note that this territorial pattern occurs in the context of a national productivity trajectory that is largely unsatisfactory and, therefore, underlines how tackling the productivity challenge as a whole is a primary policy task.

Acknowledgments We would like to thank the Editor and two anonymous referees for their insightful comments, which have significantly contributed to improving an earlier version of this manuscript. This work results from a collaboration developed within the project "*Business demography during the great recession: patterns of resilience and productivity dispersion*", which is part of a three-year strategic research plan promoted by the Italian National Institute of Statistics (ISTAT). The empirical part of this research was carried out at the ISTAT Regional Offices in Milan and Genoa. We would like to acknowledge the support of Roberto Monducci, former Director for Economic Statistics. The views expressed in this paper are those of the authors and do not necessarily reflect the views of ISTAT. We would also like to thank participants at the AISRE Annual Conference, Lecce, Italy, 2-4 September 2020, and the workshop "*Progetti di ricerca tematica all'Istat*", 28 April 2022, for their comments and suggestions. Finally, we also wish to acknowledge the help rendered by Marco Pesce and Silvio Vitale for their assistance in data management. Maurizio Baussola gratefully acknowledges financial support from Università Cattolica del Sacro Cuore, grant D.1. 2019-2020, Dinamica della produttività del lavoro nell'economia italiana: impatti territoriali e mercato del lavoro.

Funding Open access funding provided by Università Cattolica del Sacro Cuore within the CRUI-CARE Agreement.

Appendix

See Table 9.

Table 9 Description of variables

Variable	Description	Source
Productivity	The ratio of value added to the total number of employees (log)	Istat—Frame territoriale SBS
Wage	The ratio of gross wages to the number of payroll registered employees (log)	
Gap	Productivity—wage	

References

Accetturo A, Dalmazzo A, de Blasio G (2019) Spatial equilibrium in deviations: an application to skill-premium and skill-mix heterogeneity. J Reg Sci 59(4):615–632

Ahlfeldt GM, Pietrostefani E (2019) The economic effects of density: a synthesis. J Urban Econ 111:93–107

Ahrend R, Farchy E, Kaplanis I, Lembcke A (2014) What makes cities more productive? Evidence on the role of urban governance from five OECD countries. OECD Regional Development Working Papers, No. 2014/05, OECD Publishing, Paris, https://doi.org/10.1787/5jz432cf2d8p-en.

Akerlof GA, Yellen JL (1986) Efficiency wage models of the labor market. Cambridge University Press, Cambridge

Angrist JD, Pischke J (2008) Mostly harmless econometrics: an empiricist's companion. Princeton, Princeton University Press

Arbia A (2012) Spatial data configuration in statistical analysis of regional economic and related problems, vol 14. Springer, New York

Arnaldi S, Baldi C, Filippello R, Mastrantonio L, Pacini S, Sassaroli P, Tartamella F (2016) The labour cost variables in the building of the "Frame SBS." Rivista Di Statistica Ufficiale 18(1):47–69

Baldwin JR, Beckstead D, Mark Brown W, Rigby DL (2008) Agglomeration and the geography of localization economies in Canada. Reg Stud 42(1):117–132

Bartoloni E, Baussola M (2021) Productivity and earnings at the firm-plant level: the case of Lombardy's urban and non-urban agglomerations. Spat Econ Anal 16(3):333–354

Bartoloni E, Baussola M, Bagnato L (2020) Waiting for Godot? Success or failure of firms' growth in a panel of Italian manufacturing firms. Struct Chang Econ Dyn 55:259–275

Beaudry C, Schiffauerova A (2009) Who's right, Marshall or Jacobs? The localization versus urbanization debate. Res Policy 38(2):318–337

Behrens K, Duranton G, Robert-Nicoud F (2014) Productive cities: sorting, selection and agglomeration. J Polit Econ 122(3):507–553

Bertola G, Boeri T, Cazes S (2000) Employment protection in industrialized countries: the case for new indicators. Int'l Lab Rev 139:57

Bosio, G. (2009) Temporary employment and wage gap with permanent jobs: evidence from quantile regression.

Brown S, Sessions JG (2003) Earnings, education, and fixed-term contracts. Scott J Polit Econ 50(4):492–506

Buzzacchi L., De Marco A. and Pagnini M. (2021) Agglomeration and the Italian North-South divide. Banca d'Italia Questioni di Economia e Finanza (Occasional Papers), n.637, October 2021.

Caju, P.D., E. Gautier, D. Momferatou and M. Ward-Warmedinger. (2008) Institutional features of wage bargaining in 23 European countries. *The US and Japan, Banque de France NER-E*, 228.

Card D, Cardoso AR, Heining J, Kline P (2018) Firms and labor market inequality: evidence and some theory. J Law Econ 36(S1):S13–S70

Carlino G, Kerr WR (2015) Agglomeration and innovation. Handb Region Urban Econ 5:349–404

Cataldi A, Kampelmann S, Rycx F (2011) Productivity-wage gaps among age groups: does the ICT environment matter? De Economist 159(2):193–221

Cheshire P (2019) The costs of containment: or the need to plan for urban growth. In CESifo Forum (Vol. 20, No. 03, pp. 10–14). München: ifo Institut–Leibniz-Institut für Wirtschaftsforschung an der Universität München

Table 9 (continued)

Variable	Description	Source
Internationalization	The plant belong to a firm that sells its products in the international market	
Size 10–19	The plant has a number of employees between 10 and 19	
Size 50–249	The plant has a number of employees between 20 and 49	
Size 20–49	The plant has a number of employees between 50 and 249	
Size 250–499	The plant has a number of employees between 250 and 499	
Size 500 +	The plant has 500 or more employees	
Urban high density	The plant is localized in a high-density area	Eurostat- DegUrba classification
Urban intermediate	The plant is localized in a interediate urban area	
Rural	The plant is localized in a rural area	
HT manuf	The plant belong to a high-tech manufacturing sector	Eurostat- High-tech aggregation by NACE Rev.2
MHT manuf	The plant belong to a medium–high-tech manufacturing sector	
MLT manuf	The plant belong to a medium–low-tech manufacturing sector	
LOT manuf	The plant belong to a low-tech manufacturing sector	
HITS	The plant belongs to the high-technology services sector	
KWNMS	The plant belongs to the knowledge intensive market services sector	
Other services	The plant belongs to other traditional services	
Household services	The plant belongs to the household services sector	
Aged 30–49	Share of employees in the 30–49 age class	Istat, Asia occupazione
Aged 50 +	Share of employees in the 50 + age class	
Males	Share of male employment	
Temporary contracts	Share of temporary employment	
Employment rate	Employment rate at the local labour market level	Istat

HT manuf. (NACE Rev.2 21 + 26 + 30.3 + 32.5); MHT manuf. (NACE Rev.2: 20 + 25.4 + 27 + 28 + 29 + 30–30.3 + 33); MLT manuf. (NACE Rev.2 19 + 22 + 23 + 24 + 25–25.4); LT manuf. (NACE Rev.2: 10 + 11 + 12 + 13 + 14 + 15 + 16 + 17 + 18 + 31 + 32–32.5); HITS (NACE Rev.2 53 + 58 + 60 + 61 + 62 + 63 + 72); KWNMS (NACE Rev.2: 50 + 51 + 68 + 69 + 70 + 71 + 73 + 74 + 77 + 78 + 80 + 81 + 82); other services (NACE Rev.2: 45 + 46 + 47 + 49 + 52 + 55 + 56 + 59 + 75 + 79); Household services (NACE Rev.2 85,86–88,90–93, 95–96)

Ciccone A (2002) Agglomeration effects in Europe. Eur Econ Rev 46(2):213–227
Ciccone A, Hall RE (1996) Productivity and the density of economic activity. Am Econ Rev 86:54–70
Colella F, Lalive R, Sakalli SO, Thoenig M (2019) Inference with arbitrary clustering. IZA Discussion Paper
Combes P, Duranton G (2006) Labour pooling, labour poaching, and spatial clustering. Reg Sci Urban Econ 36(1):1–28
Combes P, Duranton G, Gobillon L (2008a) Spatial wage disparities: sorting matters! J Urban Econ 63(2):723–742
Combes P-P, Duranton G, Gobillon L, Puga D, Roux S (2012) The productivity advantages of large cities: distinguishing agglomeration from firm selectio. Econometrica 80(6):2543–2594
Combes P-P, Gobillon L (2015) The empirics of agglomeration economies. In: Duranton G, V. Henderson and W. Strange (Eds.) *Handbook of Urban and Regional Economics* vol.5, Elsevier.
Combes, P., Duranton, G., Gobillon, L., Roux, S. (2008b) Estimating Agglomeration Economies with History, Geology, and Worker Effects. CEPR Discussion Papers 6728.
Comi S, Grasseni M (2012) Are temporary workers discriminated against? Evide Eur Manchester Sch 80(1):28–50
Conley TG (1999) GMM estimation with cross sectional dependence. Journal of Econometrics 92(1):1–45
Costa S, Carmine P, Vicarelli C (2017) Internationalization choices and Italian firm performance during the crisis. Small Bus Econ 48:753–769
Crépon B, Deniau N, Pérez-Duarte S (2003) Wages, productivity and worker characteristics: A French perspective: INSEE
De la Rica S (2004) Wage gaps between workers with indefinite and fixed-term contracts: the impact of firm and occupational segregation. Moneda y Crédito 219(1):43–69
De la Roca JD, Puga D (2017) Learning by working in big cities. Rev Econ Stud 84(1):106–142
Dearden L, Reed H, Van Reenen J (2006) The impact of training on productivity and wages: evidence from British panel data. Oxford Bull Econ Stat 68(4):397–421
Di Giacinto V, Gomellini M, Micucci G, Pagnini M (2014) Mapping local productivity advantages in Italy: industrial districts, cities or both? J Econ Geogr 14:365–394
Di Giacinto V, Micucci G, Tosoni A (2020) The agglomeration of knowledge-intensive business services firms. Ann Reg Sci 65(3):557–590
Duranton G, Puga D (2004) Micro-foundations of urban agglomeration economies. In: Anonymous (Ed) Handbook of Regional and Urban economics (pp 2063–2117), Elsevier.
Ellison G, Glaeser EL, Kerr WR (2010) What causes industry agglomeration? Evidence from coagglomeration patterns. Am Econ Rev 100(3):1195–1213
European Union, UN-Habitat, OECD and World Bank (2021) Applying the Degree of Urbanisation: A Methodological Manual to Define Cities, Towns and Rural Areas for International Comparisons (2021 Edition)
Foster L, Haltiwanger J, Syverson C (2008) Reallocation, firm turnover, and efficiency: selection on productivity or profitability? Am Econ Rev 98(1):394–425
Fryges H, Wagner J (2008) Exports and productivity growth: first evidence from a continuous treatment approach. Rev World Econ/Welwirtschaftiliches Archiv 144(4):695–722
Garnero A, Kampelmann S, Rycx F (2014) Part-time work, wages, and productivity: evidence from Belgian matched panel data. ILR Rev 67(3):926–954
Gaubert C (2018) Firm sorting and agglomeration. Am Econ Rev 108(11):3117–3153
Hægeland T, Klette TJ, Salvanes KG (1999) Declining returns to education in Norway? Comparing estimates across cohorts, sectors and over time. Scand J Econ 101(4):555–576
Heckman J (1979) Sample selection bias as a specification error. Econometrica 47:153–161
Hellerstein JK, Neumark D (1999) Sex, wages, and productivity: an empirical analysis of Israeli firm-level data. Int Econ Rev 40(1):95–123
Hellerstein JK, Neumark D, Troske KR (1999) Wages, productivity, and worker characteristics: evidence from plant-level production functions and wage equations. J Law Econ 17(3):409–446
Hellerstein JK, Neumark D (2009) Production Function and Wage Equation Estimation with Heterogeneous Labor: Evidence from a New Matched Employer-Employee Data Set: University of Chicago Press
Henderson JV (1974) The sizes and types of cities. Am Econ Rev 64(4):640–656
Henderson JV (2003) Marshall's scale economies. J Urban Econ 53(1):1–28
Istat, (2006) Distretti industriali e sistemi locali del lavoro. Istat, Rome

Jennen M, Verwijmeren P (2010) Agglomeration effects and financial performance. Urban Stud 47(12):2683–2703
Kampelmann S, Rycx F, Saks Y, Tojerow I (2018) Does education raise productivity and wages equally? The moderating role of age and gender. IZA J Labor Econ 7(1):1
Konings J, Vanormelingen S (2015) The impact of training on productivity and wages: firm-level evidence. Rev Econ Stat 97(2):485–497
Martin P, Mayer T, Mayneris F (2011) Spatial concentration and plant-level productivity in France. J Urban Econ 69(2):182–195
McGuinness S, Kelly E, P.J. O'CONNELL. (2010) The impact of wage bargaining regime on firm-level competitiveness and wage inequality: the case of Ireland. Ind Relat J Econ Soc 49(4):593–615
Melo P, Graham D, Noland R (2009) A meta-analysis of estimates of urban agglomeration economies. Reg Sci Urban Econ 39(3):332–342
OECD (2018) Decoupling of wages from productivity: what implications for public policies?, in *OECD Economic Outlook*, Volume 2018 Issue 2. OECD Publishing, Paris,. https://doi.org/10.1787/eco_outlook-v2018-2-3-en
OECD (2019) Digitalisation and productivity: A story of complementarities, in *OECD Economic Outlook*, Volume 2019 Issue 1. OECD Publishing, Paris,. https://doi.org/10.1787/5713bd7d-en
Oi WY (1962) Labor as a quasi-fixed factor. J Polit Econ 70(6):538–555
Rice P, Venables AJ, Patacchini E (2006) Spatial determinants of productivity: analysis for the regions of Great Britain. Reg Sci Urban Econ 36(6):727–752
Rosenthal SS, Strange WC (2001) The determinants of agglomeration. J Urban Econ 50(2):191–229
Rosenthal SS, Strange WC (2004) Evidence on the nature and sources of agglomeration economies. Elsevier, In Anonymous Handbook of regional and urban economics, pp 2119–2171
Rusinek M, Tojerow I (2014) The regional dimension of collective wage bargaining: the case of Belgium. Reg Stud 48(2):301–317
Shapiro C, Stiglitz JE (1984) Equilibrium unemployment as a worker discipline device. Am Econ Rev 74(3):433–444
Sharpe A, Uguccioni J (2017) Decomposing the productivity-wage nexus in selected OECD countries, 1986–2013. Int Product Monit 32:25
Stavropoulos S, Skuras D (2016) Firm profitability and agglomeration economies: an elusive relationship. Tijdschr Econ Soc Geogr 107(1):66–80
Stock J, Yogo M (2005) Testing for Weak Instruments in Linear IV regression. In: Andrews DWK (ed) Identification and inference for econometric models, Chapter 5. New York, Cambridge University Press, pp 80–108
Sweet M (2014) Traffic congestion's economic impacts: Evidence from US metropolitan regions. Urban Stud 51(10):2088–2110
Syverson C (2011) What determines productivity? J Econ Lit 49(2):326–365
Van Biesebroeck J (2014) How tight is the link between wages and productivity? A survey of the literature: ILO
Van Ours JC, Stoeldraijer L (2011) Age, wage and productivity in Dutch manufacturing. De Economist 159(2):113–137
Webber D, Curry N, Plumridge A (2009) Business productivity and area productivity in rural England. Reg Stud 43(5):661–675
Wooldridge JM (2010) Econometric analysis of cross section and panel data. MIT Press, New York
Zheng X (2001) Determinants of agglomeration economies and diseconomies: empirical evidence from Tokyo. Socioecon Plann Sci 35(2):131–144