CESifo WORKING PAPERS

7253
2018

Original Version: September 2018
This Version: November 2019

Urban Wage Premia, Cost of Living, and Collective Bargaining

Marianna Belloc, Paolo Naticchioni, Claudia Vittori

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 7253*
Category 4: Labour Markets

Urban Wage Premia, Cost of Living, and Collective Bargaining

Abstract

In this paper, we estimate nominal and real (in temporal and spatial terms) urban wage premia (UWP) in Italy, with its economy characterized by the interplay between collective wage bargaining and spatial heterogeneity in the cost of living. Our dataset for the 2005-2015 period includes, for workers' characteristics, unique administrative data provided by the Italian Social Security Institute and, for the local CPI computation, housing prices detailed at a fine level of spatial aggregation delivered by the Italian Revenue Agency. For employees covered by collective bargaining, we find a zero UWP in nominal terms and a negative and non-negligible UWP in real terms (-2.6% when all controls are included). To capture the role played by centralized wage settings, we consider various groups of self-employed workers, who are not covered by national labour agreements, while living in the same locations and enjoying the same amenities as employees. We find that, differently from employees, self-employed workers enjoy a positive UWP in nominal terms, and do not suffer from urban real wage penalties. Results hold under a large array of robustness checks.

JEL-Codes: R120, R310, J310.

Keywords: urban wage premium, cost of living, wage setting.

Marianna Belloc
Department of Economics and Law
Sapienza University of Rome / Italy
marianna.belloc@uniroma1.it

*Paolo Naticchioni**
Department of Political Science
University of Roma Tre & INPS / Italy
p.naticchioni@gmail.com

Claudia Vittori
National Institute for Public Policy Analysis (INAPP) / Rome / Italy
c.vittori.ext@inapp.org

*corresponding author

November 2019
We thank Gaetano Basso, Tito Boeri, Simon Burgess, Lorenzo Cappellari, David Card, Luca Citino, Pierre-Philippe Combes, Edoardo Di Porto, Gilles Duranton, Maurizio Franzini, Pietro Garibaldi, Andrea Garnero, Laurent Gobillon, Paul Gregg, Andrea Lamorgese, Claudio Lucifora, Michele Raitano, Paolo Sestito, Roberto Torrini, Eliana Viviano, Josef Zweimüller, and participants in seminars at IZA, Bank of Italy, INPS, MILLS-University of Milan, University of Ancona, University of Bristol, in yearly conferences of AIEL and ECINEQ, and in the workshop on "Spatial Dimension of Inequality" at ZEW for comments and suggestions. Data access was provided through the VisitINPS Scholars programme, whose support we gratefully acknowledge. Paolo Naticchioni is also grateful to VisitINPS programme for financial support. Usual disclaimers apply. The views expressed in the article are those of the authors and do not involve the respective institutions.

1 Introduction

This paper contributes to the literature on the urban wage premium (UWP, hereafter), a well-established issue in urban economics (Marshall, 1890; Rosenthal and Strange, 2004; Duranton and Puga, 2004 among others). To this end, we emphasize the role played by the interplay of two labour market features, somewhat neglected by previous research: collective wage bargaining and spatial variability in the cost of living.

Collective bargaining, the first aspect, plays a crucial role in the labour markets of many developed countries: in the OECD economies, 155 million workers are covered by collective agreements concluded at the national, regional, sectoral, occupational, or firm level. Its functioning is set by the International Labour Organization (ILO Convention No. 98) and its impact on economic performance, such as recruiting and dismissal strategies, wage structure and inequality, has been analyzed in an extensive economic literature (see, e.g., OECD, 2017). The role of collective bargaining schemes in shaping firms' responses to economic crises has been examined by international institutions such as the ECB, the European Commission, the IMF, and the OECD.[1] However, to the best of our knowledge, the relevant literature has not investigated the effects of collective bargaining on the spatial distribution of real wages and on the estimation of the UWP. This might depend on the fact that previous research has mainly focused on the US, a country characterized by flexible wage setting institutions, or on the lack of proper data when European countries have been considered.

The spatial variability in the cost of living, the second aspect, has also been largely neglected in the existing literature for both the US and Europe, again due to the paucity of data on prices at the local level. This is a serious limitation because price levels vary greatly across regions of the same country and between urban and rural areas.

Collective bargaining tends to make wages uniform along the space dimension, while the cost of living is highly heterogeneous across locations. Hence, the interplay between these two phenomena might lead to differences in UWP between real and nominal terms. This is likely to affect location choices and sorting of firms and workers and, in turn, have an impact on firms' dynamics and economic growth, especially when cities are characterized by urbanization externalities.

As far as we know, there is only one paper, by Ichino et al. (2019), combining collective bargaining and local variability in cost of living, from both the theoretical and empirical points of

[1] Mario Draghi, in an official speech in 2015, underlined that "firms with flexibility at the plant-level have reduced employment less during the crisis than those bound by centralised wage bargaining agreements, partly because they have been more able to adjust wages to economic conditions", a statement based on ECB research (Draghi, 2015, Di Mauro and Ronchi, 2016). In its 2016, 2017, and 2018 Country Reports for Italy the European Commission also expressed concern regarding collective bargaining. Similar conclusions emerge from IMF (2016) and from independent observers such as Ichino et al. (2019).

view. One of the predictions of this work is that, in equilibrium and in the presence of collective bargaining (equal nominal wages across locations), higher productivity areas are characterized by higher costs of living, lower real wages, and lower unemployment. This prediction is confirmed by the empirical analysis applied to macro-regional unbalances, comparing the Italian and the German cases. For the Italian case, conditionally on being employed, an individual working in the lower productivity South has higher real wages than the analogous individual working in the (higher productivity) North. In Germany, where the wage setting system is more decentralized, the authors do not find real wage penalties when moving from the lower productivity East to the higher productivity West.

In this paper, we also focus on Italy, but consider the agglomeration dimension (rural versus urban areas) rather than comparisons across macro-regions. Agglomeration is a key feature in investigating the variance of wages: in Italy 95.3% of the variance on nominal wages and 95.7% of the variance of real wages is within regions (VisitINPS data for employees in 2005, see Section 4).[2]

Italy represents a perfect case to study in this context. Differently from the US, where high productivity areas tend to bear high costs of living and to enjoy high salaries (see Hornbeck and Moretti, 2018), in Italy substantial differences in terms of costs of living might not be associated with appropriate wage differentials. Indeed, as in many other European countries, a crucial component of Italian wages is set at the industry level and, hence, is uniform in the spatial dimension (between the North and the South of the country and across big and small cities). The Italian collective agreements, which in addition to the wage also set a wide range of working conditions, are characterized, *de facto*, by an *erga omnes* implementation, leading to a virtual 100% coverage (Garnero, 2018). The two-tier system also allows a second-level bargaining at the firm level or within local industrial clusters. It is subject to the *in melius* or favourability principle: workers' wages and conditions cannot be worse than those agreed at the industry level. This decentralized negotiation, however, only plays a minor role and is limited to few large firms, as well documented, for instance, by Boeri (2014), Brandolini et al. (2004), Ichino et al. (2019), and Devicienti et al. (2019).[3]

Another reason to consider Italy is the type and the quality of the data we can employ. First, for workers, we can make use of unique administrative archives from the Italian National Social

[2]Note that an additional advantage of investigating the within rather than the between regional dimension is that it allows us to neglect structural and cultural differences across units of observation that are instead relevant when comparing the South and the North of Italy, as in Ichino et al. (2019). By contrast, within-region local structural and cultural heterogeneity can be assumed as negligible.

[3]See Appendix A for a more detailed description of the Italian collective bargaining system and Section 8 for some evidence on the incidence of the second level of collective bargaining.

Security Institute (INPS), issued through the VisitINPS programme. These data cover the universe of Italian workers, including both an employer-employee dataset, with detailed individual and firm characteristics, and datasets on various groups of self-employed workers. Moreover, in most of the INPS archives it is possible to recover the information on the municipality where the job is located. Second, we make use of a very rich database on housing prices (*Osservatorio del Mercato Immobiliare*, OMI hereafter) provided by Italian Revenue Agency (2015), which delivers information on housing transactions detailed at the municipality level. Our dataset covers the 2005-2015 period.

Our empirical analysis is based on the well-established two-stage approach adopted by Combes et al. (2008) and Combes et al. (2010).[4] Using this approach, we aim at identifying the impact of our covariate of interest, the (log) of population density of the local labour market (LLM) where the job is located, on individual nominal and real wages.[5] Individual real wages are computed by deflating nominal wages by a local consumer price index (CPI) obtained using local variation in housing prices, after removing composition effects, according to a procedure developed by Moretti (2013).[6]

When considering nominal wages, our results confirm the presence of a negligible nominal UWP. The wage elasticity with respect to population density is 2.2% when no controls are included, which turns out to be rather small in international comparison. Once individual and firm level controls are introduced, the elasticity drops to zero and becomes not statistically significant, i.e. the positive estimated UWP is only driven by the employment and firm composition along the density dimension. The estimate of the UWP in real terms is, instead, negative and non-negligible: urban workers suffer a penalty in terms of real wages. In the specification that includes all possible controls for individuals' and firms' characteristics, the wage elasticity with respect to population density is -2.6%.

Consistently with the recent relevant literature, we address sorting of workers and firms into denser areas. While sorting of workers is an established fact in urban economics (Mion and Naticchioni, 2009, Combes et al., 2008), the importance of firm-fixed effects in explaining wage differentials is more recent (Card et al., 2013; Dauth et al., 2018). Our results show that sorting of individuals only slightly dampens the UWP estimate, and adding firm-fixed effects does not substantially change it. We also check the robustness of our results with respect to endogeneity issues due to reverse causality problems possibly associated with productivity shocks occurring in a given

[4]For a discussion of the advantages of this strategy with respect to reduced form regressions see also Combes et al. (2011), and Combes et al. (2019).

[5]Local labour markets are self-contained areas defined with respect to daily commuting patterns. According to the Italian National Institute of Statistics (Istat) classification, 611 LLM were identified in 2011.

[6]From now on, we will refer to real wages as wages deflated by considering both time and spatial variation in the cost of living.

location which may affect wages and, in turn, worker's location choices. This is ensured by implementing a well-established instrumental variable approach, which uses historical series of the Italian population as an instrument (as by Ciccone and Hall, 1996; Mion and Naticchioni, 2009; Combes et al., 2011).

Following the theoretical prediction by Ichino et al. (2019), we also verify whether the estimated urban real wage penalty reflects a higher probability of being employed in more densely populated areas. Our qualitative conclusions remain substantially unchanged when the unemployment rate at the local level is included in the regressions, suggesting that in an agglomeration dimension the trade-off between real wages and unemployment rate does not play a major role.

The urban economics literature suggests that the presence of urban amenities or the idiosyncratic preferences for locations are possible explanations for the existence of real wage penalties (Moretti, 2011). In our analysis, we investigate the distinct role played by collective bargaining. To isolate this effect, we consider various groups of self-employed workers, because, differently from employees, they are not subject to centralized wage settings, while living in the same locations and enjoying the same local amenities as employees.

Our preferred self-employed group is that of independent contractors (so-called *collaborazioni* or *parasubordinati*), since, as for employees, they are always closely associated with a firm, such as external staff and/or consultants, and may be either skilled or unskilled. But, unlike those of employees, their wages are the result of market forces and individual bargaining between employer and worker, with no institutional constraints. We find that the estimated UWP for this group of workers, including controls for workers' and firms' characteristics, is equal to 5.4% in nominal terms and to 2.6% in real terms, always statistically significant at least at the 5% level. Hence, urban independent contractors are able to offset the higher cost of living with higher wages. Interestingly, once worker fixed effects are included, a more pronounced drop in the UWP is detected for independent contractors with respect to employees, with a negative UWP that is however not statistically different from zero (results are confirmed by adding firm fixed effects and by instrumental variable estimation). This evidence of spatial sorting represents another interesting finding of the paper: earnings for self-employed workers react much more in the space dimension, providing higher incentives for individuals to sort into agglomerated areas.

To sum up, our findings suggest that there are significant differences in estimated UWP between employees (covered by collective bargaining) and self-employed workers: in particular, employees suffer a negative UWP in real terms; the penalty disappears when self-employed workers are considered. These results hold under a large array of robustness checks, such as considering subsamples of workers, considering alternative measures of earnings, agglomeration, and local prices, changing the spatial unit of observation, and adopting different time intervals.[7]

As an additional check, the differences in UWP estimates between employees and self-employed workers are widely confirmed when considering the standard self-employed workers, who are not subject to collective bargaining and not associated with a firm (business consultants, lawyers, physicians, architects, journalists, surveyors, and accountants). Overall, UWP estimates for this group of workers are also positive and statistically significant in both nominal and real terms. As an additional finding, we claim that, to the best of our knowledge, also the rigorous UWP estimate comparing employees and various groups of self-employed workers is something rather neglected in the literature, and represents another valuable contribution of the paper.

Finally, we investigate whether previous results change along the wage distribution, by employing unconditional quantile regressions (UQR). On the one hand, high-skilled employees might be able to extract higher wages in decentralized negotiations, and this could be more likely to occur in agglomerated areas where firms are more productive. On the other hand, if unions at the firm/local level were mostly interested in the welfare of low-wage workers, second-level bargaining could increase wages for unskilled workers (Matano and Naticchioni, 2017). Estimation output shows that this is not the case, since UWP differences between employees and self-employed workers do not differ much along different percentiles of the wage distribution.

The reminder of the paper is organized as follows. In the next section, we briefly discuss the theoretical and empirical literature relevant to our analysis. Section 3 describes the data and the variables used in the empirical investigation, while Section 4 offers some descriptive evidence. Section 5 presents the econometric strategy. Sections 6 and 7 illustrate results, respectively, for employees and independent contractors. Section 8 shows the robustness checks, whereas Section 9 draws concluding remarks.

2 Related literature

The theoretical literature has identified various possible explanations for the presence of a (mainly nominal) wage premium to urban workers. First, the so-called urbanization externalities are the

[7] The identification of the distinct role of collective bargaining is based on the assumption that employees and self-employed workers live in the same locations, enjoy the same amenities and share the same idiosyncratic preferences for locations. Nonetheless, the way such factors enter the utility functions of, respectively, employees and self-employed workers might have a role in explaining the different spatial equilibria for the two groups. For instance, the fact that real UWP are negative and substantial for employees suggests that employees with strong preferences for urban amenities are willing to suffer a wage penalty to live in cities. A different equilibrium applies to self-employed workers, for whom weaker preferences for urban facilities are required, given the higher UWP, to induce the marginal worker to move to a city.

result of efficiency gains and cost savings for firms located in areas of dense economic activity due to proximity to consumers and suppliers and to knowledge and technology spillovers among firms (Marshall, 1890; Glaeser, 1998; Kim, 1987; Ciccone and Hall, 1996). Second, learning mechanisms are related to the fact that in cities human capital accumulation tends to be faster than in rural areas (Shapiro, 2006; Glaeser, 1999; Moretti, 2004). Third, matching effects could arise by the fact that location in cities enhances the probability of a better match between workers and firms (Zenou, 2009; Kim, 1990), this probability increasing with the time spent in cities (Yankow, 2006). Finally, sorting effects imply that the best workers and firms tend to have higher probabilities of locating in urban areas (Combes et al., 2008; Mion and Naticchioni, 2009; De La Roca and Puga, 2017.

As already mentioned, due to the paucity of data on prices at the local level, most of the existing empirical literature has analyzed the UWP in nominal terms, and hence assuming a uniform cost of living across locations. But, for many countries, this assumption is unrealistic, due to the considerable variability in the cost of living across regions in the same country, between cities and countryside, and across groups of municipalities of different sizes. There are few exceptions in the literature. For the US, some papers have shown, mainly with descriptive evidence, that when taking into account the spatial heterogeneity in the cost of living the UWP decreases or even tends to zero. Dumond et al. (1999), employing the cost of living indexes provided by the American Chamber of Commerce Research Association (ACCRA), estimate the real wage differentials across cities of different sizes. Glaeser and Mare (2001), also using ACCRA indexes, document that the UWP declines markedly in real spatial terms, as also confirmed by Yankow (2006). However, the ACCRA indexes entail a number of practical difficulties and limitations (they include only six components of the cost of living, are collected for metropolitan areas but not for rural areas, and are not available for all metropolitan areas).

Also considering the US economy, Moretti (2013) derives a finer local consumer price index to study the evolution of real wage inequalities between college and high school graduates. His methodology consists in exploiting variation in housing costs across metropolitan areas to compute the local CPI that is used to deflate wages. In his paper, however, the author does not investigate agglomeration dynamics and does not address the issue of collective bargaining. [8]

As for Europe, only Blien et al. (2009) compare nominal and real wage differentials between cities and rural areas in Western Germany making use of estimated price levels at the regional level.

[8] Using this price index by Moretti (2013), Hornbeck and Moretti (2018) find that when a city experiences productivity gains in manufacturing, there are substantial local increases in employment and average earnings. Furthermore, they estimate that 38% of the overall increase in workers purchasing power takes place outside cities directly affected by local TFP growth, and this indirect effects are substantially greater for more-skilled workers.

They find that, when controlling for a large set of variables, the real urban premium disappears. However, their paper, too, fails to explicitly consider the role of collective bargaining.

As far as we know, the only attempt to investigate the interaction between spatial heterogeneity in the cost of living and collective bargaining in the determination of the UWP is offered by Ichino et al. (2019). Their theoretical model is used to derive a series of implications that are empirically confirmed for Italy and Germany. First, where nominal wages cannot fully adjust, due to stronger centralized wage agreements, provinces with low productivity should have higher non-employment rates; this is indeed true for Italy, where collective agreement is more binding, while being less relevant for Germany. Second, for both countries, a positive relationship is expected (and empirically found) between productivity and housing costs. As a consequence, where the nominal wage setting is more centralized, a negative relationship exists between real wages and value added (Italy); this relationship disappears when nominal wage settings are more flexible (Germany). Thus, in the presence of collective agreements, the winners are workers employed in low productivity locations and house owners in high productivity locations. Although Ichino et al. (2019)'s paper is strongly related to our work, it does not deal with agglomeration issues, being focused on comparisons across macro-regions.

Our paper also contributes to the debate on the degree of centralization of collective bargaining and its effects on labour market performance (OECD, 1994). More recently, some papers have questioned the superiority of two-tier multi-employer systems with respect to centralized ones in coping with the impact of the Great Recession (see Di Mauro and Ronchi, 2016; Draghi, 2015; and country reports by European Commission and IMF).

Finally, one more strand of the literature relevant to this paper concerns the role of urban amenities, which may be unbalanced in the space dimension (Roback, 1982 and Albouy, 2016). Albouy (2016) reports adjusted amenity-value estimates for US cities and shows that more valuable cities are those on the coast, with a sunny and mild weather, populated by educated citizens, and large. Preferences for artificial amenities are in general weak, but are relatively stronger for culture, restaurants, and clean air. Similar arguments apply to the case of Italy. One might argue that larger cities are characterized by well-known monuments, beautiful city centres, and attractive entertainment services (restaurants, theaters, cinemas). Besides, one might expect the quality of certain public goods, such as education and health care, to be higher in cities than in rural areas. Hence, workers living in cities could be willing to pay a cost to enjoy better amenities. This finding emerges in standard models of agglomeration economies, as for instance shown by Moretti (2011). Similar arguments could apply when taking into account idiosyncratic preferences for locations.

3 Data and variables' description

In order to estimate UWP in Italy, we exploit three different set of information: the first concerns workers' wages and personal characteristics as well as firm level variables; the second covers prices at the local level; and the third is about urban agglomeration variables.

3.1 Individual and firm level variables

Individual and firm level data are drawn from administrative employer-employee datasets collected by INPS and, more specifically, obtained through the VisitINPS programme, which allows selected researchers to access richer data, with respect to previous INPS data releases, on the universe of Italian workers.[9] These can be distinguished in two macro-groups: employees and self-employed workers. The former are employees subject to national collective labour agreements (additional details on the role of collective bargaining in Italy and its functioning are provided in Appendix A). The latter are not subject to centralized wage negotiation and comprise two heterogeneous groups of individuals: independent contractors (so-called *collaborazioni*, such as external staff and/or consultants) and standard self-employed workers belonging to professional associations (those with tertiary degree education, such as business consultants, lawyers, physicians, architects, and those with upper secondary education, such as journalists, surveyors, accountants). The data structure for these groups is similar to that for employees, and the few differences will be highlighted in the relevant Sections, respectively, 7 and 8.

For employees, at the individual level, we can exploit a wide range of variables: weekly wages, age, gender, occupation, and information on the type of contract (part-time versus full-time, temporary versus permanent). In the case of multiple contracts associated with the same individual, we keep the contract which pays the highest earnings. At the firm level, we have the municipality where the job is located (since 2005), firm size, and type of National Collective Labour Agreement (NCLA, hereafter).

We focus on males only, to overcome issues with labour market participation. Our sample reaches around around 77 million observations of males aged 15-64 over the 2005-2015 period for a total of around 10 million workers.[10]

[9] For further information about the program see http://www.inps.it/nuovoportaleinps/default.aspx?itemdir=47212.

[10] We crop outliers in the wage distribution: for each year, we drop observations in the two 0.5% tails. Furthermore, we drop municipalities that represent outliers with regard to prices, such as particularly touristic places (e.g. Cortina D'Ampezzo, Capri, etc.). Finally, we do not consider the few cases regarding national contracts with less than 500 employees in the overall 2005-2015 period, for which identification of the national contract dummy would be imprecise.

We employ both nominal and real wages. The former is the gross weekly wage at the individual level (full-time equivalent wage for part-timers);[11] the latter is defined as the nominal wage deflated by the local CPI disaggregated at the local level computed as described in Subsection 3.2 (further details are also provided in Appendix C). In our main analysis, the spatial unit of observation is the LLM (in Italy there are 611 LLM and 8,000 municipalities). This choice is guided by the fact that we have knowledge of the municipality where people work only, but not of the municipality where they live. The workers' cost of living depends on prices of the municipality of work, on prices of that of residence, as well as on prices of the surrounding municipalities where, for instance, they go to restaurants and cinemas, and use health and educational services. Hence, by employing the municipality of work as statistical unit of observation would lead us to estimate the local cost of living imprecisely. On the contrary, by adopting the LLM where the job is located as spatial unit of observation, we maximize the probability that individuals work and live in the same spatial unit, in such a way assigning to the worker a more appropriate measure of local prices.

3.2 Local consumer price index

To obtain a CPI that varies at the local (municipality or LLM) level, we combine price data available from Istat for the capital cities of 80 out of 103 Italian provinces (Istat, 2016b) and detailed information on housing prices at the municipality level (Italian Revenue Agency, 2015).[12]

Our approach follows Moretti (2013)'s two-step methodology (for details, see Appendix C) that exploits two empirical facts. First, housing prices are one of the main drivers of the spatial variation in the local cost of living. Second, the costs of housing have, in addition to their direct role in the Italian families' consumption basket (cost of having an accommodation and maintaining it), also an indirect attraction effect on the prices of the other goods and services (e.g., a slice of pizza or a haircut are likely to be more expensive in Rome than in a rural village, because operating a pizza restaurant or a barbers shop in Rome is more expensive); hence, the dynamics of non-housing prices can be split into two components, one that varies with housing prices and one orthogonal to it.

In the first step of the procedure, the "adjusted" weight of housing costs, λ, which includes both direct and indirect (attraction) effects, is estimated by computing the correlation between Istat CPI and a housing price index at the provincial level:

[11] INPS archives provide information concerning the individual percentage of part time with respect of full time workers.

[12] A province is an intermediate administrative division between a municipality and a region.

$$CPI_{pt} = d_p + \lambda HPI_{pt} + \varepsilon_{pt}, \tag{1}$$

where d_p are the province fixed effects, CPI_{pt} is Istat provincial consumer price index, with base year t=2005, and HPI_{pt} is the province average housing price index, also with base year t=2005. For the 2004-2015 period, $\hat{\lambda}$ turns out to be equal to 33% (while the official direct weight attached to housing costs by Istat in the families consumption basket is 10% in the same period).

Then, in the second step, the CPI at the local level is computed as:

$$\hat{CPI}_{mt} = \hat{\lambda} HPI_{mt} + (1 - \hat{\lambda}) NHPI_t, \tag{2}$$

where HPI_{mt} is the housing price index in municipality m and in year t; $NHPI_t$ is the component of the national non-housing price index that is orthogonal to the dynamics of housing prices (that, consistently with Ichino et al., 2019, is assumed uniform across space). Both HPI_{mt} and $NHPI_t$ have base year in 2005. The CPI at the LLM level is obtained as a population-weighted average of the municipality index.

Housing price indexes and non-housing price indexes in (2) are drawn from two different sources. The former (HPI_{mt}) are obtained from the OMI (*Osservatorio del Mercato Immobiliare*) dataset, collected by Italian Revenue Agency (2015).[13] This very rich source of information provides house property prices detailed by semester (January-June/July-December), city district (central, semi-central, peripheral, suburban, and extra-urban), type of house (cottage, expensive house, standard house, cheap house, typical house), and house status (good, standard, poor). The housing price data employed in our CPI computation procedure are then purged of composition effects and converted into index numbers with base year 2005, as explained in greater detail in Appendix C. A discussion on the Italian housing market and the role of public housing policies is also provided in Appendix B. The latter ($NHPI_t$) is computed using Istat (2016b) data and is obtained as the weighted average of the national prices included in Istat CPI at the one-digit COICOP classification level, excluding housing, weights being taken from the Istat weighting scheme.[14]

[13]The adoption of house transaction prices in our analysis is supported by the fact that the house ownership rate is very high in Italy. Eurostat (2019) reports it was about 73% in the whole period between 2005 and 2015. Nonetheless, for robustness, we also estimate λ by employing Istat data on housing rents and other housing costs (Istat, 2016b) rather than house property prices as a measure of housing costs. In this exercise, the estimated λ turns out to be 54%, which is even greater than the "adjusted" weight found for housing property prices.

[14]Istat price data are aggregated according to COICOP (Classification of Individual Consumption According to Purpose) industrial classification, according to which the one-digit level identifies macro-categories of goods and services: food and non-alcoholic drinks (01), alcoholic drinks and tobacco (02), apparel and shoes (03), housing (04), and so on. Each category is weighted according to the Istat weighting scheme which reflects the relative importance of the various goods and services in the families' consumption basket. In particular, we use the index computed for blue and white-collar households (FOI). The results would not change were we to use the index for the whole nation (NIC)

Our local CPI index is not an official measure of the local cost of living, and its variability depends mostly on the housing cost heterogeneity in the space dimension. It could be argued that this index represents an imprecise measure of local prices and could either underestimate or overestimate the actual unobserved local CPI. On the one hand, it may be that individuals working in cities react to the higher cost of living consuming less than (or in different ways from what) they would do if working in rural areas (e.g. living in a more modest flat). In this case, our local CPI might overestimate the actual cost of living. On the other hand, Combes et al. (2019) point out that the elasticity of urban costs increases more than proportionally with city population. Should the CPI measure not capture this non-linearity, it would underestimate the real cost of living in cities. Hence, it is not clear whether our CPI local index could represent a biased estimate of the true one, and the possible direction of the bias. For these reasons, we carry out several robustness checks on the local CPI. Among other things (see Appendix D), we also employ, as an alternative measure of prices, Istat official absolute poverty thresholds (Istat, 2017).

3.3 Urban agglomeration variables

Our proxy of urban agglomeration is the population density at the local level: LLM in the baseline estimates and municipality in a robustness check. This is defined as population per square kilometre (in line with, among others: Combes, 2000, Mion and Naticchioni, 2009, Matano and Naticchioni, 2012, and De Blasio and Di Addario, 2005). Alternatively, we also employ as robustness check the employment density (see Section 8 and Appendix D). Data on municipality population and city size in square kilometres are from Istat (2016c). Finally, in the instrumental variable estimation (discussed in Section 5), we also employ population back in 1871 at the LLM level, provided by Istat (2016c).

4 Descriptive analysis: local CPI, prices, and wages

This section presents some descriptive evidence on the three key variables of the paper - local CPI, agglomeration density, and (nominal and real) wages - and how they relate to one other.

To start with, Figure 1 displays local (LLM level) CPI on the vertical axis and (log) population density on the horizontal axis, whereas local population size is represented by the dimensions of the circles. The evidence gathered is clear: the CPI increases with population density and larger cities (greater circles' dimension) are characterized by higher CPI.

instead.

Figure 1 about here.

Figure 2 shows the distribution of local CPIs along the Italian map by LLM, using quintiles of the local CPI indexes. As can be seen, the local CPI is highly heterogeneous in the spatial dimension, with a clear North-South divide, local CPI being higher in the North than in the South. This is a well-known phenomenon in Italy, addressed by Ichino et al. (2019). Nonetheless, a clear agglomeration pattern also emerges: in highly agglomerated areas, the local CPI is greater and included in the highest quintile. This occurs in both the North and the South and, namely, in the LLMs of Rome, Naples, Milan, Bologna, Florence, Venice, Palermo, Cagliari, and Bari.

Figure 2 about here.

Figure 3 shows the spatial distribution of population and employment density across LLMs, also in quintiles. From this figure, we can see that the distributions of the two variables are very similar. It follows that using either of the two indexes as agglomeration measure is likely to lead to similar results (as we show in Appendix D).

Figure 3 about here.

We now move to the descriptive analysis of our main wage variables, weekly nominal and real wages of employees.

Table 1 shows the within-between variance decomposition of nominal wages in 2005. It turns out that, both for nominal and real wages, around 95% of the variance is due to a within dimension, regardless of whether the within dimension refers to macro-regions (5), regions (20), or provinces (103). This evidence brings out the importance of analyzing the variability of wages across local areas (municipalities or LLMs).

Table 2 summarizes the distribution of nominal and real wages by quintiles of population density in 2005. It is worth noting that, while the overall means of nominal and real wages are the same, the distributions by quintiles are strikingly different. Nominal wages are increasing along the quintile distribution, as expected, while real wages are much more compressed: low population density areas are characterized by higher real than nominal wages, while the opposite holds for areas of high population density. For instance, in the top quintile, the weekly wage amounts to 511 euro in nominal terms and 443 euros in real terms. By contrast, in the lowest quintile real wages are 446 euro while nominal wages are 401 euros.

Table 1 about here.

Table 2 about here.

Figure 4 depicts the distribution of nominal (left panel) and real (right panel) wages in 2005. The North-South divide is clear-cut as far as nominal wages are concerned. Again, we find a marked agglomeration pattern: higher wages tend to characterize LLMs where big cities are located (Milan, Rome, Naples, Bari, Palermo, Florence, Bologna, etc.). On turning to real wages, the North-South divide is less evident and higher wages (darker areas) are more often located in the Center and in the South than when nominal wages are considered. Interestingly, LLMs including large cites tend to display much lower real than nominal wages, suggesting a reduction in purchasing power for workers in highly agglomerated areas (for instance, Rome, Naples and Florence).

Figure 4 about here.

5 The econometric strategy

To estimate the wage elasticity to urbanization, we employ a well-established two-stage approach, as in Combes et al. (2008) and Combes et al. (2011), that allows us to take into account local level unobserved heterogeneity.

In the first stage, we estimate:

$$log(w_{it}) = \alpha_{a(it)t} + \eta_1 X_{it} + \eta_2 Z_{j(it)t} + \theta_i + \theta_{j(it)} + \varepsilon_{it} \quad (3)$$

where $log(w_{it})$ is the logarithm of the real or nominal wage of individual i at time t. X_{it} is the vector of individual controls, which includes, for worker i at time t: having a part time contract, having a fixed-term contract, occupation dummies (blue collar, white collar, manager, apprentice, executive), and dummies for age groups. $Z_{j(it)t}$ is the vector of firm-level controls, which includes, for firm j where worker i works at time t, firm size (in log) and dummies for the associated NCLA (around 250 dummies), capturing the enforcement of the centralized national wage setting.[15] $\alpha_{a(it)t}$ are area-year effects, where $a(it)$ stands for the area where individual i works at time t; θ_i and $\theta_{j(it)}$ are worker and firm fixed effects included only in some specifications, as explained below.

In the second stage, we then regress the area-year effects estimated in the first stage, $\widehat{\alpha}_{at}$, on a measure of urbanization:

$$\widehat{\alpha}_{at} = \beta log(dens_{at}) + \delta_t + \zeta_{at} \quad (4)$$

[15]Using industry dummies instead of NCLA dummies does not alter our main findings. Results are available on request.

where $log(dens_{at})$ is the logarithm of the population density in area a at time t, δ_t are year fixed effects, which control for business cycle shocks affecting all the areas, as in Combes et al. (2008).

The model (3)-(4) is estimated by OLS, with residuals clustered at the local level, i.e. LLMs or municipalities. $\widehat{\beta}$ is the estimate of the wage elasticity to agglomeration, i.e. the urban wage premium: doubling population density increases wages by β. As discussed in Section 3, in our main analysis, the spatial unit for the first stage fixed effects is the LLM where the individual works; in a robustness check, we replicate our results employing the municipality, instead, and our conclusions remain substantially unaltered (see Section 7 and Appendix D).

The two-stage approach has several advantages with respect to other estimation strategies employed by the related empirical literature and, in particular, allows us to overcome some limitations of the one-stage regression estimation (see Combes et al., 2011).[16] First, in a one-stage approach, the time-invariant area effects, that one should include to control for possible confounders at the area level, would be identified only by the workers that move between areas in the sample period: these workers might be a few and only a selected sample of the population. Second, in the one-stage approach, when including time-invariant area fixed effects, the wage elasticity to agglomeration for stayers would be identified from the time variation in the area population density: this variation (that in any case is likely to be small) could be driven by time changes in unobserved area characteristics (possibly correlated with individual wages) and, hence, the estimated elasticity could turn out biased. These problems are attenuated when using the two-stage procedure which includes, in the first stage, time-varying area effects: it allows one to identify separately the effects of worker's characteristics versus the effects of area's characteristics and, hence, to properly estimate the wage elasticity to urbanization.[17]

In estimating equations (3)-(4), in addition to the possible confounding effect of area unobserved characteristics, we also worry about other sources of biasness of the estimated elasticity, possibly stemming from unobservable workers' and firms' heterogeneity. Hence, as well-established in the urban economics literature (e.g. Glaeser and Mare, 2001, Combes et al., 2008), in some specifications we include individual fixed effects, θ_i, in the first stage. The importance of firm fixed

[16] In the robustness checks, we replicate our results employing the one-stage approach. While qualitative conclusions remain unaltered, the estimated elasticities turn out to be larger in absolute values. Results are available upon request.

[17] The area-year effects might be estimated imprecisely in the first stage. Hence, the fact that they are used as dependent variable of the second stage might cause some econometric complications. We face this possible concern by adopting the correction procedure suggested by Combes et al. (2008) and Combes et al. (2019) in our baseline estimates, to compute the covariance matrix of the aggregated error terms and verifying that our results are not substantially altered. This may be due to the fact that we observe the universe of Italian workers, i.e. a huge number of observations of our dataset, that allows us to estimate the area-year effects in the first stage quite precisely. We do not extend this procedure to all estimates since due the size of our sample is highly computationally intensive. Econometric output of the baseline estimates is available from the authors upon request.

effects in explaining wage differentials has also been emphasized by the related literature (Card et al., 2013, Dauth et al., 2018). Accordingly, we also add firm fixed effects, $\theta_{j(it)}$, in some of our first-stage regressions as in the two-way fixed effects AKM models (Abowd et al., 1999).

A further point is that population density could be endogenous to wages. For example, there could be a matter of reverse causality: a positive productivity shock in an LLM increases wages and this increase in wages attracts workers, thereby boosting population density. To address this, we adopt an instrumental variables approach. Following a well established methodology adopted by the relevant literature (see Ciccone and Hall, 1996 and Mion and Naticchioni, 2009), we instrument, in the second stage of our model, the current population density with the historical local population dating back to 1871 (Istat, 2016c). The underlying intuition is that these values are closely correlated with current population density, but are likely to be uncorrelated with current local shocks. Nonetheless, recent surveys in the urban economics literature, such as Combes et al. (2011), have shown that, from an empirical point of view, endogeneity is not a major issue in this context, since IV estimates usually confirm OLS results.

6 Main regression results

Table 3 shows our baseline regression output for the universe of male Italian private-sector employees. The top panel reports the wage elasticity obtained in the second stage (LLM-year effects estimated in the first stage regressed on the logarithm of the LLM population density and time fixed effects, as explained in Section 5); the bottom panel displays, for each column, the controls included in the first stage, as in equation (3). Estimates obtained using nominal and real wages respectively are shown in columns (1)-(3) and (4)-(6).

The elasticity reported in column (1), that refers to a first stage that only includes the LLM-year fixed effects, amounts to 0.022, which means that doubling the population density increases nominal wages by 2.2%. This suggests that the agglomeration effect in the Italian labour market is positive and statistically significant, but very mild, since a comparable estimate for the US labour market ranges between 20% and 35% (Glaeser and Mare, 2001, and Yankow, 2006).

In column (2), we show the results after introducing in the first stage workers' observable characteristics: age group dummies, dummies for part time contract, dummies for fixed term contract, and occupation dummies. As can be seen, controlling for worker observable characteristics dampens the elasticity to 0.004, which turns out not statistically significant. In column (3), we add firm level observable characteristics, such as (log) firm size and contract dummies for NCLAs (FULL OLS from now on). The wage elasticity to urban agglomeration becomes economically and statistically equal to zero.[18] This suggests that the positive estimated UWP shown in column (1) is only driven by the composition of the labour force and firm population along the space dimension.

Turning to real wages, the UWP estimated without including in the first stage controls for firms' and workers' characteristics, displayed in column (4), is very close to zero (equal to -0.003) and not statistically significant. When adding individual controls in the first stage regression, as we do in column (5), the estimated UWP is negative (-0.022) and highly statistically significant: employees located in more agglomerated areas suffer a wage penalty in real terms. Introducing observable firm's characteristics in the first stage, column (6), increases the wage penalty to 2.6%. It is also interesting to underline that the R-squared of the second stage for nominal wage regressions is always much higher with respect to the ones in real wage regressions, suggesting that population density does not explain much of the variability for real wages. This is consistent with the evidence reported in Table 2: a clear correlation between nominal wages and population density, which is not observed when real wages are considered.

Table 3 about here.

One might think that sorting of workers and firms matters in driving the above results. Hence, in Table 4, we report our estimation output after introducing, in the first stage, individual fixed effects (hereafter FE) and, in addition, firm fixed effects (AKM). The estimated wage elasticity, in nominal terms, as respectively shown in columns (1) and (2), is still very close to zero and not statistically significant at any conventional level. When we consider real wages, instead, we find again a urban wage penalty equal to 2.7% in both specifications (columns (4) and (5)).

We, then, turn to addressing the matters of endogeneity by implementing IV estimates. As an instrument for population density, we employ LLM population in 1871. The results obtained after including in the first stage regressions individual fixed effects are displayed in column (3) and (6) of Table 4 respectively for nominal and real wages (IV-FE). The Kleibergen-Paap Wald rk test (see Kleibergena and Paap, 2006) always induces us to reject the null hypothesis of weak instruments. Again, the elasticity of nominal wages to the agglomeration measure is very close to zero (0.007), whereas that of real wages becomes even greater in absolute value, equal to -0.032, pointing at a non-negligible real penalty for urban employees.

Table 4 about here.

An additional concern is related to worker heterogeneity, mainly in terms of the worker's occupation, that could play a role in the UWP estimation. Consistently with the literature, UWP are

[18] We also implement a check without including the firm size, which might be considered as endogenous, in the regressions. Results remain substantially unaltered.

supposed to be larger for high-skilled workers, who are more likely to benefit from agglomeration externalities. To investigate on this, Table 5 shows results in FULL OLS and FE specifications, from baseline regressions estimated separately for blue-collar employees, white-collar employees, and managers/executives. Interestingly, for blue collars we find zero nominal premium when workers' and firms' characteristics are included in the first stage (column (1)), and a positive but rather small wage elasticity for white collars and managers and executive, respectively equal to 0.8% and 1.6%. In the FE estimates, the UWP gets to zero for basically all categories (slightly negative for managers). When looking at real wages, our results show that, as for the full population of employees, the UWP is negative and substantial for all groups, ranging between -0.026 and -0.010 in the FULL OLS specification, and between -0.030 and -0.026, in the FE regressions. Hence, we can confidently state that our main findings (small nominal wage premia and real wage penalties for urban employees) do not depend on the workers' occupations, since they consistently apply to the various subgroups of workers.[19]

Table 5 about here.

Finally, we argue that our findings are consistent with the intuitions presented in Ichino et al. (2019). However, in their model the real penalty suffered by employees in high productivity areas is offset by a higher employment probability: the North (South) of Italy is associated with lower (higher) real wages and higher (lower) employment rates. To verify whether this mechanism also applies to areas with different agglomeration size, we collect data on employment and unemployment rates provided by Istat (2016a) at the LLM level for the period 2006-2015.[20] Figure 6 displays the correlation between the unemployment rate in 2006 and (log) population density: the unemployment rate is slightly increasing with population density, suggesting that workers in cities could be penalized in terms of both real wages (as found before) and (in addition) probability of finding a job. One might think that inactivity rates differ between agglomerated and non-agglomerated areas, for several reasons (for instance, universities are only located in medium-big cities). This is confirmed by the relationship between employment rate and (log) population density, displayed in Figure 7: the employment rate is flat along the space dimension, i.e. inactivity is lower in cities. Hence, even considering the employment rate, there is no evidence that the negative UWP computed in real terms for employees can be offset by higher (lower) employment (unemployment) probabilities.

[19]For the managers and the white collars, moving to fixed effects estimates reduces the absolute size of the penalty, suggesting that for these categories of workers there is some evidence of sorting. The previous result found for the overall group, is due to the fact that the blue collars are the majority of workers.

[20]Data for 2005 are not available for the classification of LLM issued in 2011, which we use here.

Figure 6 about here.

Figure 7 about here.

Table 6 shows results from estimation of our models (FULL OLS, FE, AKM, and FE-IV) for nominal and real wages, respectively, after introducing the LLM unemployment rate as an additional control in the second stage. As one can see, our previous conclusions are not significantly altered, suggesting that the unemployment rate is not strongly correlated with our covariate of interest (population density) and, hence, lower unemployment probability does not offset the real wage penalties suffered by individuals working in more densely populated areas.[21]

Table 6 about here.

7 Identifying the role of collective bargaining: the independent contractors

Pulling together our results so far, we can conclude that, as far as employees are concerned, Italian nominal UWP are very close to zero, whereas real UWP are negative and substantial in magnitude. These findings are at odds with the previous evidence found in the existing literature for the US (again, see Glaeser and Mare, 2001 and Yankow, 2006). Possible explanations are: the role of the collective bargaining system coupled with the local variation in the cost of living, the presence of urban amenities, and the idiosyncratic preferences for locations.

To identify the role played by the national wage setting, we then turn to alternative categories of workers, who live in the same locations as employees, while being subject to a different wage-setting scheme. Our preferred comparison group is that of the independent contractors (the relevant labour contract was introduced in the Italian labour market in 1997) since they are self-employed workers always associated with a firm, as the employees. Independent contractors are temporary and may perform a wide range of tasks, either unskilled or skilled (e.g. statutory auditor, company administrator, legal representative of the firms, external staff and/or consultancy). Furthermore, and importantly for our purposes, there are no institutional constraints to wage setting, which turn out

[21]Note that another adjustment channel in the Italian labour market is the presence of informal workers. Nonetheless, while in the North-South divide this issue can be crucial, in the agglomeration (within province) dimension its role may be deemed marginal. Indeed, we might well assume that, within a province, the incentives by firms to resort to informal workers are homogeneous across space, depending also on additional factors working at the local level, such as social norms, monitoring activities by the police and tribunals at the provincial level, etc. The province fixed effects included in our regressions should largely capture this phenomenon.

to largely depend on bargaining between employer and worker. To maximize comparability with employees, we exclude independent contractors related to the public sector (in which case the firm could not be identified).

Our analysis applied to this group of workers also considers the 2005-2015 period and covers male workers aged 18-64, for the sake of comparison with the results shown for employees. For independent contractors, the INPS archives provide us with information on gross yearly worker compensation and length of contract (in days). We can then compute a standardized measure of earnings, the daily wage. In addition, the archives include information on the worker age, the industrial sector corresponding to the firm with which the worker is associated, the municipality, and the LLM where he/she works. We select one observation per worker per year and, as in the case of employees, we choose to keep the contract with the highest earnings paid to the worker. We also crop the tails of the distributions of, respectively, earnings and contract lengths (1% at the bottom and at the top of each distribution).

In Table 7, we present the relevant estimation output. For the sake of comparison, we employ specifications as close as possible to those used for employees (see Table 3). The dependent variable is either nominal or real daily wage, the latter being computed with the same local CPI adopted for employees.

Table 7 about here.

Column (1) displays our results for nominal wages when, in the first stage of our model, for our preferred specification (FULL OLS), which includes (log) firm size, industry dummies (associated with industrial sectors according to 2-digit Ateco classification), dummies for categories of independent contractors, in addition to the LLM-year dummies: the estimated wage elasticity turns out to be equal to 5.4% and statistically significant at any confident level.[22]

To consider the role of workers' sorting, in column (2) we display the estimated elasticity from a model that includes, in the first stage regression, in addition to controls for workers' and firms' characteristics, worker fixed effects. We obtain a substantial drop in the point estimate that comes to 0.006 and turns out not statistically different from zero. Hence, sorting of workers plays a role for this group of workers, contrary to what we found for employees. This is an interesting result:

[22] Ateco classification (ATtivit ECOnomiche) is the industrial classification adopted by Istat and represents the Italian version of NACE industrial classification employed by Eurostat. The current version is Ateco 2007 that corresponds to NACE Rev. 2. They are statutory auditors (company administrators, and legal representatives), editorial collaborators, member of boards and commissions, administrators of local authorities, project contractors (external staff of private employers), door-to-door salesmen, temporary (casual) contractors, independent temporary (casual) contractors, project contractors for retired workers, external staff for public employers, extended contracts for external staff, profit sharing contractors, specialized trainees, coordinated and continuing collaboration.

the higher spatial variability of earnings for self-employed workers provides a higher incentive for individuals to sort in space, with respect to workers subject to the collective bargaining. These results are confirmed when we also consider the AKM and the FE-IV models (columns (3)-(4)).

We, then, move on to the analysis of real wages. As shown in column (5) of Table 7, in the FULL OLS specification the estimated UWP turns out to be positive (0.026) and statistically significant. This is remarkable given the substantial real penalty that was found for employees when adopting the same specification (see column (6) of Table 3). Again, introducing individual fixed effects, as in column (6), substantially dampens the point estimate of the wage elasticity, that becomes negative but not statistically significant. The AKM and the IV-FE estimates, shown in columns (7) and (8) respectively, confirm previous conclusions.

Figure 5 gives a visual and synthetic representation of our results when comparing employees and independent contractors. We run a regression of, respectively, the logarithm of the nominal wages (for employees and independent contractors) on individual controls (specified in the footnote of the Figure) and LLM fixed effects for 2005. Then, we plot the estimated trend of LLM fixed effects derived in these two separate regressions, for the two groups of workers, on the logarithm of the (log) population density by LLM. We also add to the graph the trend of our local measure of CPI index computed at the LLM level.[23] As one can see, the curve referred to the earnings of the independent contractors and that referred to the local CPI move together (earnings and prices increase with the population density), while the line referred to the employees is much flatter. This clearly shows that, where there are no institutional constraints to wage bargaining, the spatial dynamics of wages tend to follow that of the local cost of living. By contrast, the collective bargaining by making, necessarily, nominal wages more homogeneous across space, weakens possible compensation mechanisms between earnings and the cost of living.

Figure 5 about here.

All in all, the evidence reported so far suggests that, in FULL OLS, there are substantial differences in results between employees and independent contractors. Since there are no compelling reasons to believe that the two types of workers differ as regards the quality of amenities that individuals enjoy and their preferences for locations, the difference in results can be associated with the different bargaining system regulating the corresponding labour market.[24] Amenities, quality

[23]To make the magnitude of the three variables comparable, we consider, for each variable, the deviation from its national mean, computed weighting by LLM population.

[24]It might be argued that even earnings of independent contractors could be at least partially anchored to wage floors for employees involved in the national collective bargaining of the corresponding industry category. Yet, under the plausible assumption that this effect is invariant across the agglomeration dimension, this issue is unlikely to affect our conclusions (and even if this was not the case, the differences in UWP between employees and self-employed should be considered as a lower bound of the collective bargaining effect).

of public goods, and idiosyncratic preferences for locations could still play a part in explaining the absolute levels of the estimated UWP in each group. For instance, the fact that, for employees, the UWP comes to -2.7% (in the FE specification) in terms of real wages would mean that, in equilibrium, the quality of the environment where individuals work has to offset the real wage penalty suffered by workers in urban areas. Hence, in cities there will be an over-representation of employees with strong preferences for these compensating factors. By contrast, as regards independent contractors, for whom we have obtained (again FE specification) a positive estimated UWP in nominal terms and a zero premium in real terms, preferences for amenities and public goods are assumed to play a weaker role in choosing where to work.

8 Robustness checks

To verify robustness of our findings, we run a vast array of checks. Some of the checks are specific to econometric issues, while others are more general: they include a different choice of the comparison group for employees (standard self-employed instead of independent contractors), an assessment of the role of the second level bargaining, and an analysis along the wage distribution.

To begin with, we verify whether our results depend on modelling and specification choices adopted to produce the main results. First, we focus on prime age male workers (aged 25-49), often used as a reference group in the urban economics literature. Second, we change our measure of earnings and use yearly wages rather than weekly wages. Third, we change the spatial unit of observation and move from LLM to municipalities. Fourth, we employ a different agglomeration measure, employment density rather than population density.[25]

Moving to the more general checks, we investigate whether our interpretation is driven by our choice of the group of workers used for comparison with employees. This might occur, for instance, should the self-selection into the two groups, employees and independent contractors, be heterogeneous across the spatial dimension. While there is no compelling reason why this should be systematically the case, to be on the safe side, we consider an additional group for comparison: the standard self-employed workers, who are (differently from independent contractors) not necessarily associated with a firm. We thus focus on workers belonging to professional associations (*ordini professionali*).[26] In Italy, for legal qualification to work in these occupations the worker needs to

[25] An alternative proxy for agglomeration could be built by adopting EU-OECD definition of "functional urban areas" (see OECD, 2012), as in Lamorgese and Petrella (2016). We have also employed this approach and previous conclusions do not substantially change.

[26] High and medium-skilled workers categories are: psychologists, nurses, industrial engineers, farmers, biologists, journalists, architects, lawyers, physicians/dentists, veterinaries, surveyors, accountants, business consultants, and a multi-category (agronomists, foresters, actuaries, chemists, and geologists). Our data do not include workers that

pass a national examination to enter the association. In turn, the professional associations have to handle the worker social security contributions directly, through specific pension funds (*casse degli ordini professionali*), which are available in INPS, even if they are formally separated from the INPS archives covering employees and independent contractors.

Data on standard self-employed workers include: individual yearly wages, period worked, occupation, gender, age, and municipality where the worker lives. The latter variable is time-invariant and, in this aspect, differs from that available for employees and self-employed workers. This is expected to make no great difference since, by carrying out the analysis at the LLM level, we maximize the probability that individuals work and live in the same spatial unit of observation. Moreover, for standard self-employed, the variable indicating the place of residence is time-constant, and this prevents us from implementing fixed effect estimation. Nonetheless, it is worth noting that we could expect spatial mobility for standard self-employed workers to be low, since their occupations are based on client networks at the local level, and changing locations would mean losing this asset. Occupations such as those of lawyers, architects, and surveyors are clear examples of such mechanisms: leaving a local market means losing the business network.

The econometric specification is chosen as close as possible to that already used for employees and independent contractors. However, a number of variables cannot be defined, for instance, industry dummies and firm size, since standard self-employed workers are not associated with a firm, and the dummies for part time and/or temporary contracts, since these variables make little sense for this group of workers and are not collected by INPS. We again focus on male workers aged 18-64, who work the full year (more than 90% of workers work for the full year). After dropping outliers (1% at the top and bottom of the distribution of yearly income), we end up with a sample of around 4,150,000 observations for the period 2005-2015. As before, our main covariate of interest is population density. The main (first step) control variables are: age group dummies, area-year dummies, and dummies for the type of self-employed subgroup type (lawyers, nurses, etc.). Table 8 shows the second stage results when the group of standard self-employed workers is considered. After including all control variables and occupation dummies in the regression (FULL OLS), the estimated wage elasticity turns out positive and statistically significant both in nominal and real terms. These results lead us to similar conclusions as those derived for the independent contractors, supporting the intuition that the selection into different groups of workers can hardly be the main driver of our findings.

Table 8 about here.

do not belong to *ordini professionali*; for this reason, we have no information on unskilled standard self-employed workers, such as plumbers.

Another issue that we need to consider is the possible further role played by the second level wage negotiation, at the local and/or firm level, which is at work in the Italian two-tier system. One might argue that, even if national collective bargaining tends to make the wage distribution flat in the space dimension, the second level bargaining is responsible for spatial variation in worker remunerations at the local level, for instance, to align wages and productivity. As already mentioned, the Italian wage collective agreements do not allow downward flexibility (actual wages cannot be lower than those negotiated at the national level). However, the *in melius* principle is applied and workers can be paid more than the wage set at the national and industry level. To investigate whether this phenomenon is relevant for our results, we rely on an Employer and Employee Survey, *Rilevazione Longitudinale su Imprese e Lavoro* (RIL), conducted by the Italian Public Policy Research Institute (INAPP) in the years 2005, 2007, and 2015 on a nationally representative sample of firms operating in the non-agricultural private sector (around 30,000 firms).[27] This dataset is also merged with annual balance sheets and income statement data from the *Analisi Informatizzata Delle Aziende Italiane* (AIDA). We construct a dummy variable which identifies firms that have signed a second level bargaining agreement that applies also to wage premia, i.e. the second level bargaining that translates into a pay increase. The percentage of these firms account for around 3-4% of the sample and this is in line with previous literature documenting that the wage increase due to decentralized negotiation is rather negligible (Boeri, 2014, Brandolini et al., 2004, Ichino et al., 2019, and Devicienti et al., 2019). Nonetheless, to make sure that we are not neglecting this important aspect of the Italian labour market, we also verify that our conclusions do not change when we estimate our baseline regressions for two distinct groups of firms: below and above the median in terms of the incidence of second level bargaining, computed at the macro-region/industry (2-digit) dimension. Results (in FULL OLS), reported in Table 9, confirm previous findings: the nominal UWP turns out to be very close to zero and not statistically significant, and the real UWP is negative and statistically significant for both groups of firms. However, the (absolute value of the) real penalty computed above the median is half than that computed below the median, suggesting that in cities the second level bargaining slightly compensates employees for the higher cost of living there. This evidence confirms that the second level bargaining does play a role, but it does not alter previous findings from a qualitative point of view our. Thus, our implicit assumption in the paper, that the first level bargaining is the major source of wage differential, is supported by the data.

Table 9 about here.

Finally, we verify robustness of our results to possible heterogeneity (in location choices and

[27]For a more detailed information on the survey please refer to https://www.isfol.it/attivita/indagini-e-ricerche/indagini-campionarie/rilevazione-longiutudinale-su-imprese-e-lavoro.

self-selection effects) across groups of workers along the wage distribution (assuming that wages can be used as a proxy for skills). On the one hand, one might expect high-wage high-skill workers to be able to extract higher wages in decentralized and individual bargaining (on top of that resulting from industry collective bargaining), and that this could be more likely in agglomerated areas where firms are more productive. In this case, differences in UWP between employees and self-employed workers detected at the conditional mean could shrink. On the other hand, it could be that local collective agreements in cities benefit most at the bottom tail of the wage distribution, for instance, when local bargaining is undertaken by local unions mostly interested in the welfare of unskilled workers (e.g., Matano and Naticchioni, 2017, considering Italian blue-collar workers, show that the extent of rent sharing decreases along the wage distribution, mainly due to the role of the unions).

In order to further investigate these issues, and complementing our findings reported in Table 5, we make use of the unconditional quantile regressions (UQR) technique, proposed by Firpo and Lemieux (2009).[28] Here, we are forced to use reduced form (one-step) specifications (log individual wages on log population density and the same controls used in the baseline estimates for each group: employees, independent contractors, and standard self-employed workers respectively), since implementing the two-step approach makes little sense in the case of quantile regressions.[29] Figure 8 shows the FULL OLS point estimates, together with the corresponding 95% confidence interval, using UQR for each group of workers (employees, independent contractors and standard self-employed) at the 10th, 50th, and 90th percentiles of the wage distribution. The upper panel refers to nominal wages. As regards employees, the estimated wage elasticities are rather low, close to zero, and slightly increasing with wages (consistently with Matano and Naticchioni, 2012): from -0.6% at the 10th percentile to 0.7% at the 90th percentile. This suggests that unskilled employees (10th percentile) are penalized in an agglomeration dimension even when considering nominal wages. As for independent contractors and standard self-employed workers, the estimated wage elasticities are positive and much larger for all percentiles, with a U-shape impact. If any, the UWP difference between employees and self-employed workers get larger along the wage distribution, from about 4% at the median to around 6% (independent-contractors) and 7% (standard self-employed), respectively, at the 90th percentile: high-skilled employees are unable to

[28]The underlying idea is to estimate a linear regression where the dependent variable is replaced by the recentered influence function (RIF) of the distributional parameter. This methodology enables us to estimate the impact of population density on quantiles of the unconditional distribution of wages, employing the same specification chosen for the conditional mean analysis. Moreover, using UQR instead of the standard conditional quantile regressions (Koenker and Bassett, 1982), allows to detect the impact of the covariates on the median of the unconditional distribution, which is what we are interested in. By contrast, conditional quantile regressions estimates, for instance at the median, refer to the median of the "error term".

[29]On average, using a one-step approach entails the same qualitative results derived with a two-step approach, although coefficients are slightly greater in magnitude.

negotiate higher returns in cities, as the self-employed workers do. The corresponding differences at the 10th percentiles are also similar in size and, again, greater than that at the median. Corresponding results obtained for real wages are reported in the lower panel. Estimated real UWP for employees are negative and substantial, and the penalty gets smaller from the 10th to the 90th percentile. By contrast, in the case self-employed workers, the high-wage individuals (90th percentile) are able to extract a positive and non-negligible wage premium, of around 2% even in real terms. Hence, also in the case of real wages, employees do not close the gap with self-employed workers: the differences in estimated UWP across groups tend to be larger for both high-skill and low-skill workers than for medium-skill ones.

Figure 8 about here.

9 Concluding remarks and policy implications

In this paper, we present a thorough investigation into estimation of the UWP in Italy, a country characterized by a two-tier collective bargaining system and a heterogeneous local cost of living, for different categories of workers: employees and self-employed workers (independent contractors and standard self-employed). We make use of unique administrative archives from INPS and a housing transactions database from Italian Revenue Agency (2015) for the period 2005-2015.

Our results, controlling for all worker and firm characteristics (FULL OLS), reveal that, when employees are taken into account, the elasticity of nominal wages with respect to urban agglomeration size is very close to zero, while the corresponding elasticity of real wages is equal to -2.6%. This result does not hold for self-employed workers (either independent contractors or standard self-employed), for whom we obtain a positive and statistically significant premium for both nominal and real wages.

These findings, which are robust to a wide range of robustness checks, suggest a role is played by the collective bargaining system in the presence of high spatial heterogeneity in the cost of living, given that employees and self-employed workers share the same local amenities and idiosyncratic preferences for locations. Our results survive also when we take into account worker heterogeneity, both in terms of occupations and considering their position along the wage distribution.

While the evidence presented in the paper seems to be relevant *per se*, our findings contribute to a lively and topical policy debate.[30] On the wave of recent reforms in European labour markets, some authors have proposed to enhance the scope for decentralization of the current Italian

[30]For discussion of some issues related to the proper functioning of the Italian collective bargaining, see D'Amuri and Nizzi, 2017). Some examples follow. First, in recent years numerous national labor contracts have been signed by "minor" employers' and workers' associations that tend to set levels of wages below the wage floors specified in the

system and, for instance, to relax the *in melius* clause associated with the second level bargaining (see, again, D'Amuri and Nizzi, 2017), conditioning on the approval of the local labour unions. Implementing of such decentralization, however, is likely to raise serious applicability issues, especially in the phase of transition to a new equilibrium. In the case of Italy, the situation is further complicated by the specificities of the productive system, where 90% of the firms have less than 15 employees, and only less than 5% of the firms with less than 15 employees have trade unions at the local level (Devicienti et al., 2018). This is relevant because in small and/or non-unionized firms, the workers' bargaining power could be much lower than that of the employers, and this could be particularly true in areas characterized by high unemployment rates. The introduction of a minimum wage imposed by the government rather than by the national agreements has also been suggested by some scholars, but this proposal has also, so far, encountered important applicability issues (for a discussion, see Ichino et al., 2019, and Garnero et al., 2015). Finally, some other authors (e.g., Garnero, 2018) have proposed actions that would improve the functioning of the current system, without significantly altering its institutional setting, such as rationalizing the number of collective contracts, acknowledging only agreements signed by the major unions and employers' associations, and improving transparency and information disclosure on negotiated wages.

The evidence presented in our work has shed some further light on the functioning of the Italian labour markets and has unveiled a previously unexplored phenomenon: the existence of real wage penalties for urban workers and their link to the collective bargaining system. This finding seems to be crucial for future research and policy discussion.

"major" national agreements of the same job-sectoral categories (D'Amuri and Nizzi, 2017). Second, non-compliance by many firms has also implied that many workers are paid less than the sectoral minimum wage set by the national agreement (according to Garnero, 2018, on average, around 10% of the workers are paid 20% less than the minimum wage). A related issue lies in the actual non-application of Article 39 of the Italian Constitution, according to which workers' and employers' associations with legal personality should be able to stipulate agreements with compulsory effectiveness for all relevant categories, establishing a "decent" level of pay for workers (Article 36 of the Italian Constitution). Application of this principle is impeded by the fact that in Italy the trade unions remain, at present, without legal personality. As a matter of fact, however, in cases of controversy, the judge has so far often taken as reference point for the "decent" level of pay, the national contracts stipulated by the most representative unions and employers associations. Third, a conspicuous share of the contracts (at present about half of the total) have expired but are still in force. The delay in renewal of the national contracts impedes or retards "national" wage adjustments over time.

References

Abowd, J., Kramarz, F., and Margolis, D. (1999). High-wage workers and high-wage firms, *Econometrica* **67**(2): 251–333.

Albouy, D. (2016). What are cities worth? Land rents, local productivity, and the total value of amenities, *Review of Economics and Statistics* **98**(3): 477–487.

Anci-Cresme (2005). *Le Politiche Abitative in Italia. Analisi e Valutazioni*, Roma: Cresme (Centro di Ricerche di Mercato, Servizi per chi Opera nel Mondo delle Costruzioni e dell'Edilizia).

Blien, U., Gartner, H., Stber, H. and Wolf, K. (2009). Regional price levels and the agglomeration wage differential in Western Germany, *The Annals of Regional Science* **43**(1): 71–88.

Boeri, T. (2014). Two-tier bargaining, *IZA Discussion Paper* **8358**.

Brandolini, A., Casadio, P., Cipollone, P., Magnani, M., Rosolia, A. and Torrini, R. (2004). Employment growth in italy in the 1990s: Institutional arrangements and market forces, *in* L. R. Acocella N. (ed.), *Social Pacts, Employment and Growth. AIEL Series in Labour Economics*, Physica-Verlag HD, pp. 31–68.

Card, D., Heining, J. and Kline, P. (2013). Workplace heterogeneity and the rise of West Germany wage inequality, *Quarterly Journal of Economics* **128(3)**: 967–1016.

Caruso, N. (2017). Housing policies in Italy: From social housing to neo-liberalism, *Policies and Practices in Italian Welfare Housing: Turin, up to the Current Neo-Liberal Approach and Social Innovation Practices*, Cham: Springer International Publishing, pp. 23–46.

Ciccone, A. and Hall, R. E. (1996). Productivity and the density of economic activity, *American Economic Review* **86**: 54–70.

Combes, P.-P. (2000). Economic structure and local growth: France, 1984-1993, *Journal of Urban Economics* **47**(3): 329–355.

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

Combes, P.-P., Duranton, G. and Gobillon, L. (2011). The identification of agglomeration economies, *Journal of Economic Geography* **11**(2): 253–266.

Combes, P.-P., Duranton, G. and Gobillon, L. (2019). The cost of agglomeration: House and land prices in cities, *Review of Economic Studies* **86**(4): 1556–1589.

Combes, P. P., Duranton, G., Gobillon, L. and Roux, S. (2010). Estimating agglomeration effects with history, geology, and worker fixed-effects.

D'Amuri, F. and Nizzi, R. (2017). I recenti sviluppi delle relazioni industriali in Italia, *Questioni di Economia e Finanza, Occasional Papers, Banca d'Italia* **416**(December): 1–27.

Dauth, W., Findeisen, S., Moretti, E. and Suedekum, E. J. (2018). Matching in cities, NBER Working Paper No 25227.

De Blasio, G. and Di Addario, S. (2005). Do workers benefit from industrial agglomeration?, *Journal of Regional Science* **45**(4): 797–827.

De La Roca, J. and Puga, D. (2017). Learning by working in big cities, *Review of Economic Studies* **84**(1): 106–142.

Devicienti, F., Fanfani, B. and Maida, A. (2019). Collective bargaining and the evolution of wage inequality in italy, *British Journal of Industrial Relations* **57**(2): 377–407.

Devicienti, F., Naticchioni, P. and Ricci, A. (2018). Temporary employment, demand volatility, and unions: Firm-level evidence, *Industrial and Labor Relation Review* **71**(1): 174–207.

Di Mauro, F. and Ronchi, M. (2016). Centralisation of wage bargaining and firms' adjustment to the great recession - A micro-based analysis, *European Central Bank, CompNet Policy Brief* **8**.

Draghi, M. (2015). *Structural Reforms, Inflation and Monetary Policy*, https://www.ecb.europa.eu/press/key/date/2015/html/sp150522.en.html.

Dumond, J. M., Hirsch, B. T. and MacPherson, D. A. (1999). Wage differentials across labor markets and workers: Does cost of living matter?, *Economic Inquiry* **37**(4): 577–598.

Duranton, G. and Puga, D. (2004). Micro-foundations of urban agglomeration economies, *in* J. V. Henderson and J. F. Thisse (eds), *Handbook of Regional and Urban Economics*, Vol. 4, chapter 48, pp. 2063–2117.

Eurostat (2019). *Distribution of Population by Tenure Status, Type of Household and Income Group - EU-SILC Survey*, http://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do.

Firpo, S., F. N. and Lemieux, T. (2009). Unconditional quantile regressions, *Econometrica* **77**: 953–973.

Garnero, A. (2018). The dog that barks doesn't bite: Coverage and compliance of sectoral minimum wages in Italy, *IZA Journal of Labor Policy* **7**(1): 1–24.

Garnero, A., Kampelmann, S. and Rycx, F. (2015). Minimum wage systems and earnings inequalities: Does institutional diversity matter?, *European Journal of Industrial Relations* **21**(2): 115–130.

Glaeser, E. (1998). Are cities dying?, *Journal of Economic Perspectives* **12**(2): 139–160.

Glaeser, E. (1999). Learning in cities, *Journal of Urban Economics* **46**(2): 254–277.

Glaeser, E. and Mare, D. (2001). Cities and skills?, *Journal of Labour Economics* **19**: 316–342.

Hornbeck, R. and Moretti, E. (2018). Who benefits from productivity growth? the direct and indirect effects of local tfp growth on wages, rents, and inequality, *NBER working paper* **24661**.

Ichino, A., Boeri, T., Moretti, E. and Posch, J. (2019). Wage equalization and regional misallocation: evidence from Italian and German provinces, *CEPR Discussion Paper* **13545**: 1–63.

IMF (2016). Italy, *IMF Country Report* **6/222**.

Istat (2016a). *Occupati Residenti e Persone in Cerca di Occupazione nei Sistemi Locali del Lavoro*, http://www.istat.it/it/archivio/178277.

Istat (2016b). *Prezzi*, http://www.istat.it/it/prezzi.

Istat (2016c). *Sistat. Sistema Informativo Storico delle Amministrazioni Territoriali*, http://sistat.istat.it/sistat/gestioneDate.do.

Istat (2017). *Soglie di Povertá Assoluta, Condizioni Economiche delle Famiglie e Disuguaglianze*, http://dati.istat.it.

Italian Revenue Agency (2015). *Osservatorio del Mercato Immobiliare, Banca Dati delle Quotazioni Immobiliari*, http://www.agenziaentrate.gov.it/wps/portal/entrate/home.

Kim, S. (1987). Diversity in urban labour markets and agglomeration economies, *Papers in Regional Science* **62**(1): 57–70.

Kim, S. (1990). Labor heterogeneity, wage bargaining, and agglomeration economies, *Journal of Urban Economics* **28**(2): 160–177.

Kleibergena, F. and Paap, R. (2006). Generalized reduced rank tests using the singular value decomposition, *Journal of Econometrics* **133**: 97–126.

Koenker, R. and Bassett, G. J. (1982). Robust tests for heteroscedasticity based on regression quantiles, *Econometrica* **50**: 43–61.

Lamorgese, A. and Petrella, A. (2016). An anatomy of Italian cities: Evidence from firm-level data, *Questioni di Economia e Finanza, Occasional Papers, Banca d'Italia* **362**(October).

Marshall, A. (1890). *Principle of Economics*, London: Macmillan.

Matano, A. and Naticchioni, P. (2012). Wage distribution and the spatial sorting of workers, *Journal of Economic Geography* **12**(2): 379.

Matano, A. and Naticchioni, P. (2017). The extent of rent sharing along the wage distribution, *British Journal of Industrial Relations* **55**(4): 751–777.

Mion, G. and Naticchioni, P. (2009). The spatial sorting and matching of skills and firms, *Canadian Journal of Economics* **42**(1): 28–55.

Moretti, E. (2004). Human capital externalities in cities, *in* J. Henderson and J. Thisse (eds), *Handbook of Regional and Urban Economics*, Vol. 4, Amsterdam: Elsevier/North-Holland, pp. 2243–2291.

Moretti, E. (2011). Local labor markets, *in* D. Card and O. Ashenfelter (eds), *Handbook of Labor Economics*, Vol. 4, Amsterdam: Elsevier/North-Holland, pp. 2119–2171.

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

OECD (1994). *The OECD Jobs Study. Part II: The Adjustment Potential of the Labour*, Paris: OECD Publishing.

OECD (2012). *Redefining Urban: A New Way to Measure Metropolitan Areas*, Paris: OECD Publishing.

OECD (2017). *OECD Employment Outlook 2017*, Paris: OECD Publishing.

Pittini, A., Ghekire, L., Dijol, J. and Kiss, I. (2015). *The State of Housing in the EU, A Housing Europe Review*, Brussels: Housing Europe, the European Federation for Public, Cooperative and Social Housing.

Roback, J. (1982). Wages, rents, and the quality of life, *Journal of Political Economy* **90(6)**: 1257–1278.

Rosenthal, S. S. and Strange, W. (2004). Evidence on the nature and sources of agglomeration economies, *in* J. Henderson and J. Thisse (eds), *Handbook of Regional and Urban Economics*, Vol. 4, Amsterdam: Elsevier/North-Holland, pp. 2119–2171.

Scanlon, K. and Whitehead, C. M. (2015). Social housing in Europe, *European Policy Analysis* **17**: 1–12.

Shapiro, J. (2006). Smart cities: Quality of life, productivity, and the growth effects of human capital, *Review of Economics and Statistics* **88**: 324–335.

Yankow, J. (2006). Why do cities pay more? An empirical examination of some competing theories of the urban wage premium, *Journal of Urban Economics* **60**: 139–161.

Zenou, Y. (2009). *Urban Labor Economics*, Cambridge: Cambridge University Press.

Figures and Tables

Figure 1: Local CPI and population density in 2005, by LLM

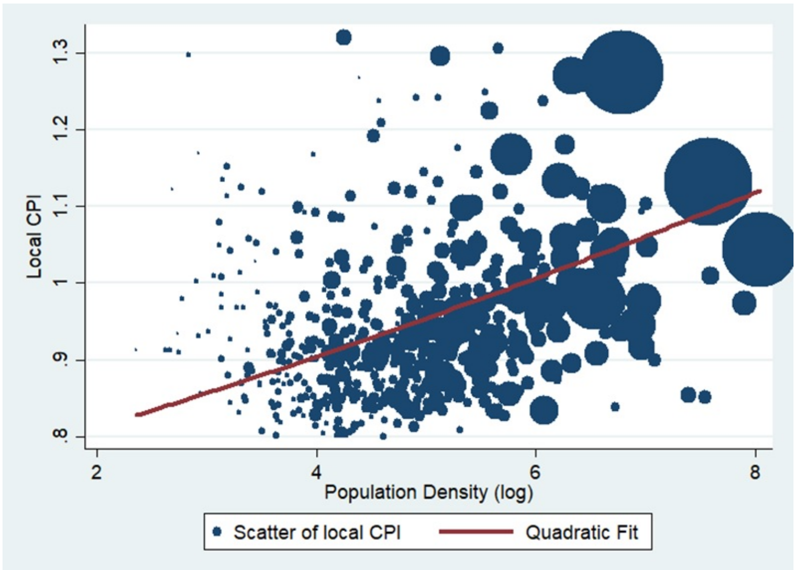


Figure 2: Distribution of local CPI in 2005, by LLM

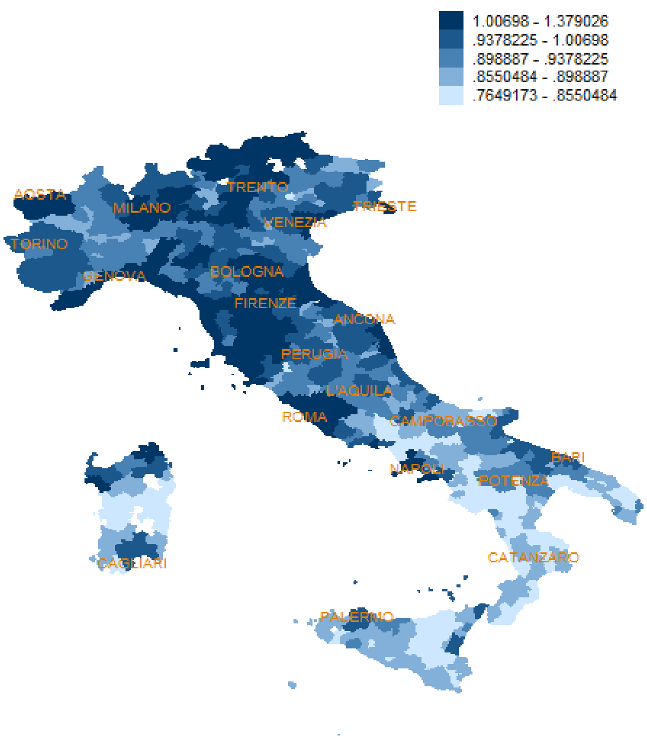


Figure 3: Distribution of population and employment density in 2005, by LLM

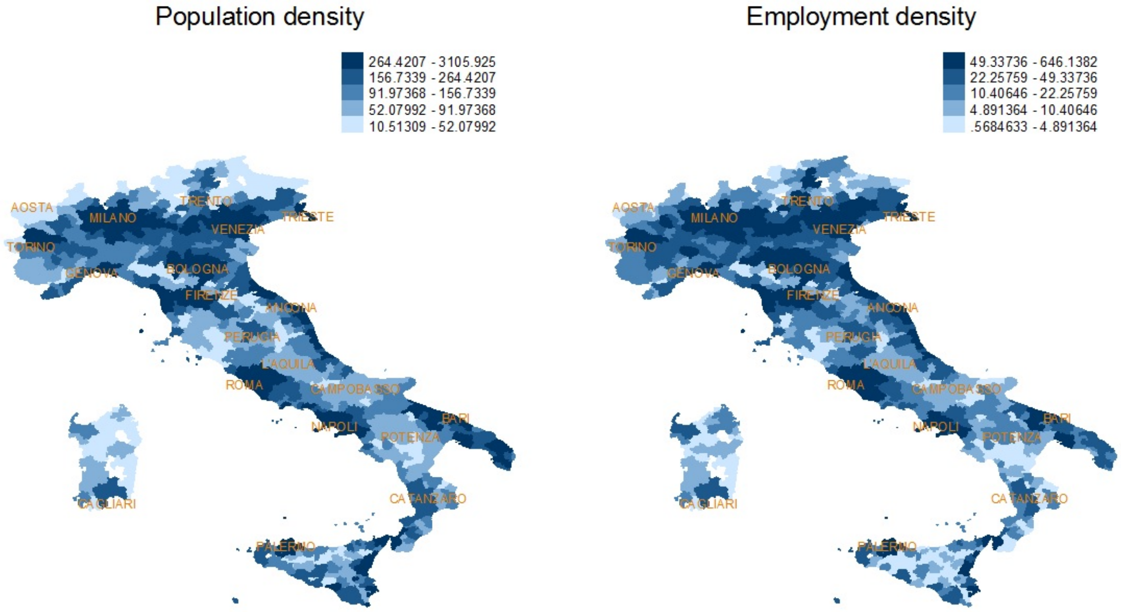


Figure 4: Distribution of nominal and real wages in 2005, by LLM

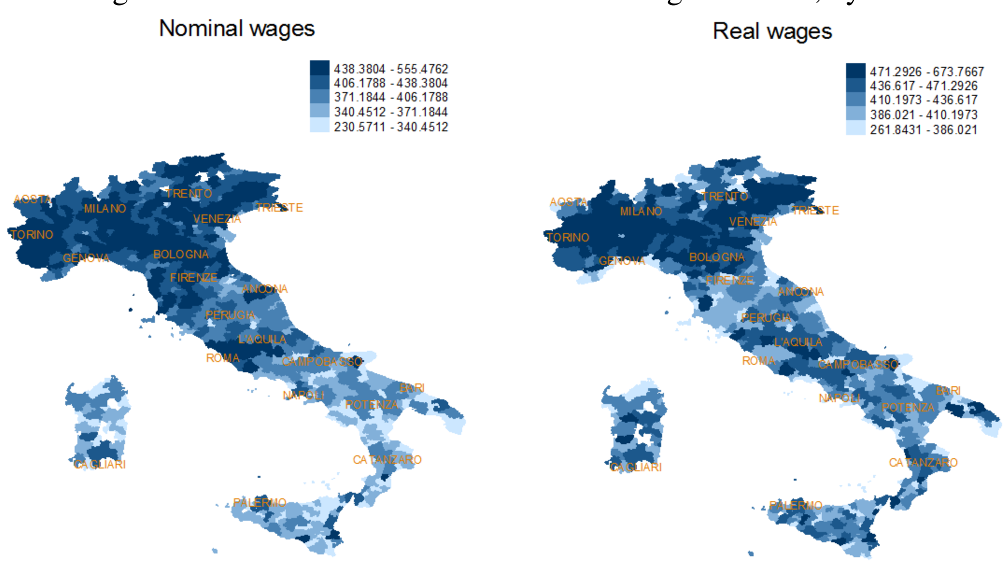
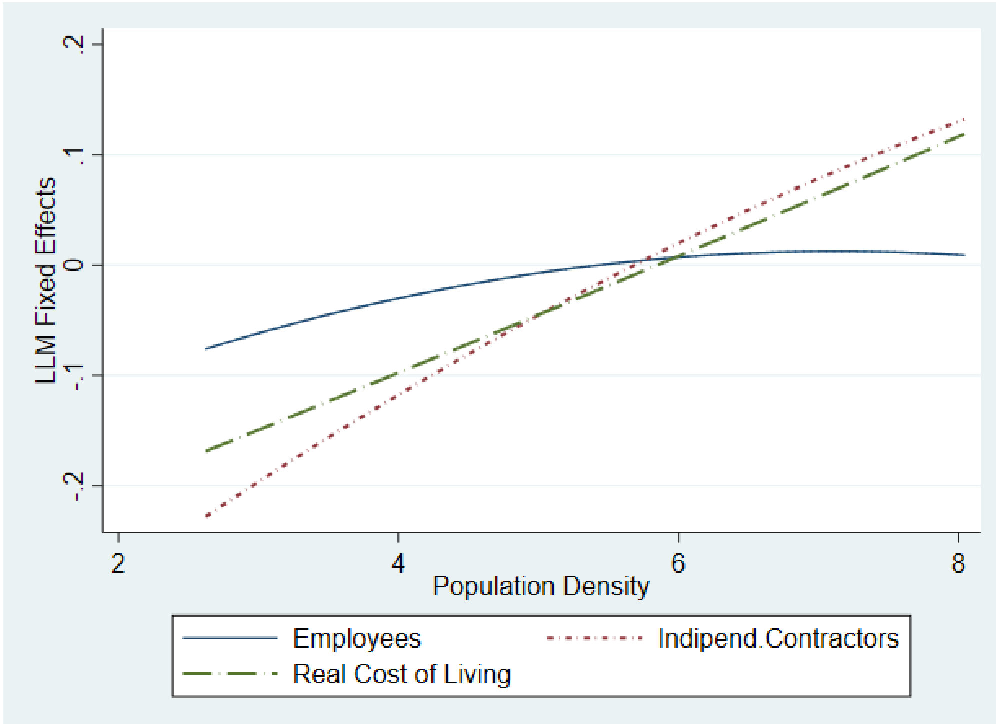


Figure 5: Relation between wage, local CPI index, and (log) population density across space



Note: LLM fixed effects on the vertical axis are estimated from a regression of, respectively, employees' (log) wage, independent contractors' (log) wage, and local CPI on worker age, (log) firm size, dummies for part time, fixed term, contractor category (for independent contractors, only), industry (for independent contractors, only), NCLA (for employees, only) and LLM dummies. To make variables comparable, we consider the deviations of each variable from its national mean, computed weighted by LLM population.

Figure 6: Unemployment rate and population density in 2005, by LLM

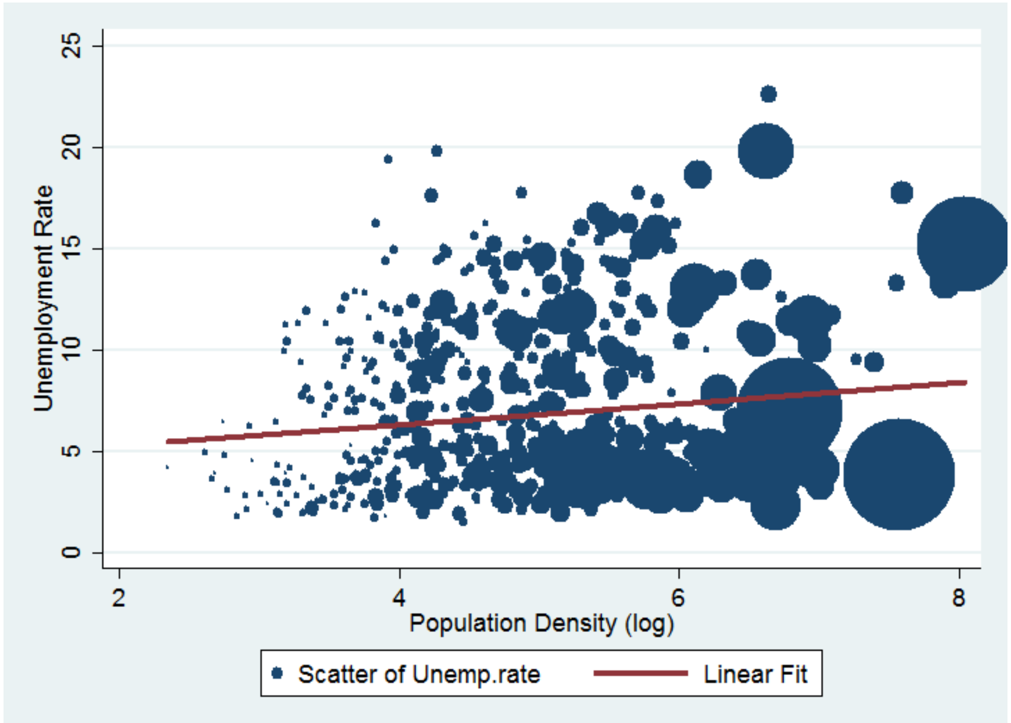


Figure 7: Employment rate and population density in 2005, by LLM

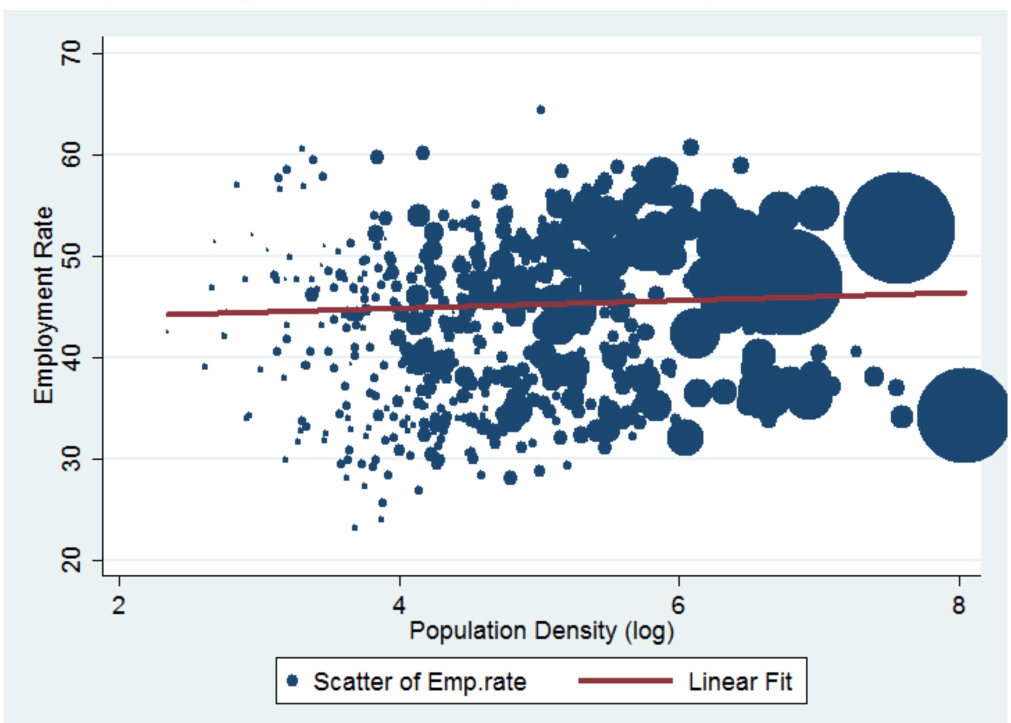
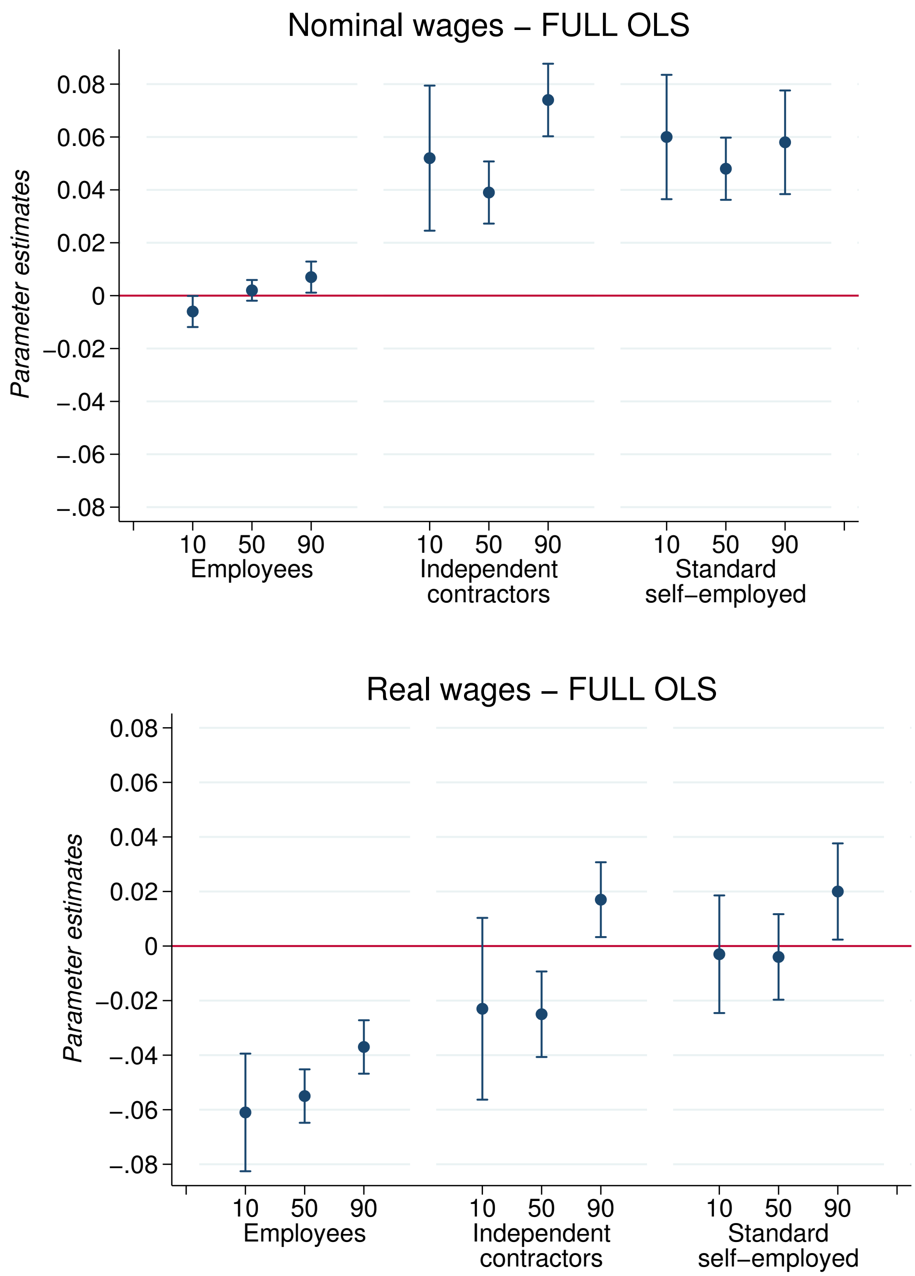


Figure 8: Unconditional quantile regressions



Note: Estimated UWP, respectively in nominal (upper panel) and real (lower panel) terms, for employees, independent contractors, and standard self-employed workers, obtained as explained in the text for UQR. Control variables are the same as in the baseline regressions for each group of workers.

Table 1: Within-between variance decomposition of nominal and real weekly wage in 2005

	Nominal Wages		
	Macro-Regions	Regions	Provinces
Between	4.4	4.7	6.2
Within	95.6	95.3	93.8
Total	100	100	100
	Real Wages		
	Macro-Regions	Regions	Provinces
Between	3.5	4.3	5.8
Within	96.5	95.7	94.2
Total	100	100	100

Note: 5 Macro-regions; 103 provinces; 20 regions. Evidence computed on the universe of private-sector employees.

Table 2: Nominal versus real weekly wage in 2005

Population density	Nominal wage	Real wage
first 20 %	401	446
20-40 %	428	459
40-60 %	444	458
60-80 %	472	465
top 20 %	511	443
total	452	452

Note: Quantiles weighted by population.

Table 3: UWP estimates. Employees: baseline specification

	Nominal wage			Real wage		
	(1) OLS	(2) + worker charact.	(3) FULL OLS	(4) OLS	(5) + worker charact.	(6) FULL OLS
	Second stage results					
log population density	0.022***	0.004	-0.000	-0.003	-0.022***	-0.026***
	(0.005)	(0.004)	(0.004)	(0.006)	(0.004)	(0.004)
year fe	yes	yes	yes	yes	yes	yes
Observations	6,558	6,558	6,558	6,558	6,558	6,558
R-squared	0.282	0.295	0.382	0.051	0.072	0.117
	First stage controls					
LLM-year dummies	yes	yes	yes	yes	yes	yes
age dummies	no	yes	yes	no	yes	yes
part time dummies	no	yes	yes	no	yes	yes
fixed term dummies	no	yes	yes	no	yes	yes
occupation dummies	no	yes	yes	no	yes	yes
contract dummies	no	no	yes	no	no	yes
firm size	no	no	yes	no	no	yes
Observations	77,015,891	77,015,891	77,015,891	77,015,891	77,015,891	77,015,891

Note: Estimation results of model (3)-(4) in the text. The first stage dependent variable in columns (1)-(3) is the (log) weekly nominal wage; in columns (4)-(6) it is the (log) weekly real wage. Age dummies refer to 9 age categories: 18/24, 25/29, 30/34, 35/39, 40/44, 45/49, 50/54, 55/59 and 60/64. Occupation dummies are for: white collar, blue collar, apprentice, manager, and executive; contract dummies are for NCLAs, which cover about 243 contracts. Standard errors are clustered at the LLM level. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table 4: UWP estimates. Employees: fixed effects and IV estimates

	Nominal wage			Real wage		
	(1) FE	(2) AKM	(3) IV-FE	(4) FE	(5) AKM	(6) IV-FE
			Second stage results			
log population density	-0.001	-0.001	0.007*	-0.027***	-0.027***	-0.032***
	(0.002)	(0.001)	(0.004)	(0.004)	(0.004)	(0.009
year fe	yes	yes	yes	yes	yes	yes
Observations	6,558	6,528	6,558	6,558	6,528	6,558
R-squared	0.627	0.890		0.121	0.100	
K-P rk Wald F statistic			79.551			79.551
			First stage controls			
LLM-year dummies	yes	yes	yes	yes	yes	yes
age dummies	yes	yes	yes	yes	yes	yes
part time dummies	yes	yes	yes	yes	yes	yes
fixed term dummies	yes	yes	yes	yes	yes	yes
contract dummies	yes	yes	yes	yes	yes	yes
occupation dummies	yes	yes	yes	yes	yes	yes
firm size	yes	yes	yes	yes	yes	yes
worker fe	yes	yes	yes	yes	yes	yes
firm fe	no	yes	no	no	yes	no
Observations	77,015,891	77,015,891	77,015,891	77,015,891	77,015,891	77,015,891

Note: Estimation results of model (3)-(4) in the text. The first stage dependent variable in columns (1)-(3) is the (log) weekly nominal wage; in columns (4)-(6) it is the (log) weekly real wage. Age dummies refer to 9 age categories: 18/24, 25/29, 30/34, 35/39, 40/44, 45/49, 50/54, 55/59 and 60/64. Occupation dummies are for: white collar, blue collar, apprentice, manager, and executive; contract dummies are for NCLAs, which cover about 243 contracts. Standard errors are clustered at the LLM level. *** p<0.01, ** p<0.05, * p<0.10.

Table 5: UWP estimates. Employees by occupation categories

	Nominal wage		Real wage	
	(1) FULL OLS	(2) FE	(3) FULL OLS	(4) FE
	Second stage results			
	Blue collars			
log population density	-0.000	-0.002	-0.026***	-0.027***
	(0.004)	(0.002)	(0.004)	(0.004)
year fe	yes	yes	yes	yes
Observations	6,558	6,514	6,558	6,514
R-squared	0.400	0.595	0.109	0.119
	White collars			
log population density	0.008*	0.000	-0.018***	-0.026***
	(0.004)	(0.002)	(0.004)	(0.004)
year fe	yes	yes	yes	yes
Observations	6,558	6,465	6,558	6,465
R-squared	0.175	0.662	0.037	0.081
	Managers and executives			
log population density	0.016***	-0.004***	-0.010*	-0.030***
	(0.004)	(0.001)	(0.005)	(0.005)
year fe	yes	yes	yes	yes
Observations	6,451	6,371	6,451	6,371
R-squared	0.275	0.732	0.027	0.096
	First stage controls			
LLM-year dummies	yes	yes	yes	yes
age dummies	yes	yes	yes	yes
part time dummies	yes	yes	yes	yes
fixed term dummies	yes	yes	yes	yes
contract dummies	yes	yes	yes	yes
firm size	yes	yes	yes	yes
worker fe	no	yes	no	yes

Note: Estimation results of model (3)-(4) in the text. The first stage dependent variable in columns (1) and (2) is the (log) weekly nominal wage; in columns (3) and (4) it is the (log) weekly real wage. Age dummies refer to 9 age categories: 18/24, 25/29, 30/34, 35/39, 40/44, 45/49, 50/54, 55/59 and 60/64. Contract dummies are for NCLAs, which cover about 243 contracts. Standard errors are clustered at the LLM level. *** p<0.01, ** p<0.05, * p<0.10.

Table 6: UWP estimates. Employees: controlling for unemployment rate

	Nominal wage				Real wage			
	(1) FULL OLS	(2) FE	(3) AKM	(4) IV-FE	(5) FULL OLS	(6) FE	(7) AKM	(8) IV-FE
	Second stage results							
log population density	0.006***	0.002*	-0.000	0.004*	-0.024***	-0.027***	-0.030***	-0.030***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.004)	(0.004)	(0.004)	(0.009)
unemployment rate	-0.013***	-0.007***	-0.002***	-0.007***	-0.004***	0.002***	0.007***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
year fe	yes	yes	yes	yes	yes	yes	yes	yes
Observations	6,558	6,558	6,528	6,558	6,558	6,558	6,528	6,558
R-squared	0.713	0.881	0.954		0.151	0.271	0.335	
K-P rk Wald F statistic				71.521				71.521
	First stage controls							
LLM-year dummies	yes	yes	yes	yes	yes	yes	yes	yes
age dummies	yes	yes	yes	yes	yes	yes	yes	yes
part time dummies	yes	yes	yes	yes	yes	yes	yes	yes
fixed term dummies	yes	yes	yes	yes	yes	yes	yes	yes
contract dummies	yes	yes	yes	yes	yes	yes	yes	yes
occupation dummies	yes	yes	yes	yes	yes	yes	yes	yes
firm size	yes	yes	yes	yes	yes	yes	yes	yes
worker fe	no	yes	yes	yes	no	yes	yes	yes
firm fe	no	no	yes	no	no	no	yes	no
Observations	77,015,891	77,015,891	77,015,891	77,015,891	77,015,891	77,015,891	77,015,891	77,015,891

Note: Estimation results of model (3)-(4) in the text. The first stage dependent variable in columns (1)-(4) is the (log) weekly nominal wage; in columns (5)-(8) it is the (log) weekly real wage. Age dummies refer to 9 age categories: 18/24, 25/29, 30/34, 35/39, 40/44, 45/49, 50/54, 55/59 and 60/64. Occupation dummies for for: white collar, blue collar, apprentice, manager, and executive; contract are dummies for NCLAs, which cover about 243 contracts. Standard errors are clustered at the LLM level. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table 7: UWP estimates. Independent contractors

	Nominal wage				Real wage			
	(1) FULL OLS	(2) FE	(3) AKM	(4) IV-FE	(5) FULL OLS	(6) FE	(7) AKM	(8) IV-FE
	Second stage results							
log population density	0.054***	0.006	0.018	0.020	0.026**	-0.019	-0.006	0.007
	(0.011)	(0.014)	(0.017)	(0.026)	(0.011)	(0.013)	(0.017)	(0.028)
year fe	yes	yes	yes	yes	yes	yes	yes	yes
Observations	6,344	5,635	5,554	5,441	6,344	5,635	5,554	5,441
R-squared	0.095	0.044	0.020		0.042	0.009	0.001	
K-P rk Wald F statistic				114.79				114.79
	First stage controls							
LLM-year dummies	yes	yes	yes	yes	yes	yes	yes	yes
age dummies	yes	yes	yes	yes	yes	yes	yes	yes
contractors dummies	yes	yes	yes	yes	yes	yes	yes	yes
industry dummies	yes	yes	yes	yes	yes	yes	yes	yes
firm size	yes	yes	yes	yes	yes	yes	yes	yes
worker fe	no	yes	yes	yes	no	yes	yes	yes
firm fe	no	no	yes	no	no	no	yes	no
Observations	4,475,347	4,475,347	4,475,347	4,475,347	4,475,347	4,475,347	4,475,347	4,475,347

Note: Estimation results of model (3)-(4) in the text. The first stage dependent variable in columns (1)-(4) is the (log) daily nominal wage; in columns (5)-(8) it is the (log) daily real wage. Age dummies refer to 9 age categories: 18/24, 25/29, 30/34, 35/39, 40/44, 45/49, 50/54, 55/59 and 60/64. Contractors are dummies for 19 categories of independent contractors; industry dummies stand for 86 2-digit Ateco industrial sectors. Standard errors are clustered at the LLM level. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table 8: UWP estimates. Standard self-employed

	Nominal wages	Real wages
	(1) FULL OLS	(2) FULL OLS
	Second stage results	
log population density	0.063***	0.037***
	(0.014)	(0.012)
year fe	yes	yes
Observations	6,558	6,558
R-squared	0.015	0.046
	First stage controls	
LLM-year dummies	yes	yes
age dummies	yes	yes
self-empl. w. dummies	yes	yes
Observations	4,154,141	4,154,141

Note: Estimation results of model (3)-(4) in the text. The first stage dependent variable in column (1) is the (log) yearly nominal wage; in column (2) it is the (log) yearly real wage. Age dummies refer to 9 age categories: 18/24, 25/29, 30/34, 35/39, 40/44, 45/49, 50/54, 55/59 and 60/64. Self-employed worker dummies are for the 14 categories of standard self-employed workers. Standard errors are clustered at the LLM level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 9: UWP: Employees and second level bargaining

	Nominal wage		Real wage	
	(1) FULL OLS below median	(2) FULL OLS above median	(3) FULL OLS below median	(4) FULL OLS above median
	Second stage results			
log population density	-0.003	0.001	-0.034***	-0.015***
	(0.004)	(0.004)	(0.004)	(0.005)
year fe	yes	yes	yes	yes
Observations	3,926	2,629	3,926	2,629
R-squared	0,368	0,410	0,156	0,090
	First stage controls			
LLM-year dummies	yes	yes	yes	yes
age dummies	yes	yes	yes	yes
part time dummies	yes	yes	yes	yes
fixed term dummies	yes	yes	no	yes
occupation dummies	yes	yes	no	yes
contract dummies	yes	yes	no	yes
firm size	yes	yes	yes	yes
Observations	23,170,816	22,926,044	23,170,816	22,926,044

Note: Estimation results of model (3)-(4) in the text. The first stage dependent variable in columns (1) and (2) is the (log) daily nominal wage; in columns (3) and (4) it is the (log) daily real wage. Age dummies refer to 9 age categories: 18/24, 25/29, 30/34, 35/39, 40/44, 45/49, 50/54, 55/59 and 60/64. Occupation dummies are for: white collar, blue collar, apprentice, manager, and executive; contract dummies are for NCLAs, which cover about 243 contracts. Columns (1) and (3) ((2) and (4)) refer to firms below (above) the median in terms of the incidence of second level bargaining associated with a wage premium, computed at a LLM-industry dimension (data source, *Rilevazione Imprese e Lavoro*, RIL). Standard errors are clustered at the LLM level. *** p<0.01, ** p<0.05, * p<0.10.

Appendix (not for publication)

A Institutional labour market framework and collective bargaining in Italy

Collective bargaining in Italy takes place between labour unions and employer organizations at the industry level. Apart from detailed working conditions, the collective national agreements settle, for each job-industry category, wage floors. This system has been introduced in the early 20th century, with the first company or territorial level collective agreements in manufacturing and agriculture, while the first nation-wide sectoral agreement goes back to the fifties. In the last decades, the number of contracts covered by collective bargaining increased over time. They almost cover the universe of private-sector employees (coverage rate is equal to 96%, employing data from the European Company Survey, and to 99%, using data from the Structure of Earnings Survey; see Garnero, 2018).

An industry collective agreement in Italy is *erga omnes*, i.e. it applies to all workers in that specific industry. This *erga omnes* extension is not stated formally in the labour law, but follows from the fact that wages set in collective agreements (*minimi tabellari*) are used by labour courts as benchmarks to establish whether firms comply with Article 36 of the Italian Constitution stating that "workers remuneration must be commensurate to quantity and quality of their work and, in any case, such as to ensure them and their families a free and dignified existence" (Garnero, 2018). Yet, non-compliance rates are non negligible: about 10% of workers are paid one fifth less than the reference minimum wage, with non-compliance being higher in the South and in micro and small firms, for women and temporary workers (again, see Garnero, 2018).

Apart from the industry/national bargaining, there is also a supplementary (non compulsory) decentralized bargaining, usually carried out at the local and/or firm level. At this level, it is possible to negotiate performance and productivity-related wage increases. In addition, the second level bargaining may address a number of additional issues, such as working hours, employment training, labor organization, and union relations. This second-level bargaining, however, is subject to the limits and provisions defined by the specific industry collective agreements: worker wages and labour conditions cannot be worse than the ones settled at the industry level. This is called *in melius* or favourability principle.

B The Italian housing market and the role of social housing

As many other European countries, Italian governments traditionally implement social housing policies that could, in principle, interfere with the spontaneous market forces of offer and demand. The European Union has no direct competence in the field, so policies of the various countries differ remarkably (see, e.g., Scanlon and Whitehead, 2015).

According to Italian national law (D.M. 22/04/2008), social housing refers to dwellings for residential use built or rehabilitated by public or private agents, also by means of public contributions or benefits (such as tax relieves, preferential treatments in city planning, etc), rented for at least eight years or sold at affordable price, to fulfill housing needs of socially disadvantaged individuals or groups of citizens.

Traditionally, social housing has consisted in three categories (see, e.g., Caruso, 2017): subsidized housing (*edilizia sovvenzionata*), that is dwellings owned by the public sector and rented at low rates to low-income individuals; assisted housing (*edilizia agevolata*), which includes dwellings provided mainly by cooperatives and given for rent or for sale to low-medium income people; agreed housing (*edilizia convenzionata*), which covers private houses offered for rent or for sale, whose price is regulated by agreements between the house owner and the municipality.

However, despite a long tradition of public intervention in the housing market, in particular, since the second world-war period, in the last 30 years Italian social housing policies have become weak and very discontinuous. Starting from the '90s, public investment in housing has fallen dramatically due to a sharp cut in public resources devoted to this aim. Furthermore, traditional housing policies have been largely substituted by subsidies to low-income families and other forms of facilitation to house owners (with almost no income limits). The number of new dwellings for assisted and agreed housing passed from 56,000 in 1984 to 11,000 in 2004, and that for subsidized housing from 34,000 (1984) to 1900 (2004) (Anci-Cresme, 2005). In 2015, social rent in Italy was about 5.5% in the total housing stock, that is a low percentage if compared, for instance, to 33% in Netherlands, 18,2% in UK, 17,5% in France (Pittini et al., 2015).

Summing up, it seems that the role of the public sector in the Italian housing market has become quite marginal and that the most recent trend in this field is that of substantially abandoning the social housing policies as they were traditionally conceived fifty years ago. Accordingly, issues related to social housing and public intervention in the housing markets do not seem to be relevant in affecting our results.

C Local CPI computation procedure

Istat CPI can be defined (Istat, 2016b) as:

$$CPI_{pt} = \sum_{n=1}^{N} \rho_{nt} PI_{npt}, \tag{A1}$$

where PI_{npt} is the price index of COICOP industrial category n for provincial capital city p in year t and ρ_{nt} is its relative weight at time t (equal across provincial capital cities). The CPI is equal to 100 in the base year, which is 2005 in the current analysis.

Aggregating the non-housing prices in a single category, Istat CPI can be seen as a weighted average of just two components, the housing price index (HPI) and the non-housing price index (NHPI), as follows:

$$CPI_{pt} = \rho_t HPI_{pt} + (1-\rho_t) NHPI_{pt}, \tag{A2}$$

where HPI_{pt} is the housing price index (correspondent to COICOP04) in province capital city p at time t, $NHPI_{pt}$ is the non-housing price index (all categories except COICOP04), and ρ_t is the official weight attached to housing prices.

This index, to be useful to our purposes, should be modified to capture local (LLM or municipality level) variability of prices. To do this, we adopt Moretti (2013)'s two-step methodology that exploits two empirical facts. First, housing prices are one of the main drivers of the spatial variation in the families' costs of living. Second, housing prices have a direct (cost of having an accommodation and maintaining it) and an indirect (attraction of housing prices on non-housing prices) role in households consumption baskets (as explained in Subsection 3.2); thus, the NHPI dynamics can be split into two components, one correlated with the dynamics of HPI and the other orthogonal to it.

Accordingly, we write:

$$NHPI_{pt} = \pi HPI_{pt} + \gamma_{pt}, \tag{A3}$$

where π captures the attraction effect of housing prices on non-housing prices and γ_{pt} is an error term.

By substituting (A3) in (A2), we get:

$$CPI_{pt} = [\rho_t + (1-\rho_t)\pi] HPI_{ct} + (1-\rho_t)\gamma_{pt}, \tag{A4}$$

that, defining $\lambda = [\overline{\rho} + (1-\overline{\rho})\pi]$, where $\overline{\rho}$ is the average of ρ_t over time, and $\varepsilon_{pt} = (1-\overline{\rho})\gamma_{pt}$, can be approximated by:

$$CPI_{pt} = \lambda HPI_{pt} + \varepsilon_{pt}. \tag{A5}$$

λ is the "adjusted" weight that incorporates the direct, $\overline{\rho}$, and the indirect, $(1-\overline{\rho})\pi$, effect of housing prices on the cost of living.

Hence, to compute the local CPI, in the first step, we estimate the "adjusted" weight by running the following equation regression:

$$CPI_{pt} = d_p + \lambda HPI_{pt} + \varepsilon_{pt}, \tag{A6}$$

where d_p are the province fixed effects, CPI_{pt} is Istat provincial consumer price index, with base year t=2005, and HPI_{pt} is the province average housing price index, also with base year t=2005, obtained as explained later in this subsection. We estimate equation (A6) by OLS for 103 provinces over the 2004-2015 period and obtain a coefficient $\hat{\lambda}$ equal to 0.33 (s.e. equal to 0.0336), which corresponds to a weight of about 33% to be compared with the official weight attached from Istat to housing costs, in the same period, that is 10%.

Then, in the second step, we employ the estimated "adjusted" weight, $\hat{\lambda}$, to compute the local CPI exploiting local variation in the housing prices as follows

$$\hat{CPI}_{mt} = \hat{\lambda} HPI_{mt} + (1 - \hat{\lambda}) NHPI_t, \tag{A7}$$

where HPI_{mt} is the housing price index in municipality m and year t and $NHPI_t$ is the national non-housing price index, with base year 2005. The price index obtained as just described is computed also at the LLM level as a population weighted average of the municipality index.

Housing data from Italian Revenue Agency (2015) (original sources of the survey data are housing agencies, estimates by Italian Revenue Agency, auctions, and courts) are detailed by semester (January-June/July-December), city district (central, semi-central, peripheral, suburban, and extra-urban), type of house ("villa" cottage, "abitazione signorile" expensive house, "abitazione civile" standard house, "abitazione economica" cheap house, "abitazione tipica del luogo" typical house), and house status (good, standard, poor). To compute the average price for a given municipality in a given year, we implement the following methodology, which is designed to purge the price data of composition effects. Accordingly, we first compute the residuals from an OLS equation regression of house prices on district, house type, status, and semester dummies; then, we take the average residuals for each year and municipality; finally, we add the average residuals to the Italian average of housing prices across municipalities and years. In the last two steps, the mean is weighted by local population size. The same procedure is repeated at the province level by year to obtain the provincial average also used in the analysis.[31]

Since consumer prices are expressed as index numbers (base year 2005, as we said), we also need to convert housing prices into indexes, HPI_{mt}. In order to exploit both the time and the territorial variation, we compute the housing price index, HPI_{mt}, in municipality m at time t as the ratio between the housing price in LLM c at time t and its population weighted average at time $t = 2005$ across LLM.

The methodology to compute non-housing price index follows Ichino et al. (2019) and, exploiting Istat (2016b) data, relies on the assumption that the price component that is orthogonal to housing prices displays uniform dynamics across different regional areas and consists in computing the weighted average of the national prices included in Istat CPI at the 1-digit COICOP classification level, excluding housing (COICOP04), weights being taken from Istat weighting scheme:

$$NHPI_t = \frac{1}{(1 - \rho_t)} \sum_{n=1}^{N} \rho_{nt} PI_{nt}, \tag{A8}$$

where PI_{nt} is the national price index of COICOP category in n and year t, and ρ_{nt} the relative weight.

[31] In a robustness check (see Appendix D), we also replicate our results employing an index of housing prices in which we do not purge the data of the composition effects as here illustrated (hence, the housing prices are population weighted averages of housing prices (*quotazioni*) as provided by OMI dataset, not the residuals).

D Robustness checks

In this appendix, we report an array of robustness checks to verify whether our conclusions are affected by our choices in the empirical strategy implemented so far. For the sake of space, we focus the FULL OLS specification for two main groups of workers: employees and independent contractors.[32]

To begin with, instead of considering the workforce aged 15-64, Table D.1 provides FULL OLS estimates, which includes both individual and worker characteristics, focusing on prime age male workers only (aged 25-50), often used as a reference group in the urban economics literature. Results are in line with those set out in Tables 3 and 7, with penalties in real terms for employees and positive premia for independent contractors, not only in terms of nominal wages but also in real terms.

Table D.1 about here.

In Table D.2 we investigate whether our results are affected by the choice of the earnings measure and use yearly wages instead of weekly/daily wages. By using yearly wages there emerge again a negative premium in real terms for employees and a positive but not significant premium for employees in nominal terms, while for independent contractors estimates of the wage elasticity to population density turn not to be significant, both in terms of nominal and real wages. The latter result could be explained by the fact that for independent contractors the employment relationship might be characterized by a discontinuity over the year that is not reflected in the yearly wage.

Table D.2 about here.

We further check whether our results hold when using more disaggregated spatial units of observation, going from 611 LLMs to more than 8,000 municipalities.[33] Results are shown in Table D.3 where we observe for employees a positive premium in nominal term and a real wage penalty. As for independent contractors, there emerge a substantial positive premium both in nominal and real terms. Thus, the conclusions drawn employing LLMs are largely corroborated.

Table D.3 about here.

In Tables D.4, we check the robustness of our previous findings to an alternative agglomeration measure, employment density. Our results are very close to those obtained when adopting population density.

Table D.4 about here.

[32] The results for standard self-employed workers are consistent with those reported in the main analysis and are available from the authors upon request.

[33] The number of individual observations can shrink in this exercise, because the number of missing data points for prices at the municipality level is slightly larger than for those at the LLM level. Indeed the number of observations for real wages are less than those for nominal wages.

In order to check whether our results depend on our measure of local CPI, we try two alternative approaches to compute it. With the former, we implement Moretti's methodology (as described in Subsection 3.2), without removing housing compositional effects. Results are displayed in Table D.5 and are consistent with previous conclusions.

Table D.5 about here.

The second measure of CPI relies on the official absolute poverty thresholds issued by Istat (2017). These thresholds are computed by macro-region (North, Center, South) and municipality size (less than 50 thousands inhabitants, between 50 and 250 thousand, and more than 250 thousand).[34] The descriptive statistics for 2015 are shown in Table D.6. In all the macro-regions, the poverty thresholds in municipalities with a population greater than 250 thousand are at least 10% higher than those in small municipalities. Of course, the basket of goods included in the computation of the absolute poverty thresholds may differ from a standard basket used to compute a CPI, although the two baskets are closely correlated.

Table D.6 about here.

We then compute an alternative CPI by computing the ratio between these poverty thresholds and the Italian average in the base year (2005). Using this index, we deflate nominal wages to derive our real wage measure. The estimation results, shown in Table D.7, are consistent with our conclusions from the baseline analysis.

Table D.7 about here.

We further explore whether our results are affected by the period considered. This might be an issue since housing prices increased in Italy until 2010 and decreased afterwards: our findings could be driven by dynamics at work in just one of these two time spans and, as a consequence, may not be generalized. Tables D.8 and D.9 respectively show for employees and independent contractors the estimated UWP for two alternative time intervals: 2005-2010 and 2011-2015. Findings are consistent across both periods for both categories of workers.

Table D.8 about here.

Table D.9 about here.

[34] Note that Istat also provides different thresholds for different households' size. Since the INPS archives do not include information on household size, we consider all workers in our dataset as one-person household. Results do not change when considering households of different size and are available upon request.

Table D.1: Prime age workers only (25-50)

	Employees		Independent Contractors	
	Nominal wages	Real wages	Nominal wages	Real wages
	(1) FULL OLS	(2) FULL OLS	(3) FULL OLS	(4) FULL OLS
	Second stage results			
log population density	-0.000	-0.027***	0.048***	0.020*
	(0.004)	(0.004)	(0.012)	(0.012)
year fe	yes	yes	yes	yes
Observations	4,675	4,675	4,188	4,188
R-squared	0.351	0.109	0.062	0.016
	First stage controls			
LLM-year dummies	yes	yes	yes	yes
age dummies	yes	yes	yes	yes
part time dummies	yes	yes	no	no
fixed term dummies	yes	yes	no	no
occupation dummies	yes	yes	no	no
contract dummies	yes	yes	no	no
firm size	yes	yes	yes	yes
contractors dummies	no	no	yes	yes
industry dummies	no	no	yes	yes
Observations	56,696,640	56,696,640	2,923,243	2,923,243

Note: Estimation results of model (3)-(4) in the text. Results for employees are shown in the first two columns, while the last two columns display those for independent contractors. The first stage dependent variable in columns (1) and (3) is the (log) of weekly/daily nominal wage; in columns (2) and (4) it is the (log) of weekly/daily real wage. Age dummies refer to 9 age categories: 25/29, 30/34, 35/39, 40/44, 45/50. Occupation dummies are for: white collar, blue collar, apprentice, manager, and executive; contract dummies are for NCLAs, which cover about 243 contracts. Industry dummies stands for 86 2-digit Ateco industrial sectors; contractors dummies are for the 19 categories of independent contractors. Standard errors are clustered at the LLM level. *** $p<0.01$, ** $p<0.05$, *$p<0.10$.

Table D.2: Yearly wages

	Employees		Independent Contractors	
	Nominal wages	Real wages	Nominal wages	Real wages
	(1) FULL OLS	(2) FULL OLS	(3) FULL OLS	(4) FULL OLS
	Second stage results			
log population density	0.005	-0.021***	0.012	-0.014
	(0.006)	(0.006)	(0.012)	(0.011)
year fe	yes	yes	yes	yes
Observations	6,558	6,558	6,344	6,344
R-squared	0.167	0.040	0.034	0.094
	First stage controls			
LLM-year dummies	yes	yes	yes	yes
age dummies	yes	yes	yes	yes
part time dummies	yes	yes	no	no
fixed term dummies	yes	yes	no	no
occupation dummies	yes	yes	no	no
contract dummies	yes	yes	no	no
firm size	yes	yes	yes	yes
contractors dummies	no	no	yes	yes
industry dummies	no	no	yes	yes
Observations	77,014,988	76,997,901	4,555,121	4,554,353

Note: Estimation results of model (3)-(4) in the text. Results for employees are shown in the first two columns, while the last two columns display those for independent contractors. The first stage dependent variable in columns (1) and (3) is the (log) of weekly/daily nominal wage; in columns (2) and (4) it is the (log) of weekly/daily real wage. Age dummies refer to 9 age categories: 18/24, 25/29, 30/34, 35/39, 40/44, 45/49, 50/54, 55/59 and 60/64. Occupation dummies are for: white collar, blue collar, apprentice, manager, and executive; contract dummies are for NCLAs, which cover about 243 contracts. Contractors dummies are for the 19 categories of independent contractors; industry dummies stand for 86 2-digit Ateco industrial sectors. Standard errors are clustered at the LLM level. *** p<0.01, ** p<0.05, * p<0.10.

Table D.3: Municipalities as spatial units of observation

	Employees		Independent Contractors	
	Nominal wages	Real wages	Nominal wages	Real wages
	(1) FULL OLS	(2) FULL OLS	(3) FULL OLS	(4) FULL OLS
	Second stage results			
log population density	0.009***	-0.015***	0.082***	0.058***
	(0.001)	(0.001)	(0.017)	(0.017)
year fe	yes	yes	yes	yes
Observations	81,912	81,912	32,923	32,923
R-squared	0.283	0.059	0.039	0.017
	First stage controls			
LLM-year dummies	yes	yes	yes	yes
age dummies	yes	yes	yes	yes
part time dummies	yes	yes	no	no
fixed term dummies	yes	yes	no	no
occupation dummies	yes	yes	no	no
contract dummies	yes	yes	no	no
firm size	yes	yes	yes	yes
contractors dummies	no	no	yes	yes
industry dummies	no	no	yes	yes
Observations	77,014,988	76,997,901	4,555,121	4,554,353

Note: Estimation results of model (3)-(4) in the text. Results for employees are shown in the first two columns, while the last two columns display those for independent contractors. The first stage dependent variable in columns (1) and (3) is the (log) of weekly/daily nominal wage; in columns (2) and (4) it is the (log) of weekly/daily real wage. Age dummies refer to 9 age categories: 18/24, 25/29, 30/34, 35/39, 40/44, 45/49, 50/54, 55/59 and 60/64. Occupation dummies are for: white collar, blue collar, apprentice, manager, and executive; contract dummies are for NCLAs, which cover about 243 contracts. Contractors dummies are for the 19 categories of independent contractors; industry dummies stand for 86 2-digit Ateco industrial sectors. The number of observations between nominal and real wages estimates slightly differs due to the presence of missing data in the cpi computation for some municipalities. Standard errors are clustered at the LLM level. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table D.4: Employment density

	Employees		Independent Contractors	
	Nominal wages	Real wages	Nominal wages	Real wages
	(1) FULL OLS	(2) FULL OLS	(3) FULL OLS	(4) FULL OLS
	Second stage results			
log of population density	0.020*** (0.002)	-0.013*** (0.003)	0.068*** (0.008)	0.035*** (0.009)
year fe	yes	yes	yes	yes
Observations	6,558	6,558	6,344	6,344
R-squared	0.436	0.075	0.133	0.044
	First stage controls			
LLM-year dummies	yes	yes	yes	yes
age dummies	yes	yes	yes	yes
part time dummies	yes	yes	no	no
fixed term dummies	yes	yes	no	no
occupation dummies	yes	yes	no	no
contract dummies	yes	yes	no	no
firm size	yes	yes	yes	yes
contractors dummies	no	no	yes	yes
industry dummies	no	no	yes	yes
Observations	77,015,891	77,015,891	4,475,347	4,475,347

Note: Estimation results of model (3)-(4) in the text. Results for employees are shown in the first two columns, while the last two columns display those for independent contractors. The first stage dependent variable in columns (1) and (3) is the (log) of weekly/daily nominal wage; in columns (2) and (4) it is the (log) of weekly/daily real wage. Age dummies refer to 9 age categories: 18/24, 25/29, 30/34, 35/39, 40/44, 45/49, 50/54, 55/59 and 60/64. Occupation dummies are for: white collar, blue collar, apprentice, manager, and executive; contract dummies are for NCLAs, which cover about 243 contracts. Industry dummies stands for 86 2-digit Ateco industrial sectors; contractors dummies are for the 19 categories of independent contractors. Standard errors are clustered at the LLM level. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table D.5: Housing prices without removing compositional effects

	Employees		Independent Contractors	
	Nominal wages	Real wages	Nominal wages	Real wages
	(1) FULL OLS	(2) FULL OLS	(3) FULL OLS	(4) FULL OLS
	Second stage results			
log population density	-0.000	-0.021***	0.054***	0.025**
	(0.004)	(0.003)	(0.011)	(0.011)
Observations	6,558	6,558	6,344	6,344
R squared	0.382	0.111	0.095	0.032
	First stage controls			
LLM-year dummies	yes	yes	yes	yes
age dummies	yes	yes	yes	yes
part time dummies	yes	yes	no	no
fixed term dummies	yes	yes	no	no
occupation dummies	yes	yes	no	no
contract dummies	yes	yes	no	no
firm size	yes	yes	yes	yes
contractors dummies	no	no	yes	yes
industry dummies	no	no	yes	yes
Observations	77,015,891	77,015,891	4,475,347	4,475,347

Note: Estimation results of model (3)-(4) in the text. Results for employees are shown in the first two columns, while the last two columns display those for independent contractors. Columns (1) and (3) replicate results for nominal wages of employees and independent contractors, respectively presented in column (3) of Table 3 and column (1) of Table 7. Columns (2) and (4) report the corresponding coefficients obtained by employing weekly/daily real wages computed without removing compositional effect from the housing prices. Age dummies refer to 9 age categories: 18/24, 25/29, 30/34, 35/39, 40/44, 45/49, 50/54, 55/59 and 60/64. Occupation dummies are for: white collar, blue collar, apprentice, manager, and executive; contract dummies are for NCLAs, which cover about 243 contracts. Contractors dummies are for the 19 categories of independent contractors; industry dummies stand for 86 2-digit Ateco industrial sectors. Standard errors are clustered at the LLM level. *** p<0.01, ** p<0.05, * p<0.10.

Table D.6: Absolute poverty thresholds (Istat) by macro-regions and municipality size (year 2015)

	North	Centre	South and Islands	North	Centre	South and Islands
Up to 50 thousands	734.74	699.49	552.39	1.00	1.00	1.00
From 50 to 250 thousands	779.97	746.44	588.52	1.06	1.07	1.07
Above 250 thousands	819.13	787.1	609.28	1.11	1.13	1.10

Note: The threshold refers to families with one member only.

Table D.7: Alternative local CPI computed using poverty thresholds

	Employees		Independent Contractors	
	Nominal wages	Real wages	Nominal wages	Real wages
	(1) FULL OLS	(2) FULL OLS	(3) FULL OLS	(4) FULL OLS
	Second stage results			
log population density	-0.000	-0.014***	0.054***	0.042***
	(0.004)	(0.003)	(0.011)	(0.011)
Observations	6,558	6,558	6,344	6,344
R squared	0.382	0.040	0.095	0.040
	First stage controls			
LLM-year dummies	yes	yes	yes	yes
age dummies	yes	yes	yes	yes
part time dummies	yes	yes	no	no
fixed term dummies	yes	yes	no	no
occupation dummies	yes	yes	no	no
contract dummies	yes	yes	no	no
firm size	yes	yes	yes	yes
contractors dummies	no	no	yes	yes
Observations	77,015,891	77,015,891	4,475,347	4,475,347

Note: Estimation results of model (3)-(4) in the text. Results for employees are shown in the first two columns, while the last two columns display those for independent contractors. Columns (1) and (3) replicate results for nominal wages of employees and independent contractors, respectively presented in column (3) of Table 3 and column (1) of Table 7. Columns (2) and (4) report the corresponding coefficients obtained by employing weekly/daily real wages computed using Istat poverty thresholds. Age dummies refer to 9 age categories: 18/24, 25/29, 30/34, 35/39, 40/44, 45/49, 50/54, 55/59 and 60/64. Occupation dummies are for: white collar, blue collar, apprentice, manager, and executive; contract dummies are for NCLAs, which cover about 243 contracts. Contractors dummies are for the 19 categories of independent contractors; industry dummies stand for 86 2-digit Ateco industrial sectors. Standard errors are clustered at the LLM level. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table D.8: Different time periods: Employees

	2005-2010		2011-2015	
	Nominal wages	Real wages	Nominal wages	Real wages
	(1) FULL OLS	(2) FULL OLS	(3) FULL OLS	(4) FULL OLS
	Second stage results			
log of population density	0.000	-0.028***	-0.000	-0.023***
	(0.004)	(0.004)	(0.004)	(0.004)
year fe	yes	yes	yes	yes
Observations	3,581	3,581	2,977	2,977
R-squared	0.249	0.110	0.053	0.079
	First stage controls			
LLM-year dummies	yes	yes	yes	yes
age dummies	yes	yes	yes	yes
part time dummies	yes	yes	no	no
fixed term dummies	yes	yes	no	no
occupation dummies	yes	yes	no	no
contract dummies	yes	yes	no	no
firm size	yes	yes	yes	yes
contractors dummies	no	no	yes	yes
industry dummies	no	no	yes	yes
Observations	42,881,163	42,881,163	34,134,728	34,134,728

Note: Estimation results of model (3)-(4) in the text. The first stage dependent variable in columns (1) and (3) is the (log) of weekly nominal wage; in columns (2) and (4) it is the (log) of weekly real wage. Age dummies refer to 9 age categories: 18/24, 25/29, 30/34, 35/39, 40/44, 45/49, 50/54, 55/59 and 60/64. Occupation dummies are for: white collar, blue collar, apprentice, manager, and executive; contract dummies are for NCLAs, which cover about 243 contracts. Standard errors are clustered at the LLM level. *** p<0.01, ** p<0.05, * p<0.10.

Table D.9: Different time periods: Independent contractors

	2005-2010		2011-2015	
	Nominal wages	Real wages	Nominal wages	Real wages
	(1) FULL OLS	(2) FULL OLS	(3) FULL OLS	(4) FULL OLS
	Second stage results			
log of population density	0.067***	0.040***	0.040***	0.016
	(0.012)	(0.012)	(0.013)	(0.013)
year fe	yes	yes	yes	yes
Observations	2,615,303	2,615,303	1,860,044	1,860,044
R-squared	0.056	0.059	0.009	0.007
	First stage controls			
LLM-year dummies	yes	yes	yes	yes
age dummies	yes	yes	yes	yes
part time dummies	yes	yes	no	no
fixed term dummies	yes	yes	no	no
occupation dummies	yes	yes	no	no
contract dummies	yes	yes	no	no
firm size	yes	yes	yes	yes
contractors dummies	no	no	yes	yes
industry dummies	no	no	yes	yes
Observations	2,615,303	2,615,303	1,860,044	1,860,044

Note: Estimation results of model (3)-(4) in the text. The first stage dependent variable in columns (1) and (3) is the (log) of daily nominal wage; in columns (2) and (4) it is the (log) of daily real wage. Age dummies refer to 9 age categories: 18/24, 25/29, 30/34, 35/39, 40/44, 45/49, 50/54, 55/59 and 60/64. Occupation dummies are for: white collar, blue collar, apprentice, manager, and executive; contract dummies are for NCLAs, which cover about 243 contracts. Standard errors are clustered at the LLM level. *** p<0.01, ** p<0.05, * p<0.10.