INPS

Istituto Nazionale Previdenza Sociale

settembre 2018 – numero 13

WorkINPS *Papers*

Urban Wage Premia, Cost of Living, and Collective Bargaining

Marianna Belloc

Paolo Naticchioni

Claudia Vittori

ISSN 2532 -8565

In copertina: uno storico "Punto cliente" a Tuscania
INPS, Direzione generale, Archivio storico

I WORK*INPS* PAPER

Le basi dati amministrative dell' *INPS* rappresentano una fonte statistica unica per studiare scientificamente temi cruciali per l' economia italiana, la società e la politica economica: non solo il mercato del lavoro e i sistemi di protezione sociale, ma anche i nodi strutturali che impediscono all'Italia di crescere in modo adeguato. All' interno dell'Istituto, questi temi vengono studiati sia dai funzionari impiegati in attività di ricerca, sia dai *VisitInps Scholars*, ricercatori italiani e stranieri selezionati in base al loro curriculum vitae e al progetto di ricerca presentato.

I WORK*INPS* hanno lo scopo di diffondere i risultati delle ricerche svolte all' intern dell'Istituto a un più ampio numero possibile di ricercatori, studenti e policy makers. Questi saggi di ricerca rappresentano un prodotto di avanzamento intermedio rispetto alla pubblicazione scientifica finale, un processo che nelle scienze sociali può richiedere anche diversi anni. Il processo di pubblicazione scientifica finale sarà gestito dai singoli autori.

Pietro Garibaldi

Urban Wage Premia, Cost of Living, and Collective Bargaining

Marianna Belloc
(Sapienza University of Rome)

Paolo Naticchioni
(University of Roma Tre and INPS)

Claudia Vittori
(Sapienza University of Rome)

Urban Wage Premia, Cost of Living, and Collective Bargaining*

Marianna Belloc† Paolo Naticchioni‡ Claudia Vittori§

July 2018

Abstract

In this paper, we estimate the urban wage premia (UWP) in Italy, with its economy characterized by the interplay between collective bargaining and spatial heterogeneity in the cost of living. We implement a reduced-form regression analysis using both nominal and real (in temporal and spatial terms) wages. Our dataset for the 2005-2015 period includes, for workers' characteristics, unique administrative data provided by Italian Social Security Institute and, for the local CPI computation, housing prices collected by 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 (-5%). To capture the role played by centralized wage settings, we also 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 the UWP for self-employed workers are up to 25 times greater than for employees. Moreover, sorting proves more notable in the case of self-employed workers, i.e. the larger UWP provide the higher incentives for high-skilled individuals and better firms to locate in cities. Our findings are confirmed on extending the analysis along the wage distribution.

JEL codes: R12, R31, J31
Keywords: Urban Wage Premium, Cost of Living, Collective Bargaining

*We thank Gaetano Basso, Tito Boeri, Lorenzo Cappellari, David Card, Luca Citino, Pierre-Philippe Combes, Edoardo Di Porto, Gilles Duranton, Maurizio Franzini, Pietro Garibaldi, Andrea Garnero, Andrea Lamorgese, Claudio Lucifora, Michele Raitano, Paolo Sestito, Roberto Torrini, Eliana Viviano, Josef Zweimller, and participants in seminars at IZA, Bank of Italy, INPS, AIEL, University of Ancona, in yearly conferences of AIEL and ECINEC, and in the workshop on "Spatial Dimension of Inequality" at ZEW for comments and suggestions. Paolo Naticchioni is also grateful to VisitINPS programme, and its sponsors, for financial support and for access to INPS administrative data. Usual disclaimers apply. The views expressed in the article are those of the authors and do not involve the respective institutions.

†Department of Economics and Law, Sapienza University of Rome. Email: marianna.belloc@uniroma1.it.

‡Department of Political Science, University of Roma Tre, and INPS. Email: p.naticchioni@gmail.com.

§Department of Economics and Law, Sapienza University of Rome. Email: claudia.vittori@uniroma1.it

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 in the relevant works on this topic: 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 wages and on the estimation of 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 the 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 UWP differing between real and spatial terms. This is likely to affect location choices and sorting of firms and workers and, in turn, have an impact on firm dynamics and economic growth, especially when cities are characterized by urbanization externalities.

As far as we know, there is only one paper, by Boeri et al. (2017), combining collective bargain-

[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 the IMF Country Report No. 16/222 on the Italian economy, and from independent observers such as Boeri (2015).

ing and local variability in cost of living, from both the theoretical and empirical points of view. In their theoretical model, workers' utility functions include nominal wages, cost of living, and unemployment. One of the predictions is that, in equilibrium and in the presence of collective bargaining (equal nominal wages across locations), the cost of living is higher in higher productivity locations and, hence, real wages are lower. In order to derive the equality of utility functions across space, the unemployment rate must be smaller in high productivity areas. This prediction is confirmed by the empirical analysis applied to regional unbalances in Italy (between the South and the North) and in Germany (between the East and the West), the former characterized by the greater rigidity of the collective bargaining system.

As Boeri et al. (2017), we also focus on Italy, considering the agglomeration dimension rather than comparisons across macro-regions. Agglomeration is a key feature in investigation of the variance of Italian wages: in this country, 95.3% of the variance of nominal wages and 95.7% of the variance of real wages is within regions (INPS data 2015, see later).

Italy represents a perfect case to study in this context. One reason is that, differently from the US, where the 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 wage differences of similar size. 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 two-tier system is also characterized by (non-compulsory) decentralized second-level bargaining, which plays only a minor role. Second-level bargaining is subject to the *in melius* or favourability principle: workers wages and conditions cannot be worse than those agreed at the industry level.

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 Security Institute (INPS), issued through the VisitINPS programme (2015). These data cover the universe of Italian workers, including both an employer-employee dataset, with detailed individual and firm characteristics, and a dataset on various groups of self-employed workers. For all groups, the INPS archives provide 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 covers information on housing transactions detailed at the municipality level.

Our empirical analysis is based on a well-established reduced-form approach for estimation of the UWP (see, for instance, Combes et al., 2011). In particular, we use a regression analysis at the individual level for the 2005-2015 period to estimate the UWP for both nominal and real wages.

Individual real wages are computed by deflating nominal wages by a local CPI obtained using local variation in housing prices, after removing composition effects, according to a procedure developed by Moretti (2013).[2] Our covariate of interest is the (log) population density of the local labour markets (LLM) where the job is located.[3]

When considering nominal wages, our results confirm the presence of a negligible nominal UWP. The wage elasticity with respect to population density is 4.6% when no controls are included, which proves very 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 composition of workers and firms differs greatly along the density distribution. 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 -5.1%.[4]

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 and Klein, 2013; Dauth et al., 2016). However, our results show that sorting of individuals only slightly dampens the UWP estimate, and adding firm-fixed effects does not substantially change it.

Following the theoretical prediction by Boeri et al. (2017), we also verify whether the estimated urban wage penalty reflects a higher probability of being employed in more densely populated areas. Our evidence does not support this conclusion: the unemployment rate does not change much along the population density distribution, while it plays a clear role in the comparison across macro-areas (Boeri et al., 2017).

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

[2]From now on, we will refer to real wages in the sense of wages deflated both temporally and spatially.

[3]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.

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

enjoying the same local amenities as employee.[5] To the best of our knowledge, our work is one of the first providing accurate separate UWP estimates for employees and self-employed workers.

Our preferred self-employed group is that of independent contractors (so-called *collaborazioni* or *parasubordinati*). This can be considered a peculiar category of self-employed workers closely associated with a firm, such as external staff and/or consultants, and may be either skilled or unskilled. 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 in nominal terms, including all possible controls, is up to 25 times greater than that found for employees (around 5%). By contrast, in real terms, the UWP falls to zero, with higher wages that, on average, offset higher costs of living. Interestingly, once worker and firm fixed effects are included, a more pronounced drop in the UWP for independent contractors is observed than for employees. 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 and firms to sort into agglomerated areas.

The differences in UWP estimates between employees and self-employed workers are widely confirmed when considering the standard self-employed workers, such as (the high-skilled) business consultants, lawyers, physicians, and architects, (and the medium-skilled) journalists, surveyors, and accountants. Overall, UWP estimates are around 5% in nominal terms and zero in real terms, very close to the estimates obtained for the group of independent contractors. More importantly, by considering various groups of self-employed workers, characterized by different self-selection processes, we provide evidence that self-selection between employees and self-employed workers can hardly constitute the main factor in differences in the UWP estimates between the two groups.

To sum up, our results suggest that there are significant differences in estimated UWP between employees (covered by collective bargaining) and self-employed workers (independent contractors and standard self-employed). When controlling for all individual and firm characteristics, the difference in UWP between the two groups, in both nominal and real terms, is around 5% (in nominal terms, the UWP is 5% for self-employed and zero for employees; in real terms, it is zero for self-employed and -5.1% for employees). Collective bargaining can be considered a driving force behind such differences in the presence of heterogeneity in the cost of living, since workers in the two groups enjoy the same amenities and live in the same locations. Nonetheless, the way such fac-

[5]Note that an additional advantage of investigating the local rather than the macro-regional dimension is that it allows us to neglect structural and cultural differences across units of observation. They would, instead, be relevant when comparing the South and the North of Italy, as in Boeri et al. (2017). By contrast, within-region local differences could be considered negligible.

tors enter the utility functions of, respectively, employees and self-employed workers might have a role in explaining the different equilibria for the two groups. For instance, the fact that real UWP are negative and substantial for employees suggests that individuals 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.

In the last part of the paper, we investigate how incentives in location choices may vary along the wage distribution, e.g. between skilled and unskilled workers, by employing unconditional quantile regressions. One might expect decentralized (second-level) bargaining at the firm/local level to affect workers at the tails of the wage distribution. 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). Our results show that this is not the case: UWP differences between employees and self-employed workers observed at the conditional mean are also found along different percentiles of the wage distribution, suggesting that second-level decentralized bargaining is not playing a major role in Italy, consistently with findings by Boeri (2015).

Finally, an array of robustness checks are performed and previous conclusions are, again, widely confirmed. First, we focus on a number of distinct categories of workers. 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 different agglomeration measures: employment density and binary dummies for urban areas. Fifth, we use two alternative measures of local CPI: a CPI computed without removing compositional housing effects, as we do in the main analysis, and a CPI based on the official absolute poverty thresholds provided by Istat (Istat, 2017).

The remainder 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 and 6 present, respectively, the econometric strategy and the empirical results. Sections 7 and 8 investigate the role of collective wage bargaining, while Section 9 offers an analysis based on unconditional quintile regressions. Section 10 provides a wide range of robustness checks and Section 11 draws concluding remarks.

2 Related literature

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

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 nation, between cities and countryside, and across groups of municipalities of different sizes. There are few exceptions. 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).

As for the US, 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. Also employing Moretti (2013)'s methodology to compute the local CPI for the US, Hornbeck and Moretti (2018) find that, when a city enjoys productivity gains in manufacturing, local employment and average earnings tend to increase. For renters, increased earnings are largely offset by higher costs of living. The combination of direct and indirect effects implies that the 1980-1990 TFP growth heightened the purchasing power of the average worker by about 0.5-0.6% per year between 1980 and 2000.

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. 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 Boeri et al. (2017). Their baseline model assumes constant labour supply, perfect labour mobility, and homogeneous propensities for locations. The utility function incorporates a trade-off between real wages, the ratio of nominal wages out of local prices, and unemployment. For simplicitys sake, it is assumed that housing prices are proportional to the region population. Finally, total factor productivity (TFP) is allowed to differ across regions. The authors show that, in the presence of competitive labour markets, there is no unemployment, and nominal wages and employment are higher (lower) in regions with higher (lower) TFP. Since the cost of living is higher where employment and TFP are higher, real wages are constant in the space dimension and this guarantees equilibrium, i.e. workers are indifferent across regions. Collective bargaining is introduced into the model by imposing equal nominal wages across locations and at a higher level than the competitive one. In this framework, employment is also higher in the high TFP regions, entailing a higher cost of living and, *ceteris paribus*, a lower real wage. Hence, to achieve the equality of utilities across locations, the unemployment rate has to adjust, i.e. it increases in lower productivity areas. According to this model, the winners are workers employed in low productivity locations and house owners in high productivity locations.

In their empirical analysis, Boeri et al. (2017) study wage differentials across macro-regions in Italy and Germany, with the former characterized by greater rigidity in the collective bargaining system. Employing a local CPI index *á la* Moretti (2013), they find that the unbalances between the North and the South of Italy are much more pronounced than those between the East and the West of Germany, due to the fact that in Germany the collective bargaining system has been reformed and relaxed over the last few decades. Consistently with the model, real wages are lower in the North of Italy and unemployment is higher in the South. However, Boeri et al. (2017) do not deal with agglomeration issues, since they are more interested in comparisons across macro-regions.

Our paper also concerns 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;

Boeri, 2015 and country reports by European Commission and IMF). Making comparison among two-tier, fully centralized, and fully decentralized regimes, Boeri (2015) stresses the role played by plant-level bargaining and the importance of wage floors provided by statutory minimum wages.

Finally, one more strand of 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 the US indicating that households pay substantially to live in areas with coastal proximity, mild temperatures, and sunshine. 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 the cost of lower real wages 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 sets 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 local economic and demographic 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, launched by Tito Boeri (INPS President) in 2015, which allows selected researchers to access richer data than previous INPS issues on the universe of Italian workers.[6] 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 a heterogeneous group 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,

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

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 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.[7]

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);[8] the latter 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 (there are 611 LLM in Italy and 8,000 municipalities). This choice is guided by the fact that we have knowledge only of the municipality where people work, 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 education services. Hence, 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 of assigning to the worker a more appropriate local cost of living.

[7]We crop outliers in the wage distribution and, 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 prove imprecise.

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

3.2 Local consumer price index

In Italy, Istat releases local consumer price index data in the capital cities of 80 out of 103 Italian provinces (see Istat 2016b).[9] Prices are aggregated according to COICOP industrial classification,[10] each category being weighted according to the Istat weighting scheme which reflects the relative importance of the various goods and services in families' consumption basket.[11]

As also stressed by Boeri et al. (2017), this price index is not adequate for our purposes to compute real wages for two reasons: first, it is measured at the province (and not LLM or municipal) level and, second, the weight attached to the housing costs in CPI aggregation only reflects their direct effect on the families' cost of living and does not take into account their attraction (indirect) effect on the dynamics of non-housing prices.

The direct effect reflects the cost of having an accommodation and maintaining it. The indirect effect captures the fact that the prices of the other goods and services can also be influenced by the housing prices. This implies that in cities where land is more expensive, the cost of many goods and services will also be higher. For example, 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 (higher rents and housing costs, for instance). In order to give local housing prices their correct weight in the family consumption basket, these two effects (direct and indirect) should be captured by the weight attached to them in the CPI aggregation. Relying on Moretti (2011)'s methodology, we implement a two-step strategy (for details, see Appendix C), as follows.

The first consists in estimating the "adjusted" weight of housing costs, including both direct and indirect effects, which we will denote by β. In our analysis, the estimated "adjusted" weight, $\hat{\beta}$, for the 2004-2015 period comes to 33% as opposed to the official weight attached to housing costs by Istat, which is 10% for the same period. The second step, employing $\hat{\beta}$, consists in computing the local CPI by exploiting the local variation in the housing prices. Accordingly, the CPI at the LLM level is defined as follow:

$$\hat{CPI}_{lt} = \hat{\beta} HPI_{lt} + (1-\hat{\beta}) NHPI_t, \tag{1}$$

where HPI_{lt} is the housing price index in LLM l and in year t and $NHPI_t$ is the national non-housing price index, both indexes with base year 2005. The CPI at the LLM level is obtained as

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

[10] For instance, the one-digit level identifies macro-categories of goods and services, such as food and non-alcoholic drinks (01), alcoholic drinks and tobacco (02), apparel and shoes (03), housing (04), and so on.

[11] 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) instead.

a population-weighted average of the municipality index. Housing price indexes and non-housing price indexes in (1) are drawn from two different sources.

The former (HPI_{lt}) are obtained from the OMI (*Osservatorio del Mercato Immobiliare*) dataset, collected by Italian Revenue Agency (2015).[12] 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.

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 CPI 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. (2016) point out that the elasticity of urban costs increases more than proportionally with a city population. Should the CPI measure not capture this non-linearity, it would underestimate the real cost of living in cities. For these reasons, we carry out several robustness checks on the local CPI. Among other things (see Section 10), 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 population density (*POPDENS*) 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: 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 and binary variables for urban areas (see Section 10). Data on municipality population and city size in square kilometres are from Istat (2016c). Finally, in the instrumental

[12]For robustness, we also estimate β 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 β turns out to be 54%, which is even greater than the "adjusted" weight found for housing property prices.

variable estimation (discussed in Section 5), we also employ municipal population back in 1871 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 start with, Figure 1 displays local (LLM level) CPI on the vertical axis and 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.

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 Boeri et al. (2017). Nonetheless, a clear agglomeration pattern also emerges: in highly agglomerated areas, the local CPI is greater and falls 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 will be shown in Section 10).

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 and real wages. It turns out that (for both) around 95% of the variance is due to a within-region 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 euro in real terms. By contrast, in the lowest quintile real wages are 446 euro, the nominal wage 401 euro.

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.

The descriptive evidence gathered so far provides a preview of the results which we will be presenting in the rest of the paper: first, the local CPI is highly heterogeneous across areas and clearly increases with population density; second, the distribution of real wages along the space dimension is much more compressed than that of nominal wages.

5 The econometric strategy

To estimate the UWP, we make use of standard reduced-form wage regressions, in line with a well-established tradition in the urban economics literature. In particular, when using the universe of employees, we regress the logarithm of weekly (nominal or real) wages on population density and a series of individual and firm-level control variables. The spatial unit of observation is the LLM, as already pointed out. In the robustness checks, we replicate our analysis also employing municipalities (see Section 10).

Our baseline equation regression is as follows:

$$log(W_{ilt}) = \alpha + \rho log(POPDENS_{lt}) + \eta_1 X_{ilt} + \eta_2 Z_{jlt} + \delta_t + \theta_p + (\lambda_i + \mu_j) + \varepsilon_{ilt} \quad (2)$$

where $log(W_{ilt})$ is the logarithm of the real or nominal wage of individual i working in LLM l at time t, $log(POPDENS_{lt})$ is the logarithm of the population density in LLM l at time t. The vector of individual controls, X_{ilt}, includes, for worker i in LLM l at time t: having a part-time contract, having a fixed-term contract, occupation dummies (blue collar, white collar, manager, apprentice, executive), and annual dummies for age. The vector of firm-level controls, Z_{jlt}, includes, for firm j in LLM l at time t: firm size (in log) and dummies for the associated NCLA (around 250 dummies), which capture the enforcement of the centralized national wage setting.[13] δ_t are year fixed effects, which control for business cycle shocks affecting all workers, and θ_p are province fixed effects, which control for unobserved time-invariant heterogeneity across provinces.[14] Regression model (2) is estimated by OLS and $\widehat{\rho}$ is the estimate of the wage elasticity to agglomeration, i.e. the urban wage premium (doubling population density increases wages by ρ).

We also consider the role of unobserved heterogeneity of workers, a well-known issue in urban economics (Mion and Naticchioni, 2009; Combes et al., 2008). To deal with this issue the literature has mainly used fixed effect estimations, which capture individual unobserved heterogeneity and its correlation with the variables of interest. Hence, in some specifications, we include individual fixed effects, λ_i. The importance of firm fixed effects in explaining wage differentials is a more recent concern. Card and Klein (2013) have pointed out that workplace heterogeneity is one of the main drivers of the increasing inequality in the case of Germany; Dauth et al. (2016) show that sorting of firms applies also in a spatial dimension and its impact on wage inequality increases over time. Consistently, we add firm fixed effects, μ_j, in some of our regressions as in the two-way fixed effects AKM models (Aboud 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. In the relevant literature, it is common to adopt historical series of population density as instruments, such as in Ciccone and Hall (1996) and Mion and

[13] Using industry dummies instead of NCLA dummies does not alter our main findings. The results are available on request.

[14] We adopt 1974 Istat classification as in INPS archives.

Naticchioni (2009). In accordance with this methodology, we use as instrument population density dating back to 1871, computed using population data in 1871 provided by Istat (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. Furthermore, 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 ones. For this reason, we will carry out IV estimates only for our baseline specification.

6 Main regression results

Table 3 shows the OLS regression results at the LLM level for the universe of male Italian employees. Estimates obtained using nominal and real wages respectively are shown in columns (1)-(3) and (4)-(6). In column (1), only the main covariate of interest, LLM population density, is included, along with time dummies that control for the business cycle. The corresponding estimated elasticity amounts to 4.6, which means that doubling the population density increases nominal wages by 4.6%. 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 worker's observable characteristics in estimation: having a part-time contract, having a fixed term contract, age, occupation, and province dummies.[15] As can be seen, controlling for worker observable characteristics dampens the elasticity to 0.6%. In column (3), we add firm level observable characteristics, such as (log) firm size and NCLA dummies (called FULL OLS specification from now on). The wage elasticity to urban agglomeration comes very close to zero, amounting to 0.2%, and not statistically significant. 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 controlling for firm and worker characteristics, displayed in column (4), is very close to zero (equal to -0.006) and not statistically significant. When adding individual controls, as we do in column (5), the estimated UWP is negative (-0.041) and highly statistically significant: individuals working in agglomerated areas suffer a wage penalty in real terms. Introducing observable firm characteristics, as shown in column (6), increases the wage penalty to 5.1%.

[15] In Section 10, we also run a robustness check in which we introduce the more demanding specification using LLM fixed effects, applying the two-step methodology suggested by Combes et al. (2008) and we obtain consistent results.

Table 3 about here.

One might think that worker heterogeneity, mainly in terms of the worker's occupation, could play a role in the UWP estimation: consistently with the literature, UWP are supposed to be larger for high-skilled workers, who are more likely to benefit from agglomeration externalities. To investigate on this, Table 4 shows results from baseline regressions estimated separately for blue-collar employees, white-collar employees, and managers/executives. Interestingly, we find, on the one hand, that in FULL OLS, the nominal UWP for white collar workers and manager/executives is higher than that obtained for blue collars (1.4% vs zero). On the other hand, when going on to real wages the UWP is still negative and substantial for all groups, and comes at about 5% (the lowest penalty is still for white collar worker, 4.5%). Hence, we can confidently state that our main findings (small nominal wage premia and strong real wage penalties for urban employees) do not depend on the workers' occupations, since they apply to the various subgroups of workers.[16]

Table 4 about here.

6.1 Worker and firm sorting and endogeneity issues

In this section, we investigate the role played by sorting of workers and firms and address matters of endogeneity. Column (1) of Table 5 reports our estimation output after introducing individual fixed effects (this is referred to as the FE specification hereafter). By comparing these results with those set out in column (3) of Table 3, we can note slight changes in the point estimates of the UWP, from 0.2% to zero, both not statistically significant. In column (2) of Table 5, we present our estimates when employing both individual and firm fixed effects (AKM specification). The estimated UWP still proves very close to zero, at 0.2%, and becomes statistically significant at the 10% level. Similar conclusions are obtained when considering real wages: introducing individual and firm fixed effects does not change the UWP point estimate much with respect to the FULL OLS specification, as can be seen in column (6) of Table 3.

The evidence offered by fixed effects and AKM estimates, for both nominal and real wages, reveals that sorting of workers and firms is not playing a major role in the determination of the UWP of employees, and that most of the wage variation across space is due to observable worker and firm characteristics.

Finally, we turn to addressing the matters of endogeneity and implement IV estimates. As an instrument for population density, we employ LLM population density in 1871. The results

[16] Heterogeneity of workers characteristics will also be addressed in Section 9 by using unconditional quantile regressions, with results consistent with those presented here.

obtained after including individual fixed effects are displayed in column (3) and (6) of Table 5 for respectively nominal and real wages (IV-FE). The instrument used in the estimation is not weak, as suggested by the Kleibergen-Paap Wald rk statistics (see Kleibergena and Paap, 2006). Again, the elasticity of nominal wages to the agglomeration measure is very close to zero, whereas that of real wages becomes greater, pointing at a non-negligible increase in the real penalty for urban employees.

Table 5 about here.

6.2 UWP and unemployment: is there a compensation mechanism?

In Boeri et al. (2017)'s 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.[17]

Figure 5 displays the correlation between the unemployment rate in 2006 and (log) population density: the unemployment rate proves 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, and 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 6: 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.

Figure 5 about here.

Figure 6 about here.

Going on, now, from descriptive evidence to more rigorous econometric specification, we estimate our models (FULL OLS and FE) for nominal and real wages, respectively, after introducing the LLM unemployment rate as an additional control. Table 6 shows that results are consistent with the descriptive evidence reported above: adding the unemployment rate to our baseline specifications does not influence our estimates of the UWP, suggesting that this variable is not correlated

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

with our covariate of interest (population density) and hence does not offset the real wage penalties suffered by individuals working in more densely populated areas. Hence, in the subsequent sections, we explore possible alternative forces driving our results.[18]

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 (-5.1%). These results are at odds with the previous evidence found in the existing literature for the US (again, see Glaeser and Mare, 2001 and Yankow, 2006). Once the role of the unemployment rate as a possible driving force has been excluded, some remaining explanations are in order: the collective bargaining system prevailing in the Italian labour market coupled with cost of living heterogeneity, the presence of urban amenities and idiosyncratic preferences for locations.

To disentangle the main mechanism underlying our results, we then turn to alternative categories of workers, who potentially live in the same locations as employees, while being subject to a different wage-setting scheme. In this section, we focus on a self-employed group of workers, namely independent contractors. The relevant labour contract was introduced in the Italian labour market in 1997, and can be considered as a self-employed worker that is associated with a firm, as in the case of employees. For this reason, this category of workers has been chosen for comparison with employees. Independent contractors are always 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 to depend largely on bargaining between employer and worker.

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

[18]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.

independent contractors, the INPS archives provide us with information on gross 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. Furthermore, for administrative reasons, it is possible to identify some subgroups of independent contractors, among whom we consider the most relevant in terms of size and type of tasks: the skilled group, which includes statutory auditors, company administrators, and legal representatives, and the external staff, which could be associated with either public or private employers.[19] 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 estimates of the UWP for the universe of independent contractors. 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) of Table 7 displays our results for daily nominal wages when age dummies, industry dummies (associated with industrial sectors according to 2-digit Ateco classification[20]), dummies for categories of independent contractors,[21] province fixed effects, and year fixed effects are included (FULL OLS): the estimated wage elasticity turns out at 4.9%, which is almost 25 times greater than the corresponding estimate for employees (see column (3) of Table 3), which was not even statistically different from zero.

To consider the role of workers' sorting, in column (2) we include, in addition to controls for workers' and firms' characteristics, worker fixed effects. We obtain a substantial drop in the point estimate that comes to 1.1%, still considerably larger than the corresponding estimate found for

[19]Note that these groups represent 90% of the total contracts of independent contractors; the biggest groups are the skilled workers (about 45%) and the external staff of private firms (around 40%).

[20]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.

[21]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. As pointed out, the three major groups cover more than 90% of the sample, the remainder being residual.

employees (see column (1) of Table 5). It is also worth noting that the drop in the point estimate from FULL OLS to FE (from column (1) to column (2)) is much greater for independent contractors than for employees, suggesting that sorting plays a more important role for the first group of workers. This is an interesting result, since higher spatial variability of earnings for self-employed workers, due to the absence of collective bargaining, provides a higher incentive for individuals to sort in space, with respect to employees. When carrying out the AKM specification, in column (3), the UWP point estimate drops even more, to 0.8%, documenting that even firm sorting is more notable for independent contractors than for employees.

We then move on to the analysis of the daily real wages. As shown in column (5) of Table 7, in the FULL OLS specification the estimated UWP proves very close to zero (equal to -0.14%) and not 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, to -5.5%, confirming that for independent contractors sorting is much more pronounced than for employees. The AKM estimate, shown in column (7), is slightly greater, amounting to -5.7%.

Finally, the IV-FE estimates, listed in column (4) and (8) respectively for nominal and real wages, follow patterns similar to those observed for employees: in nominal terms the point estimate is slightly greater than that obtained after including worker and firm fixed effects, while for real wages the negative point estimate becomes larger in absolute terms. The regressions shown in Table 5 do not include firm size, since this variable is missing in around 20% of the sample of independent contractors. Consistent conclusions are found when we also control for firm size on the restricted sample, as we do in Table 8.

Table 8 about here.

Table 9 presents the results obtained by replicating our regressions for distinct subgroups of independent contractors. This is relevant to our analysis because of the high heterogeneity in terms of skills within this group of workers heterogeneity that can only partially be captured with the introduction of subgroup fixed effects.

Table 9 about here.

The top panel considers workers performing high-skill job tasks (statutory auditors, company administrators, legal representatives). Interestingly, as far as nominal wages in FULL OLS are concerned (column (1)), the estimated UWP is even larger (around 6%) than that found for the overall group. The middle panel refers to estimates obtained when we consider the group of external staff of private firms, with a point estimates of 3.1%. Finally, the bottom panel focuses on the group of workers with a contract as external staff for the public administrations. These workers are usually involved in medium-high skill occupations, with a non-negligible share of graduates and post-graduates. The wage elasticity estimated by FULL OLS is, in this case, even higher, coming to 7.7%.

For all groups, when taking into account real wages, we observe a sharp drop in the estimated wage elasticity, which comes close to zero and is not statistically significant (with only one exception for external staff of private firms, in which case it is negative and statistically significant). When including worker fixed effects, for all categories, we confirm that workers' sorting plays a considerable role for each subgroup of independent contractors.

For final verification, in Table 10, we investigate whether controlling for unemployment rate could affect the size and statistical significance of the estimated UWP, as predicted by Boeri et al. (2017). What emerges, again, also for independent contractors, is that introducing the unemployment rate as additional explanatory variable does not change the baseline estimates. This suggests that the unemployment rates are not correlated with population density within provinces.

Table 10 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.[22]

Nonetheless, amenities, quality 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 -5.1% 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 those 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 a positive estimated UWP in nominal

[22]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).

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 Additional control groups: standard self-employed workers

Our choice of independent contractors as preferred comparison group to identify the role of collective bargaining in the estimation of the UWP is motivated by the fact that these workers are always associated with a firm like employees in this respect. Yet one might argue that the self-selection into the two groups (employees and independent contractors) is responsible for our findings, under the assumption that self-selection is heterogeneous across the spatial dimension. While there is no compelling reason why this should be systematically the case, in this section we consider additional groups of standard self-employed, who are not necessarily associated with a firm.

In particular, we consider the universe of self-employed belonging to professional associations (*ordini professionali*). In Italy, for legal qualification to work in these occupations the worker needs to 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.[23]

These self-employed workers may be either high-skilled or medium-skilled. We take high-skilled self-employed workers to be those employed in occupations that require a tertiary university degree, such as business consultants, lawyers, architects, and physicians/dentists. Among the medium-skilled self-employed, we include those with upper secondary education, such as journalists, surveyors, and accountants.[24]

The data on standard self-employed workers include: individual yearly wages, period worked (more than 90% of workers work for the full year), occupation, gender, age, and municipality where the worker lives (time-invariant). The latter information differs from that available for employees and self-employed workers, for whom we can make use of the time-varying information on where the job is located. 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

[23] Our data do not include workers that do not belong to *ordini professionali*. For this reason, we have no information on standard unskilled self-employed workers, such as plumbers.

[24] These categories are the most representative: they cover almost 90% of observations in the INPS archives of standard self-employed. The full list is: psychologists, nurses, industrial engineers, farmers, biologists, journalists, architects, lawyers, physicians/dentists, veterinaries, surveyors, accountants, business consultants, and multi-category (agronomists, foresters, actuaries, chemists, and geologists).

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 employed in our analysis for standard self-employed workers is chosen as close as possible to that already used for employees and independent contractors. A number of variables cannot, however, be defined for standard self-employed workers, such as industry dummies and firm size, since there is no firm associated, and the dummies for part time and/or temporary contracts, since these variables make little sense for a standard self-employed worker and are not recorded by INPS. We again focus on male workers aged 18-64, who work the full year. As before, our main covariate of interest is population density. The main control variables are: age dummies, province dummies, and year fixed effects. 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.

Table 11 shows the results when the whole group of standard self-employed workers is considered. As for nominal wages, OLS estimation that only includes year fixed effects gives an estimated UWP amounting to 11.4% (reported in column (1)), which is much greater than the corresponding figure found for employees (equal to 4.6%). After including all control variables in the regression and occupation dummies (FULL OLS), the wage elasticity becomes 5.3% (see column (2)), which is very close to that estimated for the independent contractors (4.8%, see column (1) of Table 8). Moving on to real wages, the wage elasticity obtained with only year fixed effects is still sizable, 5.8%, and statistically significant (column (3)), while after including all control dummies the UWP comes very close to zero and is no longer statistically significant (column (4)). Again, it is worth noting that these findings are similar to those obtained for the group of independent contractors.

Table 11 about here.

Table 12 shows the results for specific subgroups (which, aggregated, account for almost 90% of the population of standard self-employed workers). In the first four columns, the high-skilled occupations (consultants, lawyers, physicians, architects) are considered, while the medium-skilled occupations (journalists, surveyors, accountants) are presented in the last three columns. Observing the nominal wages (top panel), we see that the estimated wage elasticity to population density is very large for some high-skilled subgroups: it comes to 12.1% for lawyers and 9.9% for consultants, including all control variables. For these subgroups, a potential worry is that the very large estimated UWP is due to sorting of workers in the space dimension. For instance, one might expect that an individual aiming to excel in the legal profession would have to move to a big city, where the most important trials take place. Unfortunately, we are unable to identify such an effect with any precision since we cannot run fixed-effects estimates due to the worker location being time-invariant, as already pointed out. Nonetheless, even for the medium-skilled occupations, the estimated UWP are sizable and statistically different from zero: 6.5% for journalists, 3.3% for surveyors, and 4.7% for accountants. For these occupations, and in particular for surveyors and accountants, self-selection of workers with respect to ability and other unobservable characteristics is expected to be less important, suggesting that their positive and large estimated UWP cannot be entirely due to worker sorting and unobserved heterogeneity.

Moving on to real wages (bottom panel), we still find a sizable estimated UWP for Consultants (5.0%) and Lawyers (7.0%), including all control variables, while a penalty, even if fairly small (between -1% and -2%), is found for physicians, architects, and surveyors. For the remaining subgroups, estimated UWP are not statistically different from zero.

Overall, the analysis on additional subgroups of self-employed confirms that, when no institutional constraints are at work for wage setting, the estimated UWP is positive and non-negligible in nominal terms, and for some subgroups the UWP are positive even in real terms. Furthermore, UWP differences between employees and all the different groups of self-employed workers can hardly be entirely due to self-selection, since the average ability of the various subgroups of standard self-employed workers differ remarkably in average ability (as we have seen, for instance, between medium- and high-skilled occupations).

Table 12 about here.

9 Beyond the mean: unconditional quantile regressions

So far our analysis has been carried out at the conditional mean, adopting standard estimation methods. In this section, we estimate the UWP at different points of the wage distribution, the 10th, the 50th, and the 90th percentiles, to capture possible heterogeneity in wage elasticities across groups of workers. The aim is to explore how incentives in location choices and worker self-selection may vary across groups of individuals depending on their earnings, which represent a proxy of their skill levels. This is also relevant considering that, in the Italian two-tier multi-employer bargaining system, high-productive firms in cities can actually pay higher wages to the worker: the *in melius* clause does not prevent firms from raising wages as much as they want at the local level, for instance using performance-pay schemes, while imposing downward wage rigidity.

On the one hand, we aim at investigating whether high-wage employees (at the 90th percentile) earn higher returns in cities than the equivalent non-urban employees, even in the presence of collective bargaining. One might expect skilled 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, one might expect local collective agreements in cities to benefit most at the bottom tail of the wage distribution, for instance in the case where local bargaining is undertaken by local unions mostly interested in the welfare of unskilled workers. 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 investigate these issues, we make use of the unconditional quantile regressions (UQR) technique, proposed by Firpo and Lemieux (2009).[25] For closer comparison of the differences across groups, we present the results by plotting the point estimates as well as the corresponding 95% confidence interval, as done in Figures 7, 8 and 9. In this exercise, we adopt the same baseline specifications adopted for the regressions recorded in the previous sections, including all control variables.

The top panels of Figure 7 show the FULL OLS estimates using UQR for each group of workers (employees, independent contractors, standard self-employed workers) corresponding to the 10th, 50th, and 90th percentiles of the wage distribution. The left panel refers to nominal wages. As regards employees, the estimated wage elasticities are rather low, close to zero, and slightly increasing along the wage distribution (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 even when considering nominal wages. As regards independent contractors and standard self-employed workers, estimated wage elasticities are very high for all percentiles, with a U-shape impact. In particular, for independent contractors, the estimated UWP ranges from 5.2% at the 10th percentile to 3.9% at the median, to 7.4% at the 90th percentile; for standard self-employed, on the other hand, the estimated UWP is fairly stable along the wage distribution, at around 5-6%.

[25] The underlying idea here 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. It is worth stressing the importance of using UQR instead of the standard conditional quantile regressions (Koenker, 1982): when using conditional quantile regressions, for instance at the median, the estimates refer to the median of the "error term", whereas by using unconditional median quantile regressions it is possible to detect the impact of the covariates on the median of the unconditional distribution, which is what we are interested in.

Figure 7 about here.

This evidence suggests that differences across worker groups prove even greater when we take into account the 90th percentile of the wage distribution than when we consider its mean. In particular, the difference in estimated UWP between employees and independent contractors rises from 5% at the mean to around 7% at the 90th percentile: high-skilled employees are unable to obtain higher returns in cities or reach the level of the self-employed workers. The differences at the 10th percentiles are also comparable to those derived at the mean: the difference between employees and independent contractors is around 5%, while that between employees and self-employed workers is even greater, around 6%.

The results for real wages are shown in the right panels of Figure 7. Estimates for employees prove negative and substantial, ranging from −6% at the 10th percentile to around −4% at the 90th percentile. By contrast, in the case of independent contractors and standard self-employed workers, the high-wage individuals (the 90th percentile) obtain higher earnings than the corresponding non-urban workers: they are able to extract a positive and non-negligible wage premium, of around 2%. Hence, also in the case of real wages, employees do not close the gap with self-employed workers, the differences in estimated UWP being of about 5% at the mean and about 6% at the 90th percentile. A similar argument applies to unskilled workers at the 10th percentile.

The bottom panels of Figure 7 illustrate the results from FE estimation to take into account worker sorting, respectively for employees and independent contractors.[26] We find that, for the former group, nominal UWP (bottom-left panel) from UQR are fairly similar to those detected using FULL OLS estimates (top-left), confirming our conclusions from the analysis on the conditional mean. As for independent contractors, on the other hand, the corresponding estimates are smaller but still positive (estimated UWP equal to 2.2% at the 90th percentile and 1% at the median, while not being statistically different from zero at the 10th percentile) than those obtained in FULL OLS, suggesting, once again, that worker sorting is playing a non-negligible role here. In terms of real wages (bottom-right panel), the estimated UWP, which do not change much with respect to the FULL OLS estimates (top-right) for employees, drop substantially for independent contractors, confirming the role of sorting. Similar evidence can be observed for unskilled workers at the 10th percentile.

Finally, Figures 8 and 9 plot the UQR estimates obtained separately for the main subgroups of, respectively, independent contractors and standard self-employed workers.

As for independent contractors (Figure 8), the estimates are consistent with those regarding the overall group with a slightly increasing trend for estimated UWP of auditors/company administra-

[26] As already pointed out, we cannot carry out fixed effect regressions for standard self-employed workers.

tors and external staff of PA and a more steeply increasing trend for external staff of private firms. For the latter, the nominal (real) estimated OLS elasticity ranges from -7% (-15%) at the 10th percentile to a very high 17% (12%) at the 90th percentile, as shown in the left (right) panel.

Figure 8 about here.

As for standard self-employed workers (Figure 9), for all groups (with the sole exception of accountants) the trend of nominal UWP along the wage distribution (left panel) is clearly increasing, with some remarkable peaks for some groups of skilled workers, such as lawyers (estimated UWP amounting to 18% at the 90th percentile). It is worth noting that also for medium-skilled occupations (journalists, accountants, and surveyors) the premia are still highly positive and sizable, ranging from 3% to 7%. Turning to real wages (right panel), we find a general drop in estimates, consistently with previous results, but we still observe increasing trends along the wage distribution for the various categories and very high estimated UWP even in real terms: 15% for lawyers at the 90th percentile, 7% for business accountant.

Figure 9 about here.

Our evidence shows that UWP differences between employees and self-employed workers do not change much along the wage distribution. For high-skilled workers, this means that performance-pay schemes do not play a distinct role in mitigating their real penalties in cities. As already pointed out, there is no limitation for productive firms in cities to pay higher wages to their workers on a performance-pay base. So why is this so? On the one hand, it could be argued that two-tier systems are associated with higher overall labour costs, and this might limit the scope for the firms to introduce performance-pay schemes (Boeri, 2015). On the other hand, it could be argued that the two-tier collective bargaining might be considered a social norm, representing a reference benchmark for all firms, even those that are more productive and could pay higher wages. Since second-level bargaining is not compulsory, firms are not used to bargaining at the local level, as in other countries, and they cannot be blamed for not using the decentralized tier. This is confirmed by the low incidence of local collective agreements in Italy. Interestingly, UWP are not detected even for unskilled workers, at the 10th percentile. As stressed, this might have been the case if the unions had been interested in resorting to local agreements to raise wages mainly for unskilled workers in cities, where their purchasing power is limited by the higher cost of living.

10 Robustness checks

In this section, we implement 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 simplicity, we focus on two main groups of workers: employees and independent contractors.[27]

To begin with, instead of considering the workforce aged 15-64, Table 13 provides OLS and FE estimates on the full model, which includes both individual and worker characteristics, focusing on prime age male workers only (aged 25-49), often used as a reference group in the urban economics literature. The results are in line with those set out in Tables 3, 5, and 7.

Table 13 about here.

In Table 14, we investigate whether our results are affected by the choice of the earnings measure and use yearly wages instead of weekly/daily wages: again we find that the previous results remain substantially unaltered.

Table 14 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.[28] The results are shown in Table 15 where we observe, for employees, a substantial real wage penalty that comes very close to zero for independent contractors. Both estimates are negative and considerable in size when we also include worker fixed effects in estimation. Thus, the conclusions drawn employing LLMs are largely corroborated.

Table 15 about here.

In Tables 16 and 17, we check the robustness of our previous findings to alternative agglomeration measures. First, we use employment density: as shown in Table 16, our results are very close to those obtained when employing population density.

Table 16 about here.

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

[28]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.

Second, we adopt a dummy variable approach to identify urban areas, consisting in making use of the EU-OECD definition of "functional urban areas" (see OECD, 2012) and following the related three-step methodology to identify 73 urban areas among the 611 LLMs (Lamorgese and Petrella, 2016, see OECD, 2016, for the list).[29] We then rerun our regressions employing as agglomeration measure a dummy variable of one if the LLM is an urban area and zero otherwise. Interestingly, the results provided in Table 17 are similar to those previously reported, although coefficients of interest are now associated to dummy variables and as such should be interpreted.

Table 17 about here.

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. The results are displayed in Table 18 and are consistent with previous conclusions.

Table 18 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).[30] The descriptive statistics for 2015 are shown in Table 19. 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 19 about here.

[29]The first step selects "agglomerated areas" as clusters of dense neighboring cells (at least 1,500 inhabitants per square km) of at least 50,000 inhabitants and "urban centers" as the aggregation of municipalities where at least half of the population resides in the urban center. In the second step, the "metropolitan cores" are identified among the previously selected "agglomerated areas" as those that meet the following three conditions: the agglomerated municipalities are linked by administrative channels; at least half of the population of the area lives in one of the urban centers; at least 75% of the population of the urban centers belonging to the area lives in the urban agglomeration. The third step defines "urban areas" as the aggregation of the metropolitan core and its commuting neighborhood. Those LLMs that contain a metropolitan core are defined "urban areas".

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

We make use of this information to assign to each municipality a cost of living given by the associated absolute poverty thresholds, matching the macro-region and the municipality size. We then compute an index (we call it Istat CPI) dividing the threshold by the mean of the Italian average threshold in the base year (2005). Using Istat CPI, we deflate nominal wages to derive our real wage measure. The estimation results, shown in Table 20, are consistent with our conclusions from the baseline analysis.

Table 20 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. Table 21 shows, respectively for employees (top panel) and independent contractors (bottom panel), the estimated UWP for two alternative time intervals: 2005-2010 and 2011-2015. As will be seen, the findings are consistent across periods.

Table 21 about here.

The last robustness check consists in controlling for a much finer spatial unobserved heterogeneity, with fixed effects at the LLM (rather than province) level. Since introduction of this set of fixed effects is somewhat demanding in our empirical specification, we implement an alternative (still well-established) two-stage approach developed by Combes et al. (2008). The first stage consists in estimating the coefficients on the interaction between LLM and year fixed effects in a wage equation at the individual level with worker and firm level controls adopted in OLS estimation.[31] These estimated coefficients are then regressed in the second stage on the (log) population density, adding only yearly time dummies to control for business cycle effects. The coefficient of population density can be compared to the one developed with our preferred methodology. The estimation results from the second stage are set out in Table 22 and amply confirm our previous conclusions. Hence we can state that the results reported in Sections 6 and 7 are not affected by the introduction of the finest possible set of dummies at the LLM level.

Table 22 about here.

[31] The first-stage equation regression is:

$$log(W_{ilt}) = \beta_{lt} + \eta_1 X_{ilt} + \eta_2 Z_{jlt} + \delta_t + \varepsilon_{ilt}$$

where, as before, $log(W_{ilt})$ is the logarithm of the real or nominal wage of individual i working in LLM l at time t; β_{lt} are LLM-year fixed effects; X_{ilt} is the vector of individual controls including dummies for being on a part-time contract, being on a fixed-term contract, occupation, and age; Z_{jlt} is the vector of firm level controls including (log) firm size and dummies for NCLAs; δ_t are the year fixed effects.

11 Concluding remarks and policy implications

In this paper, we present 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, self-employed workers, and subgroups of them. 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, 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 -5.1%. This result does not hold for self-employed workers (either independent contractors or standard self-employed), for whom we obtain a corresponding estimate of about 5% for nominal wages and an estimated elasticity that is non-statistically different from zero for 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 tend to share local amenities and public goods. Indeed, since we consider various subsamples of self-employed workers, both skilled and unskilled, the difference in estimated UWP found for the two groups can hardly be put down to self-selection. Our results apply with similar magnitude also along the wage distribution, using unconditional quantile regressions, suggesting that decentralized collective bargaining is not playing a major role in Italy - neither for unskilled nor skilled workers.

While the evidence presented in the paper seems to be relevant *per se*, our findings contribute to a lively and topical policy debate, which concerns Italy as well as many other European countries. We can organize the major issues raised by debate in Italy in two groups.

We will begin with issues related to the proper functioning of centralized bargaining (see, for a review, D'Amuri and Nizzi, 2017). 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 "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).[32] Third, a conspicuous share of the contracts (at present about half of the total) have expired

[32] 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,

but are still in force. The delay in renewal of the national contracts impedes or retards "national" wage adjustments over time.

More relevant to the focus of the present paper, there are issues concerning the infrequent use of the second level of bargaining. Formally, there is no institutional constraint that stops firms from raising wages if located in cities with a high cost of living. Actually, there is some anecdotical evidence that the second level of bargaining is more at work in the high productivity areas, such as the North of Italy and the cities. But this has proven insufficient to offset the higher cost of living. This is an issue that needs a deeper investigation. Apart from cases where firms are constrained by higher labour costs due to a two-tier system, a speculative explanation might also consider the case where collective bargaining becomes a social norm, in the sense that in countries where collective bargaining is particularly widespread and well-established, it represents a reference benchmark for all firms, even those that are more productive and that could pay higher wages.

To align wages and productivity performance better, some authors have also proposed relaxing the *in melius* clause associated with the second level of bargaining (see, again, D'Amuri and Nizzi, 2017) and making this possibility dependent on the approval of the local labour unions. Other proposals refer to a possible overall change in the current system, such as replacement of the national bargaining structure with the adoption of a system based on decentralization at the firm level. Implementation 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 of areas characterized by high unemployment rates. The introduction of a minimum wages imposed by the government rather than by the collective bargaining has also been suggested by some scholars (for a discussion, see Boeri, 2009, and Garnero et al., 2015). Finally, some other authors (e.g., Garnero, 2018) propose 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 casts light on a specific, but highly significant aspect of this thorny question, concerning workers' remuneration in real terms in the presence of high variability

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.

in the costs of living across areas. Additional research effort and policy debate are required to shed further light on the functioning of the collective bargaining system and on its interaction with the features of the Italian productive system and labour markets.

References

Aboud, 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. (2009). Setting the minimum wage, *IZA Discussion Paper Series* **4335**: 1–28.

Boeri, T. (2015). Perverse effects of two-tier wage bargaining structures, *IZA World of Labor* **101**.

Boeri, T., Ichino, A., Moretti, E. and Posch, J. (2017). Unintended consequences of nominal wage equality across regions, Mimeo.

Card, D., H. J. and Klein, 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. (2016). The cost of agglomeration: House and land prices in cities, *CEPR Discussion Papers* **9240**.

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. (2016). Spatial wage disparities - Workers, firms and assortative matching, Mimeo.

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

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.

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**.

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., B. 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 (2016). *Functional Urban Areas in OECD Countries: Italy*, 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

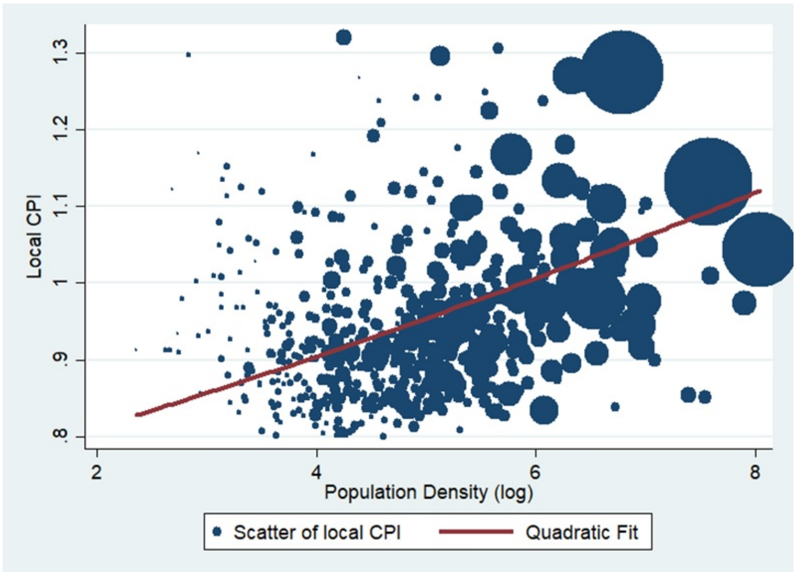

Figure 2: Distribution of local CPI in 2005, by LLM

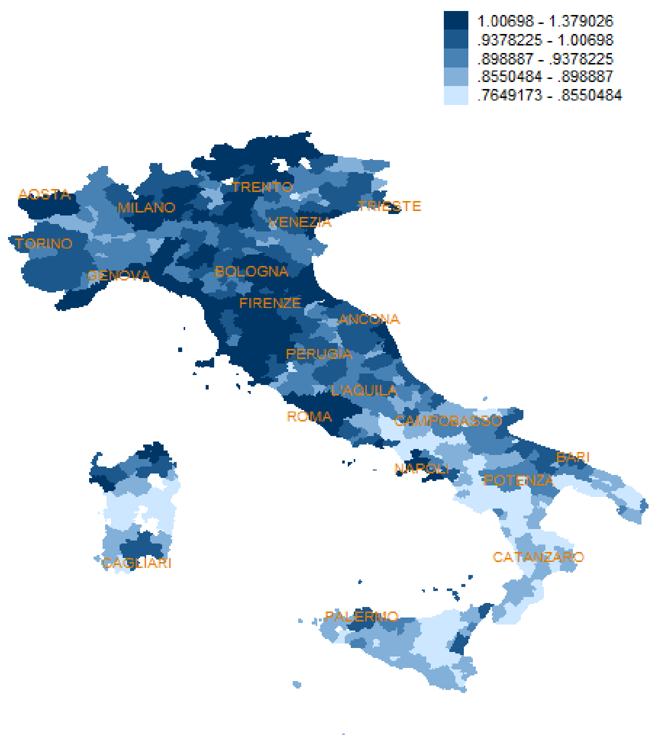

Figure 3: Distribution of population and employment density in 2005, by LLM

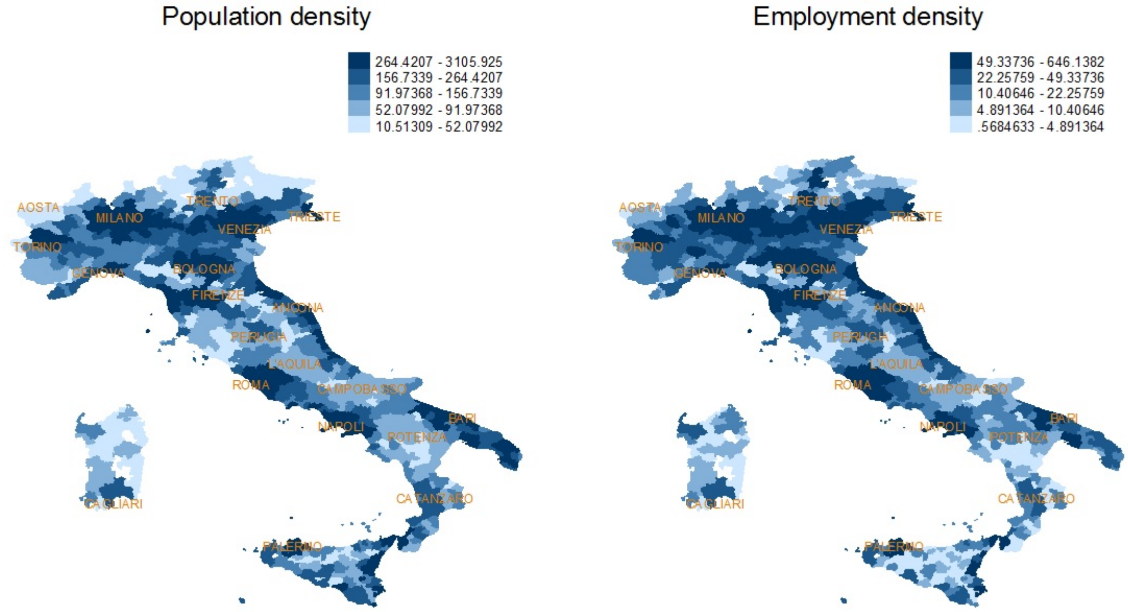

Figure 4: Distribution of nominal and real wages in 2005, by LLM

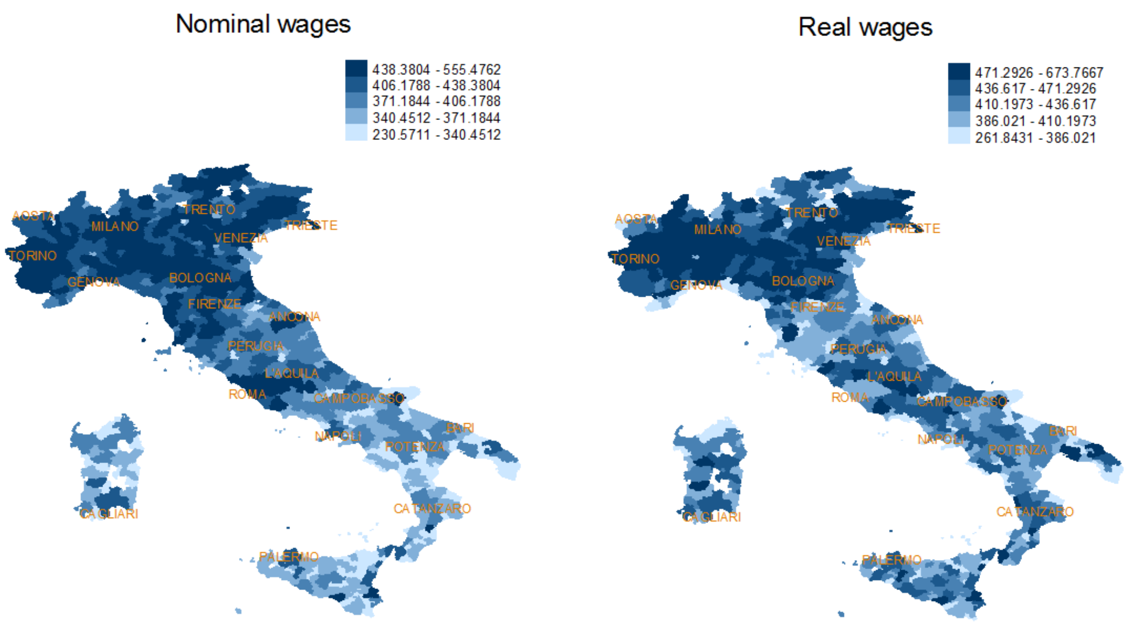

Figure 5: Unemployment rate and population density in 2005, by LLM

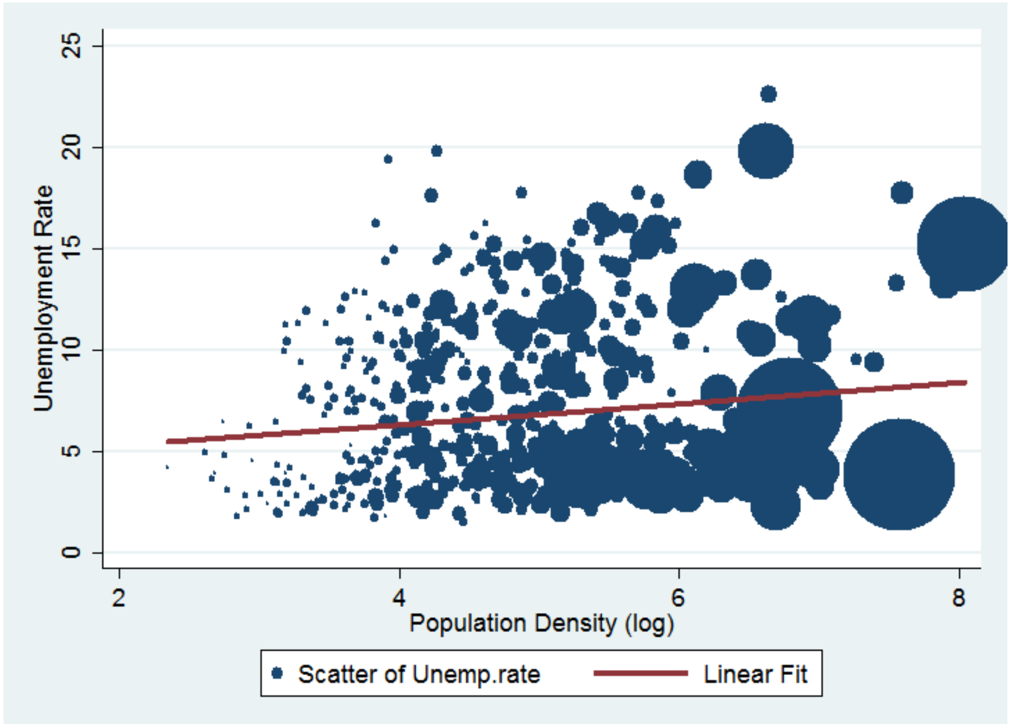

Figure 6: Employment rate and population density in 2005, by LLM

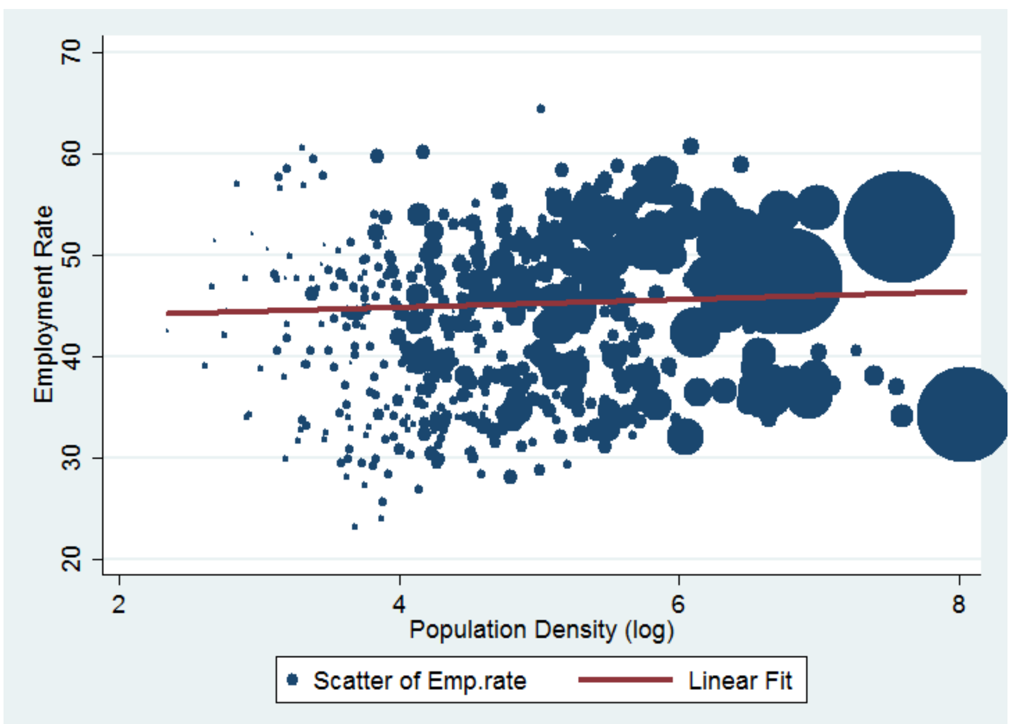

Figure 7: Unconditional quantile regressions: all workers

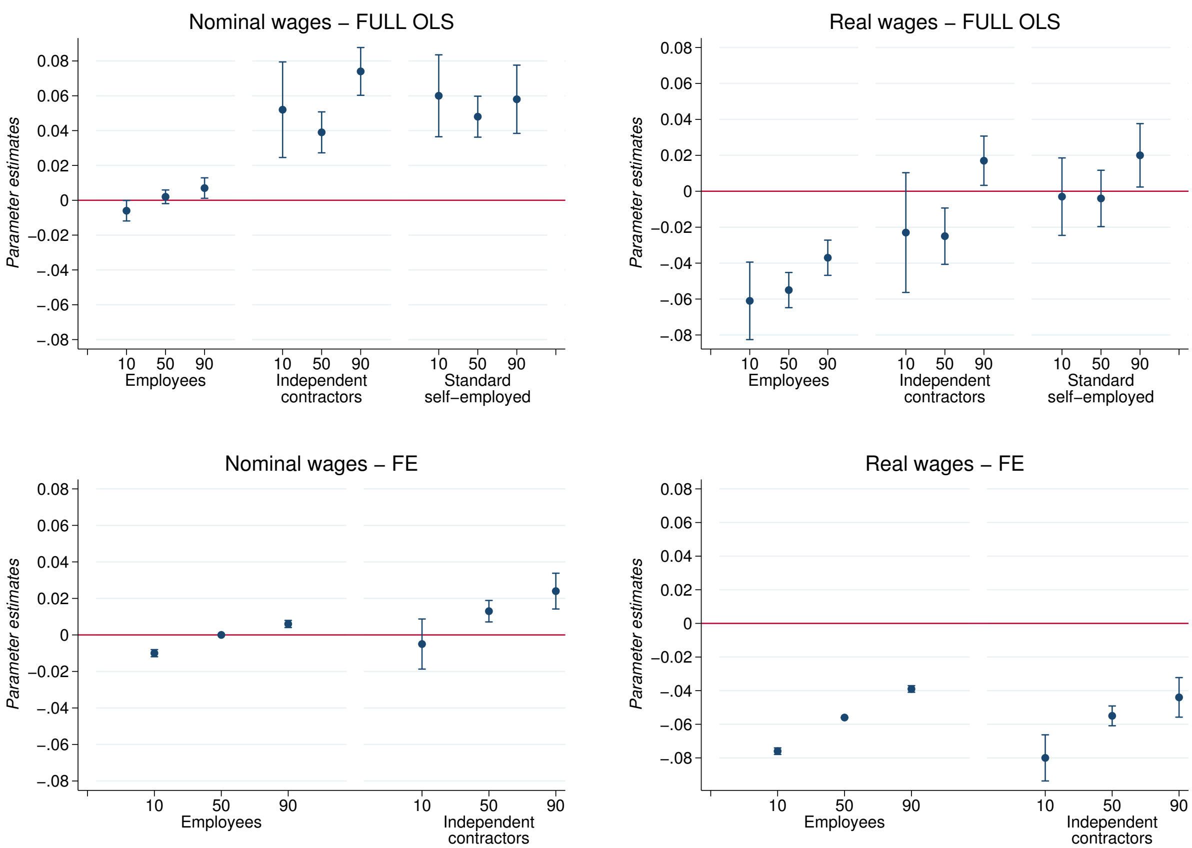

Figure 8: Unconditional quantile regressions: Independent contractors

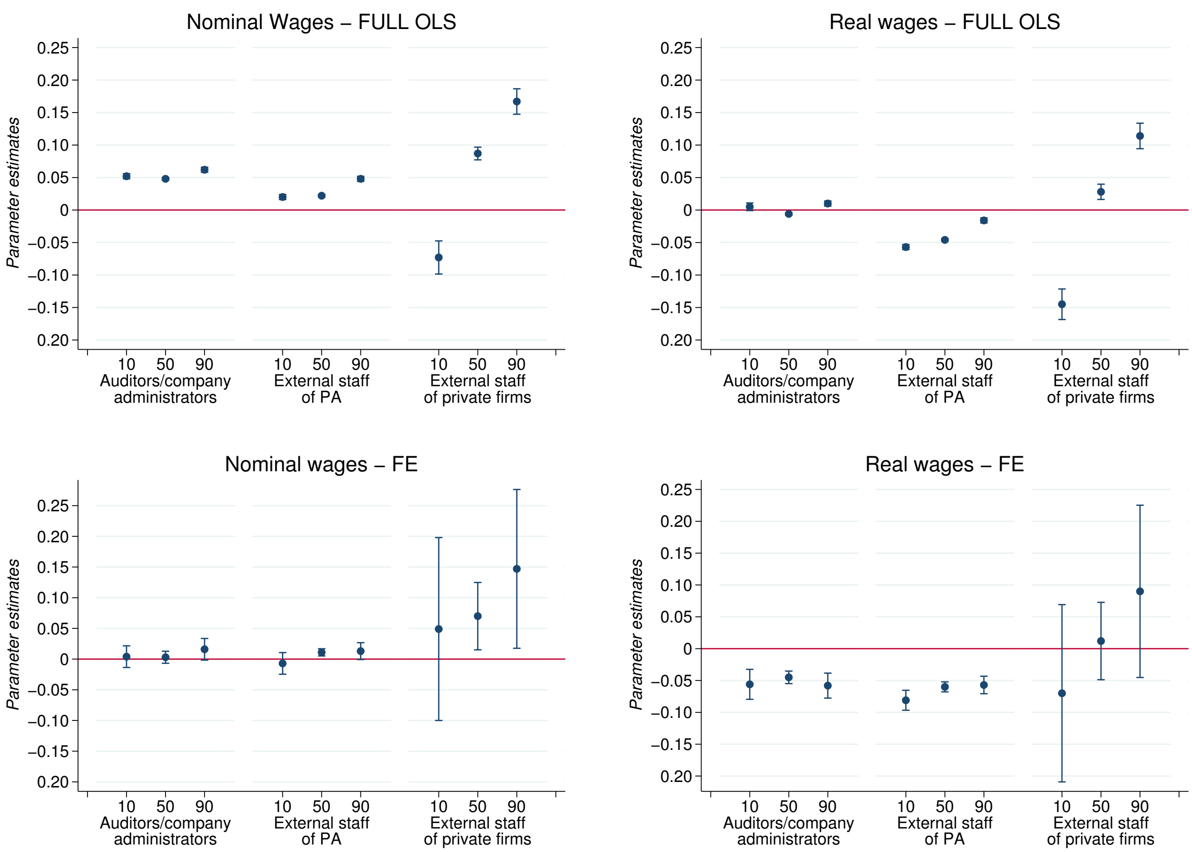

Figure 9: Unconditional quantile regressions: Standard self-employed

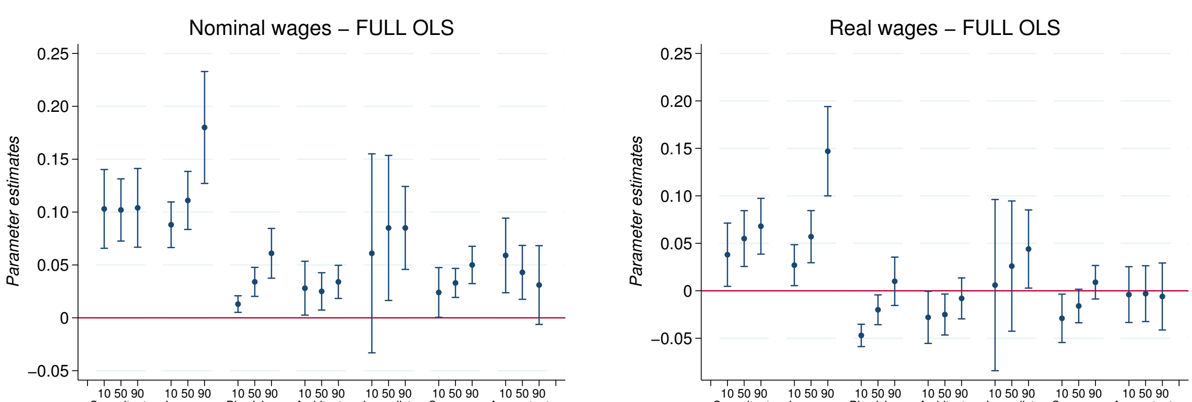

Table 1: Within-between variance decomposition of nominal and real weekly wages in 2005

	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 employees.

Table 2: Nominal versus real weekly wages in 2005

Population density	Nominal wages	Real wages
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 wages			Real wages		
	(1) OLS	(2) + worker charact.	(3) FULL OLS	(4) OLS	(5) + worker charact.	(6) FULL OLS
log population density	0.046***	0.006***	0.002	-0.006	-0.041***	-0.051***
	(0.017)	(0.002)	(0.002)	(0.017)	(0.012)	(0.004)
part time		-0.154***	-0.072***		-0.185***	-0.070***
		(0.007)	(0.006)		(0.010)	(0.006)
fixed term		-0.103***	-0.119***		-0.120***	-0.124***
		(0.008)	(0.004)		(0.008)	(0.004)
log firm size			0.023***			0.022***
			(0.002)			(0.002)
age dummies	no	yes	yes	no	yes	yes
occupation dummies	no	yes	yes	no	yes	yes
contract dummies	no	no	yes	no	no	yes
area dummies	no	yes	yes	no	yes	yes
year fe	yes	yes	yes	yes	yes	yes
Observations	77,015,891	77,015,891	77,015,891	77,015,891	77,015,891	77,015,891
R-squared	0.041	0.515	0.608	0.005	0.444	0.591

Note: The dependent variable in columns (1)-(3) is weekly nominal wage; in columns (4)-(6) it is weekly real wage. Occupation dummies stands for, namely, white collar, blue collar, apprentice, manager, and executive; contract dummies are dummies for NCLAs, which cover about 243 contracts. Standard errors are clustered at the LLM level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 4: UWP estimates. Employees by occupation categories

	Nominal wages		Real wages	
	(1) FULL OLS	(2) FE	(3) FULL OLS	(4) FE
	Blue collars			
log population density	-0.000	-0.002**	-0.052***	-0.056***
	(0.002)	(0.001)	(0.004)	(0.004)
Observations	49,559,314	49,247,306	49,559,314	49,247,306
R-squared	0.358	0.801	0.378	0.807
	White collars			
log population density	0.014***	0.005***	-0.045***	-0.055***
	(0.003)	(0.001)	(0.004)	(0.004)
Observations	20,193,761	19,952,336	20,193,761	19,952,336
R-squared	0.395	0.894	0.402	0.893
	Managers and executives			
log population density	0.014***	-0.003***	-0.051***	-0.065***
	(0.003)	(0.001)	(0.005)	(0.005)
Observations	3,343,657	3,343,657	3,343,657	3,343,657
R-squared	0.279	0.869	0.318	0.872
	Controls used in the different columns			
control variables	yes	yes	yes	yes
age dummies	yes	yes	yes	yes
contract dummies	yes	yes	yes	yes
area dummies	yes	yes	yes	yes
year fe	yes	yes	yes	yes
worker fe	no	yes	no	yes

Note: The dependent variable in columns (1)-(2) is yearly nominal wages; in columns (3)-(4) it is yearly real wage. Contract dummies are dummies for NCLAs, which cover about 243 contracts. Standard errors are clustered at the LLM level. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: UWP estimates. Employees: fixed effects and IV estimates

	Nominal wages			Real wages		
	(1) FE	(2) AKM	(3) IV-FE	(4) FE	(5) AKM	(6) IV-FE
log population density	-0.000	0.002*	0.000	-0.056***	-0.054***	-0.102***
	(0.001)	(0.001)	(0.002)	(0.004)	(0.004)	(0.011
part time	0.046***	0.065***	0.046***	0.046***	0.065***	0.047***
	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)	(0.003)
fixed term	-0.048***	-0.054***	-0.048***	-0.049***	-0.055***	-0.050***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003
log firm size	0.016***	0.017***	0.016***	0.016***	0.018***	0.017***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
occupation dummies	yes	yes	yes	yes	yes	yes
age dummies	yes	yes	yes	yes	yes	yes
contract dummies	yes	yes	yes	yes	yes	yes
area dummies	yes	yes	yes	yes	yes	yes
year fe	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	76,755,407	77,015,891	77,015,891	77,015,891
R-squared	0.892	0.892	0.9172	0.886	0.885	0.9131
K-P rk Wald F statistic			126.307			126.307

Note: The dependent variable in columns (1)-(3) is weekly nominal wage; in columns (4)-(6) it is weekly real wage. Occupation dummies stands for, namely, white collar, blue collar, apprentice, manager, and executive; contract dummies are dummies for NCLAs, which cover about 243 contracts. Standard errors are clustered at the LLM level. *** p<0.01, ** p<0.05, * p<0.1.

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

	Nominal wages		Real wages	
	FULL OLS	FE	FULL OLS	FE
log population density	0.003	0.000	-0.052***	-0.055***
	(0.002)	(0.001)	(0.004)	(0.004)
part time	-0.072***	0.046***	-0.070***	0.046***
	(0.006)	(0.003)	(0.006)	(0.003)
fixed term	-0.119***	-0.048***	-0.124***	-0.049***
	(0.004)	(0.003)	(0.004)	(0.003)
log firm size	0.023***	0.016***	0.022***	0.016***
	(0.002)	(0.001)	(0.002)	(0.001)
unemployment rate	-0.001***	-0.002***	0.001	-0.002
	(0.001)	(0.000)	(0.001)	(0.001)
age dummies	yes	yes	yes	yes
occupation dummies	yes	yes	yes	yes
contract dummies	yes	yes	yes	yes
area dummies	yes	yes	yes	yes
year fe	yes	yes	yes	yes
worker fe	no	yes	no	yes
Observations	77,015,891	77,015,891	77,015,891	77,015,891
R-squared	0.608	0.892	0.591	0.886

Note: The dependent variable in columns (1)-(3) is weekly nominal wages; in columns (4)-(6) it is weekly real wage. Occupation dummies stands for, namely, white collar, blue collar, apprentice, manager, and executive; contract dummies are dummies for NCLAs, which cover about 243 contracts. Standard errors are clustered at the LLM level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 7: UWP estimates. Independent contractors: baseline specification

	Nominal wages				Real wages			
	(1) FULL OLS	(2) FE	(3) AKM	(4) IV-FE	(5) FULL OLS	(6) FE	(7) AKM	(8) IV-FE
log population density	0.049***	0.011***	0.008**	0.009**	-0.014	-0.055***	-0.057***	-0.083***
	(0.009)	(0.003)	(0.004)	(0.004)	(0.012)	(0.009)	(0.007)	(0.013)
age dummies	yes	yes	yes	yes	yes	yes	yes	yes
industry dummies	yes	yes	yes	yes	yes	yes	yes	yes
contractors dummies	yes	yes	yes	yes	yes	yes	yes	yes
area dummies	yes	yes	yes	yes	yes	yes	yes	yes
year fe	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	5,825,362	5,825,362	5,825,362	5,825,362	5,825,362	5,825,362	5,825,362	5,825,362
R-squared	0.210	0.785	0.828	0.784	0.216	0.786	0.829	0.786
K-P rk Wald F statistic				298.009				298.009

Note: The dependent variable in columns (1)-(4) is daily nominal wage; in columns (5)-(8) it is daily real wage. Industry dummies stands for 86 2-digit Ateco industrial sectors; contractors dummies are dummies for 19 categories of independent contractors. Standard errors are clustered at the LLM level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 8: UWP estimates. Independent contractors: controlling for firm size

	Nominal wages				Real wages			
	(1) FULL OLS	(2) FE	(3) AKM	(4) IV-FE	(5) FULL OLS	(6) FE	(7) AKM	(8) IV-FE
log population density	0.048***	0.005	0.003	0.004	-0.013	-0.060***	-0.063***	-0.087***
	(0.010)	(0.004)	(0.005)	(0.006)	(0.012)	(0.010)	(0.009)	(0.015)
log firm size	0.066***	0.038***	0.054***	0.038***	0.065***	0.037***	0.054***	0.037***
	(0.012)	(0.009)	(0.012)	(0.009)	(0.012)	(0.009)	(0.012)	(0.009)
age dummies	yes	yes	yes	yes	yes	yes	yes	yes
industry dummies	yes	yes	yes	yes	yes	yes	yes	yes
contractors dummies	yes	yes	yes	yes	yes	yes	yes	yes
area dummies	yes	yes	yes	yes	yes	yes	yes	yes
year fe	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	4,475,419	4,475,419	4,475,419	3,898,533	4,475,419	4,475,419	4,475,419	4,475,419
R-squared	0.218	0.797	0.834	0.796	0.227	0.799	0.836	0.798
K-P rk Wald F statistic				303.140				303.140

Note: The dependent variable in columns (1)-(4) is daily nominal wage; in columns (5)-(8) it is daily real wage. Industry dummies stands for 86 2-digit Ateco industrial sectors; contractors dummies are dummies for 19 categories of independent contractors. Standard errors are clustered at the LLM level. *** p<0.01, ** p<0.05, * p<0.1.

Table 9: UWP estimates. Independent contractors by category

	Nominal wages				Real wages			
	(1) FULL OLS	(2) FE	(3) AKM	(4) IV-FE	(5) FULL OLS	(6) FE	(7) AKM	(8) IV-FE
	Statutory auditors, company administrators, and legal representatives							
log population density	0.060***	0.007	0.001	0.005	0.003	-0.054***	-0.059***	-0.080***
	(0.011)	(0.006)	(0.005)	(0.007)	(0.011)	(0.009)	(0.008)	(0.014)
Observations	2,729,133	2,729,133	2,729,133	2,729,133	2,729,133	2,729,133	2,729,133	2,729,133
R-squared	0.159	0.781	0.815	0.779	0.158	0.781	0.815	0.778
	External staff of private firms							
log population density	0.031***	0.006	0.003	0.009	-0.038***	-0.063***	-0.067***	-0.087***
	(0.007)	(0.005)	(0.005)	(0.006)	(0.011)	(0.007)	(0.010)	(0.012)
Observations	2,410,877	2,410,877	2,410,877	2,410,877	2,410,877	2,410,877	2,410,877	2,410,877
R-squared	0.164	0.797	0.840	0.797	0.166	0.798	0.841	0.798
	External staff of public administration							
log population density	0.077***	0.081***	0.135**	0.069*	0.012	-0.000	0.039	-0.039
	(0.028)	(0.025)	(0.055)	(0.037)	(0.033)	(0.030)	(0.058)	(0.035)
Observations	188,328	188,328	188,328	188,328	188,328	188,328	188,328	188,328
R-squared	0.139	0.773	0.786	0.775	0.134	0.771	0.785	0.774
age dummies	yes	yes	yes	yes	yes	yes	yes	yes
industry dummies	yes	yes	yes	yes	yes	yes	yes	yes
area dummies	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

Note: The dependent variable in columns (1)-(4) is daily nominal wages; in columns (5)-(8) it is daily real wages. Industry dummies stands for 86 2-digit Ateco industrial sectors. Standard errors are clustered at the LLM level. *** p<0.01, ** p<0.05, * p<0.1.

Table 10: UWP estimates. Independent contractors: controlling for unemployment rate

	Nominal wages		Real wages	
	(1)	(2)	(3)	(4)
	FULL OLS	FE	FULL OLS	FE
log population density	0.050***	0.010***	-0.013	-0.055***
	(0.008)	(0.003)	(0.012)	(0.009)
unemployment rate	-0.010***	-0.004***	-0.004**	-0.001
	(0.002)	(0.001)	(0.002)	(0.002)
age dummies	yes	yes	yes	yes
industry dummies	yes	yes	yes	yes
contractors dummies	yes	yes	yes	yes
area dummies	yes	yes	yes	yes
year fe	yes	yes	yes	yes
worker fe	no	yes	no	yes
Observations	5,825,362	5,825,362	5,825,362	5,825,362
R-squared	0.210	0.785	0.216	0.786

Note: The dependent variable in columns (1)-(3) is daily nominal wages; in columns (4)-(6) it is daily real wages. Industry dummies stands for 86 2-digit Ateco industrial sectors; contractors dummies are dummies for 19 categories of independent contractors. Standard errors are clustered at the LLM level. *** p<0.01, ** p<0.05, * p<0.1.

Table 11: UWP estimates. Standard self-employed

	Nominal wages		Real wages	
	(1) OLS	(2) FULL OLS	(3) OLS	(4) FULL OLS
log population density	0.114***	0.053***	0.058*	0.002
	(0.040)	(0.004)	(-0.034)	(0.005)
age dummies	no	yes	no	yes
area dummies	no	yes	no	yes
self-empl. w. dummies	no	yes	no	yes
year fe	yes	yes	yes	yes
Observations	4,154,141	4,154,140	4,154,141	4,154,140
R-squared	0.014	0.267	0.005	0.253

Note: The dependent variable in columns (1)-(2) is yearly nominal wages; in columns (3)-(4) it is yearly real wages. Self-employed workers dummies for 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 12: UWP estimates. Standard self-employed: high-skilled versus medium-skilled workers

	High skilled/Graduates				Medium skilled/Non graduates		
	(1) Consultants	(2) Lawyers	(3) Physicians	(4) Architects	(5) Journalists	(6) Surveyors	(7) Accountants
				Nominal wages			
log population density	0.099***	0.121***	0.036***	0.030***	0.065***	0.033***	0.047***
	(0.009)	(0.010)	(0.005)	(0.006)	(0.021)	(0.006)	(0.010)
Observations	329,709	813,197	839,317	838,825	102,666	598,215	158,975
R-squared	0.303	0.186	0.080	0.129	0.073	0.204	0.250
				Real wages			
log population density	0.050***	0.070***	-0.019***	-0.022***	0.013	-0.015**	-0.004
	(0.009)	(0.011)	(0.006)	(0.006)	(0.022)	(0.006)	(0.010)
Observations	329,709	813,197	839,317	838,825	102,666	598,215	158,975
R-squared	0.283	0.167	0.121	0.079	0.060	0.187	0.227
age dummies	yes	yes	yes	yes	yes	yes	yes
area dummies	yes	yes	yes	yes	yes	yes	yes
year fe	yes	yes	yes	yes	yes	yes	yes

Note: The dependent variable is yearly (either nominal or real) wages. Standard errors are clustered at the LLM level. *** p<0.01, ** p<0.05, * p<0.1.

Table 13: Robustness checks: Prime age workers only (25-49)

	Nominal wages		Real wages	
	(1) FULL OLS	(2) FE	(3) FULL OLS	(4) FE
	Employees			
log population density	0.002	-0.000	-0.053***	-0.057***
	(0.002)	(0.001)	(0.004)	(0.004)
part time	-0.080***	0.043***	-0.078***	0.043***
	(0.007)	(0.003)	(0.007)	(0.003)
fixed term	-0.126***	-0.050***	-0.131***	-0.051***
	(0.005)	(0.003)	(0.004)	(0.003)
log firm size	0.023***	0.015***	0.022***	0.015***
	(0.002)	(0.001)	(0.002)	(0.001)
age dummies	yes	yes	yes	yes
occupation dummies	yes	yes	yes	yes
contract dummies	yes	yes	yes	yes
area dummies	yes	yes	yes	yes
year fe	yes	yes	yes	yes
worker fe	no	yes	no	yes
Observations	56,696,640	56,696,640	56,696,640	56,696,640
R-squared	0.575	0.888	0.560	0.883
	Independent contractors			
log population density	0.068***	0.016***	-0.001	-0.053***
	(0.006)	(0.004)	(0.006)	(0.006)
age dummies	yes	yes	yes	yes
industry dummies	yes	yes	yes	yes
contractors dummies	yes	yes	yes	yes
area dummies	yes	yes	yes	yes
year fe	yes	yes	yes	yes
worker fe	no	yes	no	yes
Observations	3,837,258	3,837,258	3,837,258	3,837,258
R-squared	0.187	0.774	0.196	0.776

Note: The dependent variable in columns (1)-(2) is weekly/daily nominal wages respectively for employees and independent contractors; in columns (3)-(4) it is weekly/daily real wages. Occupation dummies stands for, namely, white collar, blue collar, apprentice, manager, and executive; contract dummies are dummies for NCLAs, which cover about 243 contracts. Industry dummies stands for 86 2-digit Ateco industrial sectors; contractors dummies are dummies for 19 categories of independent contractors. Standard errors are clustered at the LLM level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 14: Robustness checks: Yearly wages

	Nominal wages		Real wages	
	(1) FULL OLS	(2) FE	(3) FULL OLS	(4) FE
	Employees			
log population density	0.003	0.005	-0.051***	-0.051***
	(0.002)	(0.003)	(0.004)	(0.005)
part time	-0.579***	-0.355***	-0.577***	-0.355***
	(0.009)	(0.004)	(0.009)	(0.004)
fixed term	-0.507***	-0.302***	-0.513***	-0.303***
	(0.005)	(0.003)	(0.005)	(0.003)
log firm size	0.042***	0.048***	0.042***	0.048***
	(0.003)	(0.003)	(0.003)	(0.003)
age dummies	yes	yes	yes	yes
occupation dummies	yes	yes	yes	yes
contract dummies	yes	yes	yes	yes
area dummies	yes	yes	yes	yes
year fe	yes	yes	yes	yes
worker fe	no	yes	no	yes
Observations	77,015,879	77,015,879	77,015,879	77,015,879
R-squared	0.564	0.800	0.557	0.797
	Independent contractors			
log population density	0.021**	0.001	-0.042***	-0.065***
	(0.009)	(0.004)	(0.013)	(0.010)
age dummies	yes	yes	yes	yes
industry dummies	yes	yes	yes	yes
contractors dummies	yes	yes	yes	yes
area dummies	yes	yes	yes	yes
year fe	yes	yes	yes	yes
worker fe	no	yes	no	yes
Observations	5,825,362	5,825,362	5,825,362	5,825,362
R-squared	0.311	0.836	0.326	0.84

Note: The dependent variable in columns (1)-(2) is yearly nominal wages respectively for employees and independent contractors; in columns (3)-(4) it is yearly real wages. Occupation dummies stands for, namely, white collar, blue collar, apprentice, manager, and executive; contract dummies are dummies for NCLAs, which cover about 243 contracts. Industry dummies stands for 86 2-digit Ateco industrial sectors; contractors dummies are dummies for 19 categories of independent contractors. Standard errors are clustered at the LLM level. *** p<0.01, ** p<0.05, * p<0.1.

Table 15: Robustness checks: Municipalities rather than LLM (Real wages)

	Employees		Independent contractors	
	(1)	(2)	(3)	(4)
	FULL OLS	FE	FULL OLS	FE
log population density	-0.054***	-0.056***	-0.003	-0.064***
	(0.006)	(0.007)	(0.006)	(0.006)
part time	-0.068***	0.047***	no	no
	(0.006)	(0.003)		
fixed term	-0.131***	-0.051***	no	no
	(0.004)	(0.003)		
log firm size	0.022***	0.016***	no	no
	(0.001)	(0.001)		
age dummies	yes	yes	yes	yes
occupation dummies	yes	yes	yes	yes
contract dummies	yes	yes	no	no
industry dummies	no	no	yes	yes
contractors dummies	no	no	yes	yes
area dummies	yes	yes	no	no
area dummies	no	no	yes	yes
year fe	yes	yes	yes	yes
worker fe	no	yes	no	yes
Observations	76,998,804	76,998,804	5,824,543	5,824,543
R-squared	0.582	0.884	0.222	0.788

Note: The dependent variable in columns (1)-(2) is weekly real wages for employees; in columns (3)-(4) it is daily real wage for independent contractors. Occupation dummies stands for, namely, white collar, blue collar, apprentice, manager, and executive; contract dummies are dummies for NCLAs, which cover about 243 contracts. Industry dummies stands for 86 2-digit Ateco industrial sectors; contractors dummies are dummies for 19 categories of independent contractors. Standard errors are clustered at the LLM level. *** p<0.01, ** p<0.05, * p<0.1.

Table 16: Robustness checks: Employment density rather than population density

	Nominal wages		Real wages	
	(1) FULL OLS	(2) FE	(3) FULL OLS	(4) FE
	Employees			
log employment density	0.004***	0.002***	-0.042***	-0.045***
	(0.001)	(0.001)	(0.003)	(0.003)
part time	0.023***	0.016***	0.023***	0.016***
	(0.002)	(0.001)	(0.002)	(0.001)
fixed term	-0.072***	0.046***	-0.070***	0.046***
	(0.006)	(0.003)	(0.006)	(0.003)
log firm size	-0.119***	-0.048***	-0.124***	-0.049***
	(0.004)	(0.003)	(0.004)	(0.003)
age dummies	yes	yes	yes	yes
occupation dummies	yes	yes	yes	yes
contract dummies	yes	yes	yes	yes
area dummies	yes	yes	yes	yes
year fe	yes	yes	yes	yes
worker fe	no	yes	no	yes
Observations	77,015,891	77,015,891	77,015,891	77,015,891
R-squared	0.608	0.892	0.591	0.886
	Independent contractors			
log employment density	0.042***	0.012***	-0.005	-0.040***
	(0.006)	(0.003)	(0.006)	(0.007)
age dummies	yes	yes	yes	yes
industry dummies	yes	yes	yes	yes
contractors dummies	yes	yes	yes	yes
year fe	yes	yes	yes	yes
worker fe	no	yes	no	yes
Observations	5,348,595	5,348,595	5,348,595	5,348,595
R-squared	0.210	0.788	0.217	0.790

Note: The dependent variable in columns (1)-(2) is weekly/daily nominal wages respectively for employees and independent contractors; in columns (3)-(4) it is weekly/daily real wages. Occupation dummies stands for, namely, white collar, blue collar, apprentice, manager, and executive; contract dummies are dummies for NCLAs, which cover about 243 contracts. Industry dummies stands for 86 2-digit Ateco industrial sectors; contractors dummies are dummies for 19 categories of independent contractors. Standard errors are clustered at the LLM level. *** p<0.01, ** p<0.05, * p<0.1.

Table 17: Robustness checks: Urban dummy computed as in Lamorgese and Petrella (2016)

	Nominal wages		Real wages	
	(1) FULL OLS	(2) FE	(3) FULL OLS	(4) FE
	Employees			
urban LLM	0.001	-0.001	-0.066***	-0.072***
	(0.003)	(0.001)	(0.006)	(0.005)
part time	-0.072***	0.046***	-0.070***	0.046***
	(0.006)	(0.003)	(0.006)	(0.003)
fixed term	-0.119***	-0.048***	-0.125***	-0.049***
	(0.004)	(0.003)	(0.004)	(0.003)
log firm size	0.023***	0.016***	0.023***	0.016***
	(0.002)	(0.001)	(0.002)	(0.001)
age dummies	yes	yes	yes	yes
occupation dummies	yes	yes	yes	yes
contract dummies	yes	yes	yes	yes
area dummies	yes	yes	yes	yes
year fe	yes	yes	yes	yes
Observations	77,015,891	77,015,891	77,015,891	77,015,891
R-squared	0.6084	0.8917	0.5903	0.8859
	Independent contractors			
urban LLM	0.102***	0.009	-0.006	-0.091***
	(0.014)	(0.007)	(0.012)	(0.010)
age dummies	yes	yes	yes	yes
industry dummies	yes	yes	yes	yes
contractors dummies	yes	yes	yes	yes
area dummies	yes	yes	yes	yes
year fe	yes	yes	yes	yes
worker fe	no	yes	no	yes
Observations	5,825,374	5,825,374	5,825,374	5,825,374
R-squared	0.206	0.785	0.214	0.786

Note: The dependent variable in columns (1)-(2) is weekly/daily nominal wages respectively for employees and independent contractors; in columns (3)-(4) it is weekly/daily real wages. Urban LLM is a dummy variable that is equal to one if the LLM is located in an urban area and to zero otherwise. Occupation dummies stands for, namely, white collar, blue collar, apprentice, manager, and executive; contract dummies are dummies for NCLAs, which cover about 243 contracts. Industry dummies stands for 86 2-digit Ateco industrial sectors; contractors dummies are dummies for 19 categories of independent contractors. Standard errors are clustered at the LLM level. *** p<0.01, ** p<0.05, * p<0.1.

Table 18: Robustness checks: Housing pricing without removing compositional effects

	Employees		Independent contractors	
	(1)	(2)	(3)	(4)
	FULL OLS	FE	FULL OLS	FE
log population density	-0.057***	-0.062***	-0.020*	-0.063***
	(0.004)	(0.004)	(0.012)	(0.008)
part time	-0.070***	0.046***	no	no
	(0.006)	(0.003)		
fixed term	-0.125***	-0.049***	no	no
	(0.004)	(0.003)		
log firm size	0.022***	0.016***	no	no
	(0.002)	(0.001)		
age dummies	yes	yes	yes	yes
occupation dummies	yes	yes	no	no
contract dummies	yes	yes	no	no
industry dummies	no	no	yes	yes
contractors dummies	no	no	yes	yes
area dummies	yes	yes	yes	yes
year fe	yes	yes	yes	yes
worker fe	no	yes	no	yes
Observations	77,015,891	77,015,891	5,825,362	5,825,362
R-squared	0.589	0.885	0.214	0.786

Note: The dependent variable in columns (1)-(2) is weekly real wages for employees; in columns (3)-(4) it is daily real wage for independent contractors. Occupation dummies stands for, namely, white collar, blue collar, apprentice, manager, and executive; contract dummies are dummies for NCLAs, which cover about 243 contracts. Industry dummies stands for 86 2-digit Ateco industrial sectors; contractors dummies are dummies for 19 categories of independent contractors. Standard errors are clustered at the LLM level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 19: 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 20: Robustness checks: Alternative local CPI computed using poverty thresholds

	Employees		Independent contractors	
	(1) FULL OLS	(2) FE	(3) FULL OLS	(4) FE
log population density	-0.026***	-0.028***	0.015*	-0.028***
	(0.003)	(0.003)	(0.008)	(0.007)
part time	-0.073***	0.046***	no	no
	(0.006)	(0.003)		
fixed term	-0.119***	-0.047***	no	no
	(0.004)	(0.003)		
log firm size	0.022***	0.016***	no	no
	(0.002)	(0.001)		
age dummies	yes	yes	yes	yes
occupation dummies	yes	yes	no	no
contract dummies	yes	yes	no	no
industry dummies	no	no	yes	yes
contractors dummies	no	no	yes	yes
area dummies	yes	yes	yes	yes
year fe	yes	yes	yes	yes
worker fe	no	yes	no	yes
Observations	77,015,891	77,015,891	5,348,595	5,348,595
R-squared	0.572	0.882	0.199	0.785

Note: The dependent variable in columns (1)-(2) is weekly real wages for employees; in columns (3)-(4) it is daily real wage for independent contractors. Occupation dummies stands for, namely, white collar, blue collar, apprentice, manager, and executive; contract dummies are dummies for NCLAs, which cover about 243 contracts. Industry dummies stands for 86 2-digit Ateco industrial sectors; contractors dummies are dummies for 19 categories of independent contractors. Standard errors are clustered at the LLM level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 21: Robustness checks: Different time periods

	Nominal wages		Real wages	
	2005-2010	2011-2015	2005-2010	2011-2015
	(1) FULL OLS	(2) FULL OLS	(3) FULL OLS	(4) FULL OLS
	Employees			
log population density	0.003	-0.001	-0.054***	-0.049***
	(0.002)	(0.002)	(0.004)	(0.004)
part time	-0.077***	-0.069***	-0.076***	-0.066***
	(0.006)	(0.007)	(0.006)	(0.007)
fixed term	-0.119***	-0.121***	-0.123***	-0.128***
	(0.004)	(0.004)	(0.004)	(0.004)
log firm size	0.023***	0.022***	0.023***	0.022***
	(0.002)	(0.002)	(0.002)	(0.002)
age dummies	yes	yes	yes	yes
occupation dummies	yes	yes	yes	yes
contract dummies	yes	yes	yes	yes
area dummies	yes	yes	yes	yes
year fe	yes	yes	yes	yes
Observations	42,881,163	34,134,728	42,881,163	34,134,728
R-squared	0.614	0.590	0.599	0.584
	Independent contractors			
log population density	0.051***	0.048***	-0.015	-0.011
	(0.007)	(0.014)	(0.010)	(0.015)
age dummies	yes	yes	yes	yes
industry dummies	yes	yes	yes	yes
contractors dummies	yes	yes	yes	yes
area dummies	yes	yes	yes	yes
year fe	yes	yes	yes	yes
Observations	3,469,655	2,355,707	3,469,655	2,355,707
R-squared	0.194	0.233	0.201	0.242

Note: The dependent variable in columns (1)-(2) is weekly/daily nominal wages respectively for employees and independent contractors; in columns (3)-(4) it is weekly/daily real wages. Occupation dummies stands for, namely, white collar, blue collar, apprentice, manager, and executive; contract dummies are dummies for NCLAs, which cover about 243 contracts. Industry dummies stands for 86 2-digit Ateco industrial sectors; contractors dummies are dummies for 19 categories of independent contractors. Standard errors are clustered at the LLM level. *** p<0.01, ** p<0.05, * p<0.1.

Table 22: Robustness checks: Two-stage procedure as in Combes et al. (2008)

	Nominal wages		Real wages	
	(1) FULL OLS	(2) FE	(3) FULL OLS	(4) FE
	Employees			
log population density	-0.000***	-0.001***	-0.026***	-0.027***
	(0.004)	(0.002)	(0.004)	(0.004)
area dummies	yes	yes	yes	yes
year fe	yes	yes	yes	yes
Observations	6,558	6,558	6,558	6,558
	Independent contractors			
log population density	0.060***	-0.012	0.038***	-0.032*
	(0.010)	(0.014)	(0.011)	(0.015)
area dummies	yes	yes	yes	yes
year fe	yes	yes	yes	yes
Observations	6,399	6,399	6,399	6,399

Note: Estimation results of the second stage of Combes et al. (2008). Columns (1)-(2) refer to daily nominal wages respectively for employees and independent contractors; columns (3)-(4) refer to daily real wages. First stage for employees includes: part time, fixed term, log firm size, age dummies, occupation dummies, contract dummies, LLM $\times$ year dummies, year fixed effects, worker fixed effects (FE only). First stage for independent contractors includes: age dummies, industry dummies, year fixed effects, worker fixed effects (FE only). Standard errors are clustered at the LLM level. ***p<0.01, ** p<0.05, * p<0.1.

Appendix

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% according to 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, 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.

For the purposes of this paper, it is crucial to note that the second-level bargaining is subject to the limits and provisions defined by the specific industry collective agreements. This means that it is subject to the *in melius* or favourability principle, i.e. worker wages and labour conditions cannot be worse than the ones settled at the industry level.

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:

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

where PI_{npt} is the price index of COICOP industrial category n for provincial capital city p in year t and w_{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:

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

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 w_t is the official weight attached to housing prices.

This index, to be useful to our purposes, should be modified in two ways: first, it should reflect the "adjusted" weight that incorporates both the direct (cost of having an accommodation) and the indirect (attraction of housing prices on non-housing prices) effect of housing prices on the cost of living (as explained in Subsection 3.2); second, it should capture local (LLM or municipality level) variability of prices. We do this in two steps following Moretti (2013).

As a first step, we need to substitute w_t for the "adjusted" weight. Accordingly, we split the NHPI dynamics in two components, one correlated to the dynamics of HPI and the other orthogonal to it:

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

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:

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

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

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

β is the "adjusted" weight that incorporates both the direct ($\overline{w}$) and the indirect ($[(1 - \overline{w})\pi$) effect of housing prices on the cost of living. It is estimated by running the following OLS equation regression:

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

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 $\hat{\beta}$ 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%.

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

$$C\hat{P}I_{mt} = \hat{\beta} HPI_{mt} + (1 - \hat{\beta})NHPI_t, \tag{A7}$$

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.[33]

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 Boeri et al. (2017) 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 one-digit COICOP classification level, excluding housing (COICOP04), weights being taken from Istat weighting scheme:

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

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

[33] In a robustness check (see Section 10), 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).