Spatial wage disparities – Workers, firms, and assortative matching *

Wolfgang Dauth
University of Wuerzburg and IAB

Sebastian Findeisen
University of Mannheim and CEPR

Enrico Moretti
UC Berkeley, NBER and CEPR

Jens Suedekum
Duesseldorf Institute for Competition Economics (DICE) and CEPR

October 2016

[PRELIMINARY AND INCOMPLETE VERSION!]

Abstract

Are higher wages in cities driven by *worker* effects, by *firm* effects, or by stronger *assortative matching*? Using rich administrative data from Germany, we decompose the wage structure into person- and establishment-effects over the 1990-2010 period. This allows us to decompose the variation of wages across space, and the evolution of spatial wage inequality over time. We find that better worker-firm matching in denser local labour markets is a key mechanism behind the urban wage premium. This holds for aggregate local labour markets, and even more so within narrowly defined occupation- and industry-specific market segments. Quantitatively, matching adds more to the understanding of the observed trend in spatial wage inequality than all firm-based explanations, and almost as much as all worker-based sources of higher urban wages taken together.

JEL-Classification: R11, R12

Keywords: Urban wage premium, spatial wage disparities, agglomeration, assortative matching, Germany

*The authors thank Gilles Duranton, Laurent Gobillon, Giordano Mion and Michel Serafinelli and participants of the urban economics conference in Barcelona 2016 for helpful comments and suggestions. We also thank Linda Borrs and Florian Knauth for excellent research assistance. We gratefully acknowledge financial support from the DFG-priority program 1764 "The German Labour Market in a Globalised World - Challenges through Trade, Technology, and Demographics". Email: wolfgang.dauth@uni-wuerzburg.de, findeisen@uni-mannheim.de, moretti@berkeley.edu, suedekum@dice.hhu.de.

1 Introduction

Why do firms pay more and workers earn more in larger cities? Economists ever since Marshall (1890) have explored this question, which is central to our understanding why agglomeration of economic activity in space persists. A large literature has documented detailed empirical patterns for various countries, focusing either on the perspective of individual workers[1] or adopting the the perspective of firms.[2] What is still not clear, however, is whether the urban wage premium is driven more strongly from the worker side, from the firm side, or from the matching of the two.

In this paper, we take a fresh look at spatial wage disparities by using rich administrative data from (West) Germany that covers a 50% sample of all full-time employees subject to social security from 1990 to 2010, on average roughly 7.5 million persons in every year. Following the seminal paper by Abowd, Kramarz and Margolis (1999) from the labour economics literature, henceforth AKM, we first run linear log wage regressions with additive person- and establishment-fixed effects over four separate time windows in those 20 years. This allows us, in a second step, to decompose the variation of wages across space, and the evolution of spatial wage inequality over time, into three main parts: *person* effects, *establishment* effects, and their correlation which captures the degree of *assortative matching* within narrowly defined local labour markets.

By applying this AKM decomposition approach in a local context, we analyze the urban wage premium from a different perspective than previous studies and address – for the first time in the literature – the relative importance of worker-based, firm-based, and matching-based explanations for the higher wages in cities. We establish five main facts about spatial wage disparities and regional wage inequality in Germany:

- Fact 1: "Good workers" and "good firms" are concentrated in denser regions.
- Fact 2: The concentration of "good workers" has become more important over time, but not the concentration of "good firms".
- Fact 3: The degree of positive assortative matching (PAM) between workers and firms is stronger in denser regions.
- Fact 4: The density elasticity of PAM has become more important over time.
- Fact 5: PAM within local labor markets for particular occupations or industries is facilitated more strongly by the size of the specific labor market than by the aggregate local labor market size.

The definition of a "good worker" in this paper refers to his or her respective person-effect, which captures all individual-specific wage components that are portable across

[1] See, e.g., Glaeser and Mare (2001), Yankow (2006), Gould (2007), Combes et al. (2008), Baum-Snow and Pavan (2012), Eeckhout et al. (2014), Davis and Dingel (2015), De la Roca and Puga (2016).

[2] See, e.g., Henderson (2003), Moretti (2004), Combes et al. (2012) or Gaubert (2015).

different jobs in a given time frame. Similarly, a "good firm" pays a proportional wage premium to all its workers, as captured by the respective establishment-effects. Since wages in the AKM model are a multiplicative combination of the two components, this allows us to decompose if wages are higher in cities because urban workers are employed in establishments that generally pay well, if urban workers are generally well paid and would also earn more in other workplaces, or if the urban wage premium mostly stems from the fact that "good workers" match with "good firms".

Fact 1 is consistent with the large literature cited above which argues that wages are higher in cities partly because productive workers and firms choose to sort into those urban locations, and partly because density makes them more productive (either instantaneously or over time) through various agglomeration economies. Our empirical exercise neither aims to disentangle sorting from agglomeration, nor to identify particular micro-foundations of the latter as surveyed in Duranton and Puga (2004), with one exception discussed shortly. Fact 1 is consistent with both explanations and cuts through the urban wage premium in another way, by asking whether firm- or worker-specific wage components are more important to understand the observed evolution of spatial wage disparities. Fact 2 provides a clear answer in this respect: It suggests that increasing wage inequality between cities and rural areas is unlikely to be driven by an increasing gap of compensation schemes of urban and rural firms, but rather by an increasing difference of worker-specific wage components. In other words, the rising wage disparity between, say, Munich and Dortmund does not stem from the fact that Munich firms pay increasingly better than Dortmund firms in general. It comes from the fact that Munich workers tend to earn increasingly better than Dortmund workers, and would keep this advantage also with other employers in different local environments.

Turning to facts 3–5, which form the main contribution of this paper, we may summarize their main insight as follows: "matching matters"! Better matching between workers and firms in dense local labour markets is one of the canonical Marshallian agglomeration forces, and it has been modelled in an urban context by Helsley and Strange (1990), Acemoglu (1997), Rotemberg and Saloner (2000), and others, who formalize different form of localized increasing returns in the matching technology (also see Moretti 2011). Labour economists have discussed the prevalence of assortative matching (Becker 1973; Shimer and Smith 2000; Shimer 2005), and the identification of PAM in AKM models (see Eeckhout and Kircher 2011; Abowd et al. 2014; Chade et al. 2016), but that literature typically ignores the local dimension.

Unlike for the other micro-foundations, our empirical analysis can provide direct evidence for this particular channel, and we can compare the quantitative importance of matching relative to other sources of the urban wage premium. Fact 3 shows that better assortative matching of workers and firms in denser labour markets is a key mechanism, which according to fact 4 has also become increasingly important. Indeed, the elasticity of PAM (measured by the correlation of firm- and worker-effects in the region) with respect to density has substantially increased over time. Moreover, a variance decomposition of the change in between-city wage inequality shows that the share which can be attributed to the rising quality of urban workers is only a bit larger than the share that is accounted for by stronger PAM in cities. When it comes to spatial wage inequality in Germany, we thus find that assortative matching alone is quantitatively almost as important as all worker-based explanations (sorting and other agglomeration effects) together, and clearly more important than firm-based explanations which add nothing to the understanding of this trend.

The existing empirical literature on assortative matching at the local level is scant and mixed. The most closely related reference is Andersson et al. (2007) who find stronger PAM in denser counties in Florida and California. Our fact 3 is consistent with this evidence. By contrast, Figueiredo et al. (2014) find little support for stronger PAM across Portuguese industrial clusters, while Mion and Naticchioni (2009) find a positive correlation of individual ability and firm size in Italian regions, but a negative correlation of assortativeness and density. One contribution of our paper is that we find stronger PAM in denser German regions not only in the cross-section (fact 3), but in particular, that PAM has become considerably more important over time (fact 4), thereby contributing substantially to the rise in spatial wage inequality.

Turning to fact 5, we not only consider aggregate regional labour markets, but also narrowly defined occupation- and industry-specific cells. Moretti (2011) was the first to raise the question which type of market thickness is most important for the matching of firms and workers in local environments. He conjectures that the specific local density for particular types of job may be more important than aggregate density. In his words, *"a bioengineer and an architect living in the same city may face different market thickness, depending on the local agglomeration of bioengineering firms and architectural firms."* Empirical evidence on the relative importance of aggregate versus specific density for matching in cities has been missing so far, however. We study this issue by moving to fine-grained labor market cells for particular occupations or industries, and fact 5 actually supports Moretti's original conjecture. That is, our findings are in line with the claim that the assortativeness of matching of bioengineering workers with bioengineering firms is better in local environments where many such bioengineers are around, and only to a lesser extent in larger cities with more workers in general.

Finally, our paper is related to the recent study by Card, Heining and Kline (2013), henceforth CHK, who have also applied the AKM model to the case of (West) Germany, and found that its strong separability assumptions are closely met in the data. They explore the contributions of worker- and firm-effects, as well as of assortative matching, for the observed rise in overall wage inequality in Germany. Moreover, they study how rising wage dispersion across occupations, industries, or education groups can be decomposed into those three categories, and find that rising workplace heterogeneity (the establishment-effects) plays a substantial role to explain those inequality trends. CHK do not analyze the local dimension of the German labour market, however, which is our main focus. For spatial wage disparities, we find that the firm-effects are of considerably lower importance for between-city wage inequality than worker-effects and PAM. This suggests that the education-, occupation-, and industry-specific inequality trends that were driven by workplace-specific characteristics (e.g., better management practices in "good firms" and the associated wage gains) occurred uniformly across space and were not primarily an urban phenomenon.

The rest of this paper is organized as follows. In Section 2, we introduce the data, present some preliminary evidence on the urban wage premium in Germany, and discuss the AKM estimation approach that we implement by closely following CHK. In Section 3, we explore the spatial patterns of worker- and establishment-effects, and in Section 4 we analyze PAM in cities. In Section we provide our decomposition of between-city wage inequality, and Section 6 concludes.

2 Data and empirical approach

2.1 Data

The estimation of AKM models requires a large and comprehensive data set that covers not only most (or, ideally, all) of the country's workforce, but also most or all firms as the heterogeneity of worker- and firm-effects is identified from mobility of individuals across establishments within a connected set. For Germany, such a dataset is provided by the Integrated Employment Biographies (IEB) of the Institute for Employment Research (see Oberschachtsiek et al. 2009). This data covers the universe of all German workers subject to social security, excluding only civil servants and self-employed persons, and allows to follow the entire job biographies of those persons over time.

While the recent study by CHK works with the full worker sample in the IEB data, we use for the moment (for computational reasons) a 50% random sample of all male full-time workers aged 20 to 60, who work in one of the West German Federal States or in Berlin. Our data set covers 35,447,834 million worker-year-observations and a total of 7,538,784 individuals in the time period 1990–2010. For every worker we record total earnings and days in employment in each year, as well further biographical information such as age, education, occupation, and detailed place of residence. To match workers and firms, we use unique identifiers at the establishment level which allow us to recover the person's entire history of different workplaces. The estimated firm-components of wages, therefore, refer to the establishment-level.

One well-known problem of this data is the top-coding of wages at the social security contribution ceiling (around 140 € per day in 2010). To tackle this issue we implement the imputation approach by CHK to prepare the final data. Furthermore, we focus on the time after the German reunification and split our data into four five-year intervals (1990-1995, 1995-2000, 2000-2005, and 2005-2010) and conduct all analyses separately for each of those intervals.

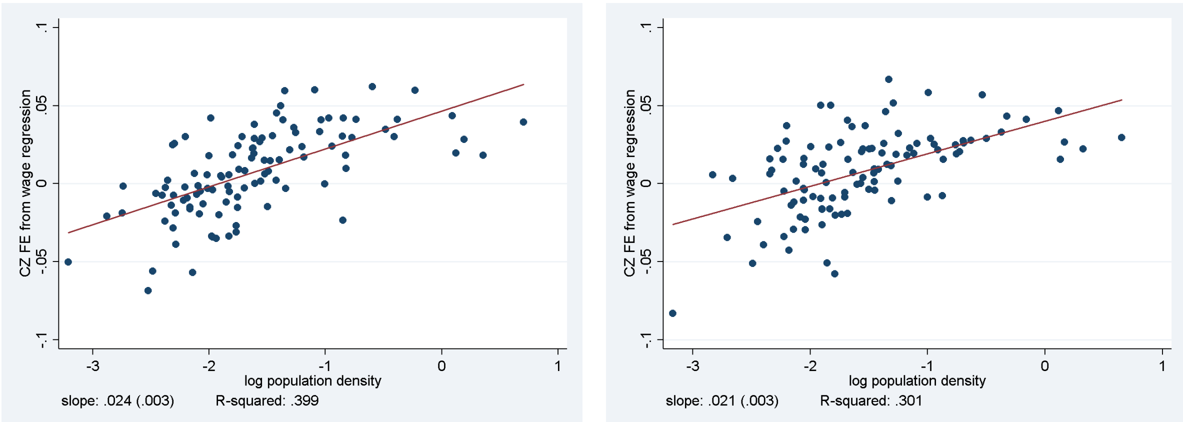
2.2 Conventional urban wage premium estimates

Since our analysis is related to previous studies on the urban wage premium, we demonstrate that our data yield comparable results when we employ the conventional estimation approach. Starting with the seminal study by Glaeser and Mare (2001), including worker-fixed effect is seen as important to tackle the issue of sorting of individuals with higher ability into cities, and to disentangle sorting from agglomeration economies which are measured by the impact of density on wages.

Following Combes et al. (2008), we use a 2-stage procedure to estimate the urban wage premium. In the first step we run the following regression:

$$\ln(\text{wage}_{it}) = \mu_i + \sigma_{c(i,t)} + \phi_{s(i,t)} + X'_{it}\gamma + \epsilon_{it}, \tag{1}$$

where μ_i is an individual fixed-effect, $\sigma_{c(i,t)}$ is a region fixed-effect for the current location of a worker, and X'_{it} is a vector of standard individual-level control variables.[3] As in Combes et al. (2008), we also include 3-digit industry fixed effects $\phi_{s(i,t)}$ in this step, to control for different sectoral compositions across local labor markets.



(a) 1990-1995

(b) 2005-2010

Figure 1: Conventional Urban Wage Premium

The identification of the fixed-effects in (1) comes from individuals who move across cities, and the common assumption in this literature is that mobility is random conditional on the included observables. A worker's μ_i should be interpreted as the value of his or her human capital (including formal education and unobserved ability) which is equally rewarded across cities, and the city fixed-effects $\sigma_{c(i,t)}$ capture the proportional wage premium paid to all workers employed in the respective city.

In the second step we then regress the estimated city fixed-effects on an intercept and log density. Figure 1 shows the results, and for brevity we focus only on the first

[3]For consistency, we use the same control variables in this model as in the AKM specification from section 2.3, namely education-specific age profiles with a cubic functional form a set of year dummies. To circumvent the age-period-cohort identification problem, we follow CHK and center the age variable around 40 and leave out the linear term to achieve identification (see Card, Cardoso, Heining and Kline (2016) for a detailed discussion of this assumption).

and the last time interval (1990–1995 and 2005–2010). In both periods, we find a highly significant coefficient of around 0.02 (std. error 0.003), implying that a doubling of population density leads to a wage increase of 2% for the average city resident. This estimate is closely in line with the findings by Combes et al. (2008) for France and De la Roca and Puga (2016) for Spain using comparable empirical specifications, and shows that the German labor market behaves similar in this respect.[4]

2.3 The AKM approach

Our main point of departure is the observation that *firms* play essentially no role in this conventional analysis of the urban wage premium. However, a recent literature in labor economics has emphasized strong heterogeneity in employer-specific wage components, i.e., that some employers generally pay better than others. That literature has, however, rarely investigated the spatial dimension of those wage differences.

The leading approach to differentiate wages into worker- and establishment-specific pay components is the AKM model (Abowd, Kramarz and Margolis, 1999). It assumes that the (log) wage of a worker can be written as:

$$\ln(\text{wage}_{it}) = \mu_i + \Psi_{\mathbf{J}(i,t)} + X'_{it}\gamma + \epsilon_{it}, \tag{2}$$

where X'_{it} are the same observable worker characteristics used before, namely a cubic term in age fully interacted with the formal education level and additional calendar year fixed effects, and where μ_i are $\Psi_{\mathbf{J}(i,t)}$ are subsequently referred to as the *person-effects* and the *establishment-effects* of individual wages, respectively.

The person-effects μ_i capture a worker's unobserved human capital components (encompassing factors such as ability, motivation, or psychological traits) that are transferable across jobs and rewarded equally by all employers in the respective period. The establishment-effects $\Psi_{\mathbf{J}(i,t)}$ can be interpreted as a proportional pay premium (or discount) paid by an establishment to all its workers in a given time interval, independently of the worker's observable or unobservable skills. Both wage components are thus time-invariant within any given time interval, but may vary across intervals for every worker and establishment when observed in more than one interval.

In a recent contribution, CHK have implemented this model (2) for the (West-)German labor market using the IEB data and provide a detailed decomposition of aggregate wage inequality over the period 1985–2009. They conduct various specification checks for the testable assumptions of the AKM approach with additive wage components, and they provide a detailed discussion about the properties of the error term ϵ_{it} that yield an unbiased estimation of the person- and establishment effects. To recap briefly, they show that workers of different skill groups receive approximately the same propor-

[4]Our findings are also close to the fixed-effects estimate of the urban wage premium in Germany by Hirsch et al. (2016) who use a 2% random sample of the IEB data. Their study focuses on differential search frictions across regions, but does not disentangle person- and establishment-specific wage components and does not address assortative firm-worker matching at the local level.

tional wage premiums at a given establishment – consistent with the simple additive structure of equation (2). Second, a fully saturated model with job-specific fixed-effects only yields a marginal improvement in terms of data fit. Finally, the match-specific component of the residual is uncorrelated with the direction of mobility between high- and low-paying firms. Taken together, the identifying assumption needed for estimation of the AKM model, thus. seem to be closely met in the German data. CHK still note that the AKM approach is not a structural model of the labor market, and that the identified person- and establishment-effects do not necessarily measure true ability or productivity; also see Eeckhout and Kircher (2011) or Abowd et al. (2014). Their descriptive analysis for aggregate wage inequality trends is then based on the assumption that this bias is similar in earlier and later time periods, and we employ a similar assumption that this measurement error of the person- and establishment-effects is constant not only over time but also across space.

The main value-added of this paper is that we explore the local dimension of the German labor market in the context of the AKM approach, more specifically, the spatial configuration of the correlation between person- and establishment effects. To do so, we first replicate the estimation approach by CHK, and apart from the fact that we work with a 50% sample instead of the full data set for computational reasons, we only deviate in how we slice the data into time intervals. In particular, we focus on the time after the German reunification and split our data into four five-year intervals (1990-1995, 1995-2000, 2000-2005, and 2005-2010). Reassuringly, we obtain very similar results to CHK when considering aggregate wage inequality trends in (West-)Germany; see, for example, Table 5 below which directly replicates one of their main results.

2.4 Exploring the variation of establishment-effects

Connecting this AKM approach with the literature on the urban wage premium raises one important conceptual issue: while there is substantial mobility of single workers across regions and industries over time, it is a salient feature of the data that existing establishments virtually never change their location or their recorded industry affiliation. All time-invariant region- and industry-specific impacts on wages of the type included in regression (1) are, thus, captured by the $\Psi_{\mathrm{J}(i,t)}$-terms. This raises the question how much of the variation in establishment-effects is driven by those components ($\sigma_{c(i,t)}$ and $\phi_{s(i,t)}$), and how much of it is due to genuine firm-specific differences in compensation schemes orthogonal to industry or local effects. To explore this question, we run an intermediate regression of the following form:

$$\Psi_{\mathrm{J}(i,t)} = \sigma_{c(i,t)} + \phi_{s(i,t)} + \alpha_1 E_{J,t0} + \alpha_2 E_{J,t0}^2 + \nu_{J(i,t)}, \tag{3}$$

That is, we regress the establishment-effects identified from (2) on a set of (2-digit) industry dummies and regional dummies, while controlling for the inital employment size of establishment J (and its squared term) to explore if establishment-effects differ systematically with firm size. We then obtain the R^2 of this regression, and decompose the overall fit of the model into the contributions of the single variable groups.[5]. Table 1 reports the results of this decomposition exercise.

Table 1: Decomposition of the AKM establishment-effects

Period	1990-95	1995-2000	2000-05	2005-10
Number of plants	488820	527711	551418	514501
R-squared	0.1311	0.1529	0.1493	0.1258
% Contribution of				
plant size	2.64	3.02	3.18	3.55
local labor market	11.19	7.99	4.92	4.24
industry	86.18	88.98	91.90	92.20

Notes: Decomposition of the R-squared of a regression of pre-estimated AKM firm-fe on plant size, 109 local labor market and 60 2-digit industry dummies.

This table shows that at most 15 % of the variation in establishment-effects can be jointly explained by the combination of firm size, industry- and local components, while the bulk of the variation seems to stem from genuine firm-specific differences orthogonal to those dimensions. This emphasizes the importance of investigating the role of *firms* for spatial wage disparities, which is missing in the conventional regression from equation (1), since individual wages seem to be affected by establishment-specific components over and above local and industry effects.

Furthermore, the other main message of Table 1 is that industry-effects have the largest explanatory power for the variation of establishment-effects among the three investigated components. To take systematic industry-differences into account in the subsequent analysis of spatial wage disparities, we construct a set of establishment-effects that are purged of those influences. In particular, we run a simplified version of equation (3) with industry-dummies only, and then recover the residuals of this regression which we henceforth label "neutralized establishment-effects". Below we consider the spatial configuration of the basic establishment-effects and these neutralized counterparts, which are orthogonal to local industry structures by construction.

3 Person- and establishment-effects in space

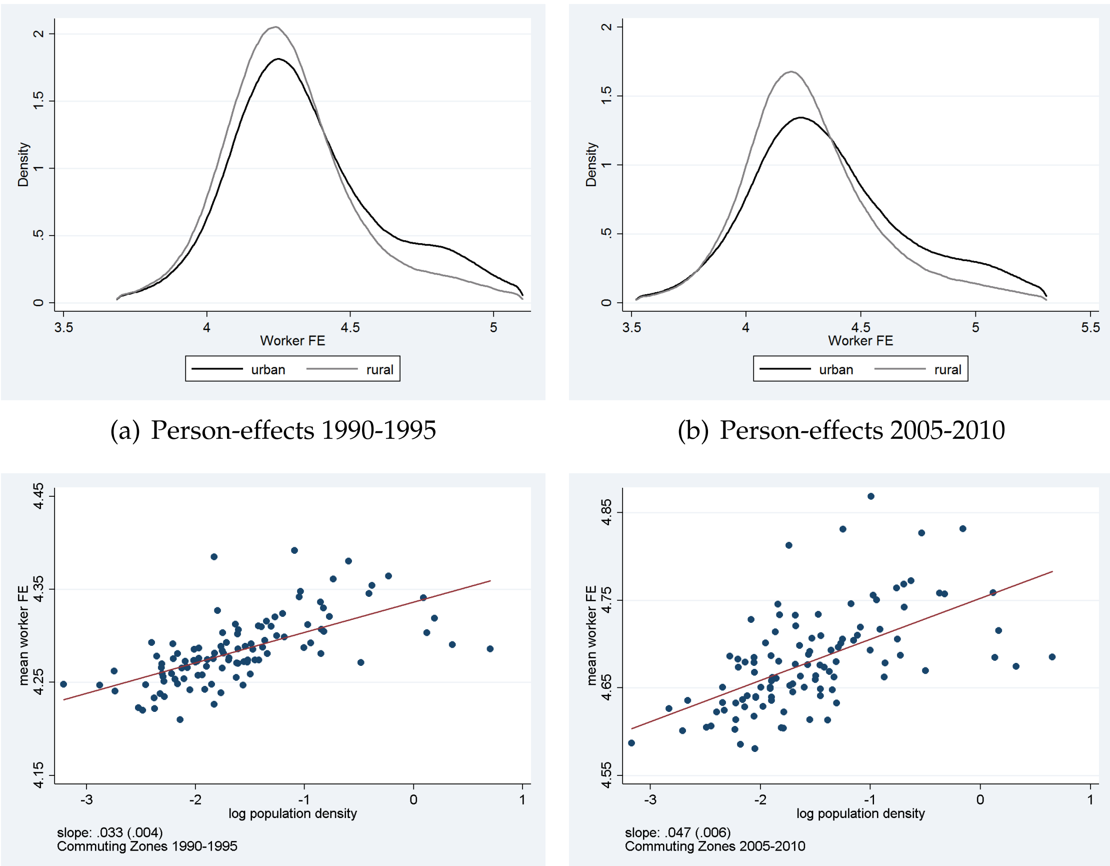
We first describe the spatial distribution of the full set of person- and establishment-effects. Our analysis refers to three different geographical levels: a) 108 consistently defined travel-to-work areas (*Arbeitsmarktregionen*), b) 325 West German administrative districts (*Kreise, without Berlin*, NUTS-3 regions comparable to US counties), and c) 8236 small-scale municipalities (*Gemeinden*). For the most part, our analysis refers to the broadest regional definition a), because we believe that those commuting zones

[5]This approach follows Huettner and Sunder (2012) implemented in the STATA package `rego`.

are the most sensible definitions of local labour markets, but in the appendix we also consider the more fine-grained regional units to check the robustness of our results.

3.1 Person-effects

Figure 2 summarizes our results for the person-effects. In panels a) and b) we first distinguish *urban* and *rural* local labour markets, following a standard classification from the German Federal Agency for Building and Regional Planning (BBR). The figure shows the distribution of person-effects among both types of local labor markets, and clearly shows a pattern of first-order stochastic dominance: Cities host workers with higher individual-specific wage components. This is already true in the first time interval (1990–1995), and even more so in the second interval (2005–2010) where the distributions have widened and the stochastic dominance features even more clearly. In other words, "good workers" with high person-effects have been over-represented in cities already in 1990–1995, but even more so in 2005–2010.



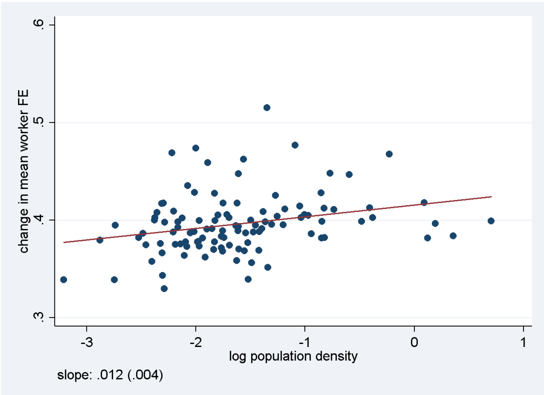
(a) Person-effects 1990-1995 (b) Person-effects 2005-2010

(c) Density and mean person-effects, 1990-1995 (d) Density and mean person-effects, 2005-2010

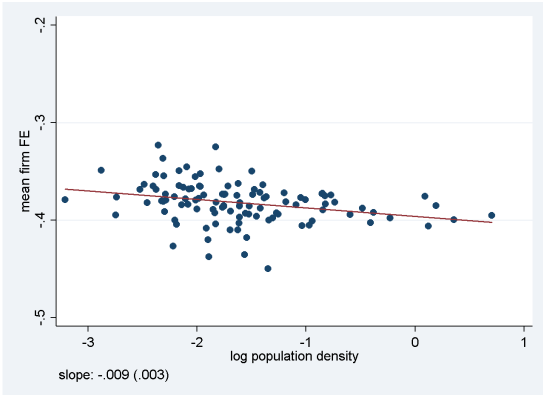
Figure 2: Worker-effects in space

In panels c) and d) we depict the relationship between initial (log) population density and the mean person-effect in the respective region for the first and the last time interval. We find highly significant elasticities of 0.033 and 0.047, respectively. That is, doubling local population density is associated with a higher mean person-effect, and this relationship has become more pronounced over time. This is a different way of saying that "good workers" with high individual-specific wage components are mostly to be found in cities, not in rural areas, and particularly so in the later time period.

In panel a) of Figure 3 we plot the *changes* of the mean person-effects from 1990–95 until 2005–10 against initial (log) population density of the respective region. We find significantly stronger *increases* in initially denser regions, which reinforces our previous finding that cities increasingly host "good workers".

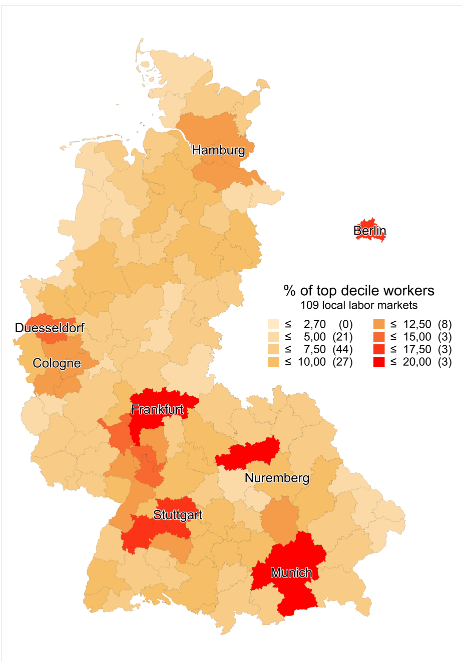


(a) Density and change of mean person-effects over time (period 1-4)

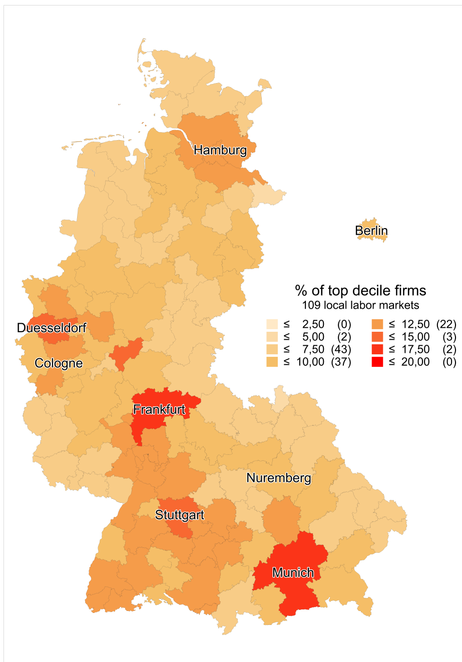


(b) Density and change of mean establishment-effects over time (period 1-4)

Figure 3: Changes of mean person- and establishment effects



(a) Share of top workers in local employment



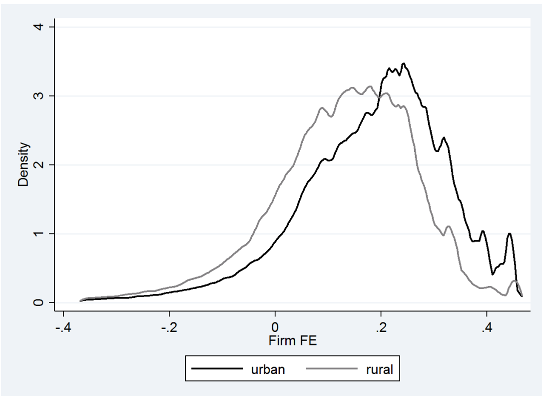
(b) Share of top firms among all local firms

Figure 4: Top workers and firms in space

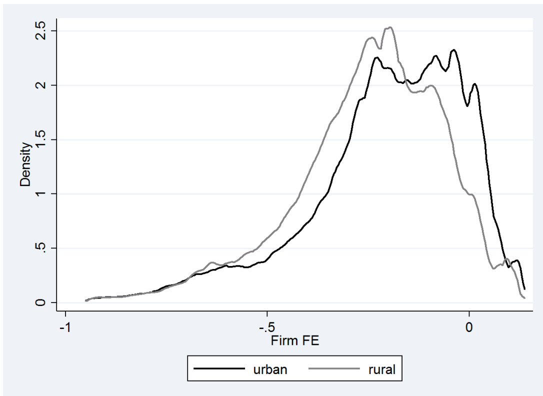
Finally, in panel a) of Figure 4 we illustrate the spatial configuration of person-effects. This map focuses on "top workers" in the local labor market, more specifically, it depicts the share of workers from the upper decile of the overall West German person-effect distribution among all employees in the respective region. In line with Figure 2, this maps show that some of the largest cities (Berlin, Munich, Frankfurt, Stuttgart, Duesseldorf, etc.) are clear hotspots for "top workers" in West Germany.

3.2 Establishment-effects (basic and neutralized)

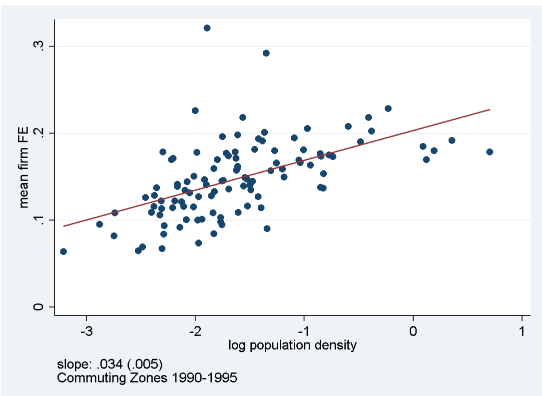
Figure 5 provides analogous illustrations for the basic establishment-effects identified from equation (2). Their distribution in urban and rural areas, as depicted in panels a) and b), appear much more bumpy than for the person-effects in Figure 2. This is due to the fact that the term $\Psi_{\mathrm{J}(i,t)}$ is identical for all workers employed in establishment J, but these distributions are shown across all workers i, which thus exhibit peaks at the values of large plants. Regardless of those bumps, we also observe a clear pattern of first-order stochastic dominance in those distributions, i.e., cities dis-proportionally host "good firms" with high establishment-specific wage components.



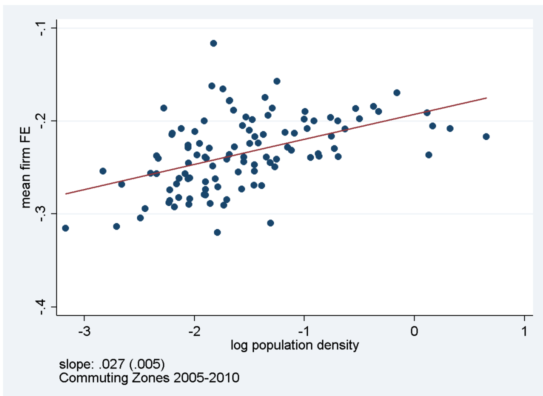
(a) Establishment-effects, 1990-1995



(b) Establishment-effects, 2005-2010

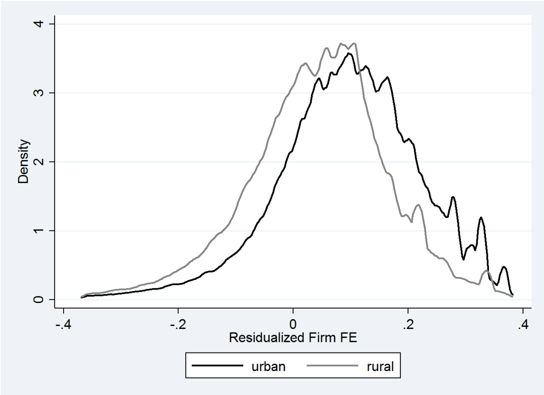


(c) Density and mean establ.-effects, 1990-1995

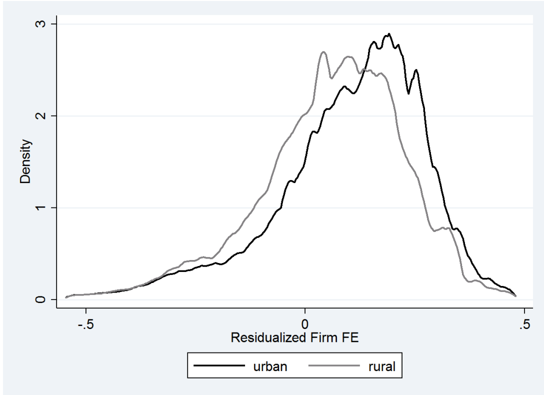


(d) Density and mean establ.-effects, 2005-2010

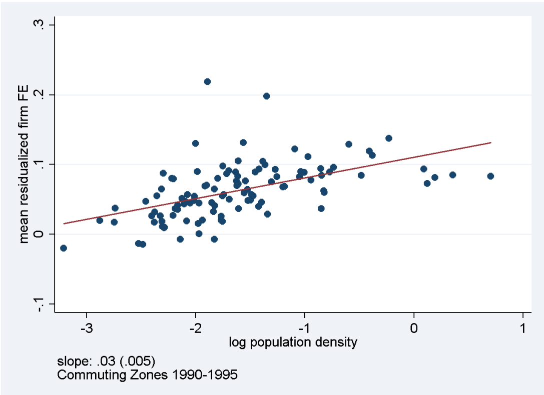
Figure 5: Basic establishment-effects in space



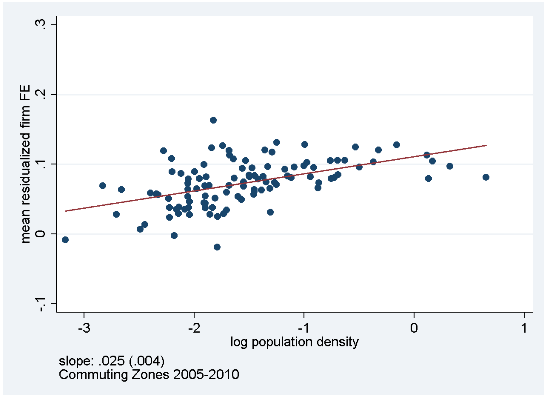
(a) Neutralized establ.-effects, 1990-1995



(b) Neutralized establ.-effects 2005-2010



(c) Density and neutralized establ.-eff., 1990-95



(d) Density and neutralized establ.-eff., 2005-10

Figure 6: Neutralized establishment-effects (purged of industry effects) in space

This pattern already prevailed in the first period, but it has not become stronger over time. Panels c) and d) of Figure 5 show that the elasticity of mean regional establishment-effects with respect to population density is positive and significant, but it has remained constant over time at around 0.023. Consistently, we find in panel b) of Figure 3 that there in no (or, if anything, a negative) relationship between density and the *change* of the mean establishment-effect across regions.

Finally, the map in panel b) of Figure 4 shows the spatial configuration of "top firms" in West Germany. It reveals that the patterns for workers and establishments are clearly correlated, but it is also evident that the degree of concentration is stronger for the former. For example, Berlin has a high density of "top workers" but not of "top firms", whereas non-urbanized regions in the South-West (Baden-Wuerttemberg) have a lower density of "top workers" than of "top firms".

Figure 6 is analogous to Figure 5 and shows the spatial pattern of the "neutralized" establishment-effects, which are purged of industry-specific wage components. Results turn out to be very similar, both for the shape of the urban versus rural distributions, and for the relationship of the mean regional effect with (log) population density. Also for the neutralized effects, we obtain an elasticity slightly above 0.02 that has not increased over time. For brevity we omit the analogue to Figure 3 for the neutralized establishment-effects; it consistently shows no relationship between density and the change in regional mean effects and an even more equal spatial pattern in West Germany when industry differences are controlled for.

3.3 Summary

Taken together, the evidence assembled in this section may be summarized as follows:

- Fact 1: "Good workers" and "good firms" are concentrated in denser regions.
- Fact 2: The concentration of "good workers" has become more important over time, but not the concentration of "good firms".

Fact 1 is consistent with a large previous literature on spatial wage disparities. Importantly, this fact also prevails after taking out systematic industry-specific wage disparities. That is, establishments in cities tend to be high-wage employers independently of their industry affiliation. Fact 2 is more novel and refers to the *dynamics* of spatial wage disparities. It suggests that increasing wage inequality between cities and rural areas is unlikely to be driven by an increasing gap of compensation schemes of urban and rural firms, but rather by an increasing difference of worker-specific wage components. Again, this is is true also after taking systematic industry-differences into acoount. We will come back to this issue in Section 5 below where we actually decompose the observed trend of spatial wage inequality along those dimensions.

Yet, what is still missing from facts 1 and 2 is the role of worker-firm matching and its contribution to spatial wage disparities. We now turn to this main issue of our empirical analysis by analyzing in detail the correlation of person- and establishment-effects in West German local labor markets.

4 Positive assortative matching (PAM) in space

In the AKM model, log wages are additive in worker and firm components but have no genuine match-specific part. Still, there can be a rationale for positive assortative matching (PAM) as "good workers" may earn relatively more than "bad workers" when working for a high-wage establishment.

The following toy model may be useful to grasp the economic intuition: Suppose the country has only two regions, a city C and a rural hinterland R, and in line with fact 1 we assume that C hosts better workers and firms. Let there be two equally large groups of workers in the city with ability $\mu_i^C = \{11, 9\}$, and two equally large groups of firms with productivity $\nu_J^C = \{110, 90\}$. In the hinterland, we also have two equally sized groups, but with $\mu_i^R = \{6, 4\}$ and $\nu_J^R = \{60, 40\}$ only. Now, suppose that a worker-firm match generates revenue $\mu_i \times \nu_J$, which is equally split.[6] With initial random matching

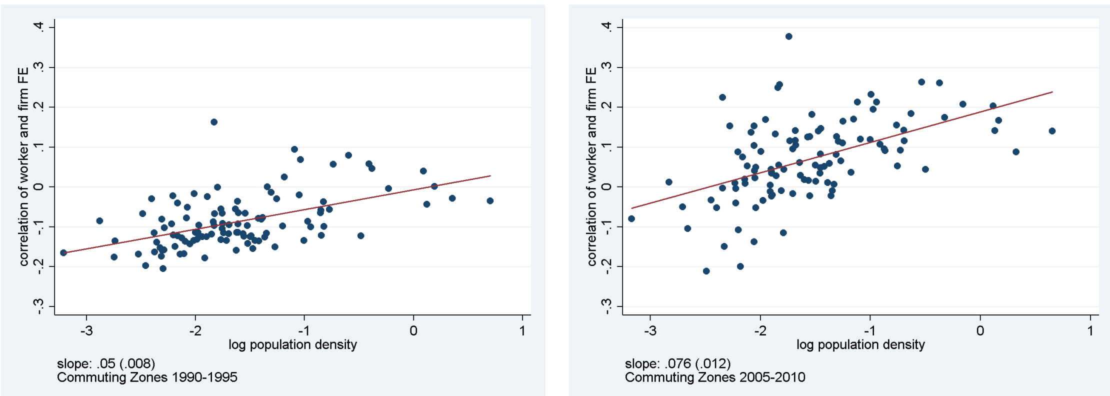
[6] Notice that this super-modular production function in levels is consistent with a log-wage equation as in (2).

within every region, the average wage (and profit) is thus $(10 \times 100)/2 = 500$€ in the city and only $(5 \times 50)/2 = 125$€ in the hinterland. Finally, close the model and assume that workers can freely send out as many job applications as they like, but only locally, and that the firms receiving those applications can switch their single employee at a fixed cost F which is the same in both regions (or at least not much higher in the city).

It is easy to see that good urban firms have the strongest incentive to re-match in this model. A good urban firm that is initially matched with a bad urban worker can gain $(110 \times 11 - 110 \times 9)/2 = 110$€ when firing their old and switching to a new guy. This is higher than for a bad urban firm (which can gain 90€) or for a good rural firm (which can gain 30€). For a certain range of F only the good urban firms will have turnover in equilibrium, and eventually all employ good workers. Hence, this simple toy model illustrates how we can have PAM despite the absence of direct match-effects in the log wage equation, and also stronger PAM in cities.[7]

4.1 Aggregate local labor markets

To measure the degree of assortativeness of worker-firm matching, we compute the correlation between all person-effects in a local labor market and the establishment-effects of the respective workplace at the end of the interval. In Figure 7 we plots the resulting region-specific correlation coefficients against (log) population density across all 108 local labour markets. Panel a) refers to the first and panel b) to the last time interval. Figure 8 is constructed analogously and uses the correlation between person- and neutralized establishment-effects purged of industry-specific wage components.



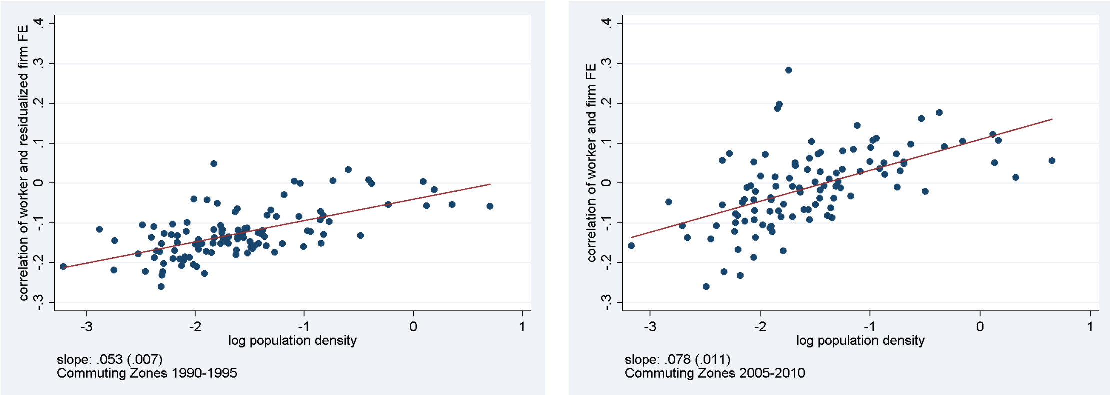
(a) 1990-1995

(b) 2005-2010

Figure 7: Density and correlation of person- and establishment-effects

We find a positive and statistically highly significant relationship with an estimated elasticity of 0.050–0.053 in the first, and 0.076–0.078 in the second interval. That is, denser local labour markets are characterized by a substantially stronger degree of

[7] This toy model only serves an illustrative purpose. See Chade et al. (2016) for a recent overview of static and dynamic search models, and a precise characterization of the conditions for PAM between heterogeneous workers and firms (though without reference to local labor markets).



(a) 1990-1995

(b) 2005-2010

Figure 8: Density and correlation of person- and neutralized establishment-effects

assortativeness in worker-firm-matches.[8] This conclusions holds for both sets of correlations (with basic and neutralized establishment-effects), which means that stronger PAM in cities does not reflect differences in local industry structures in combination with cross-industry differences in the quality of matching, but rather suggest that denser local labor markets exhibit a better worker-firm matching independently of local industrial structures. These results are similar for the smaller-scaled regional unit (see Appendix A.2), and in all cases we also find that the relationship between density and the degree of assortativeness has become stronger over time.

Figures 7 and 8 suggest that stronger PAM in denser local labor markets in an important underlying source of the higher wages in cities, particularly in the more recent time period. One potential problem, however, is that population density may at least partly be driven by the expectation of better matching, and the resulting sorting of productive workers and firms into cities. To address this concern of reverse causality, we consider a simple instrumental variable (IV) strategy to strengthen the causal interpretation of our results. In particular, in a similar vein as Combes et al. (2012), we use historical population data as instruments for current city sizes.

The earliest year where such data is available is from the Statistical Yearbooks of the German Federal Statistical Office in 1952 which allow us to construct population levels for all but one local labor markets in our data (the Saarland was not part of Germany at that time). Assuming that a) people in 1952 have not anticipated future differences in local matching technologies, and b) that there is no omitted variable that simultaneously affects the lagged population and the current density, those are valid

[8]Notice that the *level* of the correlation coefficient on the vertical axis is not particularly high, for some regions it is even negative. This is a common pattern in applications of the AKM approach, however, that is scrutinized in a substantial literature which argues that the measured correlation understates the true degree of PAM (e.g., Andrews et al. 2008; Eeckhout and Kircher 2011; Abowd et al. 2014). Also see the applications by Card et al. (2013) and Andersson et al. (2007) who find similar magnitudes and domains for the correlations as in our Figure 2. For this paper, we are less interested in the *level* of the correlation, but rather how it varies with population density and over time, and this inference is unaffected by those identification problems when the bias is constant.

instruments since they are not directly correlated with current assortativeness. Table 2 presents the regression results, with the upper panel referring to the first and the lower panel to the last time interval.

Table 2: Assortativeness and the size of the aggregate local labor market

	(1) OLS	(2) 1^{st} stage	(3) red.form	(4) 2SLS	(5) OLS
Period 1990-95					
log population in LLM	0.0524***			0.0559***	
	(0.007)			(0.008)	
log 1952 pop. in LLM		0.8397***	0.0469***		
		(0.041)	(0.007)		
log total workers in LLM					0.0473***
					(0.007)
log area	-0.0141	0.1654**	-0.0079	-0.0171	-0.0115
	(0.011)	(0.064)	(0.011)	(0.012)	(0.011)
Constant	-0.6550***	7.0034***	-0.2854***	-0.6769***	-0.5492***
	(0.084)	(0.413)	(0.073)	(0.087)	(0.076)
R^2	0.371	0.869	0.343	0.369	0.381
Period 2005-10					
log population in LLM	0.0827***			0.0919***	
	(0.011)			(0.013)	
log 1952 pop. in LLM		0.7920***	0.0728***		
		(0.040)	(0.010)		
log total workers in LLM					0.0771***
					(0.009)
log area	-0.0243	0.2274***	-0.0110	-0.0319*	-0.0220
	(0.017)	(0.062)	(0.016)	(0.017)	(0.015)
Constant	-0.8173***	6.8862***	-0.2457**	-0.8786***	-0.6508***
	(0.129)	(0.413)	(0.107)	(0.134)	(0.111)
R^2	0.378	0.860	0.369	0.375	0.433

Notes: 108 observations (columns 1 and 5), 107 observations (columns 2-4). Dependent variables: Correlation of worker-FE and firm-FE on local labor markets (columns 1, 3, 4, 5), log current population (column 2).

The first column reports the ordinary least squares (OLS) results when regressing the degree of assortativeness on current (log) population and (log) area size instead of (log) population density. The second column reports the first-stage results and shows that lagged population is, indeed, a strong predictor of current population levels, and column 3 shows the corresponding reduced-form results. In column 4 we then present the instrumental variable results from the two-stage least squares (2SLS) estimation. The estimated elasticity is 0.0559 in the first and 0.0919 in the last time period, which is close to the OLS coefficients from column 1. In sum, this IV approach suggests that a thicker local labor market actually leads to an improved worker-firm matching.

Finally, we consider the total number of workers (in logs) in column 5. This specification is useful for the further analysis below, because it shows that the elasticity of assortativeness with respect to this total number of workers in the local labor market is very similar to the elasticity with respect to total local population.

4.2 Specific Local Labor Markets

So far, we have investigated the impact of density on PAM at the level of aggregate local labor markets. However, localized increasing returns in the matching technology may not arise at the city level, but may be specific to narrow market segments, as also suggested by the quote from Moretti (2011) in the introduction. We now move to fine-grained market cells for particular occupations or industries and provide novel evidence for the density elasticity of PAM at a narrowly disaggregated level.

Table 3: Assortativeness and the size of the occupational local labor market

	(1)	(2)	(3)	(4)
Period 1990-95				
log total workers in LLM	0.0316***	0.0131***	0.0100**	
	(0.003)	(0.003)	(0.004)	
log workers in spec. LLM		0.0280***	0.0255***	0.0280***
		(0.002)	(0.004)	(0.002)
log area	-0.0152***	-0.0143***	-0.0129***	
	(0.005)	(0.005)	(0.004)	
Constant	-0.3970***	-0.3833***	-0.3407***	-0.3386***
	(0.038)	(0.038)	(0.033)	(0.016)
Occupation-FE	-	-	Yes	-
LLM-FE	-	-	-	Yes
R^2	0.016	0.039	0.390	0.066
Period 2005-10				
log total workers in LLM	0.0364***	0.0075*	0.0152***	
	(0.004)	(0.004)	(0.005)	
log workers in spec. LLM		0.0477***	0.0323***	0.0481***
		(0.003)	(0.005)	(0.003)
log area	-0.0209***	-0.0195***	-0.0186***	
	(0.006)	(0.006)	(0.005)	
Constant	-0.3678***	-0.3744***	-0.3631***	-0.4379***
	(0.046)	(0.045)	(0.039)	(0.018)
Occupation-FE	-	-	Yes	-
LLM-FE	-	-	-	Yes
R^2	0.014	0.066	0.444	0.108

Notes: 6469 observations (upper panel); 6099 observations (lower panel). Dependent variable: Correlation of worker-FE and firm-FE on local labor markets. Drop cells with less than 50 workers or 5 firms.

First we look at the occupational dimension. We distinguish 85 different (2-digit) occupations, and observe all existing jobs (worker-establishment matches) at the end of the interval (in 2010, or respectively, in 1995). We then compute the correlation of person- and establishment-effects separately for each occupation within each region. This is, of course, quite noisy as some occupational cells are empty or very small in some regions. We therefore restrict our attention to cells with at least 50 workers and 5 firms, and then regress those assortativeness measures on the size of the aggregate and the specific local labor market. Table 3 reports the results.

Table 4: Assortativeness and the size of the industry-specific local labor market

	(1)	(2)	(3)	(4)
Period 1990-95				
log total workers in LLM	0.0437***	0.0355***	0.0088**	
	(0.003)	(0.004)	(0.004)	
log workers in spec. LLM		0.0120***	0.0353***	0.0119***
		(0.002)	(0.004)	(0.002)
log area	-0.0146***	-0.0144***	-0.0130***	
	(0.005)	(0.005)	(0.005)	
Constant	-0.5290***	-0.5239***	-0.3945***	-0.2131***
	(0.041)	(0.041)	(0.039)	(0.018)
Industry-FE	-	-	Yes	-
LLM-FE	-	-	-	Yes
R^2	0.042	0.048	0.272	0.081
Period 2005-10				
log total workers in LLM	0.0466***	0.0315***	0.0060	
	(0.004)	(0.005)	(0.005)	
log workers in spec. LLM		0.0224***	0.0438***	0.0229***
		(0.003)	(0.004)	(0.003)
log area	-0.0192***	-0.0187***	-0.0169***	
	(0.007)	(0.007)	(0.006)	
Constant	-0.4698***	-0.4636***	-0.3385***	-0.2413***
	(0.051)	(0.050)	(0.045)	(0.021)
Industry-FE	-	-	Yes	-
LLM-FE	-	-	-	Yes
R^2	0.030	0.045	0.369	0.074

Notes: 4057 observations (upper panel); 4031 observations (lower panel). Dependent variable: Correlation of worker-FE and firm-FE on local labor markets. Drop cells with less than 50 workers or 5 firms.

In column 1 we regresses the occupation-specific PAM only on the aggregate size of the local labor market and area size, while we add the size of the specific labor market as measured by the total number of workers in the respective local occupation cell in column 2. As can be seen, in line with the intuition by Moretti, we find that PAM is facilitated much more strongly by the size of the latter. In other words, the assortativeness of the matching of bioengineering workers with bioengineering firms is better in local environments where many such bioengineers are around, and only to a lesser extent in larger cities with more workers in general.

This result is robust to including occupation-fixed effects as in column 4, i.e., when only exploiting the variation within particular occupations across local labor markets. Furthermore, when including region-fixed effects in column 5, we obtain a very similar coefficient as in column 2. This corroborates our finding that the size of the specific labor market facilitates better worker-firm matching also within particular locations. Finally, comparing the upper and the lower panel, we again tend to find higher elasticites for the time perion 2005–2010 than for 1990-1995. This is consistent with our previous results for aggregate labor markets, and shows that stronger PAM in denser specific local labor markets also has become increasingly important over time.

Last, Table 4 turns to the industry dimension, and repeats the analysis at the level of 60 different (2-digit) industries in the 109 local labor markets. The results are similar, although the aggregate labor market size plays a relatively more important role in this case. But when adding industry-fixed effects in column 3, we consistently find that PAM is driven more strongly by the size of the specific, rather than the aggregate local labor market. Furthermore, in column 5 we find that larger industry-labor market exhibit stronger PAM also within particular locations, and more generally we conclude that also this PAM elasticity has become more important over time.

4.3 Summary

Summing up the evidence from this Section, we state the following facts 3–5:

- Fact 3: The degree of positive assortative matching (PAM) between workers and firms is stronger in denser regions.
- Fact 4: The density elasticity of PAM has become more important over time.
- Fact 5: PAM within local labor markets for particular occupations or industries is facilitated more strongly by the size of the specific labor market than by the aggregate local labor market size.

Fact 3 is related to a small literature on PAM in cities, and our findings for West Germany are most closely related to (and consistent with) the cross-sectional evidence by Andersson et al. (2007) for Florida and California. Fact 4 is more original and suggests that PAM is not only an important source of the urban wage premium in a static sense, but that it also plays an important role for the evolution of spatial wage disparities over time. Finally, our novel fact 5 is – at least to our knowledge – the first piece of empirical evidence for PAM in specific, fine-grained local labor markets.

5 Decomposing the trend of between-city inequality

In this final section, we explore the contributions of worker-, firm-, and assortativeness-effects for the trends of between-city wage inequality that has been observed in West Germany over the 1990-2010 period, and thereby shed light on the importance of PAM for the dynamics of spatial wage disparities.

To provide some more context for this analysis, we first replicate the main decomposition exercise by CHK for overall wage inequality at the national level using our data set and time intervals. Specifically, using the AKM wage equation from (2), the variance of wages for workers in any time interval can be decomposed as:

$$\begin{aligned} Var(y_{it}) &= Var(\alpha_i) + Var(\psi_{\mathbf{J}(i,t)}) + Var(x'_{it}\beta) \\ &+ 2Cov(\alpha_i, \psi_{\mathbf{J}(i,t)}) + 2Cov(\psi_{\mathbf{J}(i,t)}, x'_{it}\beta) \\ &+ 2Cov(\alpha_i, x'_{it}\beta) + Var(r_{it}), \end{aligned} \tag{4}$$

by replacing each term with its respective sample analogue. In Table 5 we report our decomposition results for the first and the last time interval, and we report the changes in the different variance components and their contributions to the overall variance change over time.

Table 5: Decomposition of the rise in overall wage inequality

	Interval 1 (1990-1995)		Interval 4 (2005-2010)		Change interval 1 to 4	
	(1)	(2)	(3)	(4)	(5)	(6)
	Variance	Share	Variance	Share	Variance	Share
Total variance log wages	0.146	100	0.259	100	0.113	100
Components of variance:						
Variance of worker-effects	0.087	59.4	0.140	54.0	0.053	47.0
Variance of firm-effects	0.029	19.9	0.053	20.5	0.024	21.4
Variance of Xb	0.008	5.7	0.011	4.2	0.002	2.2
Variance of residual	0.011	7.3	0.013	4.9	0.002	1.8
2 Cov(worker, firm)	0.001	0.7	0.031	11.8	0.030	26.1
2 Cov(worker, Xb)	0.008	5.2	0.006	2.2	-0.002	-1.7
2 Cov(firm, Xb)	0.003	1.8	0.006	2.4	0.004	3.2

Notes: Replication of Table IV in CHK for our data and time intervals.

These results are quantitatively very similar to Table IV in CHK, which is reassuring since we cut the data into different time intervals and work with a smaller connected set of workers and establishments. In particular, consistent with their original study, we find a notable increase in wage inequality in West Germany since re-unification, as measured by the rising variance in raw (log) wages over time. As can be seen in column 6 of Table 5, most of this increase can be attributed to the rising variance in worker-effects (47%), followed by rising assortativeness (26%) and rising workplace-heterogeneity (21%). The observable worker characteristics and their covariances, by contrast, do not add much to the explanation of this inequality trend.[9]

In Table 6 we replicate another exercise by CHK with our data. More specifically, following their Table VI, we distinguish 344 different occupations and trace in panel a) the variance of mean (log) wages between those groups for our four intervals. We observe an increase in the variance from 0.055 to 0.083, and then quantify the contributions of the worker-, firm-, and assortativeness-component to this observed trend in between-group inequality. In panel b) of Table 6 we conduct an analogous exercise for 222 (3-digit) industries. Consistent with the original CHK study, we find that all three components add notably to the explanation of these wage inequality trends. In particular, for both occupations and industries, we find that rising workplace heterogeneity (the firm-effect) explains almost 20% of the increase in the variance, while rising worker heterogeneity and assortativeness are about twice as important.[10]

In panels c) and d) of Table 6 we turn to the regional dimension, which is ignored in CHK, and explore the components of between-city inequality. In panel c) we consider the 326 districts in order to have a total number of groups roughly comparable to the occupational and the industry dimension, while in panel d) we move back to the 109 local labor markets (travel-to-work areas) used so far.

Two important observations can be made from these two panels. First, both for districts and commuting zones, we find that increasing dispersion of the firm-effects (i.e., rising workplace-heterogeneity) seems to play no role for recent spatial wage inequality trends. If anything, firm-effects have even dampened this trend. This conclusion is very different than for occupational or industry-specific inequality trends. As shown in panels a) and b), notable parts of the rise in wage dispersion can be explained by workplace-specific factors in those cases. For example, some "good" firms may have further improved their management practises and their firm-specific wage premia relative to other employers in the middle or the left tail in the firm-effects distribution, and certain occupations and industries are over-represented in those "good firms", so that the rising firm heterogeneity also fuels the between-group inequality along those dimensions. We find no such evidence at the local level, however, which implies that the rising workplace-heterogeneity was not strongly spatially concentrated in cities, but occurred more uniformly across regions.

The second main observation from Table 6 c) and d) is that around half of the rise in the between-group variance in wages can be explained by rising assortativeness

[9] The original study by CHK finds contributions of 39% for person-, 25% for establishment-effects, and 34% for the covariance of the two, with negligible contributions of the other components. The small role of observable worker characteristics for the explanation of inequality trends is also consistent with the study by Baum-Snow and Pavan (2013), who do not take into account workplace- or firm-specific components, however.

[10] We disregard the other components (observable worker characteristics), which is why the shares in column 6 do not add up to one. Taken together, those other components do not add much to the variance decomposition. Moreover, we bear in mind that this decomposition is not exact, as workers may have changed occupations or industries within each interval. See footnote 38 in CHK. Our results are again similar to those in their Table VI. For occupations, CHK estimate contributions of 28-28-42 from worker, firm and matching components, and for industries they obtain 44-19-42.

Table 6: Decomposition of changes in between-group inequality

	Interval 1 1990-1995	Interval 2 1995-2000	Interval 3 2000-2005	Interval 4 2005-2010	Change int. 1 to 4	Share of total
Panel b) 344 occupations (3-digit)						
Variance mean log wages	0.055	0.063	0.076	0.083	0.028	100
Variance mean worker-effects	0.037	0.043	0.050	0.050	0.014	47.8
Variance mean firm-effects	0.011	0.010	0.014	0.016	0.005	19.0
2 Cov(mean worker,firm)	0.002	0.007	0.010	0.015	0.012	43.6
Panel b) 222 industries (3-digit)						
Variance mean log wages	0.039	0.046	0.053	0.067	0.029	100
Variance mean worker-effects	0.016	0.021	0.025	0.027	0.012	41.5
Variance mean firm-effects	0.011	0.011	0.013	0.016	0.005	17.9
2 Cov(mean worker,firm)	0.009	0.011	0.015	0.022	0.013	46.0
Panel c) 326 districts						
Variance mean log wages	0.007	0.007	0.009	0.011	0.005	100
Variance mean worker-effects	0.002	0.003	0.004	0.005	0.003	65.1
Variance mean firm-effects	0.003	0.002	0.002	0.002	0.000	-3.0
2 Cov(mean worker,firm)	0.002	0.002	0.003	0.004	0.002	51.0
Panel d) 109 travel-to-work areas						
Variance mean log wages	0.005	0.005	0.006	0.008	0.003	100
Variance mean worker-effects	0.001	0.002	0.003	0.003	0.002	79.3
Variance mean firm-effects	0.002	0.001	0.001	0.002	0.000	-14.4
2 Cov(mean worker,firm)	0.002	0.002	0.002	0.003	0.001	53.6

Notes: Replication of Table VI in CHK for our data and intervals, and analogous decomposition of between-city wage inequality for 326 *Kreise* (panel c) and 109 *Arbeitsmarktregionen* (panel d).

of workers and firms in the local labour markets. This clearly suggests that PAM is a quantitatively important mechanism for the urban wage premium. In fact, when it comes to explaining the trend in between-city wage inequality, it is only trumped by the rising heterogeneity in worker-effects over time. Recall that this relative rise in urban worker quality (that was also shown in Figure 3 above) can stem from increased sorting of "good workers" into cities, from more rapid gains in individual-specific wage components of urban workers (e.g., through learning externalities), or any combination of the two. Our evidence thus suggests that assortative matching is almost equally important for spatial wage inequality than the combination of all those worker-specific sources, particularly at the district level, where the edge of the worker-effect over the PAM-effect is not very large.

Summing up, the evidence from Table 6 conveys a consistent message with our previously stated facts: worker-based and matching-based explanations for spatial wage disparities have become considerably more important over time, and they drive spatial wage inequality trends, but firm-based explanations do not.

6 Conclusion

To be done

References

Abowd, John, Francis Kramarz, and David Margolis. 1999. High Wage Workers and High Wage Firms. Econometrica 67(2): 251ÂŰ-333.

Abowd, John, Francis Kramarz, Sebastien Perez-Duarte and Ian Schmutte. 2014. Sorting within and between industries: A testable model of assortative matching. NBER Working Paper 20472. Cambridge (Mass.)

Acemoglu, Daron. 1997. Training and innovation in imperfect labor markets. Review of Economic Studies 64(3): 445–464.

Andersson, Fredrik, Simon Burgess, and Julia Lane. 2007. Cities, matching and the productivity gains of agglomeration. Journal of Urban Economics 61: 112–128.

Andrews, Martyn J., Leonard Gill, Thorsten Schank, and Richard Upward. 2008. High Wage Workers and Low Wage Firms: Negative Assortative Matching or Limited Mobility Bias?. Journal of the Royal Statistical Society 171(3): 673–679.

Baum-Snow, Nathaniel and Ronni Pavan. 2012. Understanding the city size wage gap. Review of Economic Studies 79(1): 88–127.

Baum-Snow, Nathaniel and Ronni Pavan. 2013. Inequality and city size. Review of Economics and Statistics 95(5): 1535–1548.

Becker, Gary S. 1973. A theory of marriage: Part I. Journal of Political Economy 81(4): 813–846.

Behrens, Kristian, Gilles Duranton, and Frederic Robert-Nicoud. 2014. Productive Cities: Sorting, Selection, and Agglomeration. Journal of Political Economy 122(3): 507–553.

Card, David, Joerg Heining, and Patrick Kline. 2013. Workplace heterogeneity and the rise of West German wage inequality. Quarterly Journal of Economics 128(3): 967–1015.

Chade, Hector, Jan Eeckhout and Lones Smith. 2016. Sorting Through Search and Matching Models in Economics. Journal of Economic Literature, forthcoming.

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

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, Diego Puga, and Sebastien Roux. 2012. The productivity advantages of large cities: Distinguishing agglomeration from firm selection. Econometrica 80(6): 2543–2594.

Davis, Donald R. and Jonathan I. Dingel. 2015. The comparative advantage of cities. Processed, Columbia University.

De la Roca, Jorge and Diego Puga. 2016. Learning by working in big cities. forthcoming: Review of Economic Studies

Duranton, Gilles and Diego Puga. 2004. Microfoundations of urban agglomeration economies. In: Vernon Henderson and Jacques-FranÃğois Thisse (eds.) Handbook of Regional and Urban Economics, Volume 4. Amsterdam: North-Holland, 2063–2117.

Eeckhout, Jan, and Philipp Kircher. 2011. Identifying Sorting – In Theory. Review of Economic Studies 78(3): 872–906.

Eeckhout, Jan, Roberto Pinheiro, and Kurt Schmidheiny. 2014. Spatial sorting. Journal of Political Economy 122(3): 554–620.

Figueiredo, Octavio, Paulo Guimaraes, and Douglas Woodward. 2014. Firm-worker matching in industrial clusters. Journal of Economic Geography 14: 1–19.

Fujita, Masahisa, Paul Krugman and Anthony Venables. 1999. The Spatial Economy. Cambridge: MIT Press.

Fujita, Masahisa and Jacques-Francois Thisse. 2002. Economics of Agglomeration. Cambridge University Press.

Gaubert, Cecile. 2015. Firm Sorting and Agglomeration. Working Paper, UC Berkeley.

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

Glaeser, Edward L. and David C. Mare. 2001. Cities and skills. Journal of Labor Economics 19(2): 316–342.

Gould, Eric D. 2007. Cities, workers, and wages: A structural analysis of the urban wage premium. Review of Economic Studies 74(2): 477–506.

Helsley, Robert and William Strange. 1990. Matching and agglomeration economies in a system of cities. Regional Science and Urban Economics 20(2): 189–212.

Henderson, J. Vernon. 2003. Marshall's scale economies. Journal of Urban Economics 53(1): 1–28.

Huettner F, Sunder M. 2012. Axiomatic arguments for decomposing goodness of fit according to Shapley and Owen values. Electronic Journal of Statistics 6: 1239–50.

Krugman, Paul. 1991. Geography and Trade. Cambridge: MIT Press.

Marshall, Alfred, 1890. Principles of Economics. London: Macmillan and Co.

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

Moretti, Enrico. 2011. Local labor markets. In Orley Ashenfelter and David Card (eds.) Handbook of Labor Economics, Volume 4. Amsterdam: Elsevier, 1237–1313.

Moretti, Enrico. 2004. Workers' education, spillovers, and productivity: Evidence from plant-level production functions. American Economic Review 94(3): 656–690

Rosenthal, Stuart S. and William Strange. 2004. Evidence on the nature and sources of agglomeration economies. In: Vernon Henderson and Jacques-FranÃğois Thisse (eds.) Handbook of Regional and Urban Economics, Volume 4. Amsterdam: North-Holland, 2119–2171.

Rotemberg, Julio and Garth Saloner. 2000. Competition and human capital accumulation: a theory of interregional specialization and trade. Regional Science and Urban Economics 30(4). 373–404.

Shimer, Robert. 2005. The assignment of workers to jobs in an economy with coordination frictions. Journal of Political Economy 113: 996–1025.

Shimer, Robert and Lones Smith. 2000. Assortative matching and search. Econometrica 68(2): 343–370.

Yankow, Jeffrey. 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.

Appendix

A Appendix Tables

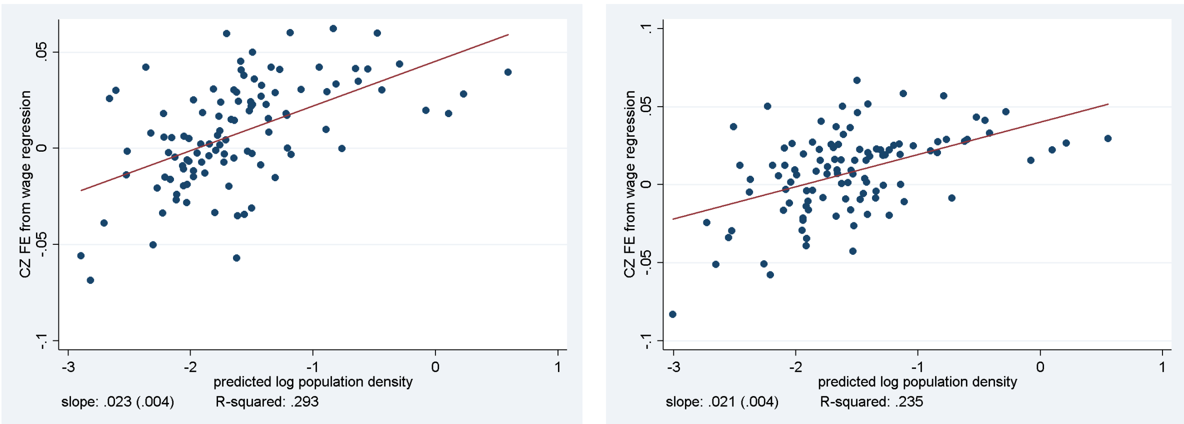
A.1 Traditional Urban Wage Premium

Table A.1: UWP Results

	(1) OLS	(2) 1^{st} stage	(3) red.form	(4) 2SLS
Period 1990-95				
log pop. dens.	0.0243***	0.8393***	0.0194***	0.0232***
	(0.0029)	(0.0403)	(0.0030)	(0.0032)
R2	0.399	0.805	0.293	0.421
Period 2005-10				
log pop. dens.	0.0209***	0.7896***	0.0164***	0.0207***
	(0.0031)	(0.0389)	(0.0029)	(0.0035)
R2	0.301	0.797	0.235	0.310

Notes: 108 observations. Dependent variables: Local labor market fixed-effect from individual level regression (columns 1, 3, 4); log population density (column 2). First stage regressions with fixed-effects for 7,477,719 workers (upper panel) or 6,378,202 workers (lower panel).

.



(a) 1990-1995

(b) 2005-2010

Figure A.1: Conventional Urban Wage Premium, density instrumented by 1952 values

A.2 Different regional aggregation levels

Table A.2: Main Results for 8236 Municipalities

Slope coefficients of outcomes regressed on ln population density			
	(1)	(2)	(3)
Outcome	1990-95	2005-10	Change
worker eff.	0.0323***	0.0527***	0.0209***
	(0.001)	(0.002)	(0.002)
establ. eff.	0.0426***	0.0406***	-0.0012
	(0.001)	(0.002)	(0.002)
corr. of effects.	0.0290***	0.0736***	0.0470***
	(0.004)	(0.004)	(0.005)
neutr. establ. eff.	0.0263***	0.0282***	-0.0004
	(0.001)	(0.002)	(0.002)
corr. of neutr. effects.	0.0330***	0.0639***	0.0334***
	(0.004)	(0.004)	(0.005)

Notes: 8236 observations. Dependent variables as noted in left column, ln population density as single regressor.

Table A.3: Main Results for 325 Districts

Slope coefficients of outcomes regressed on ln population density			
	(1)	(2)	(3)
Outcome	1990-95	2005-10	Change
worker eff.	0.0272***	0.0411***	0.0119***
	(0.002)	(0.003)	(0.002)
establ. eff.	0.0196***	0.0166***	-0.0038***
	(0.002)	(0.002)	(0.001)
corr. of effects.	0.0455***	0.0722***	0.0235***
	(0.003)	(0.005)	(0.004)
neutr. establ. eff.	0.0172***	0.0151***	-0.0031**
	(0.002)	(0.002)	(0.001)
corr. of neutr. effects.	0.0452***	0.0756***	0.0272***
	(0.003)	(0.005)	(0.004)

Notes: 325 observations. Dependent variables as noted in left column, ln population density as single regressor.

Table A.4: Main Results for 108 Local Labor Markets

Slope coefficients of outcomes regressed on ln population density			
	(1)	(2)	(3)
Outcome	1990-95	2005-10	Change
worker eff.	0.0326***	0.0470***	0.0119***
	(0.004)	(0.006)	(0.004)
establ. eff.	0.0343***	0.0270***	-0.0087***
	(0.005)	(0.005)	(0.003)
corr. of effects.	0.0495***	0.0763***	0.0224**
	(0.008)	(0.012)	(0.009)
neutr. establ. eff.	0.0297***	0.0247***	-0.0064**
	(0.005)	(0.004)	(0.003)
corr. of neutr. effects.	0.0535***	0.0777***	0.0206**
	(0.007)	(0.011)	(0.009)

Notes: 108 observations. Dependent variables as noted in left column, ln population density as single regressor.