Good Firms, Worker Flows and Productivity *
Very Preliminary and Incomplete

Michel Serafinelli†

August 21, 2012

Abstract

I evaluate the extent to which worker flows can explain existing evidence on productivity advantages of agglomeration. I employ a matched worker-firm dataset from the Veneto, a region of Italy characterized by the presence of successful industry clusters, to identify the high-wage firms (HWFs). Using balance-sheet data, I show that the HWFs are more productive and have higher intangible capital per worker. For each non-HWF in the region, I then construct a measure for the number of workers with experience gained in HWFs. I find that the effect of the recruitment of a HWF worker on a firm's productivity is an increase between 2.3 and 4.4 percent. This suggests that knowledge is embedded in workers and diffuses when workers move between firms. The number of knowledge carriers observed at a non-HWF is increasing in the number of good firms in the same local labor market and same industry. Hence geographic and economic proximity play a role in the process of knowledge diffusion. My estimates suggest that worker flows explain between 12 and 22 percent of the productivity advantages of agglomeration. Overall, the findings in this paper are consistent with the idea that when similar firms cluster in the same local labor market, their productivity benefits from better access to knowledge carriers.

Keywords: linked employer-employee data, productivity, labor mobility, agglomeration advantages JEL codes: J24; J31; J61; R23.

*I am indebted to David Card, Pat Kline and Enrico Moretti for many helpful discussions and invaluable advice. I also thank Miguel Almunia, Vladimir Asriyan, Pamela Campa, Francesco Devicienti, Yuriy Gorodnichenko, Tadeja Gracner, Bronwyn Hall, Robert Helsley, Jonas Hjort, Agata Maida, Steven Raphael, Ana Rocca, Fabio Schiantarelli, Victoria Vanasco, and seminar partecipants at Berkeley, RAND All-California Labor Conference, MILLS Milan, Stockholm University, EIEF MOOD, San Francisco WEAI and Toronto RCEF for comments. I acknowledge financial support from the Center for Equitable Growth at UC Berkeley. I am deeply indebted to Giuseppe Tattara for making available the Veneto Work History Data. I am also grateful to Bocconi University library staff for help with the AIDA data. Errors are mine and follow a random walk.

†serafine@econ.berkeley.edu

1 Introduction

Localities in many countries are characterized by important differences in productivity. For instance in the United States, total factor productivity (TFP) of manufacturing firms in areas at the top of the TFP distribution is three times larger than TFP in areas at the bottom of the distribution. [1]Another prominent feature of the economic landscape is represented by industry clustering, whereby firms tend to cluster near other "similar" firms (for example: firms that sell similar products). The concentrations of high-tech industries in Silicon Valley, biomedical research in Boston, biotech in San Diego and San Francisco are some famous examples of succesful geographic agglomeration of firms in a single industry. In addition, the large increase of multinational corporation activities in recent decades has led to the emergence of new industrial clusters around the world. Firms that agglomerated in, for example, Silicon Valley and Detroit now have subsidiaries clustering in Bangalore and Slovakia (Alfaro and Chen, 2010). Researchers have long speculated that both the large spatial heterogeneity in productivity and the success of many industrial concentrations may be due the presence of agglomeration economies. In the past twenty years, a significant amount of work has been devoted to studying the importance of these economies, which exist when productivity rises with density. Despite the difficulties involved in estimating the exact magnitude, economists seem to accept that important productivity advantages of agglomeration exist for many industries. However, the field has still not reached a consensus on the relative importance of different explanations of these advantages (Glaeser and Gottlieb, 2009). The potential sources of productivity advantages of agglomeration include technological spillovers, labor market pooling and availability of specialized intermediate inputs. In particular, localized technological spillovers is a widespread explanation for productivity advantages of agglomeration. Nevertheless, as pointed out by Combes and Duranton (2006), if information can flow easily out of the estabilishments, it must be clarified why the effects of spill-overs are localised. Building on the work of Fujita and Ogawa (1982), Helsley (1990) proposes a model where the knowledge produced in a location is a by-product of output, and diffuses through contacts between firms whose cost rises with distance. However the precise nature of the frictions associated with the transmission of knowledge over space remains unclear. For what concerns labor market pooling, the argument is that agglomeration allows a better match between an employer's needs and a worker's skills (Kim (1989), Helsley and Strange (1990)), and the higher quality of the worker-firm match may result in higher productivity. In addition, large cities or industrial concentrations, by hosting a large number of potential partners, can help mitigate hold-up problems that plague bilateral relationships between employers and employees. For instance, in Rothenberg and Saloner (2000) competition between firms to hire skilled workers makes it easier for skilled workers to recoup the cost of acquiring industry-specific human capital. While labor market pooling is a potentially promising explanation for productivity advan-

[1]Moretti (2011).

tages of agglomeration, the existing evidence is still very limited, and rather indirect. More recently, Combes and Duranton (2006), have reconsidered knowledge spillovers and labor market pooling from a theoretical standpoint. One of the main implications of their model is that spillovers and labor market pooling should not be viewed as separate sources for productivity advantages of agglomeration since the labor market at the local level can function as a conduit for the diffusion of information.

In this paper I empirically examine the role of labor mobility as a mechanism for transfer of efficiency-enhancing knowledge and I evaluate the extent to which labor mobility can explain the exisiting evidence on productivity advantages of agglomeration. The underlying idea is that knowledge may be embedded in workers and may diffuse when workers move between firms. The strong localized aspect of knowledge spillovers discussed in the agglomeration literature may then arise from the propensity of workers to change jobs within the same local labor market. Identifying the micro-mechanism behind the productivity advantages discussed in the literature is crucial for obtaining a convincing picture of the agglomeration phenomenon. First, without understanding the precise dimension of the interaction of firms and workers that generate these advantages, it is difficult to be confident about their existence. Furthermore, pinpointing the ultimate causes of agglomeration advantages is helpful for understanding differences in productivity across localities. This is not only of interest for urban and regional economists, but also for growth economists. As pointed out by Moretti (2011)

> "within country differences in productivity [...] are possibly even more remarkable than cross-country differences, since the mobility of labor and capital within a country is unconstrained and differences in institutions and regulations are small relative to cross-country differences. As a consequence, it is difficult to understand why some countries are poor and other countries are rich without first understanding why some cities within a country are poor and others are rich".

Finally, understanding the nature of productivity advantages of agglomeration is also key for understanding the economic rationale for location-based policies (Kline, 2010; Kline and Moretti, 2011).

In order to empirically assess whether knowledge spillovers penetrate through the labor market, I use a matched worker-firm dataset for the Veneto Region of Italy. Employing the method in Abowd, Creecy and Kramarz (2002), I estimate wage equations where both firm and worker effects can be identified and I define good firms as high-wage-firms (HWFs), i.e. those establishments with top values of the estimated firm effects. Then, I construct firm-specific measures for the number of workers in Venetian firms with experience gained at good firms. This is a measure of the explicit contact between good firms and other local firms. By using this measure within a productivity regression framework, I can evaluate whether employees trained at good firms who later join other local firms bring with them some of the knowledge that they have acquired. If labor mobility is to act as a channel for productvity advantages of agglomeration, one would imagine the following to be observed. First, HWFs should have a firm-specific advantage that could be the basis for knowledge transfer. Second, non-HWFs that hire workers with previous experience from HWFs should benefit in terms of increased productivity. Third, geographic proximity should play a role in the process of knowledge diffusion. I use social security earnings records for employees and balance sheet data and location information for employers in order to evaluate the evidence on all three points for Veneto manufacturing during the 1990s. While the issues analyzed in this paper are of general interest, the case of Veneto is important because this region is part of a larger economic area, distinct from the older industrial triangle (between Turin, Milan and Genoa) and the less developed South, where, like in Silicon Valley, networks of specialized firms, frequently organized in districts, have been most effective in promoting and adapting to technological change during the last three decades. This so-called "Third Italy" has received a good deal of attention by researchers, also in the United States (Piore and Sabel, 1984; Piore, 2009).

As a first exercise to assess the potential for knowledge transfer in the region, I look for evidence of an HWF advantage using the detailed firm financial information at my disposal. I show that the HWFs are more productive and have higher capital (in particular intangible capital) per worker. I progress to examine the extent to which non-HWFs benefit from hiring workers from HWFs. I enter annual firm-level measures of the number of workers with recent HWF experience in a Cobb-Douglas production function, and I find that non-HWFs which hire workers with previous experience from HWFs benefit in terms of increased productivity. The productivity effect attributed to workers with experience at good firms is not associated with recently hired workers in general; I do not find a similar productivity effect for recently hired workers without experience at good firms. A problem arises for the estimation from the suggestion that firms decide on their choice of inputs based on a realized shock to productivity which they only observe. Since the shock 'transmit to' input choices, this is known as 'transmission bias' (Eberhard and Helmer, 2010). I employ the standard productivity literature's techniques to control for the endogeneity of inputs (Olley and Pakes ,1996; Levinsohn and Petrin, 2003) to assess this issue's relevance in my paper's setting. I conclude that the effect of the recruitment of a HWF worker on a firm's productivity is an increase between 2.3 and 4.4 percent. I show that these results are not likely to be driven by positive selection of movers from HWF to non-HWFs compared to the other workers in non-HWFs. Having found evidence in favor of an important role of labor turnover as a mechanism of knowledge transfer, I then turn to the question of the importance of geographical proximity in the process of knowledge diffusion. Exploiting information on the location of firms, I show that for a non-HWF, the probability of hiring a worker with experience at good firms is increasing in the number of good firms in the local labor market where the non-HWF is located. This suggest that firm location is of importance, likely because distance acts as a barrier for workers' job mobility: the propensity of workers to change jobs in the same local labour market is greater than their propensity to move between local labor markets (Combes and Duranton, 2006). In general, one might expect labor mobility to also be affected by economic proximity. Therefore I explore whether worker flows from good firms to other firms in a local labor market are larger within an industry. I show a statistically significant relation between the number of good firms in the same local labor market and same industry and the number of knowledge carriers. I do not find a statistically significant relation when I consider good firms in same local labor market but different industry. My findings suggest that when similar firms cluster in the same local labor market, their productivity benefit froms better access to knowledge carriers.

I further relate my findings to the existing evidence on the productivity advantages from agglomeration, in particular the study by Greenstone, Hornbeck and Moretti (2010, henceforth GHM). GHM find that after the opening of a large manufacturing establishment, total factor productivity (TFP) of incumbent plants in US counties that were able to attract one of these large plants is significantly higher than the TFP of incumbent plants in counties that survived a long selection process, but narrowly lost the competition. The increase in TFP that they observe is (a) increasing over time and (b) larger if incumbent plants are in the same industry of the large plant. These two facts are consistent with the presence of intellectual externalities that are embodied in workers who move from firm to firm. However, while the sharp research design allows Greenstone, Hornbeck and Moretti (2010) to obtain credible estimates of the effect of the entry on TFP, with their data it is not possible, as recognized by the authors, to draw definitive conclusions regarding the exact mechanism at work. In oder to evaluate to what extent worker flows explain evidence on productivity advantages of agglomeration, I simulate in my framework the event in GHM. More specifically, I first predict the change in the number of workers from good firms that each of the non-HWF in the region would experience if one were to observe an increase in output similar to the one considered by GHM. I then multiply the predicted change in the number of workers from good firms by the coefficient on workers from good firms in my productivity regression. This yields the predicted change in productivity for a given firm if its locality and industry were to experience an increase in output similar to the one considered by GHM. This predicted change in productivity is found to be equal to a fraction between 12 and 22 percent of the effect found in GHM. My estimates then suggest that worker flows explain a relevant portion of the productivity advantages from agglomeration. The remainder of this paper is structured as follows. In Section 2 I relate the paper to the existing literature. Section 3 presents a simple framework that guides the empirical exercise and aids in interpreting the results. Section 4 presents the empirical model and discusses relevant estimation issues. In Section 5 I describe my data and provide a descriptive overview. The regression results along with various extensions and robustness checks are presented in Section 6 . Section 7 concludes the paper.

2 Relation to Previous Research

My paper adds to a growing literature on productivity advantages of agglomeration, which is critically surveyed in Rosenthal and Strange (2004) and Moretti (2011). The closest part in the literature to my paper is the work on micro-foundations for agglomeration advantages based on knowledge spillovers and on labor market pooling. I already discussed in the Introduction the studies by Fujita and Ogawa (1982) and Helsley (1990) on the former motive for agglomeration advantages, and Kim (1989), Helsley and Strange (1990) and Rothenberg and Saloner (2000) on the latter. The theoretical analysis in Combes and Duranton (2006) is similar in spirit to the empirical exercise in my paper, and certainly deserves further discussion in this context. Their main idea is that firms clustering in the same locality face a trade-off between the advantages of labor pooling (i.e. access to knowledge carriers) and the costs of labor poaching (i.e. loss of some keys employees to competitors and a higher wage bill to keep the others). In the context of a duopoly game, they illustrate how the strategic decisions of firms regarding locations, wages, poaching and prices, depend on market size, on the degree of horizontal differentiation between goods, and on worker heterogeneity in terms of knowledge transfer cost. [2] An additional contribution is Acemoglu (1997), who maintains that in large cities or industrial concentrations, firms invest in new technologies because they know that they can find specialized employees.[3] Further, Duranton and Puga (2004) present a model, inspired by Jovanovic and Rob (1989), Jovanovic and Nyarko (1995), and Glaeser (1999), where proximity to individuals with greater skills or knowledge facilitates the acquisition of skills and the exchange and diffusion of knowledge. Finally, Saxenian (1994) claims that the proximity of high-tech establishments in Silicon Valley is associated with a more efficient flow of new ideas. The hopping of workers from firm to firm, which is also facilited by the California culture, yields productivity benefits for the cluster as a whole. More specifically, she argues [p. 37] that

> "The decentralized and fluid environment accelerated the diffusion of technological capabilities and know-how within the region... When engineers moved between companies, they took with them the knowledge, skills, and experience acquired at their previous jobs"

I contribute to this literature by empirically evaluating the extent to which worker flows can explain existing evidence on productivity advantages of agglomeration. My results may also help explaining the findings in Henderson (2003), Cingano and Schivardi (2004), Moretti (2004b) that productivity advantages from agglomeration are increasing in economic proximity.

[2] See also Fosfuri and Rønde (2004)

[3] In the model by Kim (1989), workers in a larger market also invest more in the depth of their human capital and less in the breadth. Zucker, Darby and Brewer (1998) argue that the presence of specialized human capital is the main determinant of the entry decisions of new biotechnology firms in a city.

Other papers outside the agglomeration literature have emphasized the fact that new workers share ideas on how to organize production or information on new technologies that they learned with their previous employer. For theoretical studies, see Fosfuri, Motta and Rønde (2001), Cooper (2001), Markusen (2001), Glass and Saggi (2002), and Dasgupta (2010). As for the empirical work, Song, Almeida, and Wu (2003) showed that turnover can explain patterns of patent citations, while Rao and Drazin (2002), Kaiser, Kongsted, and Rønde (2008) and Maliranta, Mohnen, and Rouvinen (2009) find that hiring knowledge labor from R&D-intensive establishments is related to a better performance by the hiring establishment. Papers on spillovers from foreign to domestic firms have expanded the scope of the spillovers through the worker flows literature by looking at broader knowledge than that possessed and transferred by R&D labor alone. Gorg and Strobl (2005) show that domestic firms established in Ghana by entrepreneurs with experience from foreign-owned companies in the same industry are more productive and more likely to survive than other firms. Using plant-level data from Colombia, Markusen and Trofimenko (2009) present evidence to support the hypothesis that "experts" hired from abroad have substantial, although not always immediate, positive effects on value added per worker. Balsvik (2011) uses matched employer-employee data and offers a detailed account of productivity gains linked to worker flows from foreign multinational to domestic firms in Norway.[4]In a similar vein, Parrotta and Pozzoli (2012) and Stoyanov and Zubanov (2012) using linked worker-firm data both show evidence for Denmark that is consistent with models of knowledge diffusion through labor mobility. The findings in my paper are related to Balsvik (2011), Parrotta and Pozzoli (2012) and Stoyanov and Zubanov (2012). However, while these authors exclusively focus on knowledge transfer, I investigate the role of geographic proximity, and the extent to which labor mobility can explain evidence on productivity advantages of agglomeration.

3 Conceptual Framework

I am interested in evaluating the role of labor mobility as a mechanism for transfer of efficiency-enhancing knowledge. In this section I present a simple framework that guides the empirical exercise and aids in interpreting the results. The main idea is that knowledge is embedded in workers and diffuses when workers move between firms. The formulation of firm level production function, which captures knowledge transfer, is a variant of the city level production function in Moretti (2004).

I assume different locations, each constituting a separate local labor market. These labor market are completely segmented with workers being immobile between them. There is a finite collection $\mathcal{J} = \{\mathcal{J}_0, \mathcal{J}_1\}$ of firms consisting of *good* firms ($\mathcal{J}_1$), which have some relevant information, and other firms ($\mathcal{J}_0$) which do not have any relevant information. The relevant information can be thought of as a new technology, a new

[4] Poole (2009) finds a positive effect on wages paid in domestic firms in Brazil of the share of new workers previously employed by foreign-owned firms.

managerial technique, a new organizational form, or a new production process. I call this information "technology" (see Fosfuri, Motta and Rønde, 2001 and Glass and Saggi, 2002). The technology is exogenously given. Workers employed by good firms acquire knowledge about some of their technology. Some of this knowledge can be *transferred* to a $j \in \mathcal{J}_0$ firm if they switch employers[5]. Workers are considered *knowledgeable* if they have knowledge of the superior technology and *unknowledgeable* otherwise. All workers employed by good firms are knowledgeable. I assume that output is sold on international markets and capital is supplied at fixed rental rate ρ to all localities and industries.

Write the production function of firm $j \in \mathcal{F}_0$:

$$F(N_{0,}N_1, K) = B(\vartheta_0 N_0)^{\alpha_o}(\vartheta_1 N_1)^{\alpha_1} K^{\beta_k} \tag{1}$$

where ϑ's are productivity shifters, N_0 is the number of unknowledgeable workers employed by the firm, N_1 is the number of knowledgeable workers (workers previously at a good firm that switched employer), K refers to total capital inputs and B is a measure of firm j's general level of efficiency[6]. Assume $\beta_k = 1 - \alpha_1 - \alpha_o$.

I allow for knowledge transfer by letting workers' productivity depend on the number of knowledgeable workers in the firm, as well on their own ability.

$$\log \vartheta_c = \phi_c + tN_1 \qquad c = 0, 1 \tag{2}$$

where ϕ_c is a group-specific effect that captures the direct effect of own innate ability on productivity.

If there is knowledge transfer, $t > 0$. This formulation captures both knowledge that is transferred through workplace interactions among employees, and knowledge about physical capital, process innovations, intermediate inputs, or export markets that can be applied *directly* by the knowledgeable worker. I now turn to examine the effect of an increase in the number of knowledgeable workers on co-workers' productivity. My analysis is similar to Moretti (2004), where the author identifies the effect of an increase in the relative supply of college educated workers in a city on wages for other college educated workers, and for not-educated workers. In his model, workers' productivity depend on the share of educated workers in the city, and the spillover is *external* to individual firms in the city.

Define z_0 as the marginal product of labor for unknowledgeable workers. It can be shown (see Appendix) that

$$\log z_0 = \log \alpha_0 + \alpha_0(\phi_0 + tN_1) + (\alpha_0 - 1)\log(N_0) + \alpha_1(\phi_1 + tN_1) + \alpha_1 \log(N_1) + \beta_K \log K \tag{3}$$

[5] Below I discuss in more details how knowledge transfer takes place once the worker joins firm j.

[6] The nature of B will be discussed in more details below.

and

$$\log z_1 = \log\left((\alpha_0+\alpha_1)t + \frac{\alpha_1}{N_1}\right) + \alpha_0(\phi_0 + tN_1) + (\alpha_0)\log(N_0) + \alpha_1(\phi_1 + tN_1) + \alpha_1\log(N_1) + \beta_K \log K \tag{4}$$

Consider what happens to the marginal product of each type when the number of knowledgeable workers increases.

$$\frac{\delta \log(z_0)}{\delta N_1} = (\alpha_0+\alpha_1)t + \frac{\alpha_1}{N_1} \tag{5}$$

The productivity of unknowledgeable workers benefits for two reasons. First, imperfect substitution: $\frac{\alpha_1}{N_1} > 0$. Second the knowledge transfer further raises their productivity $(\alpha_0+\alpha_1)t > 0$.[7]

As for knowledgeable workers:

$$\frac{\delta \log(z_1)}{\delta N_1} = -\frac{\alpha_1}{N_1}\left(\frac{1}{(\alpha_0+\alpha_1)t + \frac{\alpha_1}{N_1}}\right) + (\alpha_0+\alpha_1)t + \frac{\alpha_1}{N_1} \tag{6}$$

The impact of an increase in the number of knowledgeable workers on their own productivity, z_1 is determined by two competing forces: the first is the knowledge transfer effect that raises productivity. The second is the conventional decreasing return effect[8].

The important feature of Eqs. (5) and (6) is that the productivity of unknowledgeable workers benefits from an increase in the number of knowledgeable workers in the firm even in the absence of any knowledge transfer ($t = 0$), but the effect on the productivity of knowledgeable workers depends on the magnitude of the transfer. If t is large enough, the net effect for knowledgeable workers should be positive although smaller than for knowledgeable workers. If $t = 0$, the net effect should be negative.

The effect of an increase in the number of knowledgeable workers on overall firm level's labor productivity can be written as

$$\frac{\delta \log(z)}{\delta N_1} = n\frac{\delta \log(z_1)}{\delta N_1} + (1-n)\frac{\delta \log(z_0)}{\delta N_1} \tag{7}$$

[7] Mas and Moretti (2009) use data from cashier clerks at a local supermarket chain to investigate the impact of high productivity co-workers. They report that social norms outweigh the potential for free-riding and suggest that having highly productive peers at work increases the marginal productivity of existing workers. The theory outlined in Antras, Garicano and Rossi-Hansberg(2006) describes a scenario in which the wages of "southern" workers unambiguously increase post-globalization as they match with better "northern" managers. The enhanced match quality allows for an increase in the marginal productivity of all workers.

[8] Rewrite $\frac{\delta \log(z_1)}{\delta N_1} = -\frac{\alpha_1}{N_1^2}\left(\frac{\delta \log(z_0)}{\delta N_1}\right)^{-1} + \frac{\delta \log(z_0)}{\delta N_1}$ to see that the sign of $\frac{\delta \log(z_1)}{\delta N_1}$ is ambiguous. If there is no knowledge transfer ($t = 0$) then $\frac{\delta \log(z_1)}{\delta N_1} = \frac{-1}{N_1} + \frac{-\alpha_1}{N_1} < 0$

where n is the share of knowledgeable workers.

Consider a firm j 's optimization problem in each time period under the assumption of perfectly competitive labor markets

$$\max \pi = pB(\vartheta_0 N_0)^{\alpha_o}(\vartheta_1 N_1)^{\alpha_1} K^{\beta_k} - w_0 N_0 - w_1 N_1 - \rho K \tag{8}$$

where w_0 and w_1 are the wages paid by firm j for unknowledgeable and knowledgeable workers respectively (both are industry-wide equilibrium wages), and $w_1 > w_0$.

The corresponding first order conditions are:

$$N_0 = \frac{\alpha_0}{w_0} pB\pi \tag{9}$$

$$N_1 = \left(\frac{w_1}{\alpha_0 \alpha_1 pB\pi} - t\right)^{-1} \tag{10}$$

$$K = \frac{\beta_k}{\rho} pB\pi. \tag{11}$$

In the remainder of the paper, I estimate the effect of a change in the number of knowledgeable workers on *overall firm level productivity* of a estabilishment $j \in \mathcal{F}_0$. The simple framework outlined here predicts that the sign of the net effect on firm level's productivity should depend on the size of the knowledge transfer, imperfect substitution and decreasing return effects, and on the share of knowledgeable workers. A small n gives less weight to the productivity of knowledgeable workers (which benefits from the transfer, but suffers from the decreasing return effect) and more weight to the productivity of unknowledgeable workers (which benefits both from imperfect substitution and the knowledge transfer).

Notice from (10) how N_1 is a function of B. If B contains a productivity term which is knows to the firm but unobserved by the researcher, a main problem arises for the estimation of the effect of a change in the number of knowledgeable workers on overall firm level productivity. The problem arises from the suggestion that firms decide on their choice of inputs based on the realized shock to B which they only observe. Since the shock 'transmit to' input choices, this is known as 'transmission bias' (Eberhard and Helmer, 2010). This estimation issue will be discussed in details in the reaminder of the paper.

As dicussed above, in this paper I also empirically evaluate the importance of geographical and economic proximity in the process of knowledge diffusion, and I relate my results to the evidence on the productivity advantages from agglomeration. Unlike in traditional analysis of agglomeration advantages, there is no real externality here. The wage of knowledgeable workers is equal to their marginal product. The strong localized aspect of knowledge spill-overs discussed in the agglomeration literature arises from the propensity of workers to change jobs within the same local labor market [9].

[9] The theoretical analysis of the spatial equilibrium in this economy is left as future work. The plan

4 Econometric Model

4.1 Identification and Characterization of Good Firms

In order to isolate the good firms, I estimate wage equations where both firm and worker effects can be identified, and then I define the good firms as the HWFs. More specifically, following Abowd, Kramarz and Margolis (1999), I specify a loglinear statistical model for wages as follows:

$$w_{ijt} = X_{it}'\beta + \theta_i + \psi_j + v_t + \varepsilon_{ijt} \tag{12}$$

where the dependent variable is the log of the average daily wage earned by worker i in firm j in year t, and is expressed as a function of individual heterogeneity, firm heterogeneity, and measured worker characteristics.[10] The firm and worker effects (ψ_j and θ_i) represent, respectively, the earnings premium that a firm pays to each worker it employs, and the earnings premium that a worker receives in each firm she works for. The firm premium may reflect rent sharing, compensating differentials, or general heterogeneity across establishments in their compensation policies. The assumptions for the statistical residual ε_{ijt} are (a) $E[\varepsilon_{ijt}|i,t,x] = 0$, (b) $Var[\varepsilon_{ijt}|i,t,x] < \infty$ and (c) orthogonality to all other effects in the model[11]; the presence of labor mobility in matched worker-firm data sets makes it possible to identify worker and firm effects.

The method in Abowd, Creecy and Kramarz (2002) identifies separate groups of workers and firms that are connected via labor mobility in the data. In my fourteen-year sample, the largest group connected via mobility contains around 99% of the observations in the dataset. I run my estimation for the largest group, and define the good firms as HWFs, i.e. those firms in the top 20% of estimated firm fixed effects. See Section 5 for more details on the procedure.

For labor mobility to be a mechanism for agglomeration advantages, we would imagine that a firm-specific advantage is observed that could be the basis for knowledge transfer from HWFs. Therefore, once I have assigned the establishments in the HWFs and non-HWFs groups, I estimate equations like

$$\ln y_{jst} = \beta_0 + \beta_1 HWF_{js} + \mu_s + v_t + e_{jst} \tag{13}$$

where the dummy HWF takes the value of 1 if firm j in industry s is classified as high-wage, and y_{jst} are different firm-level outcomes, such as output per worker, value

is to adopt a framework which is a variant of the classic model of Rosen (1979) and Roback (1982), extended to allow for taste heterogeneity and commuting costs (as in Busso, Gregory and Kline, 2012; see also Moretti (2011)), multiple industries, heterogeneity in firm technology, and knowledge transfer.

[10] Worker characteristics are tenure, tenure squared, age, age squared, a dummy for manager, a dummy for white collar, and interaction terms between gender and individual characteristics.

[11] See Abowd and Kramarz (1999a and 1999b) for a complete discussion of the exogeneity assumption for the residual.

added per worker and tangible and intangible capital per worker.[12] The results are discussed in the descriptive overview of Section 5.

4.2 Workers Flows and Productivity

With linked worker-firm data, I can establish measures of explicit contact between good firms and other local firms by constructing establishment-specific measures for the number of workers in Venetian firms with experience gained at good firms. Then, I use this measure within a productivity regression framework in order to evaluate whether employees trained at good firms who later join other local firms bring with them some of the knowledge that they have acquired.

More specifically, I estimate:

$$y_{jst} = \beta_k k_{jst} + \beta_m m_{jst} + \beta_l l_{jst} + \beta_H H_{jst} + \mu_{st} + v_t + u_{jst} \tag{14}$$

where y_{jt} is the log value of total production (in real euros) at firm j in industry s in t, and $l_{jst}, m_{jst}, k_{jst}$ are the log values of labor, material, and capital inputs. The main variable of interest is H_{jst}, i.e. the number of workers in $t-1$ with recent experience from HWFs. [13] For a worker to be counted as having HWF-experience in year t, the worker must be observed in a HWF for one or more of the years $t-k$ to $t-1$. Given that I use social security data from 1987 and I run production function regressions starting from 1995 (see Section 5), in the baseline specification I set $k = 8$ in order to have as many events of mobility out of HWFs as possible. [14]The inclusion of year dummies controls for the overall business cycles, while the large number of industry-year interaction terms controls for industry-level business cycles, and also time-varying variables such as profit margins, industry concentration, and import competition.[15] Given the way I have created H, this measure captures the recently hired workers with HWF experience. If recent hires are systematically correlated with time-varying unobservables at the firm level, my estimate of β_H will be biased upward.[16] In order to address this issue, I include N_{jst}, i.e. the number of recently hired workers in t without experience from good firms. Therefore, the possible identification of knowledge transfer

[12] I could have defined the good firms as the highly productive ones and detected them using balance sheet data. There are three reasons why I do not pursue this strategy, and instead define the good firms as HWFs. First, the use of social security data allows the introduction of measured individual characteristics and worker effect. Second, social security data are available for a longer period of time than the balance sheets, and therefore they allow me to detect more episodes of labor mobility out of good firms.

[13] Notice that L includes H. Both measures are constructed from head counts in the matched employer-employee data.

[14] I experimented with $k = 5, 6, 7$ and the results largely remained unchanged.

[15] The industry-year interaction terms are based on 21 industry dummies corresponding to the two-digit ISIC level.

[16] Moreover, if workers who change establishments are more productive than stayers in general, the effect of newly hired workers with HWF experience may equally apply to newly hired employees without HWF experience.

relies on the differential effect of hiring an employee with HWF experience over hiring an employee from another non-HWF[17]. Once I have included both H_{jst} and N_{jst}, the correlation between time-varying unobserved productivity shocks and hiring in general would not cause any bias in the difference between the impact of new hires from HWFs and non-HWFs. However, time-varying shocks that are correlated with the propensity to hire workers from good firms give rise to an upward bias in the differential effect of H_{jst}. One can imagine a situation where an establishment experiences a positive productivity shock and responds by hiring workers from good establishments whom it can now better afford (Stoyanov and Zubanov, 2012). Then, in addition to the effect of the recently hired workers from HWF, H_{jst} will carry the receiving firm's own productivity shocks of $t-1$ or earlier. The fact that I lag the number of workers from HWF only partially addresses this concern. To estimate consistently the effect of hiring from a good firm, I must ensure that H_{jst} is uncorrelated with unobserved shocks to the receiving firm's productivity coinciding with, or preceding, the hiring of workers from good firms.

Section 6.1 also employs the productivity literature's techniques to control for the endogeneity of inputs to assess this issue's relevance in this paper's setting. In particular I apply the estimator developed in Levinsohn and Petrin (2003) which uses intermediate inputs as proxies for these unobservable shocks. As a further control, I apply the Olley and Pakes (1996) approach that uses investment as a proxy variable. See the Appendix for a brief summary of the in-depth discussion of 'structural' estimators in Eberhard and Helmer (2010).

4.3 Labor Mobility and Productivity in the Local Labor Market

If labor mobility is to act as a channel for agglomeration advantages, the probability of hiring from a good firm should be higher in localities with a higher share of good firms, i.e. firm location should be of importance. Firm location may matter because the relocation costs for workers or the informational cost of identifying the "right' worker for firms are large across localities. On the former dimension, Combes and Duranton (2006) show that labor flows in France are mostly local: about 75% of the skilled workers remain in the same employment area when they switch firms. The degree of geographical mobility implied by this figure is small, since the average French employment area is comparable to a circle of radius 23 km. In Dal Bó, Finan and Rossi (2011), randomized job offers allow causal estimates of the effect of commuting distance on job acceptance rates. Distance is found to be a very strong (and negative) determinant of job acceptance: applicants are 33 percent less likely to accept the position if the municipality to which they are assigned is more than 80 kilometers (the median distance) away from their home municipality.

[17] Balsvik (2011) uses a similar approach by dividing workers newly hired by Norwegian firms into two groups: those with experience from multinational enterprises, and those without any such experience.

As concerns Veneto, in January 2012, I visited several firms and interviewed proprietors and workers about the history of their enterprises and their current operations. I have also interviewed employers' associations and chamber of commerce officials. The anecdotal evidence that I collected supports the idea that firm location is of importance.[18]

Exploiting information on the location of firms, I can empirically investigate this issue by estimating:

$$(hire\_from\_HWF)_{jlt}=\beta_1(share\_HWFs)_{lt-2}+F'_{jt}\beta_2+v_t+u_{jlt} \tag{15}$$

where the dependent variable is a dummy which takes the value of one if a non-HWF j in locality l is hiring from a good firm at time t, $(share\_HWFs)_{lt-2}$ is the lagged share of good firms in location l and F'_{jt} contains the observable firm characteristics (number of employees, share of female workers, share of white collar workers, share of managers, share of workers aged 30 or less and share of workers aged 40) The results are discussed in Section 6.2

5 Data Sources and Descriptive Overview

The data set is for Veneto, a administrative region in the North-East of Italy, which is the third Italian region by GDP and has a population of around 5 million people, around 8 percent of the country's total. The region underwent deep economic changes in the last few decades. Until after World War II, the economy was largely based upon farming and the region experienced large out-migration to Germany, Switzerland, United States, Canada and Australia. The 1960s and 1970s were characterized by intense economic development. Starting from the mid-nineties until the most recent recession, Veneto has been a full employment region with a positive rate of job creation in manufacturing, compared to a negative national rate and positive migration flows (Tattara and Valentini, 2007). Veneto is a dynamic, "manucentric" region characterized by a large presence of flexible establishments, frequently organized in districts with an industrial value added higher than the national average, and a very strong vocation for exporting their products..[19]Metal-engineering is the largest industry (electromechanical, precision machining, etc.); important industries are also goldsmiths, mechanical goods, furniture and plastics, garments, textiles, leather and shoes.[20]

The data set pools three kinds of information: individual earnings records, firm balance sheets, and information on local labor systems. The first two kinds of infor-

[18]As expressed by Federico Callegari, of the Treviso Chambre of Commerce: "When losing their job, workers tend to look for another job with a commuting time of 20-30 minutes. Why? Because they want to go home during the lunch break!"

[19]The most famous example of industrial district is probably the eyeglasses district in the Province of Belluno. This district hosts the largest world manufacturer Luxottica, whose brands include Ray-Ban eyeglasses.

[20]Benetton, Sisley, Geox, Diesel, Replay are Venetian brands.

mation (combined for the period 1995-2001) have been used by Card, Devicienti and Maida (2011).

The earnings records result from the Veneto Workers History (VWH) dataset, which was assembled by Giuseppe Tattara and collaborators at University of Venice with administrative archives of the Istituto Nazionale per la Previdenza Sociale (INPS), which is the main public institute of social security in Italy. [21] The VWH has data on private sector personnel in the Veneto region over the period 1975 to 2001. Specifically, it contains register-based information for virtually any job that lasts at least one day. The whole employment history has been reconstructed for each worker. On the employee side, the VWH contains overall earnings during the calendar year for every job, the amount of days worked during the year, the proper national contract and the level within that contract (i.e., a "job ladder" code), and the employee's age, gender, region (or country) of birth, and tenure with the employer. On the firm side, the VWH contains industry (categorized by five-digit code), the dates of "birth" and closure of the establishment (if applicable), and the location of the establishment. Balance sheet records starting from 1995 were obtained from AIDA (analisi informatizzata delle aziende), a database circulated by Bureau Van Dijk which contains information for incorporated non-financial Italian establishments with annual revenues of at least 500,000 Euros.[22] AIDA has the official balance sheet records for these firms, which contain revenues, total wage bill, the book value of capital (broken into a number of subgroups), value added, the overall number of workers, materials and industry (classified by five-digit code).

Information on local labor systems (LLMs) was obtained from the National Institute of Statistics (ISTAT). The LLMs are territorial groupings of municipalities characterized by a certain degree of working-day commuting by the resident population. ISTAT conducted three studies on LLMs: in 1981, 1991 and 2001. In 2001 686 LLMs were listed in Italy, and the median number of employees was 10763. [23]Since I use data for the period 1987-2001 (see below) I exploit the LLM classification in 1991, were the 518 municipalities (comuni) in Veneto are divided into 51 LLMs, 48 of which have their centroid in the region. I use firm identifiers to match job-year observations for workers aged 16 to 64 in the VWH to firm financial data in AIDA for the period 1995 to 2001. The match rate is fairly large: at least one observation in the VHW was found for over 95% of the employers in the AIDA sample. Around 50% of all employees observed in the VWH between 1995 and 2001 (the period of overlap with the balance sheet data) can be matched to an AIDA firm. Most of the non-matches seem to be workers of

[21] I am extremely grateful to Giuseppe Tattara for making the dataset available and to Marco Valentini and Carlo Gianelle for assistance in using it. Additional information on VHM is available in the Appendix, which draws on the descriptions in Battisti (2012), Tattara and Valentini (2011) and Card, Devicienti and Maida (2011)

[22] See http://www.bvdep.com/en/aida.html. Only a small portion of establishments in AIDA is publicly traded. I eliminate these establishments and those with consolidated balance sheets (i.e., holding companies).

[23] The average was 30576, the minimum 1251 and the maximum 1321564.

small firms which are omitted from AIDA. I could match at least one employee for around 18,000 firms, or around 10% of the entire universe of employers contained in the VWH.[24]

From this set of potential matches, I made a sequence of exclusions to obtain my estimation sample for equation (12). One, I removed all workers outside manufacturing. Two, I excluded job-year observations with remarkably high or low values for wages (I trim observations outside the 1% - 99% range). Four, I dropped observations where firms only had one employee in VWH. I use VHW data from 1987 to 2000 for firms that could be matched in AIDA. As explained above, the method in Abowd, Creecy and Kramarz (2002) identifies in the data separate groups of workers and firms that are connected via mobility. I run the grouping algorithm separately, and then use the created group variable to choose the largest group as a sample for the fixed-effects estimation. I identify the HWFs as those in the top 20% group of ranked values for the fixed effects.[25] Table 1 shows that for Veneto manufacturing, there are clear differences between HWFs and non-HWFs in labor productivity (output per worker), value added per worker and firm size – see equation (13). The table further shows differences in capital per worker (Column 4) and in particular in intangible fixed assets (intellectual property, accumulated research and development investments, goodwill). This evidence is important for assessing the potential for knowledge transfer in the region. Overall, these descriptive results point to an HWF advantage. Since labor productivity is 17% higher in HWFs, and intangible capital per worker 31% larger, we can also think of HWFs as high-productivity firms, or high-intangible-capital firms.

For labor mobility to be a mechanism for agglomeration advantages, we must observe some workers moving from HWFs to other firms. The wage premium in HWFs may encourage employees to stay in HWFs instead of switching to non-HWFs, and worker flows may therefore be small. In terms of the potential for knowledge transfer, the interesting question is how workers with recent HWF experience spread across the group of non-HWFs. In a typical year between 1995 and 2001, 3.6% of non-HWFs employ workers with recent HWF experience. If we consider all non-HWFs that hire workers from HWFs (regardless of how recently they arrived), this percentage is equal to 4.3%.[26]If I define HWFs as those in the top 50% group of ranked values for the fixed effects (instead of 20%) 8% of non-HWFs employ workers with HWF experience.

The presence of a firm-specific advantage, together with evidence of actual worker flows, can only suggest that a potential for knowledge transfer through labor turnover does exist, while a productivity benefit at the firm level due to turnover is consistent

[24] Card, Devicienti and Maida (2011) show that the average firm size for the matched jobs sample (36.0 workers) is considerably larger than the average for total employers in the VWH (7.0 workers). Mean daily wages for the matched observations are also greater, while the fractions of younger (age 30 or less) and female employees are lower.

[25] I use the a2reg Stata routine developed by Ouazad (2007) to implement the approach in Abowd, Creecy and Kramarz (2002).

[26] Overall, 839 workers switch from HWFs to non-HWFs; 797 of these workers have recent HWF experience.

with labor turnover actually working as a channel for knowledge transfer. In the next section, I discuss the extent to which non-HWFs benefit from hiring workers from HWFs. In order to obtain the estimation sample for the firm-level analysis – equation (14) – I first removed the HWFs. Then, among the non-HWF group I removed (a) firms that closed during the calendar year, and (b) firm-year observations with remarkably high or low values for several key firm-level variables, such as total value of production, number of employees, capital stock and materials (I trim observations outside the 1% - 99% range). Since I am using AIDA data (from 1995 to 2001) for the firm-level analysis, in order to reduce the influence of false matches (particularly for larger firms) I followed Card, Devicienti and Maida (2011) and eliminated the "gross outliers", a minor number of matches (less than 1% of all employers) for which the absolute gap between the number of workers reported in the balance sheet and the number found in the VWH is larger than 100.[27]The resulting firm level sample is summarized in Table 2.

6 Empirical Evidence

6.1 Evidence on Labor Mobility and Productivity

Table **??** shows the results of estimating equation (14). In column 1, the coefficient on H_{jst} is positive (0.051) and highly significant. The impact of the recruitment of a HWF worker on a firm's productivity is an increase of 5.1 percent. Columns 2 and 3 report the results where I add fixed effects for province-by-year, and province-by-industry-by-year.[28] The results largely remain unchanged. In Column 4, I include N_{jst} in order to address the issue of a potential correlation between recent hires and time-varying unobservables – see Section 4.2. The coefficient on newly hired workers without HWF experience is positive and highly significant but small (0.004). The difference in the productivity effects associated with the two types of newly hired workers is highly significant.

Table **??** presents results where I try to address the issue of transmission bias, i.e. the presence of unobservable shocks that 'transmit to' input choices, using the productivity literature's techniques. In the first 3 columns, I compare parameter estimates from OLS (Column 1), the Levinsohn and Petrin (2003, henceforth LP) estimator (Column 2) and the within estimator (Column 3). In Column 2, H_{jst} is treated as a freely

[27]I also assessed the quality of the matches by comparing the overall number of individuals in the VWH who are reported as working for a certain employer (in October of a certain year) with the overall number of employees recorded in AIDA (for the same year). When I eliminated the "gross outliers" ,the correlation between the number of workers in VHM and the figure found in the balance sheet was 0.996. Card, Devicienti and Maida compared total wages and salaries for the calendar year as recorded in AIDA with overall wage payments recorded for workers in the VWH. The two measures are extremely correlated (correlation > 0.98) and the median ratio between them is close to 1.0.

[28]A province is an Italian administrative division. There are 7 provinces in Veneto.

variable input.[29] The coefficient for H_{jst} in Column 2 is positive (0.037) and highly significant; it is lower than the OLS estimate, confirming both the theoretical and empirical results on freely variable inputs discussed in LP. [30] The estimates for H_{jst}differs quite substantially from the OLS estimate, and is closer to LP estimates. It is positive (0.026) and highly significant [31]As a further check, in Column 4 I report results using the Olley and Pakes (1996, henceforth OP) estimator. [32] The coefficient is positive (0.048) and highly significant. However, I do not observe investment, and hence I have to derive the proxy variable in t as the difference in the reported book value of capital in $t + 1$ and the value in t. This exacerbates the measurement error problems associated with the proxy variable approach. In addition it reduces my sample size quite substantially because (a) one year of data is lost when I take the difference in reported book values and (b) only observations with positive values for the proxy variable can be used in the OP approach, and a large portion of my sample does not satisfy this condition. [33]

Taken together, the findings shown here suggest that non-HWFs which hire workers with previous experience from HWFs benefit in terms of increased productivity. The productivity effect attributed to workers with HWF experience is not associated with recently hired workers in general; I do not find a similar productivity effect for recently hired workers without HWF experience. Moreover, these results do not seem likely to be driven by unobservable productivity shocks.

I subjected the sensitivity of my findings to three additional robustness checks. In the first robustness check, I entered H_{jst-1} instead of H_{jst}. The results, shown in the Appendix, are very similar. In the third robustness check, I eliminated the unobserved firm effect in equation (14) by first differencing the model. In the third robustness check, I allowed the effect of materials to differ by 2-digit industry For all robustness checks, the results were largely unchanged (These estimates not shown here but available upon request).

The estimates shown so far cannot dismiss an alternative reason for the positive $\widehat{\beta_H}$, namely that employees with HWF experience are positively selected compared to the other workers in non-HWFs. In constructing the ranking of firms in Section 4.1 I

[29] I use the levpet Stata routine developed by Levinsohn, Petrin, and Poi (2003).

[30] As for the other freely variable input, L_{jst} the OLS estimate also exceeds the LP estimates. The results for capital are also consistent with LP, which show that if capital is not correlated with this period's transmitted shock (but variable inputs are), or capital is much less weakly correlated with the productivity shock than the variable inputs are, the OLS estimate on capital is likely to be biased downward

[31] Note that the within estimator removes all permanent productivity differences among firms that might be correlated with the propensity to hire workers with HWF-experience. For instance, non HWFs may target sending HWFs with particular characteristics (e.g. some domestic HWFs may prefer to hire from multinationals) and the (long-term) stable preferences in hiring may reflect certain management practices.

[32] I use the opreg Stata routine developed by Yasar, Raciborski, and Poi (2008).

[33] Firms will have their non-positive proxy variable observations truncated from the estimation routine because the necessary monotonicity condition does not hold for these observations. See Eberhard and Helmer (2010) for more details

take into account individual heterogeneity and measured worker characteristics (see eq. (12)). However movers might be selected. As for the individual characteristics, in all year movers from HWFs are significantly more likely than stayers at non-HWFs to be young and male. In most years they are also significantly more likely to be white collars (this category also includes middle managers). In some years they are significantly more likely to be managers[34]. Given these differences in observable chracteristics, I add to equation (14) variable inputs such shares of managers, females, white collars, and differently aged workers. The results largely remained unchanged (the estimates are not shown here but are available upon request). As for unobserved ability, I compare the distributions of the worker fixed effects in non-HWFs from estimating equation (12) for stayers and workers coming to non-HWFs from good firms. A mean-comparison test of $\widehat{\theta_i}$ fails to suggest that workers coming to non-HWFs from good firms are positively selected. I also compared the overall distribution in Figure 1, which shows the quantile-quantile plot, i.e. a plot of the quantiles of the distribution of $\widehat{\theta_i}$ for the stayers at non-HWFs against the quantiles of the distribution of $\widehat{\theta_i}$ for the workers coming from good firms, for the most recent year, 2001. Figure 2 shows the quantile-quantile plot for the year 2000. Points on the right-hand side of the 45-degree line mean that the values of the distribution on the x-axis are higher than those of the distribution on the y-axis. If most points are on the right-hand side of the main diagonal, we will conclude that workers coming to non-HWFs are positively selected on unobservable ability.[35] However, Figure 1 and 2 show no evidence of this. Despite some differences in the two graphs, points on the x-axis are not usually higher than points on the y-axis[36]. Therefore this exercise again fails to suggest that worker selection is the source of the estimated productivity benefits. As a further check on this issue, I add to equation (14) $\overline{\theta}_j$, the firm level average of worker estimated person effects. The coefficient of $\overline{\theta}_j$ is positive and highly significant, but $\widehat{\beta_H}$ is very similar (the estimates are not shown here but are available upon request).

6.2 Evidence on worker flows and Productivity in the Local Labor Market

Having found evidence in favor of an important role of labor turnover as a mechanism of knowledge transfer, I then turn to the question of the extent to which worker flows can explain evidence on agglomeration advantages. As discussed in Section 4.3: if labor mobility is to act as a channel for agglomeration advantages, the probability of hiring from a good firm should be higher in localities with a higher share of good firms, i.e. firm location should be of importance. The distribution of good firms across locations

[34] The Appendix shows descriptive statistics for movers from HWFs and stayers at non-HWFs for the most recent year.

[35] Both axes are in units of the estimated θ_i from equation 12 (vertical axis for stayers and horizontal axis for the hires from good firms). For a given point on the q-q plot, we know that the quantile level is the same for both points.

[36] The same conclusion holds for the other years in the 1995-2001 period.

is not uniform. Some LLMs have a higher share of good firms than others (see Table 5). Since the region-wide share of good firms is fixed (equal to 25%, see Section 4.1) each firm should, in principle, face the same probability of hiring a worker from a good Venetian firm, regardless of location, unless geography plays a role.

Column 1 of Table **??** shows the results of estimating equation 15 using the probit method (The reported estimated coefficients are semi-elasticities; standard errors are clustered by firm). The coefficient on the lagged share of good firms in the LLM is positive (2.154) and highly significant.[37]

Columns 2 reports the results where I add fixed effects for province-by-year.[38] The coefficient on the lagged share of good firms in the LLM is again positive, even though somewhat smaller (1.706) and highly significant. A standard deviation (0.097) increase in the share of good firms in the LLM is linked to a 17% increase in the probability of hiring from a good firm. Columns 3 shows the estimates when I add fixed effects for province-by-industry-by-year. The results are unchanged. This evidence suggests that firm location is of importance because distance acts as a barrier for workers' job mobility: the propensity of workers to change jobs in the same local labor market is greater than their propensity to move between local labour markets. I subjected the sensitivity of these findings to a number of robustness checks. First, I estimated equation (15) using the linear probability model with LLM fixed effect. Second, I entered firm fixed effects. Third, I used different lags of the explanatory variables. Third, I clustered the standard error by local labor market instead of by firm. For all robustness checks, the results largely remained unchanged (the estimates are not shown here but are available upon request). Overall, my results are then consistent with the model in Combes and Duranton (2006) where firms that cluster in the same local labor market benefit from better access to workers whose knowledge enhances efficiency.

In general, one might expect labor mobility to also be affected by economic proximity. Therefore in Table (**??**) I explore whether worker flows from good firms to other firms in a local labor market are larger within an industry. I show a statistically significant relation between the number of good firms in the same local labor market and same industry and the number of knowledge carriers. I do not find a statistically significant relation when I consider good firms in same local labor market but different industry. These results may help explaining the findings in Moretti (2004b) and Greenstone, Hornbeck and Moretti (2010) that productivity advantages are increasing both in geographic and economic proximity.

[37] Recall the description of LLMs in Section 5. Here, I only consider LLMs whose centroid is in Veneto.

[38] A LLM can span more than one province

7 Conclusions

The evidence provided in this paper is consistent with labor turnover working as a channel for agglomeration advantages. First, I showed that HWFs feature higher labor productivity, higher value added per worker and higher intangible fixed assets per worker. This suggests that HWFs have a firm-specific advantage and hence, that there is a potential for knowledge transfer. Second, non-HWFs which hire workers with previous experience from HWFs benefit substantially in terms of increased productivity. These results are not likely to be driven by unobservable productivity shocks. Third, the probability of hiring from a good firm is higher in localities with a higher share of good firms (i.e. firm location matters). Thus, labor turnover seems to be a mechanism for agglomeration advantages in Veneto manufacturing.

A Transmission bias and Structural Estimators of production functions.

A standard Cobb-Douglas production function is given by

$$Y_j = A_j K_j^{\beta_k} M_j^{\beta_m} \underline{N}_j^{\beta_{\underline{n}}} \overline{N}_j^{\beta_{\overline{n}}} \tag{16}$$

where $\underline{N}_j$ and $\overline{N}_j$ represent workers without HWF experience and those with HWF experience, respectively. The empirical equivalent (denote log values with lower case letters) is

$$y_{jt} = \beta_k k_{jt} + \beta_m m_{jt} + \beta_{\underline{n}} \underline{n}_{jt} + \beta_{\overline{n}} \overline{n}_{jt} + \beta_o + \varsigma_{jt} \tag{17}$$

in equation (17) $\ln(A_j)$ is decomposed into β_o and ς_{jt}, where the constant β_o represents mean efficiency across all firms and ς_{jt} represents deviations from this mean.

Define ς_{jt} as

$$\varsigma_{jt} = \omega^*_{jt} + \nu_{jt} = \eta_j + \omega_{jt} + \gamma_t + \nu_{jt} \tag{18}$$

which indicates that ς_{jt} contains measurement error and a productivity term ω^*_{jt} (TFP) which is known to the firm but unobserved by the researcher.

Rewrite equation (17) to yield

$$y_{jt} = \beta_k k_{jt} + \beta_m m_{jt} + \beta_{\underline{n}} \underline{n}_{jt} + \beta_{\overline{n}} \overline{n}_{jt} + \beta_o + \eta_j + \omega_{jt} + \gamma_t + \nu_{jt} \tag{19}$$

The main problem for estimation of specification such as equation (19) arises from the suggestion that firms decide on their choice of inputs based on the realized shock ω_{jt} which only they observe. Since ω_{jt} is suggested to 'transmit to' input choices, this is known as 'transmission bias' (Eberhard and Helmer, 2010).

In order to assess this issue's relevance in this paper's setting consider a firm's optimization problem in each time period under the assumption of perfectly competitive input and output markets

$$\max \pi_j = pA_j K_j^{\beta_k} M_j^{\beta_m} \underline{N}_j^{\beta_{\underline{n}}} \overline{N}_j^{\beta_{\overline{n}}} - w_{\underline{N}} \underline{N}_j - w_{\overline{N}} \overline{N}_j - rK_j - qM_j$$

where prices of output and inputs are industry-wide equilibrium prices. The corresponding FOC wrt to labor with HFW experience is

$$\overline{N}_j = \left(\frac{\beta_{\overline{n}} p A_j}{w_{\overline{N}}}\right)^{\frac{1}{1-\beta_{\overline{n}}}} K_j^{\frac{\beta_k}{1-\beta_{\overline{n}}}} M^{\frac{\beta_m}{1-\beta_{\overline{n}}}} \underline{N}_j^{\frac{\beta_{\underline{n}}}{1-\beta_{\overline{n}}}} \tag{20}$$

rewritten in logs it can be seen that $\overline{n}_j$ is a function of ω_{jt}

$$\overline{n}_j = \frac{1}{1-\beta_{\overline{n}}}(\ln \beta_{\overline{n}} + \ln p + \beta_o + \varsigma_{jt} - \ln w_{\overline{N}} + \beta_k k_{jt} + \beta_m m_{jt} + \beta_{\underline{n}} \underline{n}_{jt})$$

Several solutions for the endogeneity of input choices with regard to unobserved productivity have been proposed in the literature What follows is a brief summary of the in-depth discussion of 'structural' estimators in Eberhard and Helmer (2010). OP address the issue of endogeneity of inputs by using information about observed investment to proxy for unobserved productivity and by applying a control function estimator. They assume that k_{jt} and ω_{jt} are firm-specific state variables in the firm's dynamic programming problem. The Bellman equation is

$$V_{jt}(k_{jt,}\omega_{jt}) = \max\{\pi_j(k_{jt,}\omega_{jt}) - c_j(i_{jt}) + \theta E[V_{t+1}(k_{jt+1,}\omega_{jt+1})|k_{jt,}\omega_{jt}, i_{jt}]\}$$

where $k_{jt+1} = (1-\delta)k_{jt} + i_{jt}$ is the law of motion for capital accumulation. Investment is chosen at time t and adds to the capital stock at time $t+1$. The solution gives an investment policy function that depends on capital and productivity $i_{jt}(k_{jt,}\omega_{jt})$. Labor is not included in the investment equation because it is assumed to be a 'non-dynamic' input: it can be adjusted after realization of ω_{jt} within the same period. A key assumption is that investment is strictly increasing in both capital stock and productivity. In addition, ω_{jt} is assumed to be the only unobservable driving the investment choice. Finally, when deciding upon investment in period $t+1$ any realizations of ω_{jt} prior to time t are not incorporated in the investment function because productivity evolves by assumption following an 'exogenous first-order Markov process': a firm builds expectations about its productivity at time $t+1$ exclusively based on its productivity levels realised at time t. Therefore one can assume most generally that productivity evolves according to $\omega_{jt} = g(\omega_{jt-1}) + \xi_{jt}$, where ξ_{jt} is the random 'productivity shock'. Provided the investment function is continuous in $k_{jt,}$ and ω_{jt}, and provided investment is positive, the investment equation can be inverted to yield $\omega_{jt} = f_t(i_{jt}, k_{jt})$. The OP estimator is implemented in two stage: first, by estimating

$$y_{jt} = \beta_l l_{jt} + \phi_{jt}(i_{jt}; k_{jt}) + \epsilon_{jt}$$

where

$$\phi_{jt}(i_{jt}; k_{jt}) = \beta_o + \beta_k k_{jt} + f_t(i_{jt}, k_{jt}) \tag{21}$$

OP propose estimation based on a third-order polynomial series expansion. In the second step, OP employ these estimates to run a regression of $y_{jt} - \widehat{\beta}_l l_{jt}$ on $\widehat{\phi}_{jt}(\cdot)$ and k_{jt}, which yields an unbiased $\widehat{\beta_k}$. From the assumption of a Markov process for productivity and equation (21) one can realise that

$$E[\omega_{jt}|\omega_{jt-1}] = g(\phi_{jt-1}(i_{jt-1}; k_{jt-1}) - \beta_o - \beta_k k_{jt-1}) + \xi_{jt}$$

This allows one to write

$$y_{jt} - \widehat{\beta}_l l_{jt} = \beta_k k_{jt} + g(\widehat{\phi}_{jt-1}(i_{jt-1}; k_{jt-1}) - \beta_o - \beta_k k_{jt-1}) + \xi_{jt} + \epsilon_{jt} \tag{22}$$

Given that β_k enters the equation twice and in combination with other parameters, euqation (22) is estimated using non-linear least squares (NLLS).

The OP model can be extended to include firm exit, in which case an extra step is added between the two described above, where a probit regression is fitted on a nonlinear function of i_{jt}, k_{jt} using the same argument of proxied productivity as in the first step. The predictions from this intermediate step are then added in the $g()$ function in the above second step.

Building on OP, LP suggested the use of intermediate input demand instead of investment demand as a proxy for productivity ω_{jt}. This means that the decision on intermediate input is made at time t once ω_{jt} is observed by the firm. The same applies to labour input choices, which in turn means that labor and intermediate inputs are chosen at the same, and labour preserves its assumed non-dynamic/flexible nature. In the LP approach, intermediate inputs (electricity, material inputs) are a function of ω_{jt} and k_{jt} similar to the use of investment in the OP procedure.

The LP strategy keeps on relying on the scalar unobservable and monotonicity requirements. The production function to be estimated by LP is

$$o_{jt} = \beta_o + \beta_l l_{jt} + \beta_k k_{jt} + \beta_m m_{jt} + \omega_{jt} + \epsilon_{jt}$$

where m_{jt} is intermediate inputs. Note that LP use (log) gross output as dependent variable, instead of value-added. LP specify demand for intermediate inputs as $m_{jt} = m_{jt}(k_{jt,}\omega_{jt})$ where demand is assumed to be monotonically increasing in ω_{jt}. This assumption allows to invert the function to obtain a proxy for unobserved productivity $\omega_{jt} = f_t(m_{jt}, k_{jt})$. The first step of the production function is then rewritten as

$$o_{jt} = \beta_l l_{jt} + \zeta_{jt}(m_{jt}; k_{jt}) + \epsilon_{jt}$$

where $\zeta_{jt}(m_{jt}; k_{jt}) = \beta_o + \beta_k k_{jt} + \beta_m m_{jt} + f_t(m_{jt}, k_{jt})$. The second step is

$$o_{jt} - \widehat{\beta}_l l_{jt} = \beta_k k_{jt} + \beta_m m_{jt} + g(\widehat{\zeta}_{jt-1} - \beta_o - \beta_k k_{jt-1} - \beta_m m_{jt-1}) + \xi_{jt} + \epsilon_{jt}$$

Since m_{jt} is not orthogonal with respect to ξ_{jt}, LP instrument current intermediate input levels through one-period lagged levels. See Eberhard and Helmer (2010) for further details.

B Derivations

B.1 Derivation of $\frac{\delta \log(z_0)}{\delta N_1}$ and $\frac{\delta \log(z_1)}{\delta N_1}$

$$z_0 = \alpha_0(\vartheta_0)^{\alpha_0} N_0 (\vartheta_1 N_1)^{\alpha_1} K^{\beta_K}.$$

It follows that

$$\log z_0 = \log \alpha_0 + \alpha_0 \log(\vartheta_0) + (\alpha_0 - 1)\log(N_0) + \alpha_1 \log(\vartheta_1) + \alpha_1 \log(N_1) + \beta_K \log K,$$

so

$$\log z_0 = \log \alpha_0 + \alpha_0(\phi_0 + tN_1) + (\alpha_0 - 1)\log(N_0) + \alpha_1(\phi_1 + tN_1) + \alpha_1 \log(N_1) + \beta_K \log K.$$

$$\frac{\delta \log(z_0)}{\delta N_1} = (\alpha_0 + \alpha_1)t + \frac{\alpha_1}{N_1}$$

$$z_1 = \left((\alpha_0 + \alpha_1)t + \frac{\alpha_1}{N_1}\right) \alpha_0(\vartheta_0 N_0)^{\alpha_0} (\vartheta_1 N_1)^{\alpha_1} K^{\beta_K}.$$

Then

$$\log z_1 = \log\left((\alpha_0 + \alpha_1)t + \frac{\alpha_1}{N_1}\right) + \alpha_0 \log(\vartheta_0) + (\alpha_0)\log(N_0) + \alpha_1 \log(\vartheta_1) + \alpha_1 \log(N_1) + \beta_K \log K,$$

so

$$\log z_1 = \log\left((\alpha_0 + \alpha_1)t + \frac{\alpha_1}{N_1}\right) + \alpha_0(\phi_0 + tN_1) + (\alpha_0)\log(N_0) + \alpha_1(\phi_1 + tN_1) + \alpha_1 \log(N_1) + \beta_K \log K.$$

$$\frac{\delta \log(z_1)}{\delta N_1} = -\frac{\alpha_1}{N_1}\left(\frac{1}{(\alpha_0 + \alpha_1)t + \frac{\alpha_1}{N_1}}\right) + (\alpha_0 + \alpha_1)t + \frac{\alpha_1}{N_1}$$

B.2 Derivation of First-Order Conditions

$$\max \pi = pB(\vartheta_0 N_0)^{\alpha_o}(\vartheta_1 N_1)^{\alpha_1} K^{\beta_k} - w_0 N_0 - w_1 N_1 - \rho K \tag{23}$$

It follows that

$$\frac{\partial \pi}{\partial N_0} = \frac{\alpha_0}{N_0} pB\pi - w_0$$

$$\frac{\partial \pi}{\partial N_1} = \alpha_0 \alpha_1 \left(t + \frac{1}{N_1} \right) pB\pi - w_1$$

$$\frac{\partial \pi}{\partial N_0} = \frac{\beta_k}{K} pB\pi - \rho.$$

Thus, we get the first order conditions:

$$N_0 = \frac{\alpha_0}{w_0} pB\pi$$

$$N_1 = \left(\frac{w_1}{\alpha_0 \alpha_1 pB\pi} - t \right)^{-1}$$

$$K = \frac{\beta_k}{\rho} pB\pi.$$

References

[1] Abowd, John M. and Francis Kramarz, "Econometric Analysis of Linked Employer-Employee Data," Labour Economics, 6 (March 1999(a)), pp. 53-74.

[2] Abowd, John M. and Francis Kramarz "The Analysis of Labor Markets Using Matched Employer-Employee Data," in O. Ashenfelter and D. Card (eds.) Handbook of Labor Economics, Volume 3(B), Chapter 26 (Amsterdam: North Holland, 1999(b)), pp. 2629-2710.

[3] Abowd,J., F. Kramarz, D. Margolis (1999) "High-Wage Workers and High-Wage Firms" Econometrica, 67, 2, 251-333.

[4] Abowd,R. Creecy, F. Kramarz (2002), «Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data, » Cornell University Working Paper

[5] Acemoglu, Daron, 1997. Training and innovation in imperfect labor markets. Review of Economic Studies.

[6] Alfaro, L. & Maggie Chen, 2010. "The Global Agglomeration of Multinational Firms," Harvard Business School Working Papers 10-043, Harvard Business School

[7] Arachi, Giampaolo and Federico Biagi (2005). "Taxation, Cost of Capital, and Investment: Do Tax Asymmetries Matter? Gionale degli Economisti e Annali di Economia 64 (2/3), pp. 295-322.

[8] Balsvik "Is labor mobility a channel for spillovers from multinationals? Evidence from Norwegian manufacturing" Review of Economics and Statistics 2011 93:1, 285-297

[9] Battisti, Michele "High Wage Workers and High Wage Peers" mimeo Simon Fraser University, 2012

[10] Becker, G., 1964. Human Capital. Columbia University Press, New York.

[11] Card, David & Devicienti, Francesco & Maida, Agata, 2011. "Rent-Sharing, Hold-up, and Wages: Evidence from Matched Panel Data," IZA Discussion Papers 6086, Institute for the Study of Labor (IZA)

[12] Cingano, F. & Fabiano Schivardi, 2004. "Identifying the Sources of Local Productivity Growth," Journal of the European Economic Association, MIT Press, vol. 2(4), pages 720-742, 06.

[13] Combes, Pierre-Philippe & Duranton, Gilles, 2006. "Labour pooling, labour poaching, and spatial clustering," Regional Science and Urban Economics, Elsevier, vol. 36(1), pages 1-28, January.

[14] Cooper, David P., \Innovation and Reciprocal Externalities: Information Transmission via Job Mobility," Journal of Economic Behavior and Organization, 2001, 45, 403{425.

[15] Dal Bo, Finan and Rossi "Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service" mimeo UC Berkeley, 2011

[16] Dasgupta, Kunal, \Learning and Knowledge Diffusion in a Global Economy," JPE

[17] Duranton, Giles, Puga, Diego, 2004. Micro-foundations of urban agglomeration economies. In: Henderson, J. Vernon, Thisse, Jacques-Francois (Eds.), Handbook of Urban and Regional Economics, vol. 4. North-Holland, Amsterdam.

[18] Eberhardt, M. and Helmer, C "Untested Assumptions and Data Slicing: A Critical Review of Firm-Level Production Function Estimators", , Oxford University, Department of Economics Discussion Paper No. 513

[19] Ellison, Glenn, Glaeser, Edward L., 1997. Geographic concentration in US manufacturing industries: a dartboard approach. Journal of Political Economy 105 (5), 889927.

[20] Ellison, Glenn, Glaeser, Edward L., 1999. The geographic concentration of industry: does natural advantage explain agglomeration? American Economic Review Papers and Proceedings 89 (2), 311316.

[21] Ellison, Glenn, Glaeser, Edward L., Kerr,William, 2010. What causes industry agglomeration? Evidence from coagglomeration patterns. American Economic Review

[22] Elston, Julie Ann and Laura Rondi (2006). "Shareholder Protection and the Cost of Capital: Empirical Evidence from German and Italian Firms." CERIS-CNR Working Paper No. 8. Montcalieri (Italy): CERIS-CNR.

[23] Fosfuri, Andrea; Motta, Massimo and Thomas Rønde (2001), 'Foreign Direct Investment Spillovers through Workers' Mobility', Journal of International Economics 53(1): 205-222.

[24] Foster, Lucia, John C. Haltiwanger, and Chad Syverson. 2008. "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" American Economic Review 98 (March): 394-425.

[25] Freedman, M. "Job Hopping, Earnings Dynamics, and Industrial Agglomeration in the Software Publishing Industry" July 2008 Journal of Urban Economics, Vol. 64, No. 3

[26] Glaeser, E.L. (1999). "Learning in cities". Journal of Urban Economics 46 (2). 254-277.

[27] Glaeser, Edward, Gottlieb, Joshua, 2009. The wealth of cities: agglomeration and the spatial equilibrium in the United States. Journal of Economic Literature.

[28] Glass, Amy J. and Kamal Saggi (2002), 'Multinational Firms and Technology Transfer', Scandinavian Journal of Economics 104(4): 495-513

[29] Görg, Holger, and Eric Strobl, "Spillovers from Foreign Firms through Worker Mobility: An Empirical Investigation," Scandinavian Journal of Economics 107 (2005), 693–709.

[30] Greenstone, Hornbeck and Moretti "Identifying Agglomeration Spillovers: Evidence from MDPs" Journal of Political Economy, 2010.

[31] Helsley, R.W. (1990). "Knowledge production in the CBD". Journal of Urban Economics 28 (3), 391-403.

[32] Henderson, J. Vernon, 2003. Marshall's scale economies. Journal of Urban Economics 55 (January), 1–28.

[33] Jovanovic, B., Nyarko, Y. (1995). "The transfer of human capital". Journal of Economic Dynamics and Control 19 (5-7), 1033-1064.

[34] Jovanovic, B., Rob, R. (1989). "The growth and diffusion of knowledge". Review of Economic Studies 56 (4), 569-582.

[35] Kaiser, U., H. C. Kongsted, and Rønde (2008). Labor mobility and patenting activity. CAM Working Papers 2008-07, University of Copenhagen. Department of Economics. Centre for Applied Microeconometrics.

[36] Kim, Sunwoong. "Labor Specialization and the Extent of the Market" The Journal of Political Economy, Vol. 97, No. 3 (Jun., 1989), pp. 692-705

[37] Kline, P. "Place Based Policies, Heterogeneity, and Agglomeration" American Economic Review: Papers and Proceedings 100 (May 2010), pp. 383-387

[38] Kline, P. and Moretti "Local Economic Development, Agglomeration Economies and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority" mimeo UC Berkeley, 2011

[39] Krugman, Paul. Geography and Trade. Cambridge, Mass.: MIT Press, 1991.

[40] Levinsohn, James, and Amil Petrin. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables." Review of Economic Studies 70 (April): 317-342.

[41] Levinsohn, J., Petrin, A. and Poi, B. P. (2003). "Production Function Estimation in Stata using Inputs to Control for Unobservables." Stata Journal 4(2): pp. 113-123.

[42] Levinsohn, James, and Amil Petrin "Measuring Aggregate Productivity Growth Using Plant-Level Data," forthcoming, RAND Journal of Economics.

[43] Maliranta, M., P. Mohnen, and P. Rouvinen (2009). Is inter-firm labor mobility a channel of knowledge spillovers? evidence from a linked employer–employee panel. Industrial and Corporate Change 18 (6),1161–1191.

[44] Markusen, James R., \Contracts, Intellectual Property Rights, and Multinational Investment in Developing Countries," Journal of International Economics, 2001, 53 (1), 189{204.

[45] Markusen, James R. and Natalia Tro menko, \Teaching Locals New Tricks: Foreign Experts as a Channel of Knowledge Transfers," Journal of Development Economics, 2009, 88, 120{131.

[46] Martins, P. S. (2011), Paying More To Hire The Best? Foreign Firms, Wages, And Worker Mobility. Economic Inquiry, 49: 349–363. Doi: 10.1111/J.1465-7295.2010.00301.

[47] Moretti "Local Labor Markets", in Handbook of Labor Economics, 2011.

[48] Olley, Steven G., and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." Econometrica 64 (November): 1263-98

[49] Ouazad A., Program for the Estimation of Two-Way Fixed Effects, available at http://personal.lse.ac.uk/ouazad/, 2007.

[50] Parrotta, Pierpaolo and Pozzoli, Dario "The effect of learning by hiring on productivity" The RAND Journal of Economics 43,1, 2012

[51] Pesola, H. "Foreign Ownership Labour Mobility and Wages" mimeo

[52] Poole, J. "Knowledge Transfers from Multinational to Domestic Firms: Evidence from Worker Mobility" forthcoming The Review of Economics and Statistics.

[53] Piore, M. and C. Sabel, "Italian Small Business Development: Lessons for U.S. Industrial Policy" in John Zysman and Laura Tyson, eds. American Industry in International Competition. Cornell University Press, 1984.

[54] Piore, M "Conceptualizing the Dynamics of Industrial Districts" in G. Becattini, M. Bellandi and L. De Propris (Eds.) The Handbook of Industrial Districts. Cheltenham, UK; Northampton, MA, USA: Edward Elgar, 2009.

[55] Rao, H. and R. Drazin (2002). Overcoming resource constraints on product innovation by recruiting talent from rivals: a study of the mutual fund industry, 1986-94. Academy of Management Journal 45 (3), 491–507

[56] Rosenthal, Stuart S., Strange, William C., 2004. Evidence on the nature and sources of agglomeration. In: Henderson, J. Vernon, Thisse, Jacques-Francois (Eds.), Handbook of

Urban and Regional Economics, vol. 4. North-Holland, Amsterdam.

[57] Rotemberg, Julio R., Saloner, Garth, 2000. Competition and human capital accumulation: a theory of interregional specialization and trade. Regional Science and Urban Economics 30 (4), 373404.

[58] Saxenian, A. (1994). Regional Advantage: Culture and Competition in Silicon Valley and Route 128. Harvard University, Cambridge, MA.

[59] Song, J., P. Almeida, and G. Wu (2003). Learning-by-hiring: When is mobility more likely to facilitate inter-firm knowledge transfer. Management Science 29 (4), 351–365.

[60] Stoyanov, A. and Zubanov, N., (2012) "Productivity Spillovers Across Firms through Worker Mobility", American Economic Journal: Applied Economics

[61] Syverson, Chad. 2004a. "Market Structure and Productivity: A Concrete Example." Journal of Political Economy 6 (December): 1181-1222.

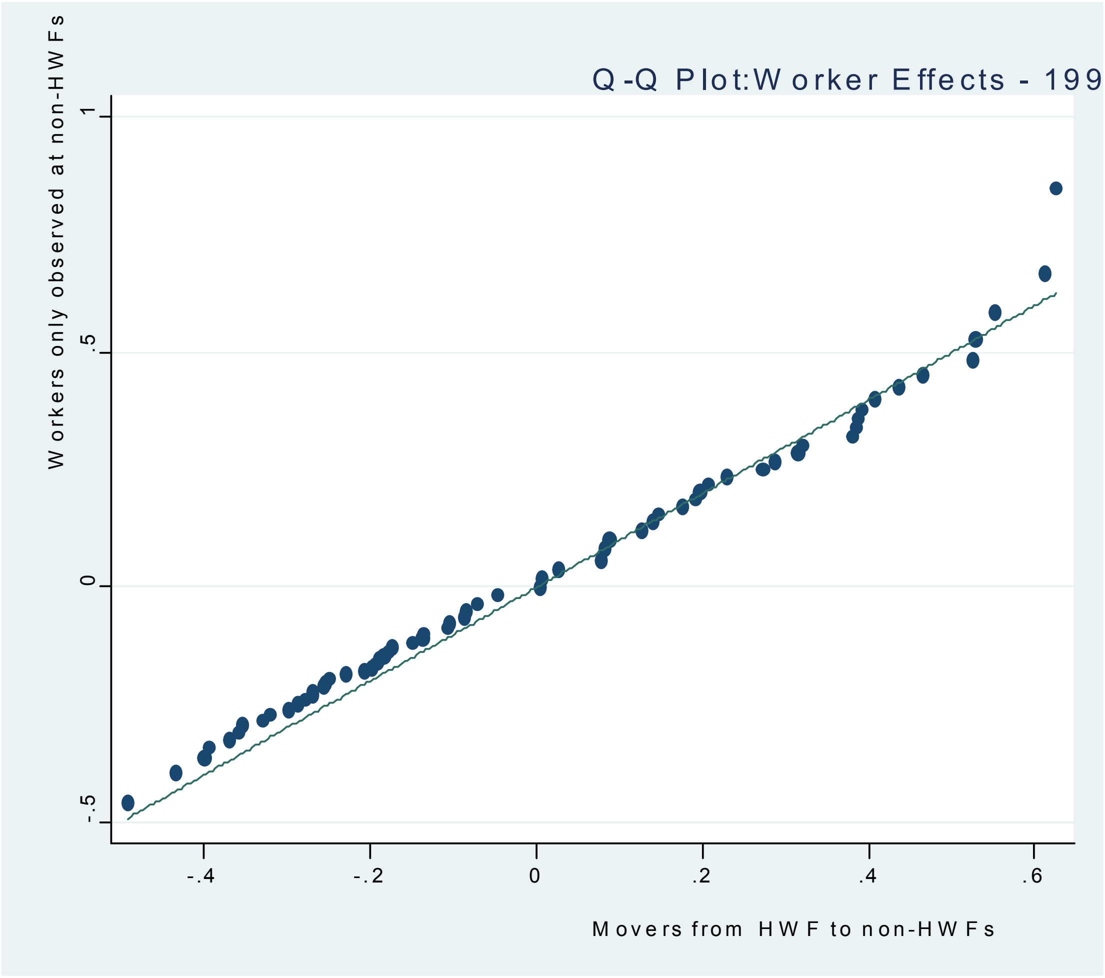
[62] Tattara, Giuseppe and Marco Valentini (2007). "The Cyclical Behavior of Job and Worker Flows." Working Paper No. 16. Department of Economics Ca' Foscari University of Venice.

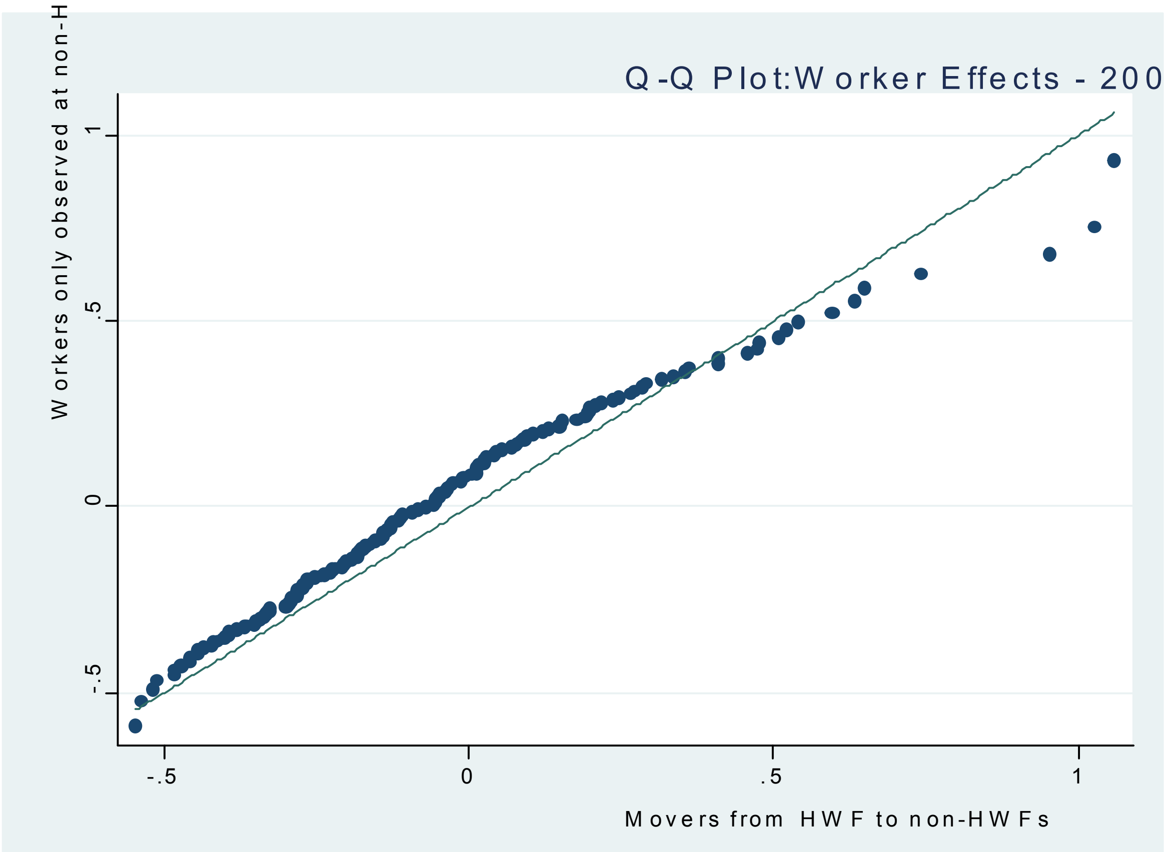
[63] Tattara, G. and M. Valentini (2010). Turnover and Excess Worker Reallocation: The Veneto Labour Market between 1982-1996. Labour 24 (4), 474 { 500.

[64] van Biesebroeck, Johannes. 2004. "Robustness of Productivity Estimates." Working Paper no. 10303 (February), NBER, Cambridge, MA.

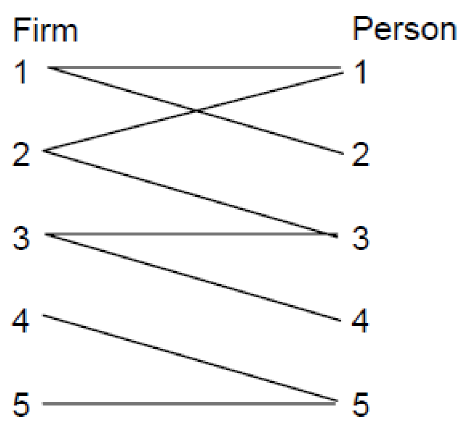
[65] Yasar, R. Raciborski, and B. Poi. 2008. Production function estimation in Stata using the Olley and Pakes method, 8(2), 221-231.

[66] Zucker, Lynne G., Darby, Michael R., Brewer, Marilynn B., 1998. Intellectual human capital and the birth of US biotechnology enterprises. American Economic Review 88 (March), 290306.


Q-Q Plot:Worker Effects - 1995
Workers only observed at non-HWFs
Movers from HWF to non-HWFs
-.5
0
.5
1
-.4
-.2
0
.2
.4
.6



Firm	Person	Group
1	1	1
1	2	1
2	1	1
2	3	1
3	3	1
3	4	1
4	5	2
5	5	2



The grouping algorithm. Source: Abowd, Creecy and Kramarz (2002)

Table 1: Firm Characteristics, 1995-2001

	(1)	(2)	(3)	(4)	(5)	(6)
	Y/L	VA/L	L	K/L	tang.K/L	intang.K/L
	b/se	b/se	b/se	b/se	b/se	b/se
HWF	0.173	0.136	0.083	0.106	0.065	0.313
	(0.021)	(0.014)	(0.034)	(0.030)	(0.032)	(0.050)
Observations	25137	25137	25137	25137	25137	25137

Dependent Variables are in Logs. All OLS regressions include year and industry dummies. See equation (2)

Output, Value Added and Capital variables are in 1000's of 2000 real euros

Standard errors (in parentheses) clustered by firm

The dummy HWF takes value 1 if the firm is classified as high-wage

Table 2: non-HWFs, 4388 Individual Firms

Variable	**Mean**	**(Std. Dev.)**	**Min.**	**Max.**	**N**
Output in 2001 (1000's of real euros)	8940.529	(10331.616)	1101.159	94712.109	3014
Capital in 2001 (1000's of real euros)	2106.03	(2904.353)	60.286	23070.875	3014
Materials in 2001 (1000's of real euros)	4443.621	(6046.294)	81.723	53299.207	3014
firm age (years) in 2001	19.574	(10.96)	1	100	3014
workers from HWF	0.038	(0.211)	0	4	21330
workers from non-HWF	3.574	(6.583)	0	177	21330
employees	51.141	(51.499)	3	455	21330

Table 3: Workers with HWF experience and Productivity, OLS Estimates

	(1) Baseline b/se	(2) Prov.-Year FE b/se	(3) Prov.-Ind.-Year FE b/se	(4) non-HWF workers b/se
log(capital)	0.096	0.096	0.098	0.094
	(0.004)	(0.004)	(0.004)	(0.004)
log(materials)	0.576	0.576	0.573	0.573
	(0.007)	(0.007)	(0.007)	(0.007)
log(employees)	0.237	0.238	0.239	0.231
	(0.007)	(0.007)	(0.007)	(0.007)
workers from HWF	0.044	0.043	0.040	0.036
	(0.010)	(0.010)	(0.010)	(0.010)
workers from non-HWF				0.003
				(0.001)
Observations	21539	21539	21539	21539

Dependent variable: Log(Output). Standard errors (in parentheses) clustered by firm
All regressions include year dummies
Column 1 reports estimates from the baseline specification, see equation ()
Column 2 includes province(7)-year interaction dummies
Column 3 includes province(7)-industry-year interaction dummies
Column 4 includes the number of newly hired workers from non-HWFs

Table 4: Workers with HWF experience and Productivity: Robustness to Specifications Adjusting for the Endogeneity of Firm Inputs

	(1) OLS b/se	(2) OP b/se	(3) LP b/se	(4) Within b/se
main				
log(capital)	0.096	0.084	0.151	0.064
	(0.004)	(0.017)	(0.009)	(0.005)
log(materials)	0.576	0.578		0.597
	(0.007)	(0.008)		(0.012)
log(employees)	0.237	0.232	0.207	0.063
	(0.007)	(0.009)	(0.005)	(0.004)
workers from HWF	0.044	0.036	0.023	0.016
	(0.010)	(0.015)	(0.009)	(0.006)
Observations	21539	8365	21539	21539

Dependent variable: Log(Output)

All regressions include year dummies

Column 1 reports estimates from OLS specification. See equation ()

Column 2 implements the procedure in Olley and Pakes (1996)

Column 3 implements the procedure in Levinsohn and Petrin (2003)

Column 4 reports within estimates

Table 5: Local Variables

Variable	**Mean**	**(Std. Dev.)**	**Min.**	**Max.**	**N**
Firms in LLM	207.073	(120.398)	8	479	34859
HWFs in LLM	35.741	(25.35)	0	109	34859

Table 6: Hiring from HWFs, Probit Regressions

	(1) Baseline b/se	(2) Prov.-Year-Ind. FE b/se	(3) by Ind. b/se	(4) Prov.-Year-Ind. FE b/se
lag5(HWFs in LLM)	0.006	0.006		
	(0.002)	(0.002)		
lag5(HWFs in LLM same IND)			0.042	0.042
			(0.010)	(0.010)
lag5(HWFs in LLM diff IND)			-0.001	-0.001
			(0.002)	(0.002)
Observations	34859	33116	34859	33116
$\Pr(\overline{x}'\widehat{\beta})$	.014	.015	.014	.015

Marginal effects

(d) for discrete change of dummy variable from 0 to 1

Dep. Variable is dummy equal to one if non-HWF hires from a HWF

Column 1 reports estimates from the baseline specification - see equation ()

Column 2 includes province(7)-industry(92)-year interaction dummies

Column 3 divides local HWFs depending on whether they belong to same industry of the non-HWF

Column 4 adds province(7)-industry(92)-year interaction dummies to Column 3

Standard errors (in parentheses) clustered by firm. Regressions include year dummies

Estimated Coefficients are semi-elasticities

Table 7: Number of HWFs in LLM by Industry and Workers with HWF experience

	(1)	(2)	(3)	(4)
	Baseline	Prov.-Year FE	Prov.-Year-Ind. FE	LLM FE
	b/se	b/se	b/se	b/se
lag5(HWFs in LLM same IND)	0.0041	0.0037	0.0041	0.0029
	(0.0009)	(0.0009)	(0.0009)	(0.0010)
lag5(HWFs in LLM diff IND)	-0.0003	-0.0002	-0.0003	0.0006
	(0.0001)	(0.0001)	(0.0001)	(0.0005)
Observations	34859	34859	34859	34859

Dependent variable: Workers with HWF experience
All regressions include year dummies
Column 1 reports estimates from baseline specification
Column 2 includes province(7)-year interaction dummies
Column 3 includes province(7)-industry-year interaction dummies
Column 4 includes LLM effects

Table 8: Workers with HWF experience from same industries and Productivity

	(1)	(2)
	OLS	LP
	b/se	b/se
log(capital)	0.089	0.151
	(0.004)	(0.009)
log(materials)	0.576	
	(0.007)	
log(employees)	0.262	0.230
	(0.008)	(0.006)
workers from HWF in same ind.	0.055	0.039
	(0.018)	(0.015)
Observations	21330	21330

Dependent variable: Log(Output)
All regressions include year dummies
Column 1 reports estimates from OLS specification. See equation ()
Column 2 implements the procedure in Levinsohn and Petrin (2003)

Table A.1: Movers from HWF to non-HWFs, 2001

Variable	Mean	(Std. Dev.)	Min.	Max.	N
age	34.503	(8.153)	18	62	372
white collar	0.395	(0.49)	0	1	372
manager	0.03	(0.17)	0	1	372
female	0.237	(0.426)	0	1	372

Table A.2: Stayers at non-HWFs hiring from non-HWFs, $2 > 001$

Variable	Mean	(Std. Dev.)	Min.	Max.	N
age	37.539	(9.602)	16	65	58654
white collar	0.272	(0.445)	0	1	58614
manager	0.014	(0.119)	0	1	58614
female	0.317	(0.465)	0	1	58654

Table A.3: Local Variables

Variable	Mean	(Std. Dev.)	Min.	Max.	N
local_HWFs	38.028	(26.781)	0	116	50918
share of HWFs in LLM	0.175	(0.075)	0	0.6	50918

Table A.4: Hiring from HWFs, Probit Regressions

	(1) Baseline b/se	(2) Prov.-Year FE b/se	(3) Prov.-Year-Ind. FE b/se
lag2 (share of HWFs in LLM)	2.154	1.706	1.799
	(0.412)	(0.477)	(0.474)
Observations	16425	16425	16425
$\Pr(\overline{x}'\widehat{\beta})$	.033	.032	.029

Dep. Variable is dummy equal to one if non-HWF hires from HWF
Column 1 reports estimates from the baseline specification - see equation (15)
Column 2 includes province(7)-year interaction dummies
Column 3 includes province(7)-industry(92)-year interaction dummies
Standard errors (in parentheses) clustered by firm. Regressions include year dummies and firm level controls
Estimated Coefficients are semi-elasticities

Table A.5: Number of HWFs in LLM and Workers with HWF experience

	(1) Baseline b/se	(2) Prov.-Year FE b/se	(3) Prov.-Year-Ind. FE b/se
lag(HWFs in LLM)	0.0012	0.0011	0.0011
	(0.0003)	(0.0004)	(0.0003)
Observations	44355	44355	44355

Dependent variable: Workers with HWF experience
All regressions include year dummies and LLM effects
Column 1 reports estimates from baseline specification
Column 2 includes province(7)-year interaction dummies
Column 3 includes province(7)-industry-year interaction dummies

Table A.6: Workers with HWF experience and Productivity, OLS Estimates using lagged values for number of workers

	(1) Baseline b/se	(2) Prov.-Year FE b/se	(3) Prov.-Ind.-Year FE b/se	(4) L from non-HV b/se
log(capital)	0.090	0.090	0.093	0.088
	(0.005)	(0.005)	(0.005)	(0.005)
[1em] log(materials)	0.570	0.571	0.569	0.569
	(0.007)	(0.007)	(0.007)	(0.007)
[1em] lag(workers from HWFs)	0.051	0.051	0.047	0.047
	(0.012)	(0.012)	(0.013)	(0.012)
[1em] log(employees)	0.264	0.264	0.264	0.254
	(0.008)	(0.008)	(0.008)	(0.008)
[1em] lag(workers from non-HWFs)				0.004
				(0.001)
Observations	19560	19560	19560	19560

Dependent variable: Log(Output). Standard errors (in parentheses) clustered by firm

All regressions include year dummies

Column 1 reports estimates from the baseline specification, see equation ()

Column 2 includes province(7)-year interaction dummies

Column 3 includes province(7)-industry-year interaction dummies

Column 4 includes the number of newly hired workers from non-HWFs